IMPLEMENTATION OF NAS USING RAID CONFIGURATION

SEMINAR-1 REPORT

Submitted by

ABDUL SAMAD (RA2111003020190)

SANTHOSH N.S (RA2111003020209)

BALAJI PANDI .R (RA2111003020205)

Under the guidance of

Mr. M. Sadhasivam
Dr. A. Manju

(Assistant Professors, Department of Computer Science and Engineering) ***In partial fulfillment for the award of the degree Of***
BACHELOR OF TECHNOLOGY
In

COMPUTER SCIENCE AND ENGINEERING
Of

COLLEGE OF ENGINEERING AND TECHNOLOGY

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY

RAMAPURAM CAMPUS, CHENNAI-600089

NOVEMBER 2023

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY

(Deemed to be University Under Section 3 of UGC Act, 1956)

BONAFIDE CERTIFICATE

Certified that the Seminar-I report titled **"IMPLEMENTATION OF NAS USING RAID"** is the bonafide work of **"SANTHOSH N.S [RA2111003020209], ABDUL SAMAD [RA2111003020190], BALAJI PANDI R [RA2111003020205]"** submitted for the course 18CSP103L Seminar – I. This report is a record of successful completion of the specified course evaluated based on literature reviews and the supervisor. No part of the Seminar Report has been submitted for any degree, diploma, title, or recognition before.

SIGNATURE
Mr.M.Sadhasivam, M.E., (Ph.D)
Assistant Professor
Computer Science & Engineering
SRM Institute of Science and Technology
Ramapuram, Chennai

SIGNATURE
Dr. K. RAJA, M.E., Ph.D.,
Professor and Head
Computer Science & Engineering
SRM Institute of Science and Technology
Ramapuram, Chennai.

Submitted for the Seminar-1 Viva Voce Examination held on …… at SRM Institute of Science and Technology, Ramapuram Campus, Chennai-600089.

EXAMINER I **EXAMINER II**

TABLE OF CONTENTS

ABSTRACT:

This paper is a work on completely backing up disk files on client server/cloud drives. This paper contains objectives, introduction, literature review, methodology, hardware, and software required, application and references. The current solution for home or small businesses to acquire extra storage is through online cloud services and solutions from the likes of big tech giants. While this is the simplest method, it is not the most efficient or cost effective, not to mention the lack of full control over what happens to your data once it is uploaded. This is where the concept of NAS solution (Network Attached Storage) proves to be a viable and advantageous alternative. NAS (Network-attached storage) is a data storage, which is connected to a computer network. NAS acts a file server in a network, offering data storage to be located in a stand-alone network, which client computers can be connected. NAS can be seen as a computer drive (via Ethernet) and as such, it can be used to save documents and files and as well as read them. NAS is specialized for serving files either by its hardware, software, or configuration. It is often manufactured as a computer appliance – a purpose-built specialized computer. NAS systems are networked appliances that contain one or more storage drives, often arranged into logical, redundant storage containers or RAID. Like any scientific advancement, NAS systems have both advantages and disadvantages. However, as the need for storing large amounts of data and accessing them quickly grows, the potential of NAS systems will continue to increase.

LIST OF FIGURES

LIST OF ACRONYMS AND ABBREVIATIONS

1. NAS	Network Attached Storage
2. RAID	Redundant Array of Inexpensive/Independent
3. NCP	NetWare Core Protocol
4. NFS	Network File System
5. IBM	International Business Machines (Company)
6. SMB	Server Message Block
7. UNIX NFS	Network File System implemented in UNIX
8. EMC Celra	Early model of NAS by Dell
9. DAS	Direct Attached Storage
10. SAN	Storage Area Network
11. USB	Universal Serial Bus
12. ATA	Advanced Technology Attachment
13. iSCSI	Internet Small Computer System Interface
14. AoE	ATA over Ethernet
15. LAN	Local Area Network
16. GB, TB, PB	Gigabyte, Terabyte. Petabyte
17. GUI	Graphical User Interface
18. IT	Information Technology
19. NIC	Network Interface Card
20. iSCSI	Internet Small Computer System Interface
21. ATX	Advanced Technology Extended
22. OMV	OpenMediaVault
23. ZFS	Zettabyte File System
24. HDD	Hard Disk Drive
25. PCIe	Peripheral Component Interconnect Express
26. BTRFS	B-TRee File System
27. AFP	Apple Filing Protocol
28. VM	Virtual Machines
29. EXT4	4 extended file system

CHAPTER 1

INTRODUCTION

NCP protocol was released in 1983. Release of NFS in 1984 allowed network servers to share their storage space with networked clients. Inspired by the success of file servers from Novell, IBM, and Sun, several firms developed dedicated file servers. A group of Auspex engineers split away in the early 1990s to create the integrated NetApp filler, which supported both the Windows SMB and the UNIX NFS protocols and had superior scalability and ease of deployment. This started the market for proprietary NAS devices now led by NetApp and EMC Celera.

- NAS (Network-attached storage) is a data storage, which is connected to a computer network. NAS acts a file server in a network, offering data storage to be located in a stand-alone network, which client computers can be connected.
- NAS can be seen as a computer drive (via Ethernet) and as such, it can be used to save documents and files and as well as read them.
- Fundamentally, a NAS is a computer, optimized in hardware and software to be a file server. Data is a critical asset for company. Without access to their data, companies may not provide their expected level of service. Poor customer service, loss of sales, or team collaboration problems are examples of what can happen when information is not available.
- Each of these issues contribute to lack of efficiency and potential loss of income if customers cannot wait for a data outage to be corrected.
- The key difference between direct-attached storage (DAS) and NAS is that DAS is simply an extension to an existing server and is not necessarily networked.
- As the name suggests, DAS typically is connected via a USB or Thunderbolt enabled cable. NAS is designed as an easy and self-contained solution for sharing files over the network.
- Both DAS and NAS can potentially increase availability of data by using RAID or clustering

- NAS provides both storage and a file system. This is often contrasted with SAN (storage area network), which provides only block-based storage and leaves file system concerns on the "client" side. SAN protocols include Fibre Channel, iSCSI, ATA over Ethernet (AoE) and Hyper SCSI
- Despite their differences, SAN and NAS are not mutually exclusive and may be combined as a SAN-NAS hybrid, offering both file-level protocols (NAS) and block-level protocols (SAN) from the same system. An example of this is Open filer, a free software product running on Linux-based systems. A shared disk file system can also be run on top of a SAN to provide file system services.

