3. storage

SAN and NAS

SAN Storage Area Network

It is a high-speed network of storage devices that provides block-level access to data for multiple servers. **mean you get space like 2tb and you can do anything install os make partition and do your work.**
SANs are typically used in enterprise environments to centralize storage resources and improve data availability, scalability, and performance.

SANs use specialized switches, host bus adapters (HBAs), and storage arrays to create a dedicated network for storage traffic. They support various protocols such as Fibre Channel, iSCSI, and FCoE (Fibre Channel over Ethernet) to transfer data between servers and storage devices.

SANs offer several advantages over traditional direct-attached storage (DAS) or network-attached storage (NAS) solutions. They provide higher bandwidth, lower latency, and better redundancy and fault tolerance, making them ideal for mission-critical applications that require high availability and fast data access. However, SANs can be more complex and expensive to set up and manage than other storage solutions.

SANs are commonly used in industries such as finance, healthcare, media and entertainment, and scientific research, where large amounts of data need to be stored, processed, and shared across multiple servers and applications.

NAS (Network Attached Storage)

NAS stands for Network Attached Storage. It is a type of storage device that is connected to a network and provides data access to multiple clients on the network.

- **In NAS you only get a just a folder to save or backup your file you can't install os and all like SAN.**
  NAS devices typically run a specialized operating system and can be accessed using standard network protocols such as NFS (Network File System) or SMB (Server Message Block). They are often used for file sharing, data backup, and media streaming in small to medium-sized businesses and homes. NAS devices come in various form factors, including standalone boxes and integrated components in servers and routers. They offer advantages such as scalability, ease of use, and

cost-effectiveness compared to traditional direct-attached storage (DAS) or storage area network (SAN) solutions.

Feature	NAS	SAN
Connection	Ethernet	Fibre Channel
Protocol	TCP/IP	SCSI
File System	Yes	No
Scalability	Limited	High
Performance	Lower	Higher
Cost	Lower	Higher
Use Case	File sharing, backup, and multimedia storage	Block-level storage for databases and high-performance applications

Structure data / unstructured data

Structured Data	Unstructured Data
Organized and stored in a predefined format	Does not have a predefined format or organization
Examples: customer information, sales transactions, financial data, inventory records	Examples: emails, social media posts, audio and video recordings, images, text documents
Stored in relational databases, spreadsheets, or other tabular formats	Stored in a more free-form format
Easily searchable, sortable, and analyzable	Requires more sophisticated tools and techniques to extract meaning and value
Used for business intelligence, reporting, and data mining applications	Used for content management, search, and analytics applications
Examples of tools: SQL, Excel, Tableau	Examples of tools: text analytics, natural language processing, machine learning
Data is organized and accessed efficiently	Data is more complex and challenging to manage

RAID system

RAID stands for **Redundant Array of Independent Disks**. RAID is a means of storing the same data in different locations on multiple hard disk drives or solid-state drives to secure the data in the event of a drive failure.

RAID TYPE:

Hardware RAID:

Hardware RAID uses a dedicated RAID controller card to manage the RAID array. This controller handles all the RAID functions independently of the operating system.

- Can support more complex RAID levels and configurations.
- Typically more expensive due to the need for specialized hardware.
- Offloads processing from the CPU, which can lead to better performance, especially in high-demand environments.

Software RAID:

