

## Create a Comparative Analysis of an Oracle Database Using NAS and SAN Storage Architectures

Sergio Hirata and Volnys Borges Bernal

EMC Proven Profesional Knowledge Sharing 2009



Sergio Hirata  
Columbia Storage

Volnys Borges Bernal  
Universidade de São Paulo/LSITec

## Table of Contents

1	Foreword.....	3
1.1	Motivation .....	5
1.2	Objective .....	6
1.3	Scope .....	7
1.4	Contribution .....	7
2	Related works .....	7
2.1	Fibre channel and iSCSI comparison .....	7
2.2	Database performance analysis.....	9
3	Application description .....	10
4	Storage Architecture .....	12
4.1	DAS (Direct Attached Storage) .....	12
4.2	Storage Area Network (SAN) .....	13
4.3	Network Attached Storage (NAS).....	14
5	Test bed planning and execution .....	15
5.1	Test bed preparation .....	15
5.2	Scenarios .....	15
5.2.1	DBMS configuration .....	15
5.2.2	Storage architecture configuration .....	16
5.3	Storage system characterization .....	17
5.4	Architecture comparative analysis.....	18
5.5	Ease of Operation Analysis .....	19
5.6	Cost Analysis.....	20
5.7	Performance Analysis .....	21
5.7.1	Test bed planning .....	21
5.8	Performance results analysis .....	23
6	Conclusion .....	24
6.1	Limitation .....	27
6.2	Suggestion for future works.....	27

*Disclaimer: The views, processes or methodologies published in this compilation are those of the authors. They do not necessarily reflect EMC Corporation's views, processes, or methodologies*

## List of illustrations

Figure 1 — Response time. (Ding, 2005) .....	5
Figure 2 — Comparative analysis architecture .....	6
Figure 3 — DAS environment.....	12
Figure 4 — SAN environment.....	13
Figure 5 — NAS environment.....	14
Figure 6 — SAN test bed.....	16
Figure 7 — iSCSI and NAS test bed environment .....	17
Figure 8 — Fibre Channel HBA installation program interaction.....	32
Figure 9 — sdb, the new disk .....	33

## List of tables

Table 1 — Query types defined by BORAL and DEWITT, (1984).....	10
Table 2 — Amazon.com visits (www.compete.com).....	11
Table 3 — Architecture maximum I/O capacity.....	18
Table 4 — Time spent to create new area in the architecture.....	20
Table 5 — Preliminary tests results .....	22
Table 6 — Collect data .....	23
Table 7 — Peak number of threads on each architecture.....	24
Table 8 — Relating the threads and monthly visits.....	25
Table 9 — Final results.....	26

## List of Abbreviations and Symbols

ANSI — *American National Standard Institute*  
Array — Physical disks group  
BI — *Business intelligence*  
DAS — *Direct Attached Storage*  
DBMS — *Database Management System*  
MFAS — *Fabric Attached Storage*  
FC-AL — *Fibre Channel Arbitrated Loop*  
FCIA — *Fibre Channel Industry Association*  
FC-SW — *Fibre Channel Switched*  
HR — *Human Resource*  
IOPS — *Input Output per second*  
lqn — Identifier queue number  
iSCSI — *Internet Small Computer System Interface*  
IX — *Information Transport*  
MBps — *MegaBytes per second*  
NAS — *Network Attached Storage*  
OE — *Order Entry*  
OLTP — *Online Transaction Processing*  
PL/SQL — *Procedural Language/Structured Query Language*  
PM — *Product Media*  
RAID — *Redundant Array of Inexpensive Disk, later Redundant Array of Independent Disk*  
RDBMS — *Relational database management system*  
RPM — *Rotations per minute*  
SAN — *Storage Area Network*  
SCSI — *Small Computer System Interface*  
SNIA — *Storage Network Industry Association*  
wwn — *WorldWide Name — fibre channel identification*

## **Abstract**

The database management system (DBMS) is a basic component for most computational systems and, in particular, the order entry application. The order entry application's performance is directly related to DBMS performance.

DBMS performance is affected by the storage architecture, storage communication data protocol, and existing components in each network. This is reflected in application response time. Today's storage architecture is divided into three groups: DAS (*Direct Attached Storage*), NAS (*Network Attached Storage*) and SAN (*Storage Area Network*). Each of these architectures has components that influence latency, performance and implementation cost.

The storage administrator sometimes has difficulty aligning application needs with the appropriate storage architecture. This drove me to compare DBMS behavior in major storage architectures.

This article uses an order entry Simulator to create I/O operations in a DBMS installed in SAN iSCSI, SAN fibre channel and NAS NFS architectures. It collects performance information such as response time and transactions per minute.

The results of this work illustrate the conditions that each storage architecture should use and the expected response time.

Keywords: performance, DBMS, fibre channel, iSCSI, SAN and NAS.

## 1 Foreword

Most DBMS-based applications need high availability to assure 24 x 7 data access. Application availability depends on the operating system, server hardware, DBMS software, network connectivity and storage systems (Microsoft, 2008).

Microsoft shows the cluster environment to leverage availability in the case of Server hardware, operational system and DBMS software failure. The cluster environment requires two or more servers with access to the same storage area, forcing the use of centralized and external storage (ORACLE, 2007b), (WRIGHT, 2008), (Microsoft, 2008). Network connectivity is achieved using the equipment with hot swap and hot plug fans and power supply. In some cases, companies need two or more pieces of equipment to create redundancy. Most of external storage implementation uses a storage system with protection against disk failure, RAID, and they are directly connected to the server or use storage architecture such as SAN fibre channel, SAN iSCSI or NAS.

The application environment becomes complex since each component has different latencies and the sum of latencies directly influences application response time. It is very important to choose the most appropriate architecture and technology to achieve the application's performance requirements. To do that, it's necessary to characterize the application by defining the access profile, volume of I/O operations, and expected response time.

It is now possible to implement each storage architecture using different technologies, among them:

- DAS: technologies SCSI (*Small Computer System Interface*) (INCITS, 1986), ATA (*Advanced Technology Attachment*) (ANSI, 1990), SATA (*Serial Advanced Technology Attachment*) (INCITS, 2003) and SAS (*Serial Attached Small Computer System Interface*) (INCITS, 1986)
- NAS: technologies SMB (*Server Message Block*) (Feigenbaum, 1980), NFS (*Network File System*) (Sun Microsystems, 1995) and Lustre (Braam, 1999).
- SAN: technologies *fibre channel* (ANSI T11,1998) and iSCSI (*internet Small Computer System Interface*) (IETF, 2004)

DAS architecture is characterized by a direct connection between the Server and storage device, such as an internal disk.

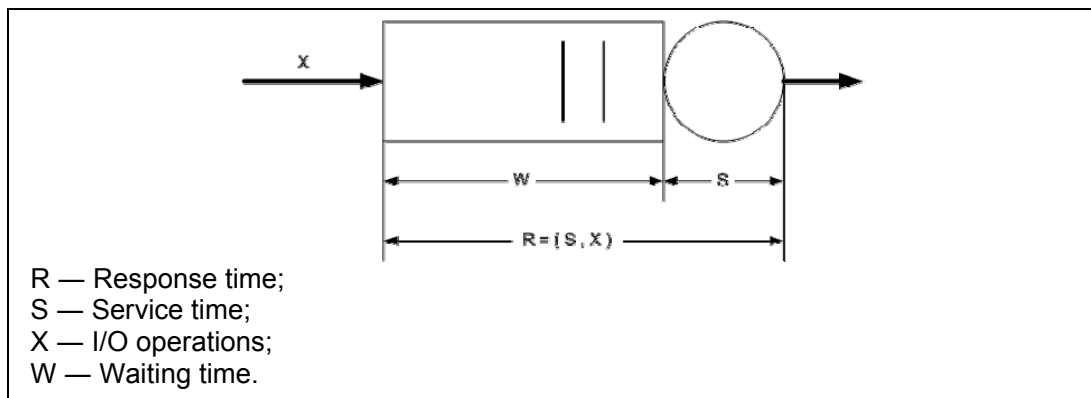
Most NAS implementations are based on a map to a server directory, a share in a Windows environment, or export to another server in the UNIX / Linux world. Some companies also install NAS appliances in their environment.

