

WELCOME TO MODULE 9

DISK MANAGEMENT



DISK MANAGEMENT

- Section I- What's Computer Storage
- Section II- Type of Computer Storage
- Section III- How to Add Disk
- Section IV- Extend and Existing Disk
- Section V- Disk Cleanup and Defragmentation.
- Section VI- RAID
- Section VI- Windows Server Backup

Section I- What's Computer Storage

- Storage is a process through which digital data is saved within a data storage device by means of computing technology. Storage is a mechanism that enables a computer to retain data, either temporarily or permanently. ... Storage may also be referred to as computer data storage or electronic data storage.

- Computer storage is a technology consisting of computer components and recording media that are used to retain digital data. It is a core function and fundamental component of computers
- Operating system of any computer is installed on a storage
- Storage is needed to store short term and long term data



We will learn...

- How to identify our computer storage
- Storage mapping and partitioning.

Section II- Type of Computer Storage

- There are 4 different type of computer storage.
- Local Storage
- Direct Attach Storage (DAS)
- Storage Area Network (SAN)
- Network Attached Storage (NAS)



1. Local storage

- Floppy disk



- CD/DVD disk



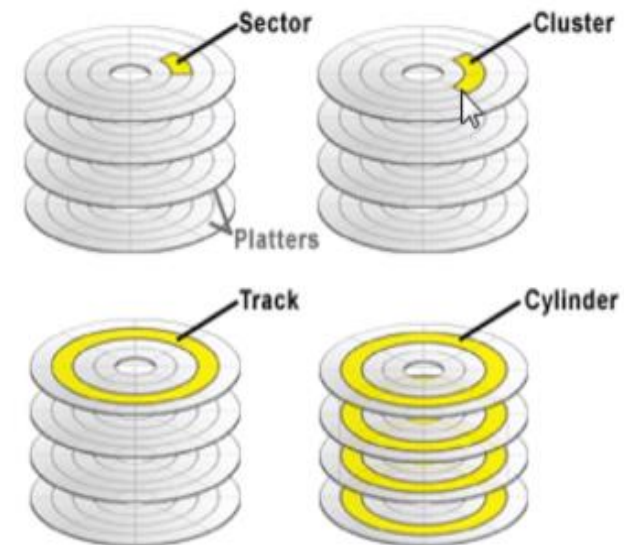
- RAM



- Hard disk



Disk is also referred to as a diskette, is a hard round, flat, and magnetic platter capable of having information read