1.1 Problem Statement:

The current scenario for the people with aspiring small businesses and startups is that during the requirement of a centralized storage for their files and data. However, traditional storage solutions like using multiple external hard drives, USB sticks and sometimes temporary subscription services. This generates a threat of a data loss in multiple ways: user may accidentally remove his/her own data and while there are no constant backups made, the data may not be recoverable. This can be very unreliable and slow for business use, not to mention the amount of devices that will be piled up as a result. This can cause a very unorganized structure of storage and managing all of it in a very professional and efficient way is not possible.

This is where NAS (network attached storage) comes into the equation which aims to solve all the above problems stated. The company can implement a NAS solution using RAID. RAID is a data redundancy technique that combines multiple physical disks into a logical unit, providing performance, fault tolerance and sometimes even both. Using this NAS solution with RAID can have benefits like increased performance, scalability, reliability, flexibility, and many more use cases which will be discussed in the upcoming topics.

1.2 Objective:

The aim of this research is to create a data storage system for the case company which utilizes and fulfills the following requirements:

- Data backups
- External access over Internet
- Access from LAN
- Optimized data security
- Maintainability
- Future expansions
- System health monitoring
- Personal server grade applications
- Smart security systems
- Personal cloud server

1.3 Scope of the Project:

The purpose of network-attached storage is to enable users to collaborate and share data more effectively. The scope of NAS storage is growing at a high rate with the emerging new technologies that are developed. NAS storage is a versatile solution that can be used for a variety of purposes. Here's an overview of the scope of NAS in brief:

1. Accessibility:

NAS are most used and preferred mainly for their ease of access over Ethernet. This means that if it is configured as per the requirements, the clients can access their files from anywhere from the NAS. This is because NAS acts basically as an independent / standalone cloud server with Ethernet connection given to it.

2. Cost-effectiveness:

NAS storage can be a cost-effective way to store data. This is because NAS servers are typically less expensive than traditional SAN

3. Collaboration:

NAS storage can be used to collaborate on projects with others on a network. This can be useful for businesses, schools, and other organizations that need to work together on projects.

4. Backup and disaster recovery:

NAS storage can be used to backup data from computers and other devices. This can help to protect data from loss or corruption in the event of a disaster.

5. File sharing:

NAS storage can be used to share files between users on a network. This can be useful for businesses, schools, and other organizations that need to share files with a large number of people.

6. Data archiving:

NAS storage can be used to archive data that is no longer needed on a regular basis. This can help to free up space on computers and other devices.

CHAPTER 2

PROJECT DESCRIPTION

2.1 Existing System:

There exist many cloud-based storage solutions provided by the tech giants like Microsoft, Google and other companies. But the issue arises when the need for the storage volume exceeds the provided limit by these providers. These cloud storage solutions have a very limited 10-15 GB base free trial to use their platform/Service and then the pricing scales up quite high for frankly limited amount of storage going up to 2TB of storage maximum from the likes of Google. 2TB to run an entire company, even if it is a small one, is very constrictive. If the company stores any big volumes of media and maybe programming based file requirements, 2TB can be filled up very quickly and hence not providing any headroom for the future. Other companies like Microsoft and pCloud offer bit more storage in the form of Business plans. But even in business plans you can see that the pricing gets more expensive for what you get and it is not quite scalable. On top of all of the pricing, you don't get much flexibility as to how you get to use your storage. The cloud storage provided by the companies can be used only based on the limited flexibility of the services provided by them. This means the storage you buy cannot be used however you deem fit to use it for. For example, it cannot be used to install programs, run virtual machines, operate as a media server. Easy file sharing, and many more use cases that you simply get with traditional physical storage.

2.2 Issues in Existing System:

1.Security concerns: Cloud-based storage solutions can be a target for cyberattacks, so it is important for small businesses to choose a provider with strong security measures in place. Additionally, small businesses should implement their own security measures, such as encrypting their data and using strong passwords.

2.Compliance challenges: Small businesses in certain industries may be required to comply with specific regulations regarding data storage and security. It is important for small businesses to choose a cloud-based storage provider that can help them meet their compliance requirements.

3.Lack of control: Small businesses may have less control over their data and applications when using cloud-based storage solutions. This can be a concern for small businesses that need to maintain strict control over their data or that must comply with regulations that require them to keep their data on-premises.

4.Hidden costs: Small businesses should be aware of the hidden costs associated with cloud-based storage solutions, such as egress fees (fees for transferring data out of the cloud) and overage fees (fees for exceeding the amount of storage space included in their plan).

5.Vendor lock-in: It can be difficult and expensive for small businesses to switch cloud-based storage providers once they have migrated their data to the cloud. It is important for small businesses to carefully choose a provider that they can trust and that will meet their needs in the long term.

2.3 Software/Hardware Requirements:

Hardware:

The hardware requirements for building a NAS can vary from one system to another just like a normal personal computer build. The hardware requirements stated below are the minimum requirements to build a decently capable NAS server.

CPU	Intel Pentium G4560, Intel Core i3-10100, Intel Core i5-12400, AMD Ryzen 3 3100G, AMD Ryzen 5 5600G
RAM	8GB, 16GB, 32GB
STORAGE DRIVES	1TB, 2TB, 4TB, 8TB, 12TB, 16TB, and so on. (Depends on how much storage your needs require)
NETWORK	Gigabit Ethernet, 10 Gigabit Ethernet

INTERFACE CARD (NIC)	
POWER SUPPLY	400W, 500W, 600W
CASE	ATX case, mini-ITX case, Rack mounted server case
MOTHERBOARD	Any Motherboard which works with the CPU platform, room for expandable RAM, 2 or more PCIe Slots, and optional integrated NIC if possible, within budget

Software:

While the hardware requirements might vary, the software required to get the NAS system up and running is the same across the board with few optional software that can be installed after the minimum software setup processes. This can be considered equal to installing a fresh new copy of windows to a new computer with optional software installed later according to our needs.