The SAN fibre channel has a dedicated network to transmit data blocks from the server to the storage system. It is the fastest and most reliable choice for high availability applications. However, the fibre channel network has a high implementation cost. A fibre channel Host Bus Adapter (HBA) installed in a server to communicate with the storage system is twenty time more expensive than an Ethernet card. A fibre channel switch is ten times more expensive than an intelligent gigabit Ethernet switch.

YAMAGUCHI (2005) is worried about the number of professionals specialized in fibre channel design and performance analysis, while the Ethernet network has many professionals. The declining number of fibre channel Professionals implies more expensive storage management due to the need to contract for the design, implementation and maintenance of storage architectures.

When the IETF released the RFC 3720 (SATRAN, SAPUNTZAKIS, CHADALAPAKA, 2004) iSCSI protocol, in which SCSI commands are transported using the Ethernet network, it became becoming an alternative to implementing a SAN architecture. Known network problems such as collisions, retransmissions and the chatty TCP/IP protocol revealed doubts about the performance and reliability of this new standard. Today, some developers already homologated their applications using iSCSI storage solutions for environments with up to 4,000 users (EMC, 2006) (Microsoft, 2004).

The DBMS can use DAS, SAN fibre channel, SAN iSCSI and NAS, so we must choose the storage architecture by considering three aspects: data throughput, Volume of I/O operations, and response time. Ding (2005) studied the workload influence in the data network and concluded that the response time is based in the physical latency and the network wait time. Figure 1 shows the workload, response time and latency relation that Ding created. The latency (service time) depends on the transmitted block size, distance between the server and storage systems, RAID level, and physical disk used.



**Figure 1 — Response time. (Ding, 2005)**

The application designer must identify a target response time; this is the premise to define the storage network.

There are some, like (Rodrigues Neto, 2004) and (Malakapalli, 2001), who used a throughput measure tool to compare fibre channel and iSCSI implementations. They concluded that the throughput of these technologies was equivalent and suggested using application workload simulators to search different parameters to define the storage infrastructure.

### 1.1 Motivation

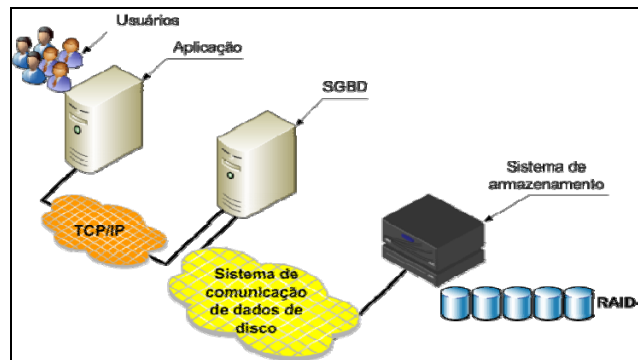
Consumers have a choice where to purchase product. Order entry management is a critical application with high availability requirements. It receives input from the purchase chain from the consumer to the deliverer, using different devices like PDAs, cell phones and computers.

The order entry applications manage customer registration, stock control, order, payment and product delivery. DBMS structures the data and eases future searches. After identifying the application need, I/O volume, response time and budget for the project, you also need to consider DBMS performance behavior on NAS, SAN fibre channel, and SAN iSCSI to compare the application parameters and define the optimal storage architecture.



## 1.2 Objective

The objective of this article is to perform a comparative analysis of performance, cost and operational ease of a storage system used by a DBMS, installed in different storage architectures, and responding to I/O requests from an order entry application.



**Figure 2 – Comparative analysis architecture**

We will assess response time, and I/O volume per second using 3 application categories of workload performance (low, medium and high).

The DAS architecture will be characterized using the server internal disk. We will use the same storage system for the NAS with NFS, SAN iSCSI and SAN fiber channel architecture evaluations. This will reduce the influence of its internal components, such as processors, buffers, memory, RAID level and physical disk.

The work load simulator will represent the I/O operations generated by an order entry application to the DBMS. The simulator charts the transactions and response time.

The analysis scenario assumes that the storage system is dedicated exclusively to the application DBMS. The architecture used in this work is represented in figure 2.

The end of this article demonstrates that the order entry application may be used on studied architectures without performance degradation. This allows us to decide on the architecture based on ease of operation and cost.

### **1.3 Scope**

This work uses a benchmark tool to generate I/O requests to the DBMS installed in a storage architecture configured in a SAN fibre channel, and execute the experiment with the order entry module. The application and DBMS communication use a gigabit Ethernet link via a cross-over cable, eliminating TCP/IP traffic influence. The DBMS server is connected to the storage system using a single port. Later, the same experiment will be repeated using the storage network based in SAN iSCSI and NAS, creating conditions to compare order entry performance using a DBMS.

The fibre channel port rate is 4Gbps, the Ethernet port rate is 1Gbps. The intent is to verify if throughput is limited by the communication port speed, fibre channel and gigabit Ethernet, and between the DBMS server and storage system.

This work does not intend to configure the storage system in a different RAID level, use the NAS architecture with CIFS protocol, servers in cluster environment, or multiple paths between the server and storage system.

### **1.4 Contribution**

This study continues the work of Rodrigues Neto (2004) and Malakapalli (2001). It strengthens the storage architecture analysis for DBMS by mounting a DBMS instance in architectures, SAN fiber channel, SAN iSCSI, and using a simulator to generate data traffic between the server for the application, the DBMS server and the storage system. This verifies the response time obtained with each of the architectures.

## **2 Related works**

### **2.1 Fibre channel and iSCSI comparison**

Rodrigues Neto (2004) described the fibre channel interface. At that time, the speed was 1Gbps and the iSCSI protocol showed the overhead and how the HBAs work and interact with the server processor. The fibre channel interface has the lowest server processor dependency due to the fibre channel protocol. SCSI commands are processed at the HBA level, the server processor only needs to address the data to the application. The

NAS architecture and the NFS are usually server processor dependent, TCP/IP is 100% processed at the server processor. The iSCSI protocol may use the standard Ethernet port built in the server chassis, a TCP Offload Engine (TOE) or an iSCSI HBA. Figure 3 illustrates server processor responsibility using each Ethernet card option. The TOE card relieves the entire TCP/IP processing from the server processor, while the iSCSI HBA only relieves the iSCSI stack.

Rodrigues Neto compared the fibre channel and Ethernet interfaces' bandwidth using the IOMETER. This tool measures I/O in the storage system for servers in a single or cluster configuration, gathering the throughput and server processor utilization. Neto concluded that iSCSI throughput using the standard Ethernet port is very close to fibre channel 1Gbps, although fibre channel has a slight advantage. Adding the server processor as an object of comparison, fibre channel uses less processor.

When Neto used an iSCSI HBA, the server processor consumption decreased but throughput was reduced. One explanation is that the HBA processor has a clock 50% less than the server processor. The throughput, when comparing iSCSI using HBA and the fibre channel, is closer than iSCSI using a standard port.

Neto compared the technologies via a direct attached connection, changing the interfaces from the fibre channel HBA and Ethernet standard port. Neto's work illustrates the need to test the behavior of the system to verify if the storage system or server limit was achieved, or if it is possible to increase the workload.