- 2nd Type of Storage

2. Direct Attached Storage

- Tape drive



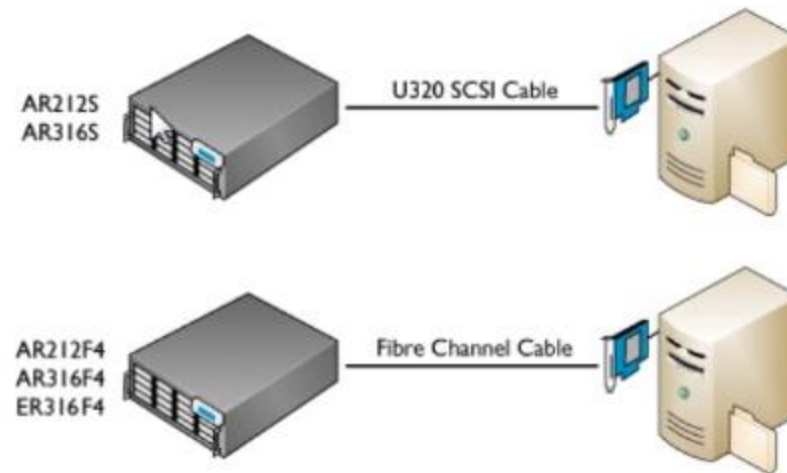
- External disk



- USB flash drive



- Storage Array



- 2nd Type of Storage

2. Direct Attached Storage

- Tape drive



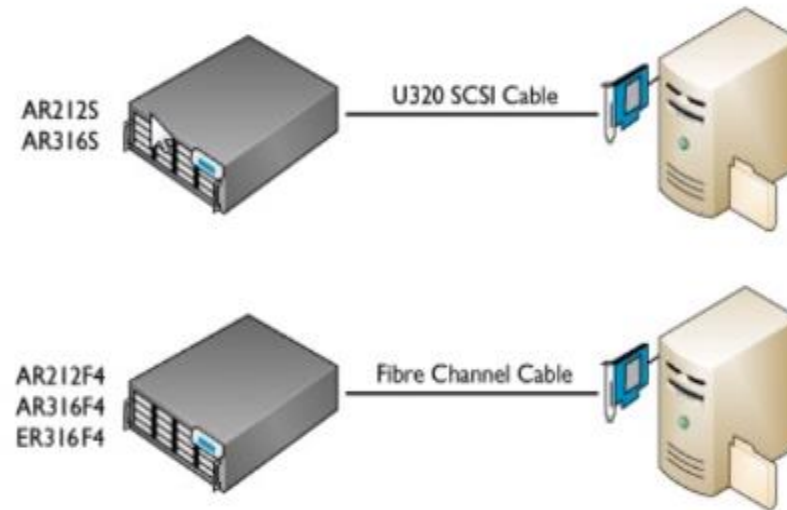
- External disk



- USB flash drive



- Storage Array



- Third Type of Storage

3. SAN (Storage Area Network)



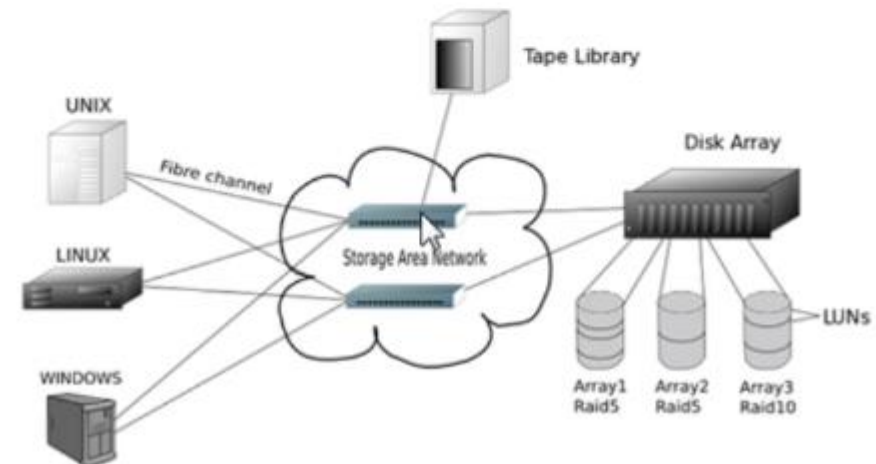
- SANs are primarily used to enhance storage devices, such as disk arrays and tape libraries, accessible to servers so that the devices appear to the operating system as locally attached devices
- Mostly the SAN storage is attached over:
 - iSCSI
 - Fiber Channel



