Disk Storage Systems

CloudPlus Ch2

Topics

- Disk Storage Systems
- Disk Types and Configurations
- RAID
- File Systems

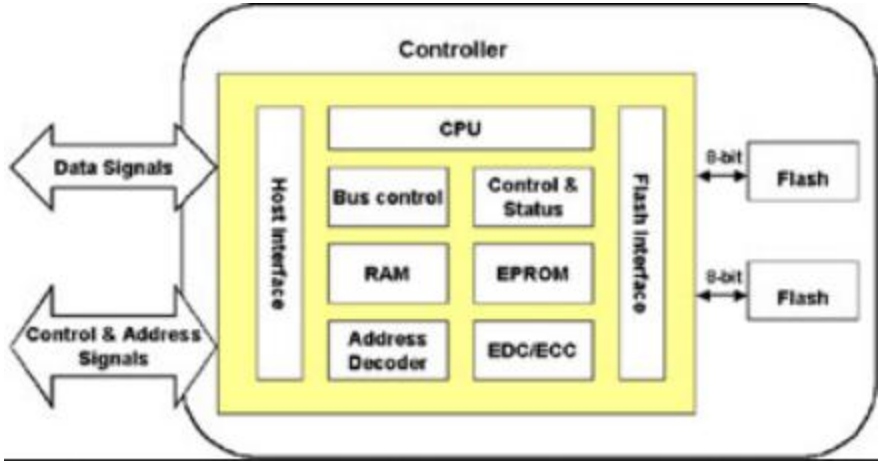
Rotational Media

- Disk storage
  - Generic term used to describe storage where data is digitally recorded by electronic, magnetic, optical, or mechanical methods on rotating media.
- May use fixed or removable media.
  - Removable media may be a compact disk, floppy disk, or USB drive.
  - Fixed or nonremovable media refers to a hard disk drive.
- Most common HDD sizes:
  - 3.5 inch
  - 2.5 inch

SSD

- Primary HDD competitors are solid state drives (SSD) and flash memory cards.
- SSDs replacing rotating hard drives in portable electronic devices because of speed and ruggedness.

Disk Controller Block Diagram



- Host interface
- Control signals
- Data signals
- Drive/SSD/Flash interface
- Related electronics

Interface Types One

- HDDs interfaces include:
  - ATA/IDE
  - SATA
  - Fibre Channel
  - SCSI
  - SAS
- HDDs connect to host bus interface adapter with a single data cable.
- Each HDD has own power cable.

Interface Types Two

- Advanced technology attachment (ATA)
  - Interface standard for connecting storage devices. (parallel ATA or PATA).
- Integrated drive electronics (IDE)
  - The integration of the controller and the hard drive itself
  - Allows the drive to connect directly to the motherboard or controller.
  - IDE also known as ATA.
- Serial ATA (SATA) is used to connect host bus adapters to mass storage devices.
  - Offers reduced cable size, lower cost, native hot swapping, faster throughput, and more efficient data transfer.

Interface Types Three

- SCSI
  - Faster and more flexible than earlier transfer interfaces.
  - Uses bus interface.
  - Every device in chain requires a unique ID.
- Serial attached SCSI (SAS) is a data transfer technology designed to replace SCSI.
  - Backward compatible with SATA drives.
- Fibre Channel
  - High-speed network technology used in storage networking.
    - Well suited to connect servers to a shared storage device such as a storage area network (SAN).

HDD Interface Types

Connector	Advantages	Disadvantages
Integrated drive electronics (IDE)	■ Lower cost ■ Large capacity	■ Only one device is able to read/write at a time if used in the typical master/slave configuration.
Serial ATA (SATA)	■ Lower cost ■ Large capacity ■ Faster transfer rates than ATA ■ Easy configuration	■ Slower transfer rates than SCSI ■ No native support in older operating systems
Small computer system interface (SCSI)	■ Faster speeds ■ Greater scalability ■ Compatible with older SCSI devices ■ Reliability ■ Appropriate for large amounts of data	■ Higher cost ■ Large variety of interfaces ■ Higher RPM, causing more noise and heat ■ More difficult configuration
Serial attached SCSI (SAS)	■ Compatibility with SATA ■ Higher transfer speeds ■ Serial communication vs. parallel ■ Increased availability	■ Higher cost ■ Use of SCSI command set

Speed

- Hard drive's speed is measured by the amount of time it takes to access data stored on drive.
- Access time is the response time of the drive and is a direct correlation of seek time and latency.
- Seek time is the measure of how long it takes the drive to find the data being accessed
- Latency is the measure of the time delay that it takes for the drive to properly position the sector under the read/write head.