Malakapalli (2001) used two servers to create an experiment. Using the iSCSI target, the iSCSI initiator was taken from the operational system source. The workload was generated using a known benchmark tool, the TTCP. An alteration was accomplished for the TTCP to generate the TCP load at the kernel level instead of user level, creating the KTTCP (Kernel-ported Test TCP). The iSCSI infrastructure achieved the 180Mbps throughput with sequential access. A fibre channel target was developed to connect both servers creating the fibre channel environment. This work concluded that there is a need to consider different parameters when comparing beyond the bandwidth. These may include reliability, traffic scalability, availability, implementation cost and total cost of ownership in the local and remote network.

Malakapalli and Neto compared fibre channel and iSCSI with the same speed rate and concluded the equivalency of the interfaces in that parameter. They signaled that implementation cost and availability are points to be explored in the future. The present work remakes the comparison with new interfaces and expands it by using an order entry application to compare the behavior with parameters like response time, operational ease and cost.

## **2.2 Database performance analysis**

Mendes, in his 2006 study, approached the performance analysis by relating the DBMS and SQL transaction processing to ensure database consistency in case of disaster. The author mentioned three techniques to evaluate system performance: analytical modeling, measurement and simulation. The evaluation metrics included service and response time that we use in this work. Mendes suggested a methodology to analyze environment performance, define the study object, identify basic components, characterize the workload, monitor the system, and collect the data. He used the TPC-C (TPC, 2007a) created by the Transaction Processing Performance Council.

This benchmark consists of transactions that simulate the activity found in an OnLine Transaction Processing (OLTP). Mendes created a mathematical model to estimate environment performance and experimented with two environments, each with different memory sizes and number of server processors. The components' behavior was close to the calculated values, validating the mathematical model.

Boral and DeWitt (1984) reinforced that their study didn't intend to create a metric to compare the performance of different systems, due to various factors that influence the application. They were concerned about the important parameters to compare systems, such as response time, transfer rate and costs. Their study points to three major factors that affect the performance in a multi-user environment: the multi-programming level, operation volume at the same data block, and the transaction mix. This requires studying four basic scenarios, which differ by processor and storage utilization, to build a reference point and evaluate the performance from a great variety of workloads.

Benchmark data entry is a synthetic database and a user-generated routine group. One advantage of using a synthetic database is control over I/O operations during the execution of an experiment. The study had two phases: first, a group of routines are executed in single user mode to get measures in a condition where there is no competition. During the second phase, a multi-user environment is used to obtain 'real' data.

Boral and DeWitt divided the queries in groups based on processor utilization and I/O operations volume at hard disk:

Type	Processor utilization	Disk Utilization
1	Low	Low
2	Low	High
3	High	Low
4	High	High

**Table 1 – Query types defined by BORAL and DEWITT, (1984)**

The study affirms that processor and disk utilization represent a multi-user environment, but may not be used as a base condition in a single user environment.

According to BORAL and DEWITT's conclusions, query types 1 and 2 are typical on any database system. However, types 3 and 4 depend on the algorithm used by the DBMS.

Boral and DeWitt intend to use the response time as a performance indicator, but their routines are not prepared to calculate the response time. Instead, they tried to get a timestamp at the start and stop of the routine. They do not trust this data, so they used the request per second as a performance indicator.

In this study, the conclusion was that the multi-user environment represents the real world more closely. However, it is very difficult to reproduce and control the I/O operations; a single user environment is very important to solve punctual performance problems.

### **3 Application description**

This work shall use a benchmark tool called the Swingbench (Giles, 2005) to generate the I/O operations at the DBMS server.

Swingbench is designed to stress test a database by simulating a workload using user-defined transactions. Swingbench requires an Oracle database version 9 or above. The user can control the number of users (threads) that attach to a database and the amount and type of work they perform. Users can dynamically monitor the response times and load which is displayed in a series of graphs.

This study will use the Swingbench order entry module to simulate the classic database order entry. It has a similar profile to the TPC-C benchmark. This version models an online order entry system with users required to log-on before purchasing goods. The benchmark uses three files names: .txt, nls.txt and productids.txt. These contain sample data used for each of the transactions.

The order entry operations are distributed as follows: 50% select, 30% insert, and 20% update. According to the Oracle manual (1999), each select operation requires 3 read operations, insert require 5 operations (3 read and 2 write) and update requires 7 operations (5 reads and 2 writes), driving us to conclude that this module will create data traffic with 70% of read I/O operations and 30% write I/O operations. Swingbench permits the creation of a DBMS with a different size, but the version 2.3.0.391, compatible with Oracle 11g, has a bug and it got an error when we created tables bigger than 2GB. Subsequently, the database used in this work will be generated with size of 2GB.

The website Alexa ([www.alexa.com](http://www.alexa.com)) rates Amazon.com as #1 in the shopping category, positioned as #33 in the most visited in the global ambit. The Alexa website does not show the number of visits to the Amazon.com website. The compete.com ([www.compete.com](http://www.compete.com)) site shows the average number of visits at Amazon.com during the last 12 months, see table 2.

Date	Visits
Jul/08	193.193.844
Jun/08	186.560.804
May/08	186.154.134
Apr/08	174.509.181
Mar/08	172.645.361
Feb/08	169.439.389
Jan/08	177.030.556
Dec/07	208.380.460
Nov/07	176.874.778
Oct/07	144.665.128
Sep/07	141.806.797
Aug/07	138.367.820
Jul/07	132.820.582

**Table 2 – Amazon.com visits (www.compete.com).**

We estimated the number of visits per second with formula (1), considering the month with higher visits from the table 2 (December 2007). We derived the value of 80,39 visits per second on average. Applying the Pareto principle, considering that 80% of the traffic occurred in 20% of the time (formula 2), we derived a peak value of 321 visits per second. This study simulates 3 application conditions: low, medium and high. The low workload application will use 3 threads, medium will use 30 threads, and high will use 300 threads.

$$\text{number of visits} = \frac{\text{Visits per month}}{\text{time in seconds}} \quad (1)$$

$$\text{Peak of visits} = \frac{\text{Visits per month} \times 80\%}{\text{time in seconds} \times 20\%} \quad (2)$$

## 4 Storage Architecture

The Server to storage system connection influences global application performance. Sharing the storage system with different server and applications profiles makes the applications more vulnerable to performance problems. Any application with higher access volume will impact other applications. The most popular storage architectures are:

### 4.1 DAS (Direct Attached Storage)

A DAS environment is a direct connection between the Server and the storage device. This scenario fits with most of existing servers, desktop and notebooks, even the server uses an iSCSI or fibre channel port to connect the server and the storage device is DAS architecture.

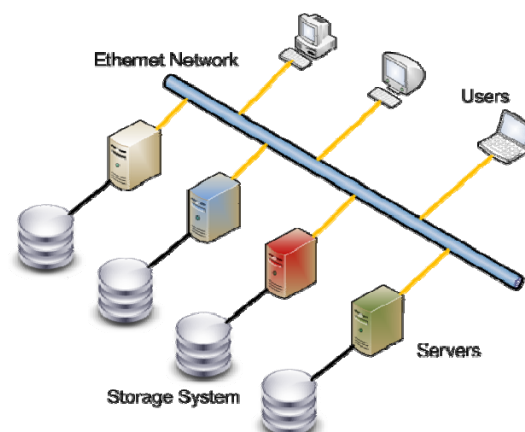


Figure 3 — DAS environment

## 4.2 Storage Area Network (SAN)

A dedicated network is created to traffic data blocks between the Server and the storage subsystem. This includes fibre channel switches or directors to expand storage system connectivity. Today, a SAN may use fibre channel or iSCSI technologies.

The IETF created the iSCSI standard in 2004, offering the market an alternative to create less expensive SANs. Some studies compared iSCSI and fibre channel using a benchmark tool, concluding that the throughput of fibre channel 1Gbps and iSCSI using a gigabit Ethernet are very close. Today, the fibre channel transfer rate is 4Gbps and the iSCSI still in 1Gbps, the market even has some 10Gbps implementations.