There are mainly two choices of OS for a NAS system available for free to use.

1. OpenMediaVault (OMV)
2. TrueNAS Scale

We will discuss in detail about which is best suited for whom in the later sections.

2.4 Literature Review:

1. Network-Attached Storage For Small Companies

Author: Koivisto, Jari-Pekka

This study focuses on finding the proper way to implement an NAS for the needs of the case company and as such to improve their data security. Since the case company is relatively small and it operates under the regulations of European Union as it's a foundation supported by the EU, it doesn't have the required budget to create IT-solutions which utilize the industry's best policies and practices. This also gives the study some of its limitations; the

best solution might not be suitable in this case as the budget for the system to be built is only approximately 500 euros. This study is narrowed down to discuss the NAS-system; not the equipment around it, such as the network and the client computers. These aspects should be discovered before the actual implementation stage of the NAS. The case company in this study has the needed network infrastructure created beforehand by Nordic Lan & Wan Communication Oy.

2. Network-Attached Storage: Data Storage Applications

Author: Ravi Kumar M G , Ayudh Nagaraj, Benjamin Paul , Sharat P Dixit

This paper is a work on completely backing up disk files on client server/cloud drives. This paper contains objectives, introduction, literature review, methodology, hardware and software required, application and references. NAS (Network-attached storage) is a data storage, which is connected to a computer network. NAS acts a file server in a network, offering data storage to be located in a stand-alone network, which client computers can be connected. NAS can be seen as a computer drive (via ethernet) and as such, it can be used to save documents and files and as well as read them.NAS is specialized for serving files either by its hardware, software, or configuration. It is often manufactured as a computer appliance a purpose-built specialized computer. NAS systems are networked appliances that contain one or more storage drives, often arranged into logical, redundant storage containers or RAID. Network-attached storage removes the responsibility of file serving from other servers on the network. As businesses and industries keep growing and technology is advancing the need for drivers for accessing data quickly has become a necessity, hence many drivers have been introduced and implemented in different fields as this paper adopts those concepts. As to any situation or a gift of science , there is always two sides to it, yes, pros and consequences. NAS systems will always have a growing and increasing potential as long as technology continues to grow for the need of storing large amounts of data and also the desire of accessing them quickly for a smooth and efficient working space.

3. Cloud computing: implementing nas in directing digital images without local storage

Author: J.C.Rosaline Christi, R.Illakia, M.Gowri

In this paper, we investigate that the photographs can be directly emerged into the cloud. Compared with the traditional software theory and application, Cloud computing have a lot of benefits cannot be replaced and representation of the revolutionary ideal. a camera should be

able to operate without local storage over a home Wi-Fi network. A local memory cache would, of course, be needed, but our tests show that a small 32 MB cache would be quite effective and anything larger would be a bonus. it is desirable that latest generation 802.11n networking technology is used. It is also true that we did not test with more than two cameras operating simultaneously, but these are issues that can be dealt with through robust systems design.

4. Harnessing RAID mechanism for enhancement of data storage and security on cloud

Author: Sudipta Sahana, Rajesh Bose and Debabrata Sarddar

As one of the most sought-after technologies in the world today, cloud computing finds a prominent place among researchers, service providers, hardware and software engineers, and consumers alike. Among the key services provided by cloud computing, the most widely used is storage on the cloud. To ensure a high degree of reliability and an almost constant state of availability, cloud storage systems rely on replication of data. Replicating data is not without its costs. In an attempt to mitigate costs of storage, designers have banked on erasure coding practices in lieu of data replication. This has resulted in development of Cloud RAID systems. An ideal Cloud RAID system strikes an optimum balance between performance, storage, and reliability. In addition to this, data security is of paramount interest. In this paper, we introduce an entirely new technique of storing data in cloud networks by file type classification. Combining RAID types with this, our paper proposes a model that is robust and can withstand instances of hard disk failures. Further, we discuss a methodology that is embedded in our proposed framework that supports user authentication for data storage and retrieval. Our Secure User Authenticated Cloud RAID model shields user data from prying eyes and even from unauthorized attempts by the cloud service provider our Secure User Authenticated Cloud RAID architecture.

5. Dataplexed Network Attached Storage System

Author: Praveen Vijai, Minnu Meria, Midhu Babu.

This paper proposes a system which aims to resolve the problems associated with traditional client/server systems, namely the high performance requirements, and costly maintenance of the computing resources. With the traditional model, information’s are stored on servers and downloaded to clients as required, where they are then accessed locally. After updation, the information’s are saved back to the centralized. However information’s stored in a centralized

system are dependent on an OS and a native file. Centralized storage always harbors a possible risk of total loss of data due to system hang-ups & disk crashes. Having to know the specific information like IP address of a system or having to create localized user accounts on each peer system to access shared resources in existing system is also pretty discouraging. We propose Dataplexed NAS System [Dataplexed Network Attached Storage System], a cross-platform distributed file system, which provides strong security and client/server heterogeneity both at hardware and operating system level

CHAPTER 3

DESIGN

3.1 Proposed System:

The NAS system will allow flexibility for multiple use cases which prove to be simply superior to the typical cloud-based storage. For starters, NAS systems can be deployed on-premises, giving businesses more control over their data security. It can help businesses to comply with regulations that require them to keep their data on-premises. They give businesses more control over their data and applications. NAS can be more cost-effective than cloud-based storage solutions in the long term, especially for businesses that have large amounts of data to store. They also include huge scalability and flexibility to meet the needs of a thriving business. The proposed NAS system also contains multiple features that cloud storage solutions cannot even come close to providing. Since a NAS is our personal cloud storage, we can treat it like the storage is natively accessible to the computer. This means the NAS system would support a variety of file sharing protocols, such as SMB, NFS, ZFS, EXT4, BTRFS and AFP. This would allow users to access their files from a variety of devices. It would include storage management tools to help users manage their files and folders. It also allows easy installation of backup software to help users back up their data. Just like a natively integrated storage, this system could be configured as a media server, allowing users to stream their media files to their devices. It can also support running of Virtual machines (VM), allowing businesses to consolidate their IT resources.