iSCSI



Fiber

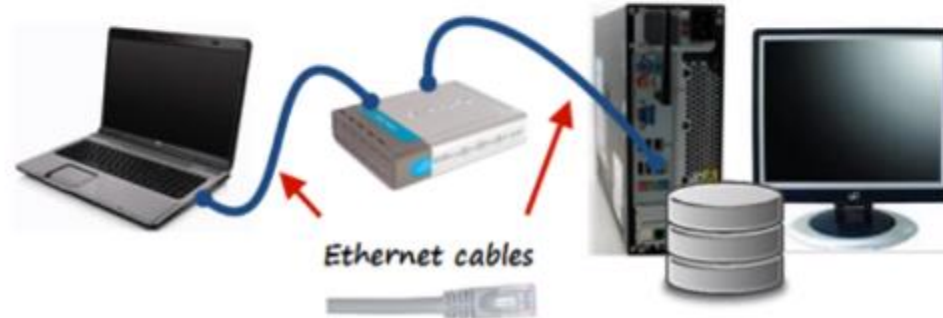


- Forth Type of Storage

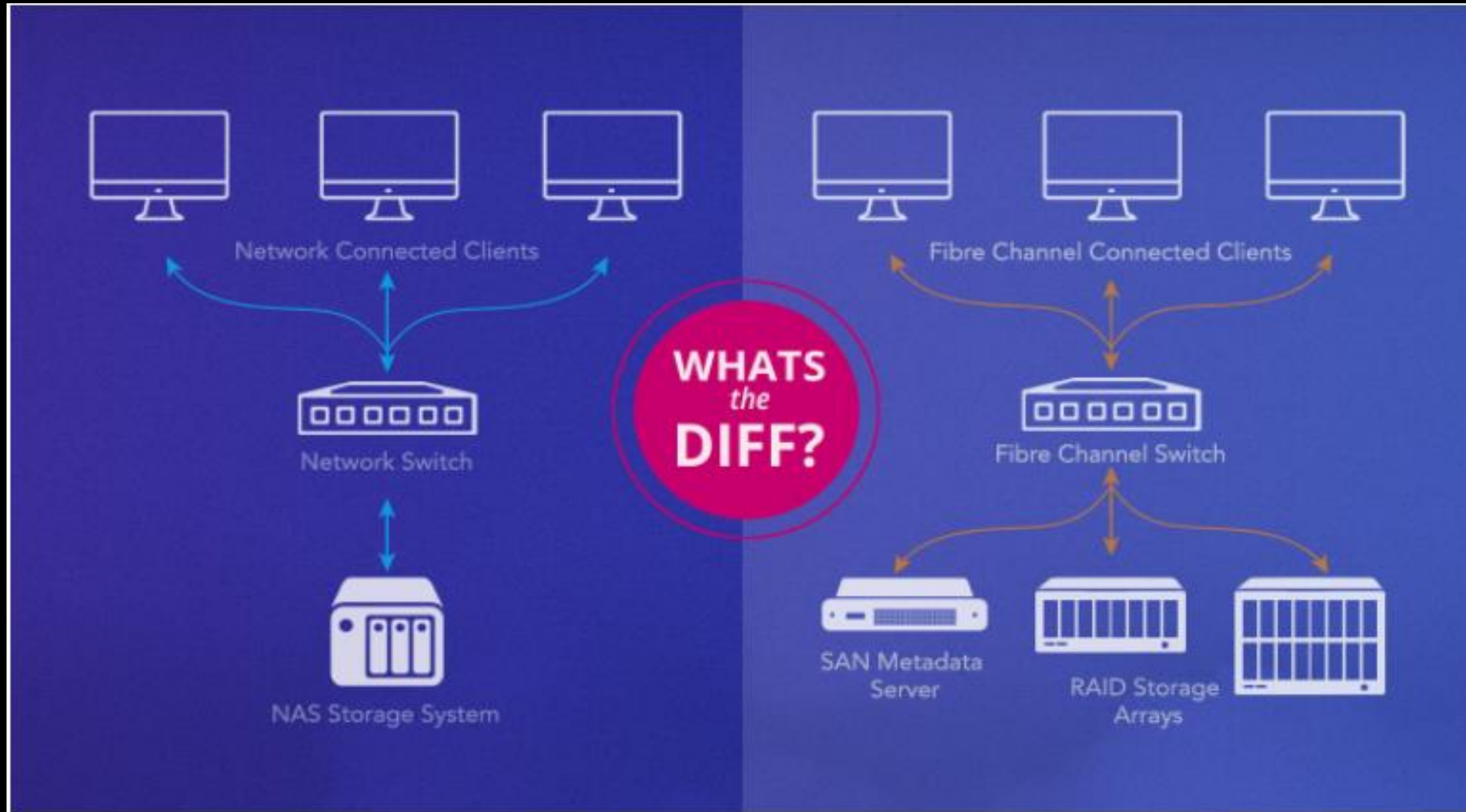
4. NAS (Network Attached Storage)



- Network-attached storage is a file-level computer data storage server connected to a computer network providing data access to a heterogeneous group of clients. NAS is specialized for serving files either by its hardware, software, or configuration
- In simple words: NAS allows computer to attach external storage over network (TCP/IP)
- Example of NAS storage:
 - Samba
 - NFS



NAS VS SAN



NAS

- NAS is a single storage device that serves files over Ethernet and is relatively inexpensive and easy to set up, while a SAN is a tightly coupled network of multiple devices that is more expensive and complex to set up and manage. From a user perspective, the biggest difference between NAS and SAN is that NAS devices deliver shared storage as network mounted volumes and use protocols like NFS and SMB/CIFS, while SAN-connected disks appear to the user as local drives.