Strand (2003) explored the usage of fibre channel and iSCSI in a Windows Server 2003 environment, using the IOMETER to generate = data traffic. The conclusion was that the server standard NIC card can be used to implement an iSCSI network with a minimal performance penalty for certain types of applications having general workloads such as a moderately loaded file server or workstation. However, if the requirements for bandwidth and large number of I/O operations per second are crucial, Fibre Channel is currently the only option for implementing a Storage Area Network. The Fibre Channel HBA was the clear winner in both small and large transfer size tests.

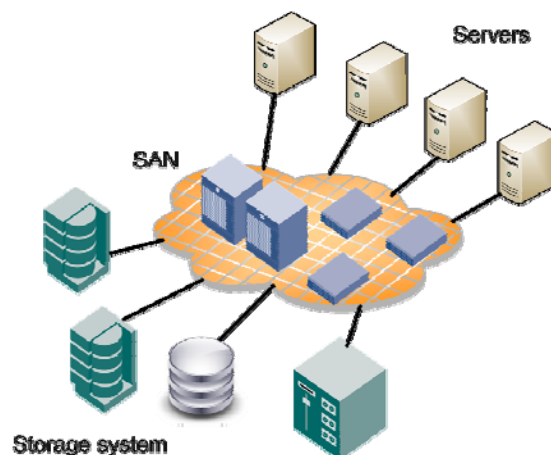


Figure 4 — SAN environment

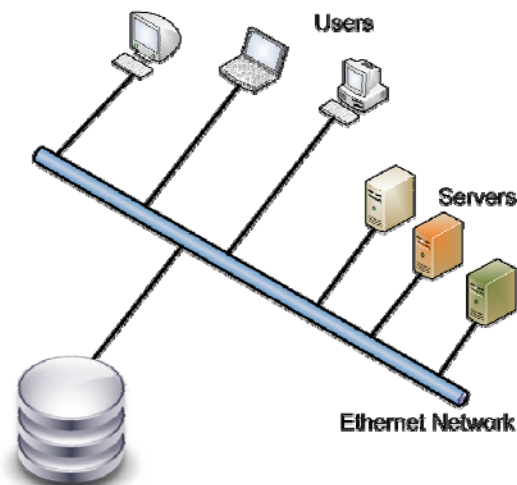


TCP/IP protocol processing yielded a performance penalty. CPU utilization rose quickly for the Fast Ethernet NIC and the Gigabit Ethernet as the load increased. The Alacritech Gigabit Ethernet card (TOE card) solved this problem by offloading TCP handling to hardware on the card, leading to better CPU utilization and better performance than the other two Ethernet cards.

(Senapathi, 2003) approaches Ethernet network speed evolution such as Gigabit Ethernet and 10Gigabit Ethernet. The host processor can become a bottleneck due to increased TCP/IP overhead. Using TCP/IP, a server write to a file in another server produces a series of interrupts to fit the data to the configured packet size and handle the acknowledgements. Every data transfer generates data copies from application buffers to system buffers, and then to network adapters. For 1500-byte packets, the host OS stack would process more than 83,000 packets per second. TOE processing dramatically reduces the network transactional load; TOEs and the host CPU can process an entire application I/O transaction with one interrupt.

#### **4.3 Network Attached Storage (NAS)**

The NAS environment is based in a server writing files on storage over an Ethernet. Users have permission to read and write files on NAS storage. Using a NAS appliance, file system control is transferred from the server to the appliance, enabling new functionality like snapshots, backup to tape drives, and file sharing over different operating systems like Windows and Linux. NAS architecture is widely used for file server consolidation. It provides better file management, simplified backup processes, and control over write file permission.



## 5 Test bed planning and execution

### 5.1 Test bed preparation

The test bed will consist of:

- An application server
- A DBMS Server
- Data network between the servers
- Storage Network: fibre channel or TCP/IP (iSCSI and NFS)
- Storage System

### 5.2 Scenarios

This work will use three scenarios. First, we will connect the storage using a fibre channel network. In the second scenario, we will connect the same storage in an Ethernet network for iSCSI connection, and in the third we will mount the storage as a NAS appliance.

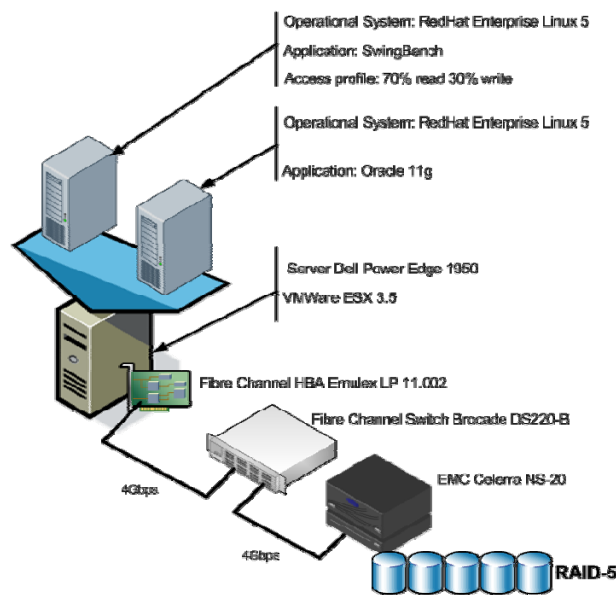
All scenarios use the same servers, Dell 1950 with 4 processors Xeon 1.66 GHz, 8GB memory. The storage system is EMC Celerra® NS-20. The server has VMware ESX version 3.5, and the application and DBMS servers will be virtual machines. The virtual machine configuration will have 2 processors, 2GB memory and 5GB disk. Red Hat Enterprise Linux version 5 is the virtual servers' operating system and Oracle version 11g is the database since they are widely deployed in corporate and education markets.

#### 5.2.1 DBMS configuration

The Oracle installation accepted the default file locations with database, logs and archive in the same folder structure. The preliminary tests showed that the default Oracle parameters do not accept the workload generated by the simulator with 300 threads, so the parameters shown in table X were changed to adjust the database to the workload.

Parameter	Initial value	Final value	Location
Processes	150	2048	Spfile
Sessions	165	5000	init.ora
shared_pool_size	52428800	3000M	
db_block_buffers	200	400000	

Changing these parameters makes the Simulator run without errors with 300 threads.



**Figure 6 — SAN test bed**

## **5.2.2 Storage architecture configuration**

### **5.2.2.1 Scenario SAN**

The ESX Server has an Emulex fibre channel HBALP11,000 installed. This HBA has a transfer rate of 4Gbps, connected to a Brocade Switch DS220-B, then connected to Celerra NS-20, showed in figure 7.

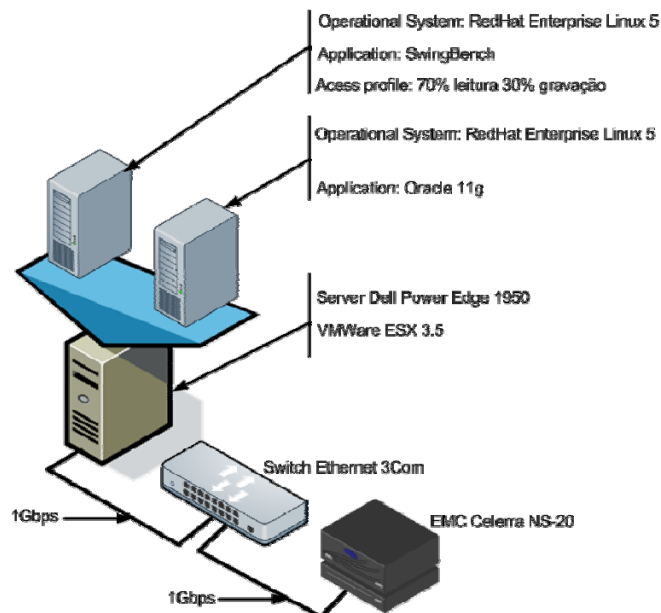
The fibre channel disk will be configured to the DBMS server, adding a hard disk to the virtual machine with the option of raw device mapping. The HBA configuration is described in appendix B.