Along with the stated proposal and benefits to using a NAS system, we can also implement RAID. RAID is traditionally implemented in businesses and organizations where disk fault tolerance and optimized performance are crucial to implement. Servers and NAS in business data centers typically have a RAID controller a piece of hardware that controls the array of disks.

These systems feature multiple SSD (solid state drive), SATA (serial advanced technology attachment) drives, or HDD, depending on the RAID configuration. Increasing storage demands of consumers, home NAS devices also support RAID. Pro-sumer, home and small

business NAS are increasingly using storage with two or more disk drive so that users can take advantage of the power of RAID just like an enterprise can.

Software RAID is the setup of RAID without the need for a dedicated hardware RAID controller.

It works by Holding the data on multiple disks and allowing (I/O) operations to overlap in a balanced way. This is done because using multiple disks increases the mean time to failure, while storing data redundantly also increases fault tolerance. It works on single logical drive on an operating system

The techniques used to achieve this is by either Disk Mirroring or Disk Striping

A. Mirroring will copy identical data onto more than one drive.

B. Striping partitions help spread data over multiple disk drives.

The stripes of all the disks are interleaved and addressed in order. Disk mirroring and disk striping can also be combined in a RAID array. storage space is divided into units ranging from 512 bytes up to several megabytes

Standard levels of RAID:

RAID 0

RAID 0 (also known as a stripe set or striped volume) splits ("stripes") data evenly across two or more disks, without parity information, redundancy, or fault tolerance. Since RAID 0 provides no fault tolerance or redundancy, the failure of one drive will cause the entire array to fail; as a result of having data striped across all disks, the failure will result in total data loss.

RAID 1

RAID 1 consists of an exact copy (or mirror) of a set of data on two or more disks; a classic RAID 1 mirrored pair contains two disks. This configuration offers no parity, striping of disk space across multiple disks since the data is mirrored on all disks belonging to the array.

RAID 5

RAID 5 consists of block-level striping with distributed parity. Parity information is distributed among the drives. It requires that all drives but one be present to operate. Upon failure of a single drive, subsequent reads can be calculated from the distributed parity such that no data is lost. RAID 5 requires at least three disks.

RAID 6

RAID 6 extends RAID 5 by adding another parity block. It uses block-level striping with two parity blocks distributed across all member disks. As in RAID 5, there are many layouts of RAID 6 disk arrays depending upon the direction the data blocks are written, the location of the parity blocks with respect to the data blocks, and whether or not the first data block of a subsequent stripe is written to the same drive as the last parity block of the prior stripe.

3.2 Architecture Diagram:

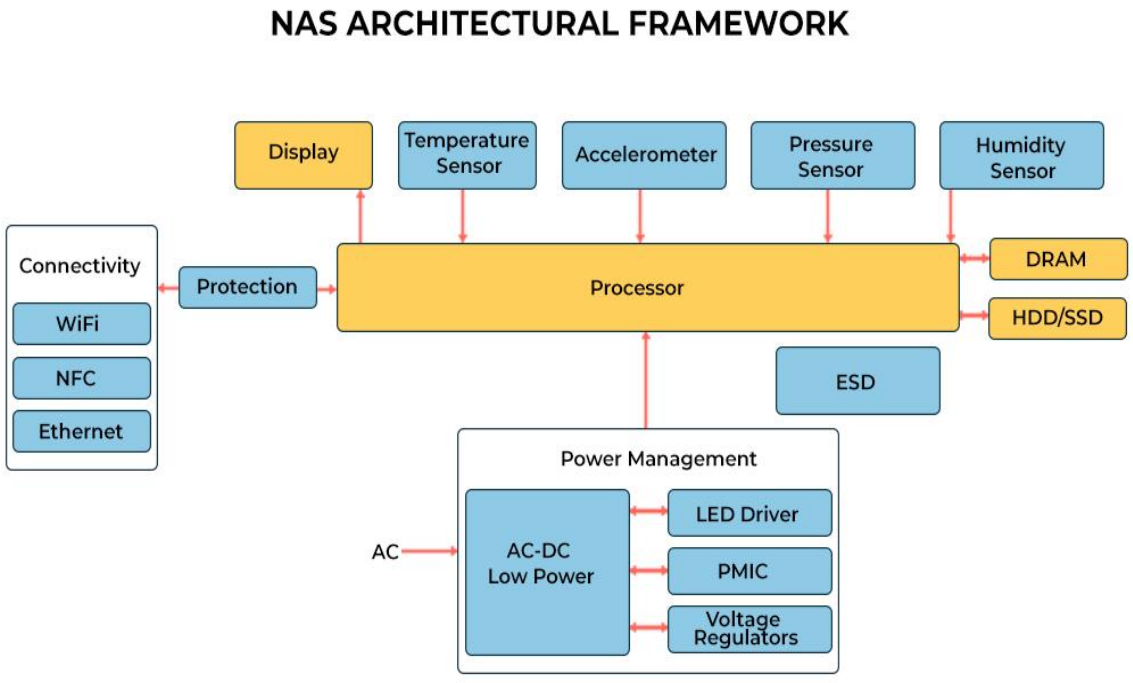

Figure 3.2.1 - NAS Architecture Diagram

3.3 Design Phase:

The Design Phase consists of the UML diagrams to design and construct the project.