Speed vs Latency

Rotational Speed (RPM)	Latency (MS)
3600	8.3
4200	7.1
5400	5.6
7200	4.2
10000	3
15000	2

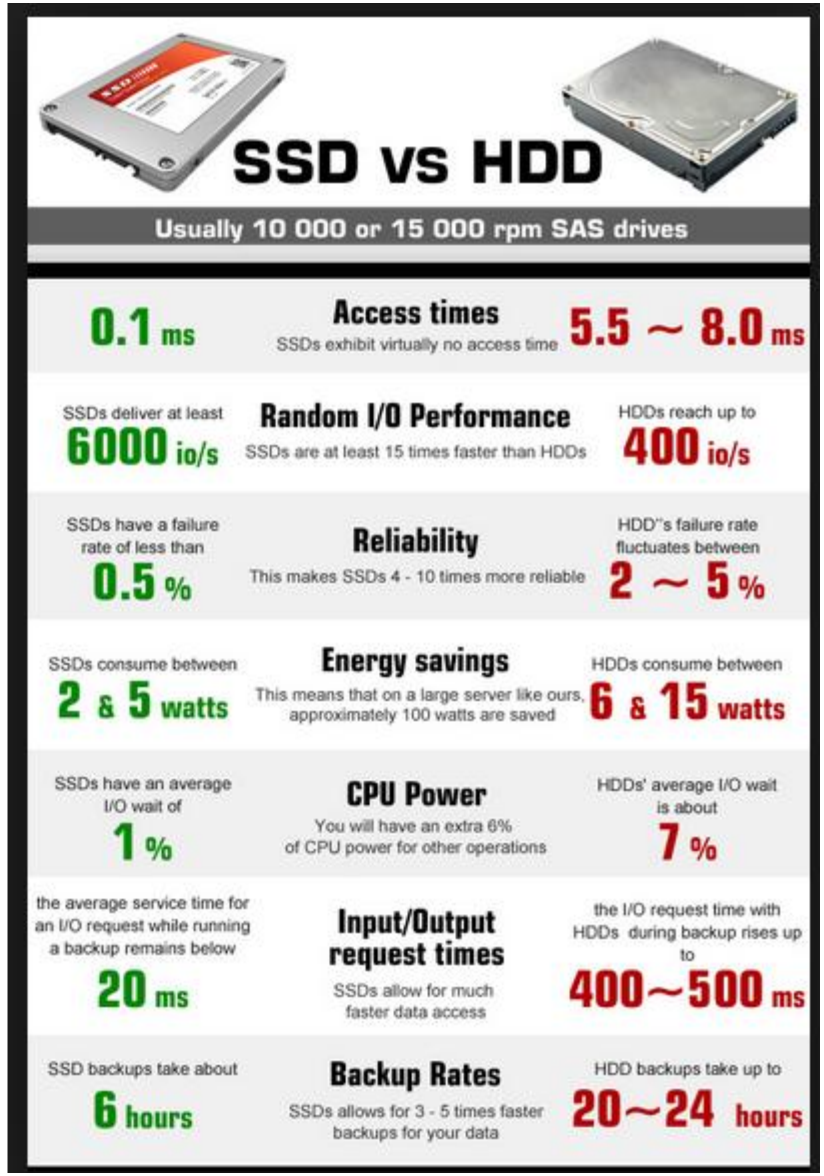
Solid State Drive (SSD)

- High-performance storage device containing no moving parts.
- Majority of SSDs use “not and” (NAND)–based flash memory.
  - Nonvolatile memory type.
- Less susceptible to shock or being dropped.
  - Much quieter
- Faster access time and lower latency than HDDs.