### **5.2.2.2 Scenario iSCSI**

The SAN iSCSI test bed, figure 8, has an iSCSI initiator from the Red Hat installation package. The storage system is connected to the gigabit Ethernet. The DBMS virtual machine is using the ESX server NIC card to connect to Ethernet switch. The procedure to configure the iSCSI disk on the Red Hat Enterprise Linux is described in appendix C.

### **5.2.2.3 Scenario NAS**

The NAS environment will be enabled creating an area at the storage system. We will assign this area to an export using the protocol NFS version 3.



**Figure 7 — iSCSI and NAS test bed environment**

### 5.3 Storage system characterization

SAN fibre channel, SAN iSCSI and NAS NFS will be the storage architectures. One unique storage system permits the connection to the fibre channel and Ethernet networks, enabling us to connect this storage system to the SAN fibre channel, SAN iSCSI and NAS NFS. The storage system's physical disks have RAID technology to protect against data loss due to disk failure. We will use the disks in RAID-5 because this is widely deployed in the industry.

We will use the `io_profile` to measure the transfer rate, and I/O operations volume to address the question about the real transfer rate of each storage architecture and technology. The `io_profile` is a file system benchmark tool that measures the time it takes to perform a large number of open-seek-read (or write)-close cycles at random locations in a set of randomly chosen files. It makes these measurements by scanning the I/O size exponentially from a minimum to a maximum value. It measures performance of the whole I/O stack from the application's perspective, not the kernel's nor device's perspective. The `io_profile` will be executed at the DBMS server with the same parameters seeking to identify the effective transfer rate and I/O operations volume on each architecture.

The Logic Storage system Inc. (LSI), (2003) classified applications based on the access pattern. Random data access is measured in I/Os per second (IOPS) and is essential for transaction-based applications, like OLTP and database, with random, small-block I/O. Sequential data access is measured in megabytes per second (MB/s) and is crucial for bandwidth-intensive application, such as rich media streaming and seismic processing, with sequential and large block I/O. The Oracle database could have a block size varying from 2KB to 32KB (Oracle, 2007), so the io\_profile results of block size 2KB and 32KB will be collected in table 3.

Technology	Block de 2KB		Block de 32KB	
	MBps	I/Ops	MBps	I/Ops
SAN Fibre Channel	34.008	520	23.341	1.980
SAN iSCSI	21.255	435	14.588	980
NAS com NFS	17.543	370	3.701	484

**Table 3 — Architecture maximum I/O capacity**

According to the Oracle database performance tuning guide (Oracle, 2007a), the I/O request maximum size is based in the parameter locate at the file init.ora

~~Maximum I/O request size = 8KB x 8 = 64KB (3)~~, where 8KB is the block size of the database on this work and 8 is the parameter from the file init.ora DB\_FILE\_MULTIBLOCK\_READ\_COUNT.

#### 5.4 Architecture comparative analysis

The storage architecture will be compared based on:

- Ease of operation
- Cost
- Performance

## 5.5 Ease of Operation Analysis

This item will describe the methodology used to compare the ease of operation on the storage architecture.

Ease of operation is measured by the time spent creating a new area at the storage system, enabling the new area to communicate with the server, associating it to the server, and testing the area in the server. We will assign the factor 1 to the architecture needing the least time; any other architecture will receive a proportional factor. Then, we will divide the time necessary to create each architecture by the least time to create the area at the storage system. Only one architecture will receive factor 1, another will receive a lower factor.

To create a new area in the fibre channel architecture and use it in a server:

- Create a new area in the storage system management console
- Assign the new area to the server identification (wwn)
- Create zoning in the fibre channel switch, associating the server HBA and the storage system
- Execute the following command sequence:
  - `echo scsi add-single-device 0 1 2 3 > /proc/scsi/scsi;`
  - Fdisk – l to localize the new disk,
  - `echo "scsi add-single-device 0 1 2 3" > /proc/scsi/scsi;`
  - Fdisk – l to find the new disk, like in Figure 10

The iSCSI architecture requires these steps to create a new area:

- Create a new area in the storage system
- Map this area to the Server identification (iqn) in the storage system
- Discover the new area in the server with the command `iscsiadm -m discovery -t sendtargets -p storage_ip_address`
- Restart the iSCSI service
- Use the `fdisk – l` to verify the new disk identification, figure 10.

Ease of operation in the NAS NFS will be measured by:

- Creating an area in the storage system
- Creating an export with this area with NFS permission
- Mapping the area on the server

The steps were repeated five times to establish a trusted pattern.

Architecture	Time
SAN <i>fibre channel</i>	18 minutes
SAN iSCSI	8 minutes
NAS	4 minutes

**Table 4 — Time spent to create new area in the architecture**

## 5.6 Cost Analysis

The TPC (TPC, 2007b) has a methodology to provide a fair and honest comparison of vendor implementations as they accomplish an identical, controlled and repeatable task. Pricing for these implementations must also allow a fair and honest comparison for customers to review, while still supporting the sales strategy of each company.

The following must exist for the TPC analysis:

- The price must be based on a pricing model that the company actually employs with customers
  - The price methodology must be the same used to create the price to customer in a similar configuration.
- The price must be verifiable, the pricing model must represent the pricing that could be obtained by any customer in a request for bid to a single vendor
  - Any special discount must be the same used to create the price to customer in a similar configuration.

The TPC methodology specifies that the price must reflect the purchase price of the tested system, and must include hardware, software and maintenance charges for 36 months.

This study considers the price of the components used in each architecture, software and maintenance charges for 36 months, but excludes the cost of servers.

The storage system configured in this test has 2TB for the NAS and iSCSI portion and 2TB for the fibre channel world, using 300GB @15,000 RPM physical disks, management software. This model has bundled the multi-path software. This configuration also includes 24 x 7 on-site support.

The fibre channel technology includes the storage system, the 16-port fibre channel switch, and the HBA to be installed at the server.

The iSCSI and NAS technologies need only the addition of the storage system, because the Ethernet interface is already installed in the server.

The price is based on the Brazilian Market. The storage system used in the tests is R\$ 528,600.00; including the fibre channel option raises the price to R\$ 535,600.00. The fibre channel HBA costs R\$ 5,500.00, and the 16 port fibre channel switch costs R\$ 73,900.00.

## **5.7 Performance Analysis**

We used the next steps to create the data to compare storage architecture performance.

### **5.7.1 Test bed planning**

This work needs to verify the behavior of a storage system under different workloads. Workloads range from server to storage systems. We aim to identify the conditions that each architecture offers to support the DBMS performance need. The number of threads changes the simulator workload. We will use 3, 30 and 300 threads as test levels. During the preliminary tests, we got the message ORA-0020 (maximum number of processes (%s) exceeded). The Oracle Technology Network (OTN) showed the need to raise the parameter process. According to Niemiec (2007), if the parameter process is set to 2,500, it permits an Oracle application with 2,000 simultaneous active users. After this change, the simulator was stable with 300 threads. As a result, we will limit the workload to 300 threads.

After the environment is prepared, initiate a routine for 15 minutes to populate the server and storage system buffers. Initial tests showed that the response time and the transactions per second change if they are less than 3 units after the 7<sup>th</sup> minute, this work will adopt 10 minutes as the collect time.

The times for each thread level will be repeated based on the Jain (1991) study that related the results reliability and error with the number of samples to be collected.



$$n = \left( \frac{100 Z S}{R \bar{x}} \right)^2 \quad (4)$$

N = Number of samples;

Z = 1,96 – standard value for 95% of reliability;

S = standard deviation;

R = error, this work uses 10%;

$\bar{x}$  = average value of reference samples.