- Use Case Diagram
- Data flow Diagram
- Deployment Diagram

3.4 Use Case Diagram:

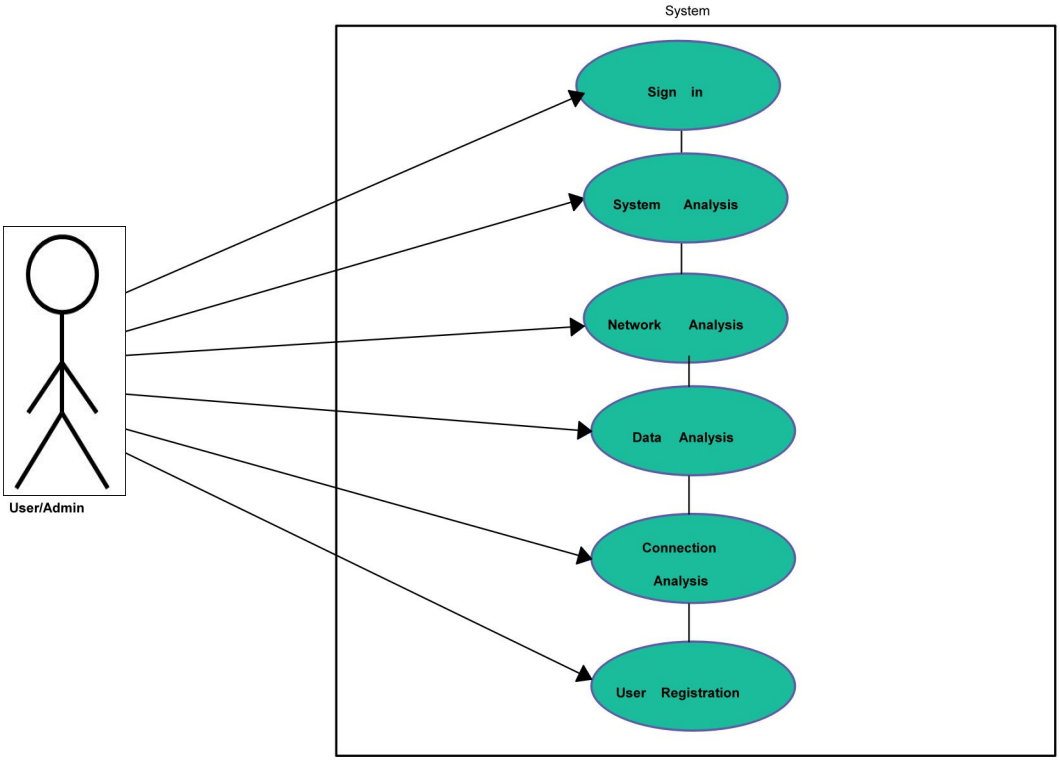

Figure 3.4.1 - NAS Simple Use Case Diagram

The use case diagram describes the basic steps of getting started to work with a NAS system starting from analysis of the system and the network conditions and functionality. The user/admin is also involved in the data present within the NAS and it's analysis too.

3.5 Data Flow Diagram:

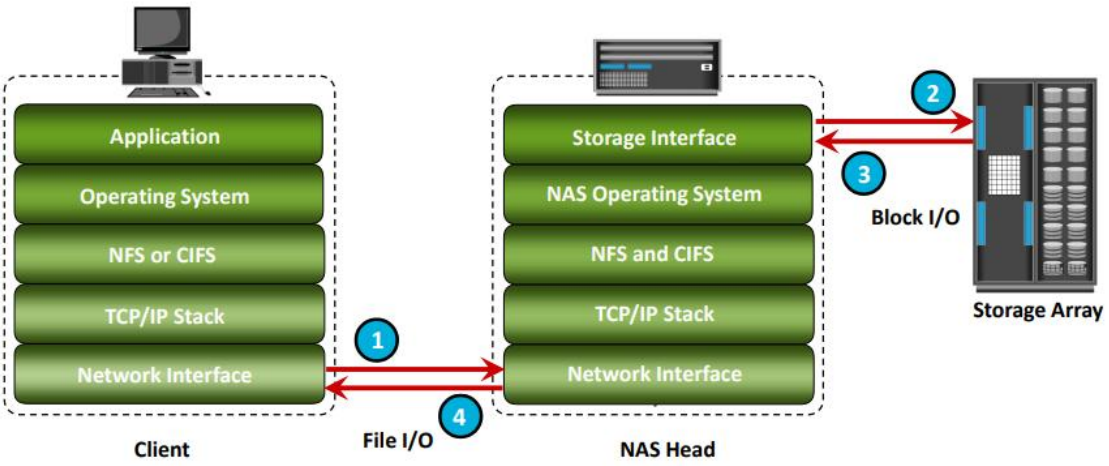

Figure 3.5.1 - Data Flow Diagram

The data flow diagram helps to identify the various fundamental layers of a NAS system including hardware, networking and the software. File transfers starts and end with the network interface which is then processed till the application layer.

3.6 Deployment Diagram:

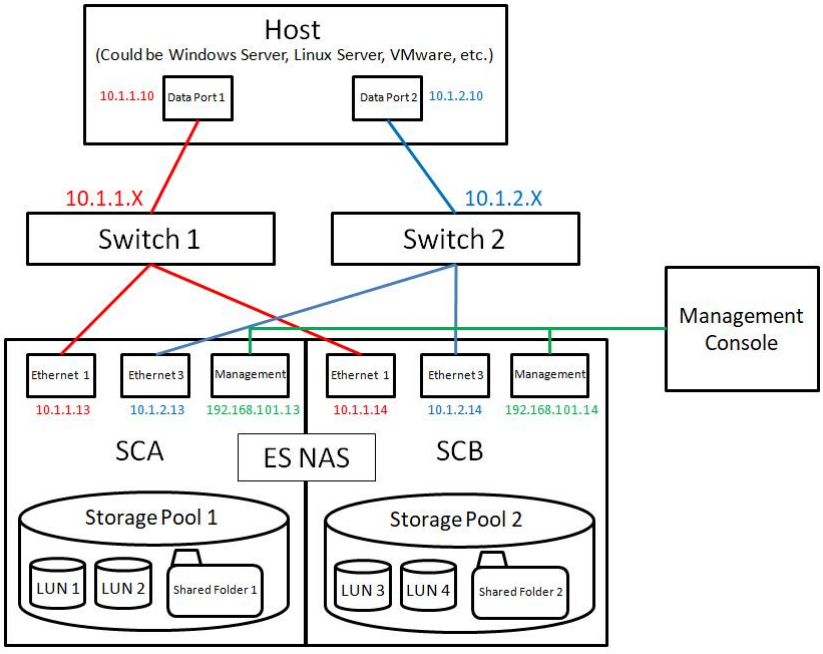

Figure 3.6.1 - Deployment Diagram

The deployment diagram showcases the multiple parts of networking needed to integrate a storage system to the internet in a functional manner with switched and various LAN ports.