Benefits of NAS

A NAS is frequently the next step up for a home office or small business that is using external hard drives or direct attached storage, which can be especially vulnerable to drive failure. The move up to NAS is driven by the desire to share files locally and remotely, having files available 24/7, achieving data redundancy, having the ability to replace and upgrade hard drives in the system, and most importantly, supports integrations with cloud storage that provides a location for necessary automatic data backups.

Summary of NAS Benefits:

Relatively inexpensive.

A self-contained solution.

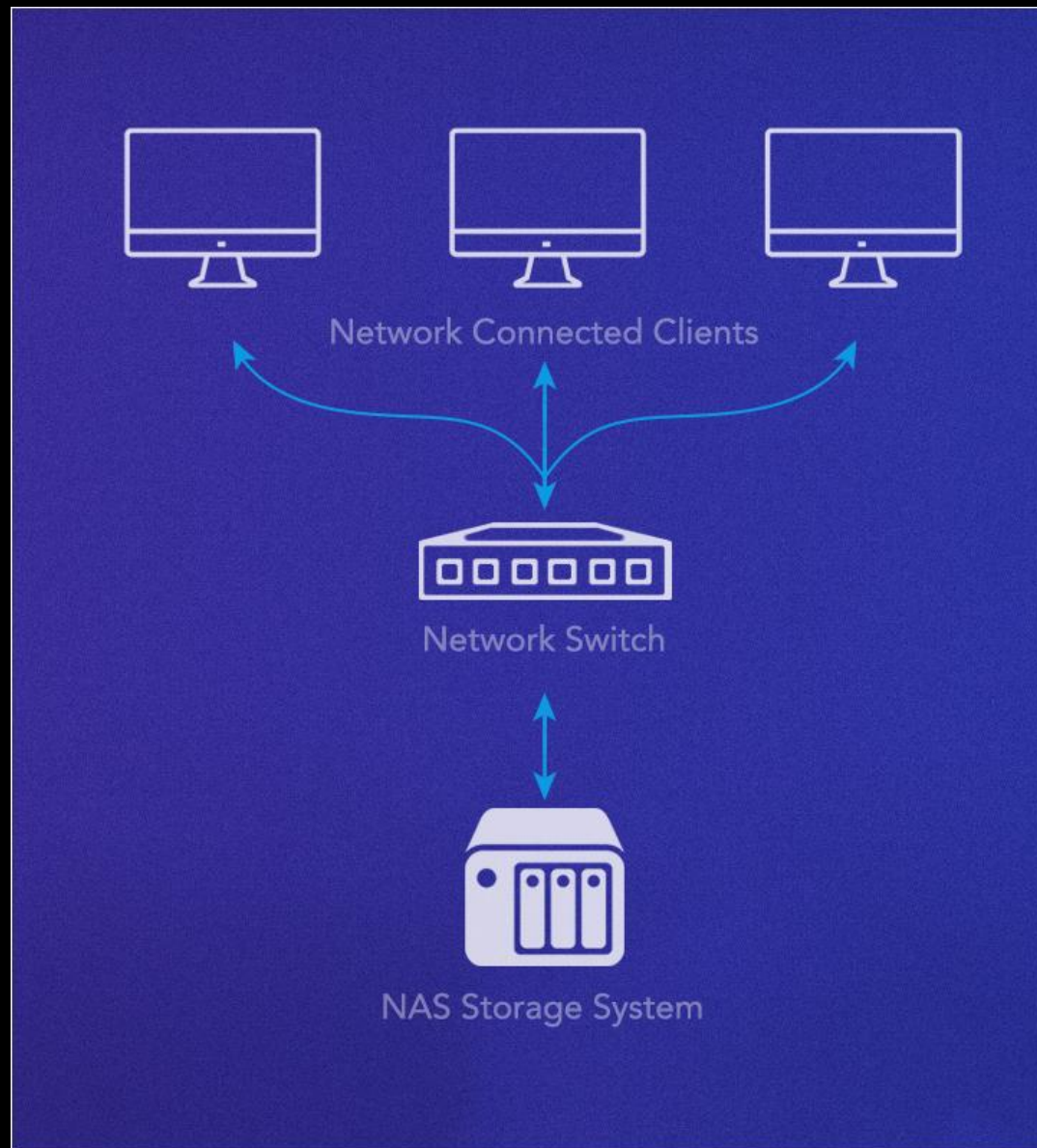
Ease of administration.