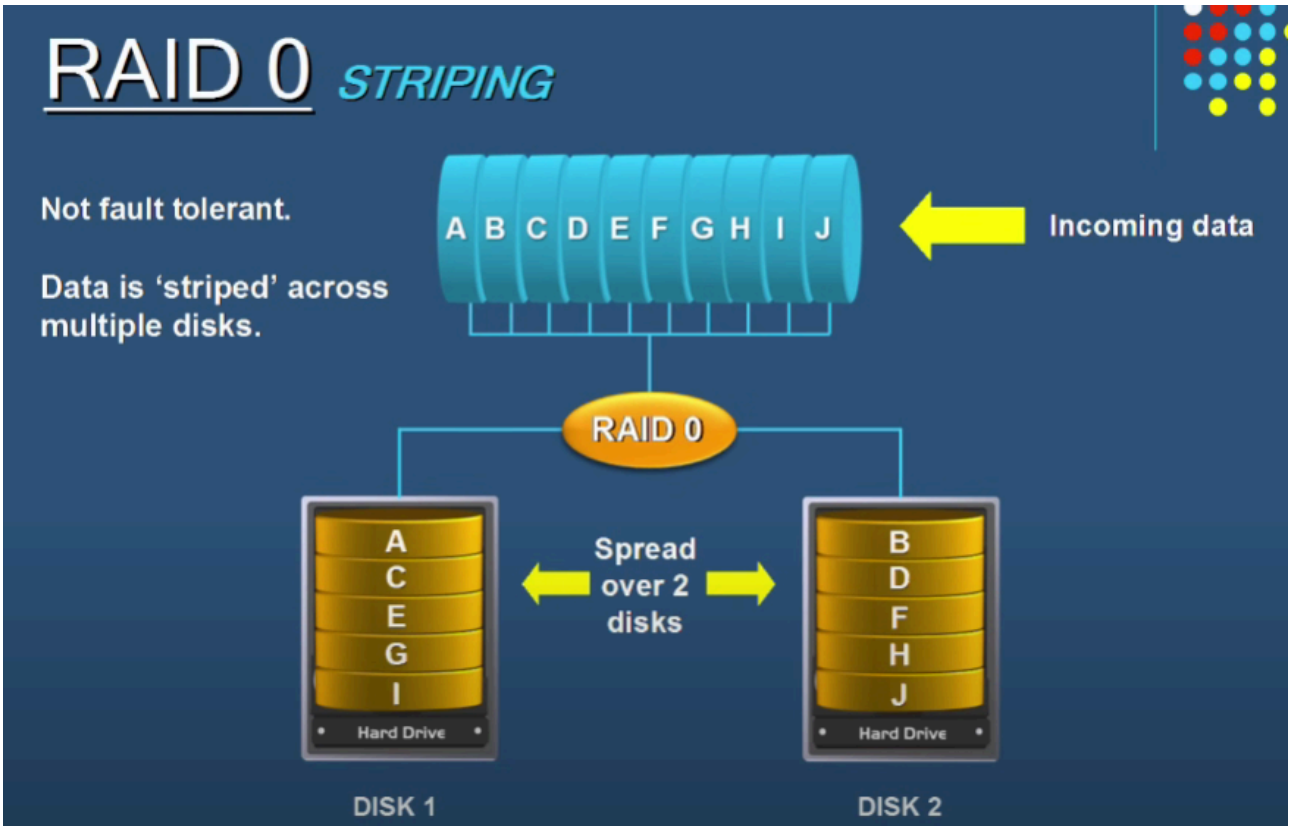
Software RAID is managed by the operating system using its built-in software tools. The OS handles the RAID functions, and no additional hardware is required.

- Can be slower than hardware RAID, as it uses system resources for RAID processing.
- May not support all RAID levels or advanced features available in hardware RAID.
- Can be more easily moved between different systems, as it relies on the OS rather than specific hardware.

RAID Levels:

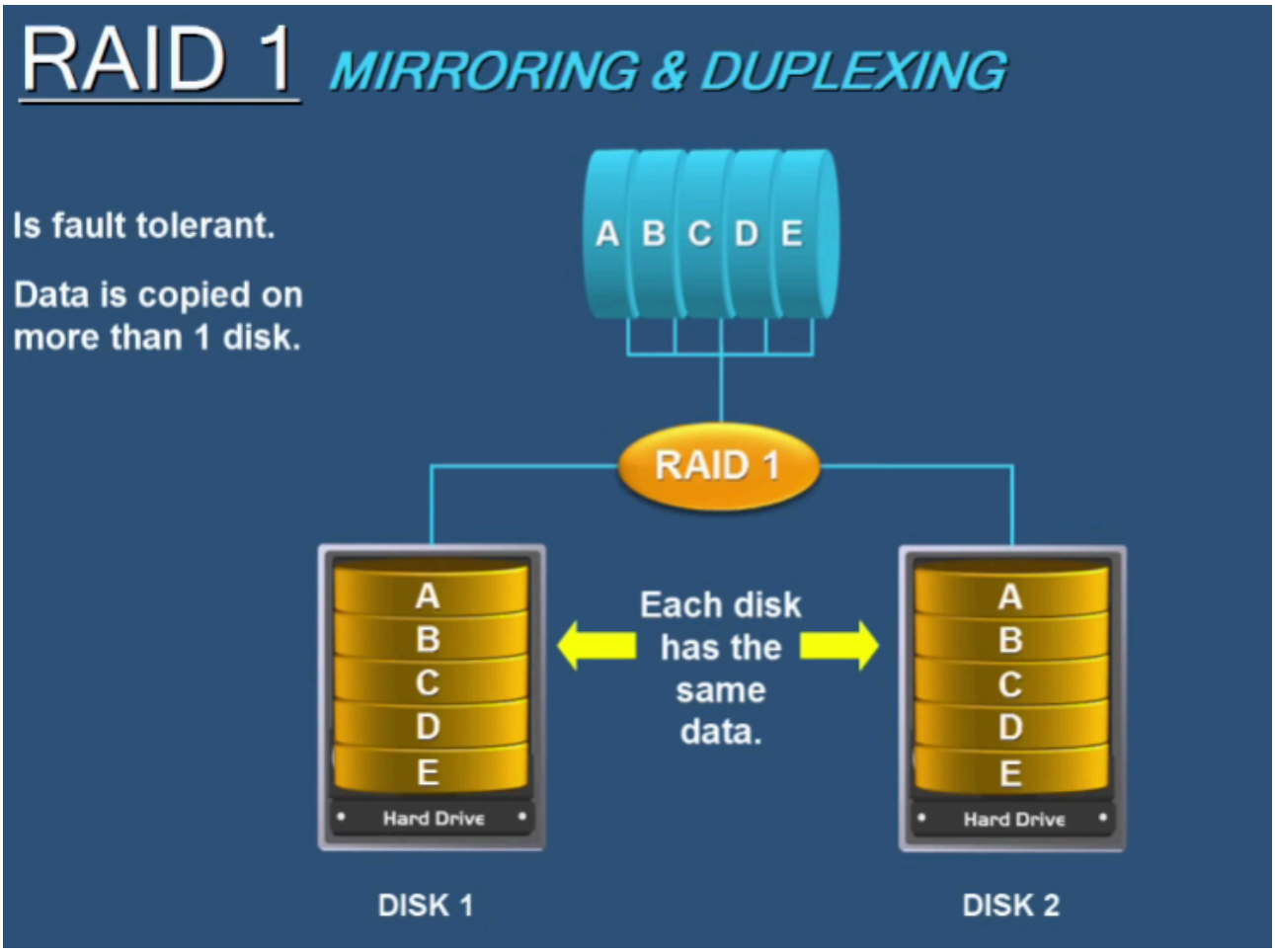
1. **RAID 0 (Striping)**:

```
- **Description**: Data is split across multiple drives for
improved performance.
- **Advantages**: Increased speed and storage capacity.
- **Disadvantages**: No redundancy; if one drive fails, **all data
is lost.**
```



2. **RAID 1 (Mirroring)**:

```
- **Description**: Data is duplicated on two or more drives.
- **Advantages**: High redundancy; if one drive fails, data is still available on the other.
- **Disadvantages**: Storage capacity is halved (only 50% usable).
```