3.7 Module Description:

The following are the main modules of this NAS based System

- Hardware Module
- Software Module

Figure 3.7.1 - enterprise NAS system

3.7.1 Hardware Module:

The hardware module consists of several sub modules which come under the hardware side of the NAS system.

1. NAS server: The NAS server is the central component of a NAS system. It is responsible for storing and sharing data. NAS servers come in a variety of form factors, from small desktop devices to large rack mount appliances.

Figure 3.7.1.1 - NAS Server

2. Storage drives: NAS systems typically use hard drives or solid-state drives to store data. The number and type of storage drives that you need will depend on the amount of data that you need to store and the performance requirements of your NAS system.

Figure 3.7.1.2 - Storage Drive for NAS

3. Network interface card (NIC): The NIC allows the NAS server to communicate with other devices on the network. NAS servers typically have at least one NIC, but some models have multiple NICs for high-availability and performance. The NIC must be capable enough to support the high volume of data transfer between nodes through the internet. It is recommended to use NIC which support Gigabit Ethernet or 10-Gigabit Ethernet depending on the internet plan purchased from the Internet Service Provider. Modern motherboards tend to come with NICs directly integrated into the board with some providing full 10-Gigabit Ethernet

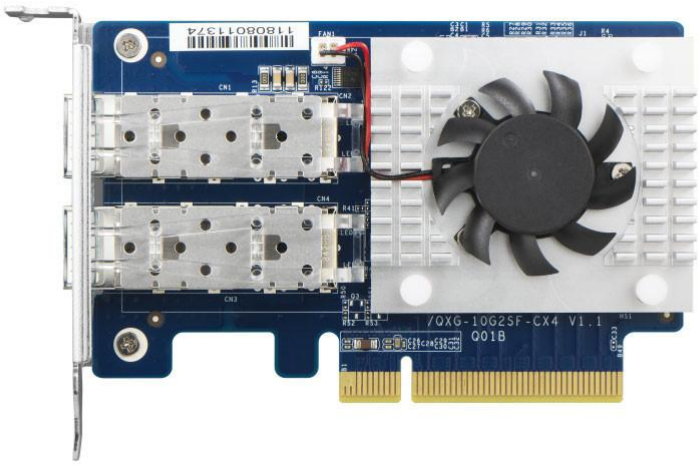

Figure 3.7.1.3 - NIC

4. Power supply: The power supply provides power to the NAS server and its components. It is important to choose a power supply that is powerful enough to meet the needs of your NAS system. A small NAS with a few TB of storage might need only around 500W or 600W power supply. Whereas a very large one with even a Petabyte of storage might need 1500W to 2000W of power to run the NAS safely

Figure 3.7.1.4 - Power Supply

5. Case: The case houses the NAS server and its components. NAS cases come in a variety of sizes and styles. It is important to choose a case that is large enough to accommodate all of the components of your NAS system and that has good airflow to prevent the components from overheating. The case of choice varies from a traditional ATX computer case to a more robust Rack mounted case which is used typically in large server grade applications.

Figure 3.7.1.5 - ATX Case for NAS

3.7.2 Software Module:

The software module can vary from each system because the use cases can differ a lot. So let us use the MoSCoW method to understand the requirements. MoSCoW prioritization is generally used in software developing but it can be used in different projects where a prioritization is important. As the budget and the time frame is limited in this research, some prioritization is needed to get the most business value out of from the project. MoSCoW – prioritization is generated from different priorities and they're as follows:

- M = Must have (Priority 4)
- S = Should have (Priority 3)
- C = Could have (Priority 2)
- W = Want to have (Priority 1)

According to the coded data, there are several demands which are graded as priority 4. These aspects are taken care of when designing the NAS-system for a small company as "Must have".

Priority 3 features and use cases are considered to be "Should have". These aspects of the system are considered to be quite essential and should be implemented to the final product if there are no good reasons why they should be left out.

Priority 2 features and use cases are considered as "Could have" meaning they're

not important aspects and can easily be left out from the final product. These aspects can be implemented if the budget and the time-frame allow.

Priority 1 features and use cases are considered as "Want to have" and they are

aspects that are not important at all and only if budget and time-frame allow, they

can be implemented.

CATEGORY	CODED DATA	PRIORITY
Use Cases	Reachable from home	4
	Reachable from LAN	4

	Reachable abroad	3
	Can be used from “My Computer” as a network drive	4
	Able to grant access for specific external clients	2
	Able to access NAS from public computers	3
Features	Constant backups from NAS to external device / location	4
	Able to have an own space in NAS	4
	Able to add users via web browser	2
	Able to remove users via web browser	2
	Able to share files in a common network drive	4
	Able to implement RAID for data safety	4

And so, when it comes to the operating system of choice, we have previously mentioned the two in the software requirements section. Both the choices mentioned here run on Debian Linux as their base because Linux is based on open source. Open-source software’s main idea is that its source code is free and, in this context, it means that it is open, public and non-proprietary. As the business needs to operate within a specific budget, there’s not much room for extra costs. Therefore, it was vital to find a solution which doesn’t only meet the requirements, but meets them in inexpensive way. It \is also vital to understand that in some use scenarios the better choice, as an operating system, would be Microsoft Windows Server. However, in this implementation there were no specific requirements which would have required Windows

1. TrueNAS Scale
2. OpenMediaVault (OMV)

Let us look at both of them in brief.

TrueNAS Scale is a free and open-source NAS operating system based on Debian Linux and the ZFS file system. It is a scale-out NAS solution that can be used to build clusters of NAS servers with capacities of up to hundreds of petabytes.

TrueNAS SCALE offers a wide range of features, including:

- Support for a variety of file sharing protocols: TrueNAS SCALE supports SMB, NFS, AFP, ZFS and iSCSI file sharing protocols.
- Advanced storage management features: TrueNAS SCALE includes a variety of advanced storage management features, such as ZFS pools, snapshots, and replication. These features allow users to protect their data from loss and corruption, and to efficiently manage their storage resources.
- Scalability: TrueNAS SCALE can be used to build clusters of NAS servers with capacities of up to hundreds of petabytes.
- High availability: TrueNAS SCALE can be configured in a high-availability configuration to ensure that data is always available, even if one of the NAS servers in the cluster fails.
- Appliance support: TrueNAS SCALE can be installed on a variety of hardware platforms, including bare metal servers, virtual machines, and pre-built NAS appliances.
- File sharing: TrueNAS SCALE can be used to share files between users on a network.
- Data backup and recovery: TrueNAS SCALE can be used to back up data from computers and other devices on a network.
- Media streaming: TrueNAS SCALE can be used to stream media files to devices on a network. This can be useful for homes and businesses that want to share their media collections.
- Virtualization: TrueNAS SCALE can be used to host virtual machines. This can help to consolidate IT resources and reduce costs.