24/7 and remote data availability.

Wide array of systems and sizes to choose from.

Drive failure-tolerant storage volumes.

Automatic backups to other devices and the cloud.



The NAS system and clients connect via your local network—all file service occurs via Ethernet.



NAS with eight drive bays for 3.5" disk drives.

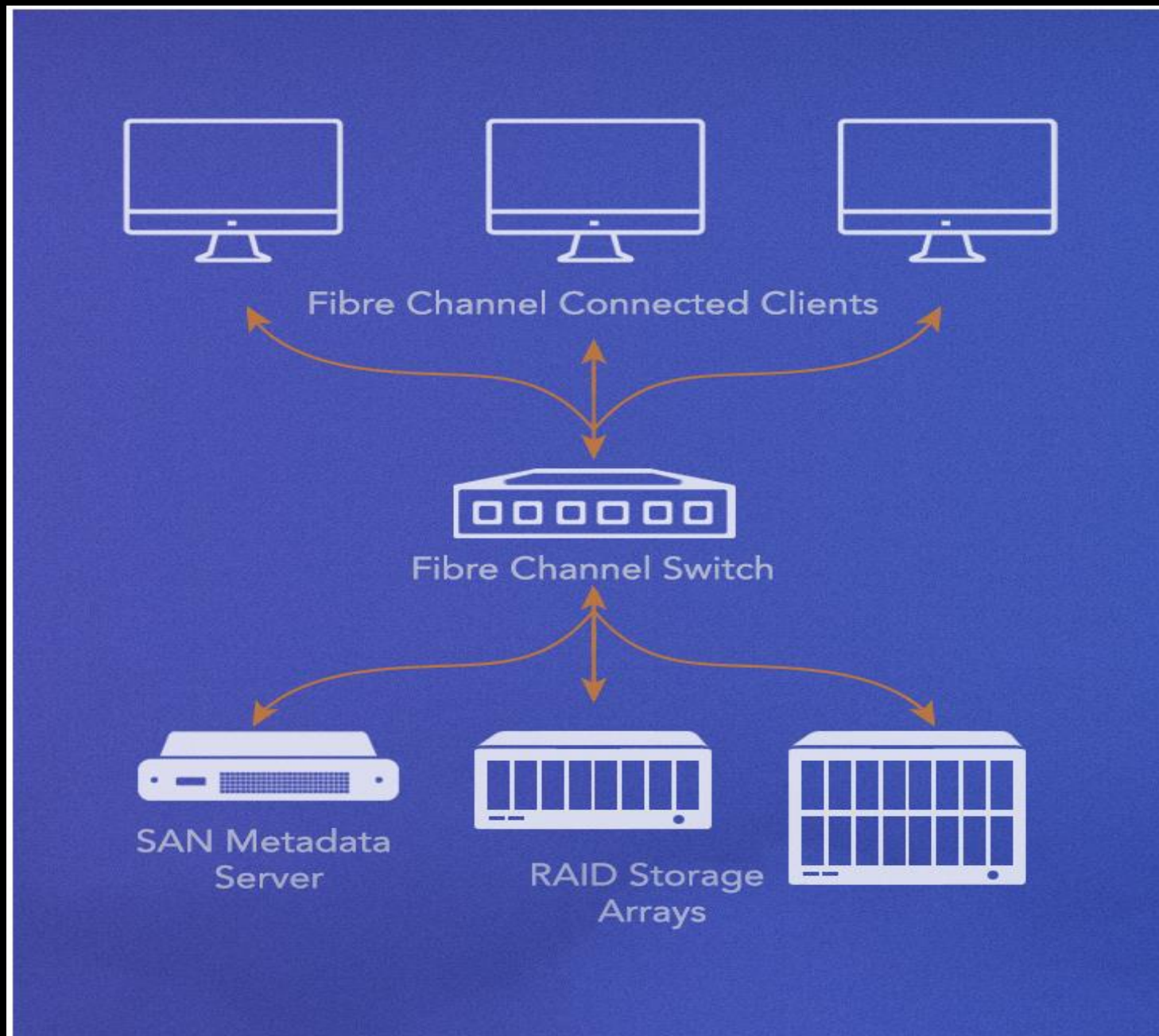
Limitations of NAS

The weaknesses of a NAS are related to scale and performance. As more users need access, the server might not be able to keep up. At this point, you might be able to expand the storage, but most likely it will need to be replaced with a more powerful system with a bigger on-board processor, more memory, and faster and larger network connections. The other weakness is related to the nature of Ethernet itself. By design, Ethernet transfers data from one place to another by dividing the source into a number of segments called packets and sending them along to their destination. Depending on existing network traffic or issues, any of those packets could be delayed or sent out of order, and the file might not be available to the user until all of the packets arrive and are put back in order.

Any latency (slow or retried connections) is usually not noticed by users for small files, but can be a major problem in demanding environments such as video production, where files are extremely large, and latency of more than a few milliseconds can disrupt production such as video editing.

SAN

- A SAN is a way to provide users high-performance, low-latency shared access to storage. A SAN is built from a combination of servers and storage over a high speed, low latency interconnect that allows direct Fiber Channel connections from the client to the storage volume to provide the fastest possible performance. The SAN may also require a separate, private Ethernet network between the server and clients to keep the file request traffic out of the Fiber Channel network for even more performance. A SAN is a flexible way to deliver shared storage for a number of users in demanding applications, like video editing or multiple application servers. By joining together the clients, SAN server, and storage on a Fiber Channel network, the SAN volumes appear and perform as if it were a directly connected hard drive. Storage traffic over Fiber Channel avoids the TCP/IP packetization and latency issues, as well as any local area network congestion, ensuring the highest access speed available for media and mission critical stored data.



