IMPLEMENTATION OF NAS USING RAID CONFIGURATION

ABDUL SAMAD (RA2111003020190)

N.S.SANTHOSH (RA2111003020209)

BALAJI PANDI. R (RA2111003020205)

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. 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.

INTRODUCTION

NCP protocol was released in 1983 . Release of NFS in 1984 allowed network servers to share their storage space with networked clients. 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.

• 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.

• Both DAS and NAS can potentially increase availability of data by using RAID or clustering

• 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.

LITERATURE REVIEW

The implementation of NAS system has been well studied upon and is being used currently by larger businesses at enterprise levels. But there has not been extensive research made on the small scale adaptation of the NAS system. Small business fear to experiment with such new storage systems because of lack of in depth research and reliability. Recent developments have been quick and the technology has been improving at a steady rate. They will prove to be significant changes in the future implementations of NAS. One such development is using multiple SSD for storage instead of HDD by using PCIe technology. This can allow for even more performance and also create new use cases for the future. This development is not being accounted for in research as well even if early adopters are quick to utilize such upgraded components. The strengths of my research accounts for the small scale adaptation possibilities and how it can be performed after researching various journals and sources across the web.

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.

1. Accessibility: NAS are most used and preferred mainly for their ease of access over Ethernet.

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.

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.

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.

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. 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. 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.

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.

2.Compliance challenges: Small businesses in certain industries may be required to comply with specific regulations regarding data storage and security.

3.Lack of control: Small businesses may have less control over their data and applications when using cloud-based storage solutions.

4.Hidden costs: Small businesses should be aware of the hidden costs associated with cloud- based storage solutions, such as egress fees and overage fees.

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.

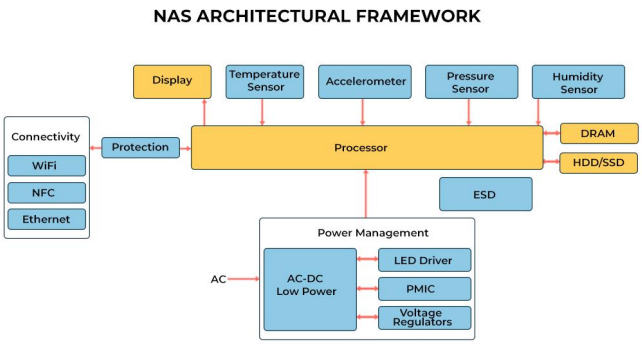
PROPOSED SYSTEM

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-effectivethan cloud-based storage solutions in the long term, especially for businesses that have large amounts of data to store.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 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. Using multiple disks increases the mean time to failure, while storing data redundantly also increases fault tolerance. It works on single logical drive on a OS. 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.



MODULE DESCRIPTION

HARDWARE:

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.

2. Storage drives: NAS systems typically use hard drives or solid-state drives (SSDs) 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.

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 ISP. Modern motherboards tend to come with NICs directly integrated into the board with some providing full 10-Gigabit Ethernet

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.

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.

SOFTWARE:

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:

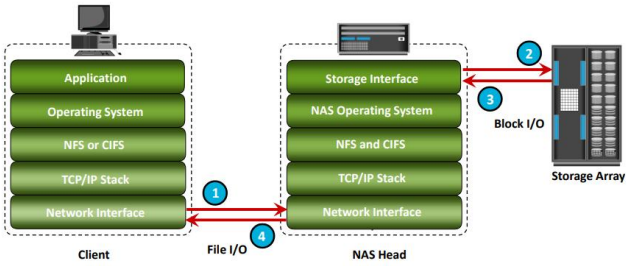
M = Must have (Priority 4)

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.



TRUE NAS SCALE AND OPEN MEDIA VAULT

True NAS Scale: It 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. 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 (HA)

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): It 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.

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.

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.

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.

ANALYSIS AND FINDINGS

Data loss can happen with any hardware device, so we implement software to work with hardware disks, cloud clients, and servers to reduce time, increase efficiency, and avoid data loss due to poor backups.Gigabit networks can become congested, so the most effective way to improve network performance is to increase bandwidth or improve network utilization. We can also strengthen backups by using cloud computing or virtual storage. A more efficient network file system can also boost NAS performance.

For a generic NAS architecture with a Gigabit network, the processing power of a CPU is not a serious issue. Hardware methods are unnecessary

to reduce CPU utilization, and can even have a negative impact on throughput due to small file accesses in NAS.

Software RAID is sufficient to meet the performance requirements of most basic NAS applications. Adding more disk drives to a NAS with a Gigabit network will only achieve limited performance improvement, except for expanding storage capacity when the network is saturated.

DISCUSSION

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. 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.

LIMITATIONS AND FUTURE SCOPE

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. 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.

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.