OpenMediaVault (OMV)is also a free and open-source NAS operating system based on Debian Linux. It is a simple and easy-to-use solution that allows users to set up and manage a NAS system without having to have any prior Linux experience.

OMV offers a wide range of features, including:

- Support for a variety of file sharing protocols: OMV supports SMB, NFS, AFP, FTP, and iSCSI file sharing protocols.
- Advanced storage management features: OMV includes a variety of advanced storage management features, such as RAID, LVM, and file system support for BTRFS, XFS, and EXT4.
- Web-based user interface: OMV has a web-based user interface that makes it easy to manage the NAS system from any device with a web browser.
- Plugin system: OMV has a plugin system that allows users to extend the functionality of the NAS system with additional features. Some popular plugins include media servers, backup software, and virtualization software.

TrueNAS SCALE is a good choice for small businesses because it offers a number of advantages over OpenMediaVault.

- OpenZFS has supported Linux well. The issue tends to be that most Linux distributions don't include OpenZFS by default. TrueNAS SCALE changes this around by making OpenZFS the default file system and supporting it via APIs and WebUI. It is using Debian Linux and so there are drivers for newer chips and NICs. The ZFS side of SCALE has been very solid and gets extensively tested.
- The power consumption and performance of SCALE seems to be very similar to CORE. Most of the software (ZFS, SAMBA, TrueNAS Middleware) is the same/similar.
- It also has a more understandable and a rather intuitive, yet comprehensive user interface for set up and management utility. Any user who is looking to get into building their first NAS or an existing user, can feel at ease while trying to use TrueNAS Scale.
- Features that might require plugins to install on OMV is natively available to use on TrueNAS Scale. And the support is also proven to be robust as this software is used by millions of users for variety of use cases.

With that said, let us discuss the Software Module for the NAS system

A. NAS operating system (NOS): The NOS is the operating system that runs on the NAS server. We have chosen to go with TrueNAS Scale for reasons discussed previously.

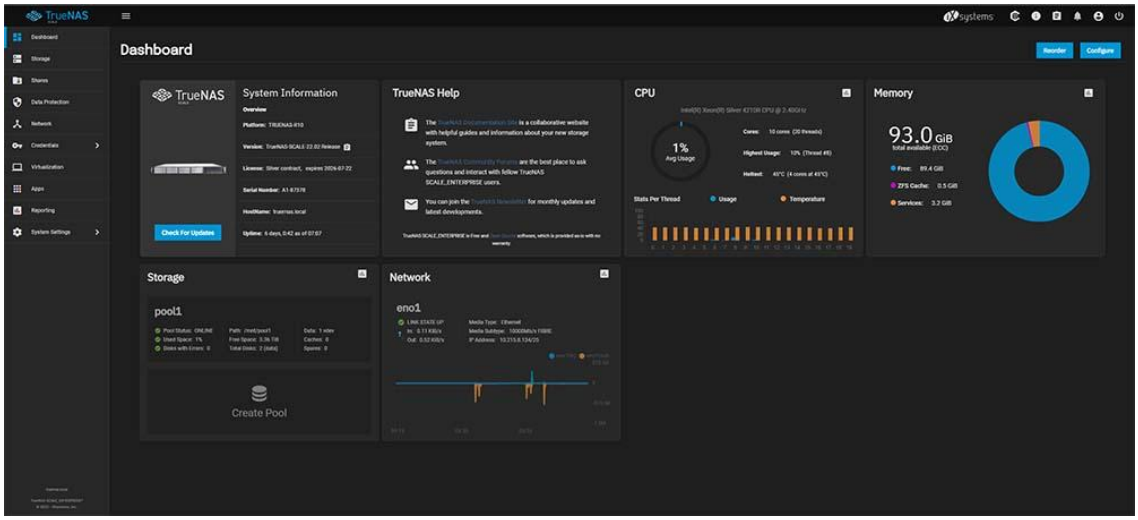

Figure 3.7.2.1 - TrueNAS Scale

B. File sharing protocols: File sharing protocols allow users to access files on the NAS system from their devices. ZFS is the file system of choice for our NAS implementation. ZFS uses a checksum algorithm to verify the integrity of all data written to the file system. This helps to protect data from corruption, even if there is a hardware failure or power outage. ZFS snapshots allow users to create point-in-time copies of their data. These snapshots can be used to restore data to a previous state or to create clones of data for testing or development purposes. ZFS replication allows users to replicate their data to another NAS system or to the cloud. This can help to protect data from loss or corruption in the event of a disaster. ZFS can compress data to save storage space. This can be especially beneficial for NAS systems that store large amounts of data.ZFS can encrypt data at rest to protect it from unauthorized access. This is important for NAS systems that store sensitive data.

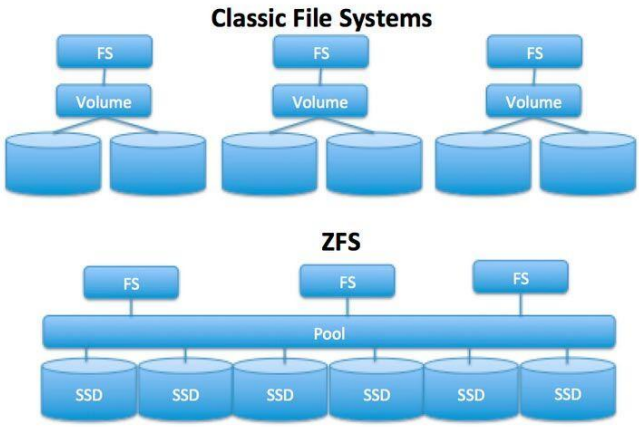

Figure 3.7.2.2 - Classical vs ZFS file system

C. Storage management tools: Storage management tools help users to manage their files and folders on the NAS system. TrueNAS has a built in Storage management system which can be used with their intuitive GUI