The preliminary test execution showed the following results:

Architecture	Threads	Transaction per minute	Response time
DAS	50 users	1639	953
		1674	801
		2167	512
		1923	445
		1646	818
		1432	771
		2236	519
		2313	489
		2571	455
		2479	435

**Table 5 — Preliminary test results**

The response time average value of the samples is 619 and the standard deviation is 172.  
The Jain formula is:

$$n = \left( \frac{100 \times 1.96 \times 172}{10 \times 619} \right)^2 \Rightarrow n = 29.66 \quad (5)$$

The Jain study shows the need to repeat each thread level 30 times to achieve 95% reliability. The value used in the architecture comparison is the average of the results. During the tests, we will monitor the DBMS and application server resources usage to assure that server processor, memory and NIC card aren't exhausted.

## 5.8 Performance results analysis

We monitored server resources during the tests using the package **dstat**. It collects usage from the processor, memory, Ethernet card and disk. The simulator was configured to stop the test after 10 minutes to keep execution time under control.

		DBMS Server			Application Server				
	Threads	CPU %	Mem%	NW%	CPU %	Mem%	NW%	TPM	Resp time
SAN fibre Channel	3	7	9	6	53	10	5	192	19
	30	28	16	8	54	10	15	1910	25
	300	100	12	8	54	11	19	2521	5380
SAN iSCSI	3	8	9	4	55	11	5	194	18
	30	20	14	5	56	11	20	1904	28
	300	100	52	20	56	12	30	1806	7321
NAS NFS	3	11	11	3	56	10	1	196	19
	30	25	18	7	55	13	8	1922	26
	300								

Table 6 — Collect data

The application server was not exhausted during the test, but the DBMS achieved 100% utilization (0% idle) during the execution of 300 threads for SAN fibre channel and iSCSI.

When the NAS environment was set up for 300 threads, the DBMS processor jumped to 100% utilization, but no traffic arrived at the storage system. The NAS environment limit was 187 threads, after this number the traffic to system storage stops.

We performed new measures to find the point where each architecture exhausts the DBMS server resources. The swingbench was set up with 100 threads and we monitored both server resources. After the routine run finished the execution time, we performed a new execution with the number of threads increased or decreased to find the exhaust point. Each architecture's peak number of transactions per minute is shown in table 7. Based on the formula (4), the number of monthly visits was calculated for each architecture.

Architecture	Threads	Transaction per minute	Response time	Monthly visits
SAN fibre channel	102	4562	312	66.096.000
SAN iSCSI	98	2832	998	63.504.000
NAS com NFS	84	4067	238	54.432.000

**Table 7 — Peak number of threads on each architecture**

SAN fibre channel achieved the higher number of transactions per minute. NAS, with lower threads, got 11% fewer transactions per minute, leveraging the need to have a server with more processor power. The iSCSI had the worst numbers for transaction per second and response time, even compared to NAS.

## **6 Conclusion**

The application workload simulated in the test bed is shown in table 8. It has the relative number of visits simulated in each thread level, permitting comparison to the real world. The test bed's levels were repeated 30 times ensuring data reliability.

The final data was inserted in table 9; now it is possible to compare the three architectures. The application response time was very close, the difference less than 1%. The server processor is not exhausted with 3 and 30 threads (as seen in table 9). The server processor was in an exhausted state, all commands going to the processor get in queue. The fibre channel architecture has a higher performance due to HBA capacity of process, fibre channel protocol, and SCSI commands. The NAS and iSCSI processing needs waited in queue.

Threads	Monthly visits
3	1.944.000
30	19.440.000
300	194.400.000

**Table 8 — Relating the threads and monthly visits**

The server configuration needs more processing capacity. The 2GB memory installed in the application and DBMS servers was enough for this test session, proven by the swap area that is used only when the server processor is exhausted. The peak of threads and transactions per minute for each architecture is shown in table 7. It shows that the fibre channel architecture achieves a higher thread level because the server processor handles only one interruption per I/O operation and the HBA process all the fibre channel layers. This avoids fibre channel commands staying in the exhausted server processor command queue. Environment monitoring verified that the network was not a bottleneck at any thread level, inspiring us to discard a possible use of link aggregation to provide more iSCSI connection throughput.

Choose fibre channel when the most important parameter is the number of transactions per minute. Comparing the environments, the number of transaction per minute in fibre channel environment is 18% higher than NAS NFS, and 4% higher than iSCSI. Including price in the comparison, fibre channel is 14% more expensive in the configuration used for this test bed environment. Fibre channel requires more time to configure a new area compared with any other architecture and in most cases, a storage vendor's professional service is required to leverage the solution.

A user environment with workloads of less than 54,432,000 visits per month achieves the best cost benefit in the NAS architecture. Performance is 11% lower than fibre channel, but the response time is 23% lower than fibre channel with a price that is 14% lower.

Storage architecture	Threads	Performance		Operational easiness	Cost
		Transactions per minute	Response time		
SAN Fibre Channel	3	192	19	4,5	R\$ 615.000,00
	30	1910	25		
	300	2521	5380		
SAN iSCSI	3	194	18	2	R\$ 528.600,00
	30	1904	28		
	300	1806	7321		
NAS com NFS	3	196	19	1	R\$ 528.600,00
	30	1912	25		
	300				

**Table 9 — Final results**

The numbers show that iSCSI is not a good solution for DBMS. The number of transactions is close to the NAS environment at the same price but with more effort to operate the environment. Some applications do not support NAS disks (mapped areas) for storing their information, so iSCSI becomes a good solution in this case.

The VMware infrastructure is able to host an Oracle server with a simulated workload of a web site with 66,096,000 monthly access, with a storage system's fibre channel disk configured as RDM (Raw Device Map).

A new test bed was created with newer servers. Their configuration is:

SUN Fire 4150 with 4 processors Xeon dual-core @2.33GHz and 20GB memory. The fiber channel HBA still the Emulex LP11000.

The new test results showed performance between the different storage architectures is pretty close. The response time and workload values differ as 1% from the lower and higher marks. The fiber channel kept the best performance, but the iSCSI and NAS are less than 1% lower.

## **6.1 Limitation**

Table 9 shows the exhaust of server processing capacity with 300 threads for SAN fibre channel SAN iSCSI. Due to the server limitation, it was not possible to show the correct behavior of the data traffic between the DBMS and the storage system under this workload. The DBMS and application servers were guests from a VMware ESX server, and the virtualization influence in the test was not a focus of this project.

## **6.2 Suggestion for future works**

This study used a virtual environment due to a resource limitation. The server processor was not powerful enough so the first suggestion is to use a more powerful server to ensure that there is data traffic limitation in the storage architecture.

This work used a virtualized DBMS server. In the future, another work should compare the virtualization influence in a DBMS environment.

The storage industry is always looking for faster interfaces using a new or improved physical connection. Fibre Channel over Ethernet (FCoE) is one of the newest options. This implements the fibre channel protocol using the Ethernet as the transport layer instead of the fibre channel physical layer.

## Bibliography

AIKEN, S. ET al. **A Performance Analysis of the iSCSI Protocol. The 2005 Symposium on Applications and the Internet, 2005. Proceedings.** 2005. Available at <<http://portal.acm.org/citation.cfm?id=824467.824988>>.

BRAAM, PETER. **Lustre File Systems.** Accessed in <[http://en.wikipedia.org/wiki/Lustre\\_\(file\\_system\)](http://en.wikipedia.org/wiki/Lustre_(file_system))> . Accessed in 01/11/2008

BORAL, H.; DEWITT, D. J. **A Methodology for Database System Performance Evaluation. Data Management international Conference.** 1984. Available at <<http://pages.cs.wisc.edu/~dewitt/includes/benchmarking/sigmod84.pdf>>

BURR, W. E.; LOHMEYER, J.B. **Small Computer System Interface.** American National Standard for Information System. Revision 17B. Dec, 10th 1985.