The SAN management server, storage arrays, and clients all connect via a Fibre Channel network—all file serving occurs over Fibre Channel.

Benefits of SAN

- Because it's considerably more complex and expensive than NAS, SAN is typically used by large corporations and requires administration by an IT staff. For some applications, such as video editing, it's especially desirable due to its high speed and low latency.
- The primary strength of SAN is that it allows simultaneous shared access to shared storage that becomes faster with the addition of storage controllers. For example, hundreds of video editors can use 10's GB/s of storage simultaneously. For this reason, SAN is widely used in collaborative video production environments.

Summary of SAN Benefits.

- Extremely fast data access with low latency.
- Relieves stress on a local area network.
- Can be scaled up to the limits of the interconnect.
- OS level (“native”) access to files.
- Often the only solution for demanding applications requiring concurrent shared access.

Limitation of SAN

- The challenge of SAN can be summed up in its cost and administration requirements—having to dedicate and maintain both a separate Ethernet network for metadata file requests and implement a Fibre Channel network can be a considerable investment. That being said, a SAN is often the only way to provide very fast data access for a large number of users that also can scale to supporting hundreds of users at the same time.

Main Difference Between NAS and SAN

NAS	SAN
Often used in homes and small to medium sized businesses.	Often used in professional and enterprise environments.
Less expensive.	More expensive.
Easier to manage.	Requires more administration.
Data accessed as if it were a network-attached drive.	Servers access data as if it were a local hard drive.
Speed dependent on local TCP/IP Ethernet network, typically 1GbE to 10GbE, and affected by the number of other users accessing the storage at the same time. Generally slower throughput and higher latency due to the nature of Ethernet packetization, waiting for the file server, and latency in general.	High speed using Fibre Channel, most commonly available in 16 Gb/s to 32 Gb/s. Fibre Channel can be delivered via high speed Ethernet such as 10Gb or 40Gb+ networks using protocols such as FCoE and iSCSI.
SMB/CIFS, NFS, SFTP, and WebDAV.	Fibre Channel, iSCSI, FCoE.
Lower-end not highly scalable; high-end NAS scale to petabytes using clusters or scale-out nodes.	Can add more storage controllers, or expanded storage arrays allowing SAN admins to scale performance, storage, or both.