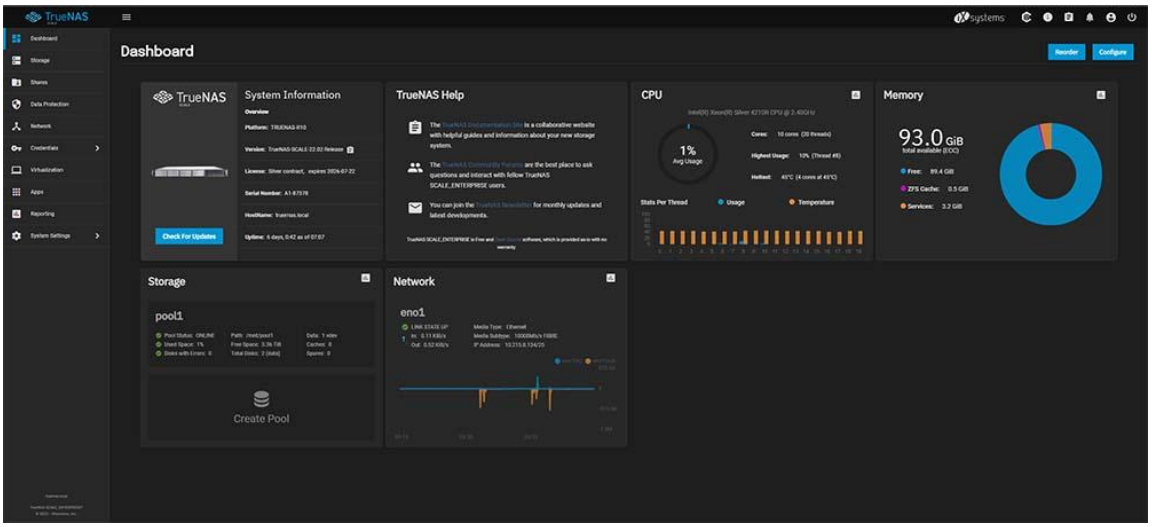

Figure 3.7.2.3 - Storage Management Utility in TrueNAS

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Discussion:

4.1.1 Scope and limitation

This study focuses on finding the best way to implement a NAS for a small company with a limited budget. This means that the best solution may not be suitable which utilizes industry leading policies, features and practices. The study discusses the NAS system itself, not the surrounding equipment such as the network and client computers. These aspects are to be discovered before the implementation of the NAS along with specific requirements regarding storage capacity, backup architecture, and more. Even the standard level of RAID to be implemented depends heavily on the user's intent and budget as it can affect the actual amount of usable storage. One option for further study is the other optional ways to implement NAS

4.1.2 Reliability and Validity:

Reliability means a statistical measurement of data about how reproducible it is. Since this research only created one type of NAS system with some alternative options for certain components and requirements, it is important to note that the reliability of the system is unknown. It is not possible to compare different hardware in real-world scenarios due to the limited time and budget. However, the facts used to support the decision in this thesis are from reliable sources. The NAS system should generate the same test results in various scenarios, depending on the intended use.

The validity in this research is confirmed by researching the findings in literature and by comparing the findings between the results in this study and the findings in literature.

4.1.3 Conclusion:

Information Technology (IT) is vital to modern businesses, even small ones. Businesses generate a lot of valuable data that needs to be stored securely. Centralized storage is a better solution than storing data locally on PCs because it offers more security and is easier to maintain. NAS is a good option for centralized storage because it is relatively inexpensive and easy to set up.

However, there may be other suitable solutions for your business, depending on your specific needs. The artifact discussed in this research is based on an industry-leading server technology that allows for more services than just file sharing.

NAS can be built on a server or on a ready-made NAS device that is designed for file sharing only. However, for the same budget, using industry-leading server technology can significantly improve fault tolerance.

In other words, if you build your own NAS using server components, you can get a more reliable and fault-tolerant system than if you buy a pre-built NAS device. This is because server components are typically designed to be more durable and reliable than consumer-grade components.

4.1.4 Future Study:

One option for further study is the other optional ways to implement NAS. This study discusses implementations which work on small scaled NAS or home-environment applications. However, with a higher price tag, there are also NAS-devices which are usually rack mountable and decently fault-tolerant. It would be a good idea to conduct a further study about the differences between pure NAS-devices and servers like in this research.

CHAPTER 5

REFERENCES

1. Network-Attached Storage For Small Companies, Case: Design Foundation Finland, Business Information Technology, Bachelor Thesis Autumn 2012, Jari-Pekka Koivisto
2. Network-Attached Storage: Data Storage Applications, Turkish Journal of Computer and Mathematics Education, V ol.12 No.12 (2021), 2385-2396 Research Article, Ravi Kumar M G, Ayudh Nagaraj, Benjamin Paul, Sharat P Dixit.
3. Cloud computing: implementing nas in directing digital images without local storage, Proceedings of International Conference "ICSEM'13", J.C. Rosaline Christi, R. Illakia, M.Gowri
4. Harnessing RAID mechanism for enhancement of data storage and security on cloud, Sudipta Sahana, Rajesh Bose and Debabrata Sarddar, Brazillian Journal of science and technology
5. Network Attached Storage Architecture, Communications Of The ACM November 2000/Vol. 43, No. 11, Garth A. Gibson and Rodney Van Meter.
6. Dataplexed Network Attached Storage System, Volume : 3, Issue : 2, February 2014. Praveen Vijai, Minnu Meria, Midhu Babu.
7. NAS (Network Attached Storage), International Research Journal of Engineering and Technology (IRJET), Volume: 07 Issue: 04 | Apr 2020, Siddh Mistry, Janki Prajapati, Mohini Patel, Ms. Swati. S. Saxena.
8. Hitachi NAS Storage for Unstructured Data, Silverton Consulting, Inc. StorInt Briefing, 2021.
9. Dell Unity: NAS Capabilities, Dell Technologies, April 2022.
10. NAS or iSCSI? Selecting a storage system, White Paper 2006, Open-E.