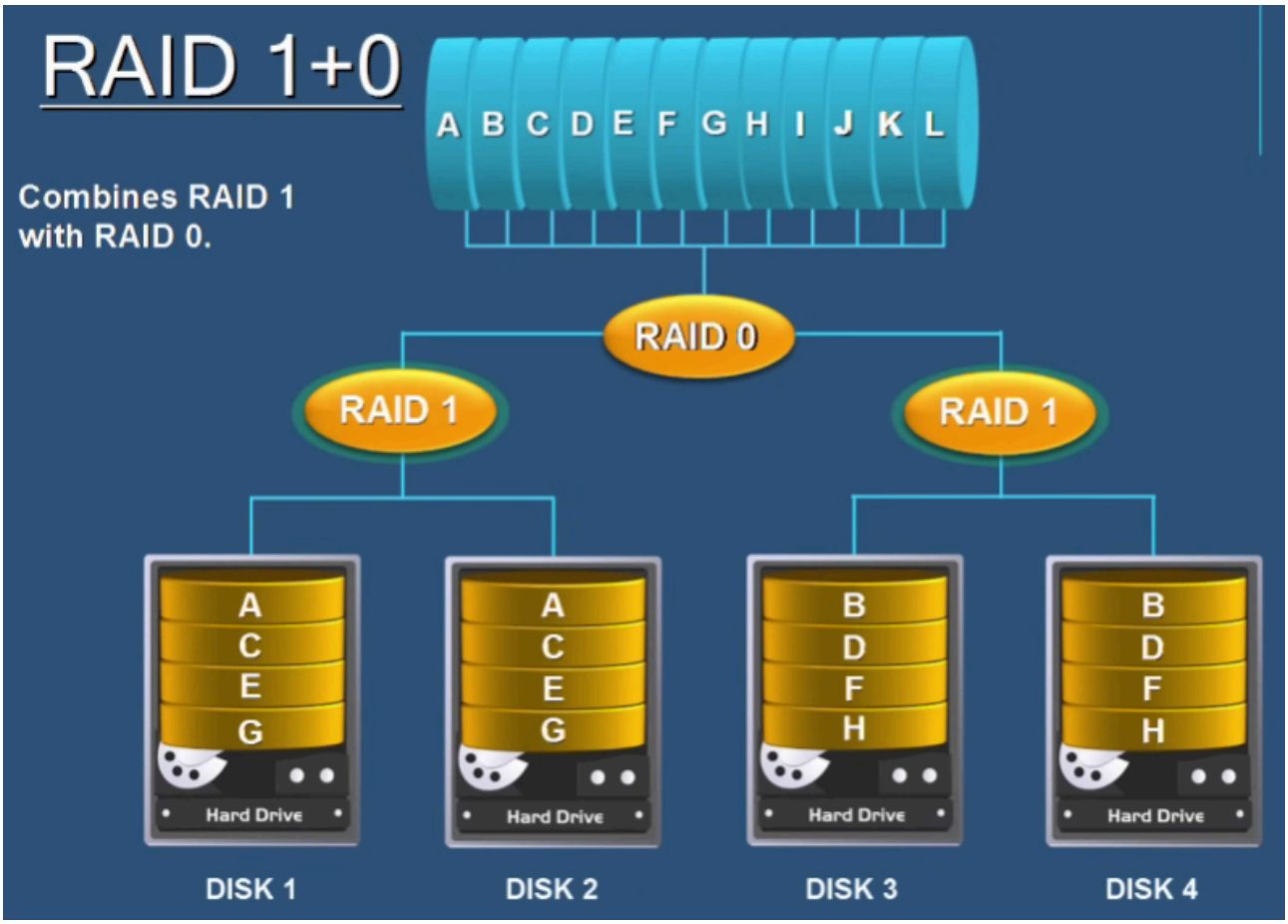


3. **RAID 5 (Striping with Parity)**:

```
- **Description**: Data and **parity** information are distributed
across three or more drives.
- **Advantages**: Good balance of performance, redundancy, and
storage efficiency.
- **Disadvantages**: Slower write speeds due to parity
calculations; can only tolerate one drive failure.
```

Example: An array of 4 disks totaling 4 terabytes, only 3 terabytes will be used for actual data storage.

4. **RAID 10 (1+0)**:
   1. RAID 10 – combining mirroring and striping
   - **Description**: Combines RAID 1 and RAID 0; data is mirrored and then striped.
   - **Advantages**: High performance and redundancy; can tolerate multiple drive failures as long as no mirrored pair fails.
   - **Disadvantages**: Requires at least four drives and has lower storage efficiency (50% usable).


RAID 1+0
Combines RAID 1
with RAID 0.
A B C D E F G H I J K L
RAID 0
RAID 1
RAID 1
A
C
E
G
A
C
E
G
B
D
F
H
B
D
F
H
Hard Drive
Hard Drive
Hard Drive
Hard Drive
DISK 1
DISK 2
DISK 3
DISK 4