SSD

- Especially valuable where I/O response time is critical.
  - Database server or a server hosting a file share…
  - Used in laptops. SSDs provide shock resistance
  - Use less power
  - Provide faster startup time than HDDs.
- Currently more expensive than traditional hard disk drives
- Less risk of failure and data loss.
- Table 2-3 (Next Slide) shows some of the differences between SSDs and traditional hard disk drives.

SSD/HDD Comparison Hosting Context



Universal Serial Bus (USB) Drive

- An external plug-and-play storage device that can be plugged into a computer's USB port.
  - Recognized by the computer as a removable drive and assigned a drive letter.
- Powered via computer's USB port.
- A USB drive is portable and retains the data stored on it as it is moved between computer systems.
  - Many external storage devices use USB
    - Hard drives
    - Flash drives
    - DVD drives.

Tape

- A tape drive is a storage device that reads and writes data to a magnetic tape.
- Tape drives provide sequential access to the data, whereas an HDD provides random access to the data.

Tiered Storage

- Can refer to an infrastructure that has a simple two-tier architecture
  - SCSI disks and a tape drive
- Can also refer to a more complex scenario of three or four tiers.
- Helps organizations
  - Plan their information life cycle management
  - Reduce costs
  - Increase efficiency.
- Tiered storage requirements can also be determined by functional differences
  - For example, the need for replication and high-speed restoration.
  - With tiered storage, data can be moved from fast, expensive disks to slower, less expensive disks.

Hierarchical Storage Management (HSM),

- Allows for automatically moving data among four different tiers of storage.
  - For example, data that is frequently used and stored on highly available, expensive disks can be automatically migrated to less expensive tape storage when it is no longer required on a day-to-day basis.
- One of the advantages of HSM is that the total amount of data that is stored can be higher than the capacity of the disk storage system currently in place.

Tier Performance Levels

Tier 1

- Mission-critical, recently accessed, or secure files.
- Expensive and highly available disks such as RAID with parity…
- Better performance, capacity, reliability, and manageability.

Tier 2

- Runs major business applications such as, e-mail and ERP.
- Balance between cost and performance.
  - Does not require sub-second response time.
  - Still needs reasonably speed.

Tier 3

- Financial data that needs to be kept for tax purposes
  - Not accessed on a daily basis.

Tier 4

- Data used for compliance requirements or for keeping e-mails or data for long time periods.
- Not needed to be instantly accessible.

Policies

- An organization can implement policies that define what data fits into each tier
  - Then manage how that data migrates between the tiers.
  - For example, when financial data is more than a year old, the policy could be to move that data to a tier 4 storage solution.
- Tiered storage offers a solution for managing organizational data while also saving time and money.

Redundant Array of Independent Disks (RAID)

- Storage technology that combines multiple hard disk drives into a single logical unit.
  - Offers improved performance and/or increased redundancy.

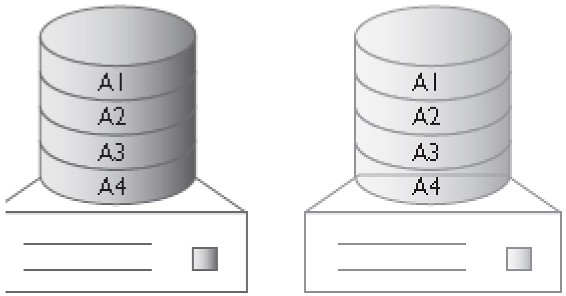
RAID 1

- Mirrored identical disks.
  - Read requests serviced by either disk 1 or disk 2.
  - Write requests always update both disks and can be accessed independently.

RAID 0

- "Stripes" writes across both disks
  - Increases performance by getting access to multiple physical spindles.
- If any of the drives fails, however, entire array is irreparably damaged.
- Typically used for noncritical data that is regularly backed up and requires high write speed.
- Low cost way to increase performance.