DEYRING, KLAUS-PETER. **High Speed Serialized Attachment.** Year 2001. Serial ATA workgroup.

DING, YIPING. Bandwidth and latency: their changing impact on performance. **Proceedings of the 35th Computer Measurement Group Conference.** December 2005. Page 3.

Dr.FEIGENBAUM, BARRY. **Server Message Block Protocol.** Available at <<http://timothydevans.me.uk/nbf2cifs/c1757.html>>.

EMC CORPORATION. **EMC CLARiiON CX3-20c iSCSI (4000 Users) Storage Solution for Microsoft Exchange Server.** Available at <<http://www.emc.com/collateral/hardware/specification-sheet/cx3-20c-storage-solution-microsoft-exchange-server.pdf>> . 31/October/2006.

FLORES, FELIPE. **Análise de desempenho do banco de dados Oracle 9i. Graduation Monograph (Science Computing Bachelor).** December/2003. Universidade Luterana do Brasil.

GILES, DOMINIC. **Swingbench.** Available at <<http://www.dominicgiles.com/swingbench.php>>.

INTEL CORPORATION. **Iometer Project.** Available at < <http://www.iometer.org/>>.

INCITS — INTERNATIONAL COMMITTEE FOR INFORMATION TECHNOLOGY STANDARDS. **Small Computer System Interface.** Year 1986. American National Standards Institute.

JAIN, R. **The Art of Computer Systems Performance Analysis.** USA. John Wiley & Sons, Inc. 1991

JIBBE, M.K.; HAMMOND-DOEL, TOM; WILSON, STEVEN. **Fibre Channel Technologies current and future.** 2007. Available at <[www.snia.org](http://www.snia.org)>.

KANCHERLA, BHARGAVA; NARAYAN, GANESH; GOPINATH, K. **Performance Evaluation of Multiple TCP Connections in iSCSI.** Year 2007. Article. Available at <<http://portal.acm.org/citation.cfm?id=1306871.1306912&coll=GUIDE&dl=GUIDE>>.

LIN, MENGJOU et al. **Performance of high-speed network I/O subsystems: case study of a fibre channel network**. December 1994. Article. Available at <<http://portal.acm.org/citation.cfm?id=602770.602804>>.

LOMEYER, J. B.; LAMERS, L. J. **AT Attachment Interface for Disk Drives**. Year 1994. Computer and Business Equipment Manufacturers Association.

LUCHESI, RAY. **Storage performance**. March 2007 Available at <<http://www.snia.org>>.

LSI – LOGIC STORAGE SYSTEMS. **Implementing SATA technology in the enterprise data center**. October/2003.

MALAKAPALLI, CHRIS; GUNTURU, VAMSI. **Evaluation of SCSI over TCP/IP and SCSI over Fibre Channel Connections**. 2001. Available at <<http://doi.ieeecomputersociety.org/10.1109/HIS.2001.946698>>.

MENDES, MARCELO RODRIGUES NUNES. **Análise de desempenho de sistemas OLTP utilizando o benchmark TPC-C**. October/2006. Graduation work. Available at <<http://www.cin.ufpe.br/~tg/2006-1/mrnm.pdf>>.

MICROSOFT CORPORATION. **Planning the Jetstress Performance Test**. Available at <<http://technet.microsoft.com/en-us/library/bb643101.aspx>>. 16/07/2007.

\_\_\_\_\_. **Windows Server 2008 High Availability Program – customer white paper**. Windows Server 2008 High Availability Program. Available at <[http://download.microsoft.com/download/8/2/f/82fa3808-7168-46f1-a07b-f1a7c9cb4e85/WS08%20High%20Availability%20Program%20customer%20whitepaper\\_FIN\\_AL.docx](http://download.microsoft.com/download/8/2/f/82fa3808-7168-46f1-a07b-f1a7c9cb4e85/WS08%20High%20Availability%20Program%20customer%20whitepaper_FIN_AL.docx)>. January/2008.

MUGOL, JEFFREY. **TCP offload is a dumb idea whose time has come**. 2003. Available at <<http://www.cs.uwaterloo.ca/~brecht/servers/readings-new/mogul-offload-2003.pdf>>.

NetApp - NETWORK APPLIANCE INCORPORATED. **FAS2050 iSCSI 5,000 User Exchange 2007 Storage Solution**. Page 8. 20/8/2007. Available at <<http://media.netapp.com/documents/fas2050-5000-user-iscsi-esrp-v2.0-storage-solution.pdf>>

\_\_\_\_\_. iSCSI Overview. Available at <[http://www.stemmer.de/service/workshops/sbb2006juli/download/netapp\\_iscsi.pdf](http://www.stemmer.de/service/workshops/sbb2006juli/download/netapp_iscsi.pdf)> Page 11. Year 2006.

NIEMIEC, RICHARD J. **Oracle Database 10g Performance Tuning Tips and Techniques**. USA. McGraw-Hill Osborne. June/2007. Chapter 4.

ORACLE CORPORATION. **Oracle® Database Performance Tuning Guide 11g Release 1 (11.1)**. July/2007a.

\_\_\_\_\_. **Oracle® Real Application Cluster 11g**. Page 4. April/2007b.

\_\_\_\_\_. **Oracle Quick Installation Guide 11g Release 1 (11.1) for Linux x86**. September/2007c.

\_\_\_\_\_. **Oracle8i Parallel Server Administration, Deployment, and Performance**. December/1999.



RODRIGUES NETO, A. J. **Um estudo do desempenho dos protocolos iSCSI e Fibre Channel**. December/2004. Dissertation (Mastering in Science Computer) — Instituto de Computação, UNICAMP.

SATRAN, J.; SAPUNTZAKIS, C.; CHADALAPAKA, M. **Internet Small Computer Interface (iSCSI)**. Abril/2004. Available at [HTTP://www.ietf.org/rfc/rfc3720.txt](http://www.ietf.org/rfc/rfc3720.txt).

SENAPATHI, SANDHYA; HERNANDES, RICH. **Introduction to tcp offload engine**. March/2004. Available at <http://www.dell.com/downloads/global/power/1q04-her.pdf>.

SHIVER, ELISABETH; HILLYER, BRUCE; SILBERSCHATZ, AVI.. **Performance analysis of storage systems**. 2000. Available at <http://www.cs.nyu.edu/~shriver/bell-labs/papers.html>

STRAND, STAFFAN; AHLSTRAND, MARKUS. **I/O Performance Characteristics of iSCSI NICs and Fibre Channel HBAs**. November/2003. Available at [http://www.cisco.com/global/SE/pdfs/storage/FC\\_vs\\_iSCSI\\_Performance.pdf](http://www.cisco.com/global/SE/pdfs/storage/FC_vs_iSCSI_Performance.pdf).  
SUN MICROSYSTEMS RFC1094. IETF 1998

TOMIYAMA, H.; ISHIHARA, T.; INOUE, A. and YASUURA, H. **Instruction scheduling for power reduction in processor-based system design**. 1998. In Proceedings of the Conference on Design, Automation and Test in Europe 98. Available at [www.acm.org](http://www.acm.org).

TPC — TRANSACTION PROCESSING PERFORMANCE COUNCIL. **TPC Benchmark™ C**. June/2007a.

TPC — TRANSACTION PROCESSING PERFORMANCE COUNCIL. **TPC Pricing Specification**. June/2007b.

TURVEY, STEVE. 2005. **Road test: Four databases tested**. 25/12/2005. Article available at <http://www.builder.au.com.au/architect/database/soa/Road-test-Four-databases-tested/0,339024547,339224962,00.htm>.

WRIGHT, MONTY. **Implementing high availability with DB2 9.5**. Page 4. IBM *development works*. 31/07/2008. Available at <http://download.boulder.ibm.com/ibmdl/pub/software/dw/dm/db2/dm-0807wright/dm-0807wright-pdf.pdf>.