Simply connects to your existing Ethernet network.	Requires dedicated Fibre Channel connections for clients, servers, and storage, and often a separate, dedicated Ethernet network for file request traffic.
Entry level systems often have a single point of failure, e.g. power supply.	Fault tolerant network and systems with redundant functionality.
Subject to general Ethernet issues.	Behavior is more predictable in controlled, dedicated environments.
A central place from which to back up or sync user files.	A place from which to archive files not immediately needed.

When considering NAS or SAN, you might find it helpful to think of it this way: NAS is simple to set up, easy to administer, and great for general purpose applications. Meanwhile, SAN can be more challenging to set up and administer, but it's often the only way to make shared storage available for mission critical and high performance applications.

Section III- How to Add Disk

- The requirement for adding a disk:
 - Disk space issue
 - High availability
 - Application requirement for additional disk



We will learn...

- How to add a disk from VirtualBox (Virtual disk)
- Format the new disk
- Map or mount the new disk
- Verification.

Section IV- Extend and Existing Disk

- Disk extension occurs when existing disk run out of disk space



We will learn...

- How to add a disk from VirtualBox (Virtual disk)
- Format the new disk
- Extend an existing disk ($5\text{G} + 5\text{G} = 10\text{G}$).

Section V- Disk Cleanup and Defragmentation

- Disk Cleanup is a computer maintenance utility included in Microsoft Windows designed to free up disk space on a computer's hard drive. The utility first searches and analyzes the hard drive for files that are no longer of any use, and then removes the unnecessary files
- Disk Defragmenter is a utility offered by Microsoft Windows to rearrange files on a disk to occupy continuous storage space

We will learn...

- How to perform disk cleanup
- How to run disk defragmentation.



Fragmented Disk



Defragmented Disk

Section VI- RAID

- RAID is a data storage virtualization technology that combines multiple physical disk drive components into one or more logical units for the purposes of data redundancy, performance improvement, or both.
- RAID is a technology that is used to increase the performance and/or reliability of data storage. The abbreviation stands for either Redundant Array of Independent Drives or Redundant Array of Inexpensive Disks, which is older and less used. A RAID system consists of two or more drives working in parallel. These can be hard discs, but there is a trend to also use the technology for SSD (Solid State Drives). There are different RAID levels, each optimized for a specific situation. These are not standardized by an industry group or standardization committee. This explains why companies sometimes come up with their own unique numbers and implementations.

Here we will cover the following RAID levels:

- RAID 0
- RAID 1
- RAID 5
- RAID 6
- RAID 10

Below is an overview of the most popular RAID levels:

- [RAID level 0](#)

In a RAID 0 system data are split up into blocks that get written across all the drives in the array. By using multiple disks (at least 2) at the same time, this offers superior I/O performance. This performance can be enhanced further by using multiple controllers, ideally one controller per disk.

- [Advantages of RAID 0:](#)

RAID 0 offers great performance, both in read and write operations. There is no overhead caused by parity controls.

All storage capacity is used, there is no overhead.

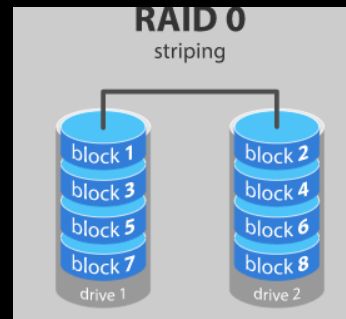
The technology is easy to implement.

- [Disadvantages of RAID 0](#)

RAID 0 is not fault-tolerant. If one drive fails, all data in the RAID 0 array are lost. It should not be used for mission-critical systems.

- [Ideal use](#)

RAID 0 is ideal for non-critical storage of data that have to be read/written at a high speed, such as on an image retouching or video editing station.



Below is an overview of the most popular RAID levels:

- RAID level 1

Data are stored twice by writing them to both the data drive (or set of data drives) and a mirror drive (or set of drives). If a drive fails, the controller uses either the data drive or the mirror drive for data recovery and continuous operation. You need at least 2 drives for a RAID 1 array.