YAMAGUCHI, SANEYASU; OGUSHI, MASATO; KITSUREGAWA, MASARU. **Trace System of iSCSI Storage Access**. 2005. Available at <http://ieeexplore.ieee.org/Xplore/login.jsp?url=/iel5/9522/30170/01386139.pdf> >.

## Appendix A

### Test Script

#### DBMS Server installation:

- Operational system install;
- The Oracle requires a minimum of 4,096MB;
- Follow the Oracle (2007c) install guide;
- During the Oracle directory creation, mount the storage area in that directory;
- Install io\_profile, move to the Oracle directory and execute it to measure the architecture limits;
- Install the database;
- Install the simulator;
- Configure the simulator;
- Create the order entry data;
- Install the dstat package;
- Monitor the resources;
- Start the simulation.

#### Application Server installation:

- Operational system installation;
- Install Java jre.1.5.0 package;
- Install and configure the Swingbench;
- Install dstat package;
- Monitor the resources.
- Execute the tests;

## Appendix B

### Fibre Channel HBA installation.

```
Download the Emulex driver;  
tar xzf lpfc_2.6_driver_kit*.tar.gz  
cd lpfc_2.6_driver_kit/  
./lpfc-install
```

```
Install the application;  
tar xvf elxlinuxapps*.tar  
cd elxLinux*  
./install
```

Accept the default options in installation process:

```
Select desired mode of operation for HBAnyware  
  
1 Local Mode : HBAs on this Platform can be managed by  
HBAnyware clients on this Platform Only.  
2 Managed Mode: HBAs on this Platform can be managed by  
local or remote HBAnyware clients.  
3 Remote Mode : Same as '2' plus HBAnyware clients on this  
Platform can manage local and remote HBAs.  
  
Enter the number '1' or '2' or '3' <2>  
You selected: 'Managed-only Mode'  
  
Would you like to enable configuration features?  
  
Enter y to allow configuration. (default)  
Enter n for read-only mode.  
  
Enter the letter 'y' or 'n' <y>  
You selected: Yes, enable configuration  
  
Do you want to allow user to change management mode using  
set_operating_mode script located in /usr/sbin/hbanyware ?  
  
Enter the letter 'y' if yes, or 'n' if no <y>  
You selected: Yes
```

**Figure 8 — Fibre Channel HBA installation program interaction**

```
Scan o bus Fibre Channel  
echo "- - -" > /sys/class/scsi_host/host0/scan  
echo "- - -" > /sys/class/scsi_host/host1/scan  
echo "- - -" > /sys/class/scsi_host/host2/scan
```

```
exec fdisk -l to localize the new disk
```

```
[root@rhel5 ~]# fdisk -l
```

```
Disk /dev/sda: 80.0 GB, 80032038912 bytes
255 heads, 63 sectors/track, 9730 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
```

Device	Boot	Start	End	Blocks	Id	System
/dev/sda1	*	1	13	104391	83	Linux
/dev/sda2		14	9730	78051802+	8e	Linux LVM

```
Disk /dev/sdb: 10.7 GB, 10737418240 bytes
64 heads, 32 sectors/track, 10240 cylinders
Units = cylinders of 2048 * 512 = 1048576 bytes
```

```
Disk /dev/sdb doesn't contain a valid partition table
```

**Figure 9 — sdb, the new disk**

Format the new disk >mkfs.ext3 /dev/sdb

Mount the new *fibre channel* disk >mkdir /mnt/fibre and mount /dev/sdb /mnt/fibre

Change the /etc/fstab to assure the disk is mounted after the reboot and insert /dev/sdb  
/mnt/fibre ext3 \_netdev 0 0

## Appendix C

### Install the iSCSI *initiator*

- Mount the RHEL5 DVD;
- Install the package `rpm -ivh iscsi-initiator-utils-4.0.3.0-7.i386.rpm`;
- List the initiator name;
  - `cat /etc/iscsi/initiatorname.iscsi`;
  - The command answer looks like;

`InitiatorName=iqn.1987-05.com.cisco:01.718963daf13`;

Configure the initiator name in the storage system;

Start the iSCSI service `/etc/init.d/iscsi start`;

Scan the iSCSI target `iscsiadm -m discovery -t sendtargets -p 10.240.0.90`;

Restart the iSCSI service `/etc/init.d/iscsi restart`;

Identify the iSCSI disk `fdisk -l`;

`mkfs.ext3 /dev/sdx`;

- Format the new disk;
  - `>mkfs.ext3 /dev/sdb`;
- Mount the new iSCSI disk;
  - `>mkdir /mnt/iscsi` and `mount /dev/sdb /mnt/iscsi`;

Execute this command for the iSCSI disk is mounted during the next boot;

- `chkconfig iscsi on`;
- Change the `/etc/fstab` file and insert `/dev/sdb /mnt/iscsi ext3 _netdev 0 0`;

## Appendix D

### Install the io\_profile

Download the io\_profile from [http://www.estorian.com/form\\_download\\_util.php?target=util\\_io](http://www.estorian.com/form_download_util.php?target=util_io);

Unzip the package io\_profile1\_0.zip;

Follow the README file instructions;

./io\_profile.

### Download the dstat package

wget <http://dag.wieers.com/rpm/packages/dstat/dstat-0.6.7-1.el5.rf.noarch.rpm>;

rpm -I dstat-0.6.7-1.el5.rf.noarch.rpm;

dstat -a --output filename;

### Download the LSHW package

wget <http://packages.sw.be/lshw/lshw-2.10-1.el5.rf.i386.rpm>;

rpm -I lshw-2.10-1.el5.rf.i386.rpm;

lshw;

## **Appendix E**

### **Parameters to find the data block size**

```
from init.ora file DB_FILE_MULTIBLOCK_READ_COUNT=8  
from spfile<SID>.ora file db_block_size = 8192 e _db_cache_size=138412032  
from sqlplus & show parameters area_size -> hash_area_size 131072; sort_area_size 65536
```

## **Appendix F**

Install the simulator Swingbench

Download the simulator:

- `wget http://www.dominicgiles.com/swingbench/swingbench230391.zip`

Customize the configuration file swingbench.env

Create the order entry schema

`>./oewizard`

Exec the load simulator

`>./swingbench`



## Appendix G

Map the NAS export in the server

```
Mkdir /mnt/nas
```

```
mount -o tcp 10.240.0.90:/vol/vol1 /mnt/nas
```

Change the fstab file

```
10.240.0.90:/vol/vol1 /mnt/nas nfs hard,intr,proto=tcp,noauto
```

## **Biography**

### **Sergio Hirata**

Sergio Hirata graduated in Electrical engineering by Faculdade de Engenharia de São Paulo and will conclude the Master's degree in Science Computing in 2009. He has 20 years experience in information technology, working with the major storage products. His expertise is storage and backup architecture and capacity planning. Sergio is also allocated in storage and backup SLA driven design projects. He has achieved a number of EMC Proven Professional Accreditations in Backup, Foundation and Storage.

### **Volnys Borges Bernal, Phd.**

Volnys Borges Bernal graduated with a degree in Science Computing from the Universidade Federal de São Carlos, and Master and PhD in Electrical Engineering from Escola Politécnica // Universidade de São Paulo. He is a Security auditor certified by CISA/ISACA and a researcher at Laboratorio de Sistemas Integráveis da Escola Politécnica da USP since 1990. Volnys worked as technical coordinator at several industrial and academic projects, and CSO (Chief Security Officer) at LSITEC. He is a Professor at IPT Science Computing Mastering with more than 10 years working at system security area. Today, digital certification, system security and forensics investigation are his primary focus. [www.lsi.usp.br/~volnys](http://www.lsi.usp.br/~volnys)