- Advantages of RAID 1

RAID 1 offers excellent read speed and a write-speed that is comparable to that of a single drive.

In case a drive fails, data do not have to be rebuild, they just have to be copied to the replacement drive.

RAID 1 is a very simple technology.

- Disadvantages of RAID 1

The main disadvantage is that the effective storage capacity is only half of the total drive capacity because all data get written twice.

Software RAID 1 solutions do not always allow a hot swap of a failed drive. That means the failed drive can only be replaced after powering down the computer it is attached to. For servers that are used simultaneously by many people, this may not be acceptable. Such systems typically use hardware controllers that do support hot swapping.

Below is an overview of the most popular RAID levels:

- RAID level 5

RAID 5 is the most common secure RAID level. It requires at least 3 drives but can work with up to 16. Data blocks are striped across the drives and on one drive a parity checksum of all the block data is written. The parity data are not written to a fixed drive, they are spread across all drives, as the drawing below shows. Using the parity data, the computer can recalculate the data of one of the other data blocks, should those data no longer be available. That means a RAID 5 array can withstand a single drive failure without losing data or access to data. Although RAID 5 can be achieved in software, a hardware controller is recommended. Often extra cache memory is used on these controllers to improve the write performance.

- Advantages of RAID 5

Read data transactions are very fast while write data transactions are somewhat slower (due to the parity that has to be calculated).

If a drive fails, you still have access to all data, even while the failed drive is being replaced and the storage controller rebuilds the data on the new drive.

- Disadvantages of RAID 5

Drive failures have an effect on throughput, although this is still acceptable.

This is complex technology. If one of the disks in an array using 4TB disks fails and is replaced, restoring the data (the rebuild time) may take a day or longer, depending on the load on the array and the speed of the controller. If another disk goes bad during that time, data are lost forever.

- Ideal use

RAID 5 is a good all-round system that combines efficient storage with excellent security and decent performance. It is ideal for file and application servers that have a limited number of data drives.

Below is an overview of the most popular RAID levels:

- RAID level 10 – combining RAID 1 & RAID 0

It is possible to combine the advantages (and disadvantages) of RAID 0 and RAID 1 in one single system. This is a nested or hybrid RAID configuration. It provides security by mirroring all data on secondary drives while using striping across each set of drives to speed up data transfers.

- Advantages of RAID 10

If something goes wrong with one of the disks in a RAID 10 configuration, the rebuild time is very fast since all that is needed is copying all the data from the surviving mirror to a new drive. This can take as little as 30 minutes for drives of 1 TB.

- Disadvantages of RAID 10

Half of the storage capacity goes to mirroring, so compared to large RAID 5 or RAID 6 arrays, this is an expensive way to have redundancy.

What about RAID levels 2, 3, 4 and 7?

These levels do exist but are not that common (RAID 3 is essentially like RAID 5 but with the parity data always written to the same drive). This is just a simple introduction to RAID-systems. You can find more in-depth information on the pages of Wikipedia or ACNC.

Section VI- Windows Server Backup

4 Different Types of Backups

1. System backup (entire image using tools such as acronis, Veeam, Commvault etc.)
 2. Application backup (3rd party application backup solution)
 3. Database backup (Oracle dataguard, SQL backup etc.)
 4. Filesystem backup (tar, gzip directoris etc.)
-
- Backup and Restore is a component of Microsoft Windows introduced in Windows Vista and included in later versions that allow users to create backups and restore from backups created earlier.
 - Backup and Restore supports two different types of backup: **file level backup** and **system image**.



Section VI- Windows Server Backup

- RAID is NOT a substitute for back-up!. Back-up will come in handy if all drives fail simultaneously because of a power spike. It is a safeguard when the storage system gets stolen.
- Back-ups can be kept off-site at a different location. This can come in handy if a natural disaster or fire destroys your workplace
- The most important reason to back-up multiple generations of data is user error. If someone accidentally deletes some important data and this goes unnoticed for several hours, days or weeks, a good set of back-ups ensure you can still retrieve those files
- Backup also provides protection against operating system failure



We will learn...

- How perform system backup
- Restore from a previous backup.