RAID Configurations



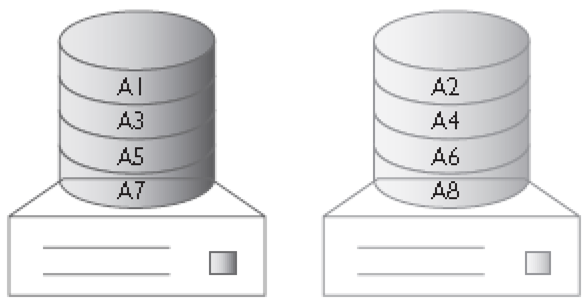
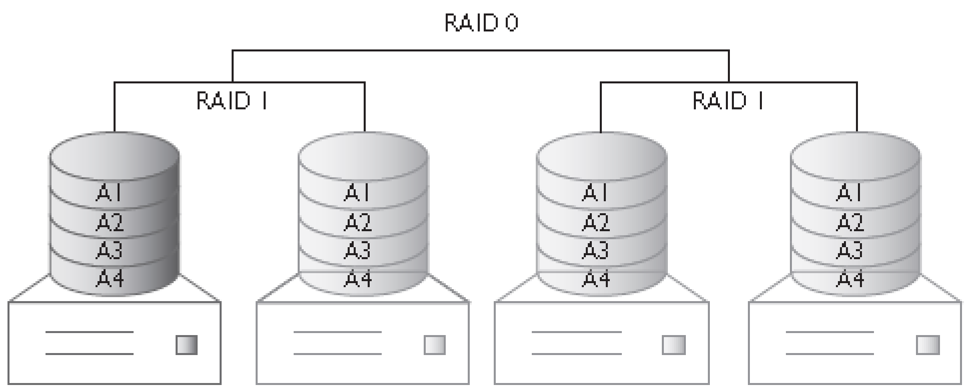


FIGURE 2-3

RAID 1+0 mirroring and striping, no parity.



RAID 10

RAID 1+0

- Raid 1+0 consists of a top-level RAID 0 array that is in turn composed of two or more RAID 1 arrays.
  - Incorporates both the performance advantages of RAID 0 and the data protection advantages of RAID 1.
  - Although its official designation is RAID 1+0, it is often referred to as RAID 10.
- If a single drive fails in a RAID 10 array, the lower-level mirrors will enter into a degraded mode while the top-level stripe can continue to perform as normal because both of its drives are still working as expected.
- RAID 10 cuts usable storage in half.
  - Could be used if an application requires both high performance and reliability.
  - Examples include enterprise servers, database servers, and high-end application servers.

RAID 5

- Historically, most commonly used RAID implementation.
  - Balances data protection, performance, and cost-effectiveness.
- Uses block-level striping for a performance enhancement with distributed parity for data protection.
  - Distributes parity and the data across all drives
  - Requires that all drives but one be present in order to operate.
- Means that a RAID 5 array is not destroyed by a single drive failure.
  - When a drive fails, the RAID 5 array is still accessible to read and write data.
- After the failed drive has been replaced, the array enters into data recovery mode, which means that the parity data in the array is used to rebuild the missing data from the failed drive back onto the new hard drive.

RAID 6

- Can be viewed essentially as an extension of RAID 5, as it uses the same striping and parity block distribution across all the drives in the array.
- The difference is that RAID 6 adds an additional parity block, allowing it to use block-level striping with two parity blocks distributed across all the disks.
- The inclusion of this second parity block allows the array to tolerate the loss of two hard disks instead of the one failure that RAID 5 can tolerate.
- RAID 6 causes no performance hit on read operations but does have a lower performance rate on write operations due to the overhead associated with the parity calculations.

RAID Level Benefits and Requirements

Level	Description	Minimum Number of Disks	Fault Tolerance	Storage Efficiency
RAID 1	Blocks are mirrored. No striping or parity.	2	1 drive	50% or *n*/2
RAID 0	Blocks are striped. No mirror or parity.	2	None	100%
RAID 1+0 (or RAID 10)	Blocks are mirrored and striped.	4	1 drive per span up to maximum of 2	50%
RAID 0+1	Blocks are striped across two disks and mirrored on the third disk.	4	1 drive per span up to a maximum of 2	50%
RAID 5	Blocks are striped. Distributed parity.	3	1 drive	Number of drives −1
RAID 6	Blocks are striped with double distributed parity.	4	2 drives	Number of drives −2

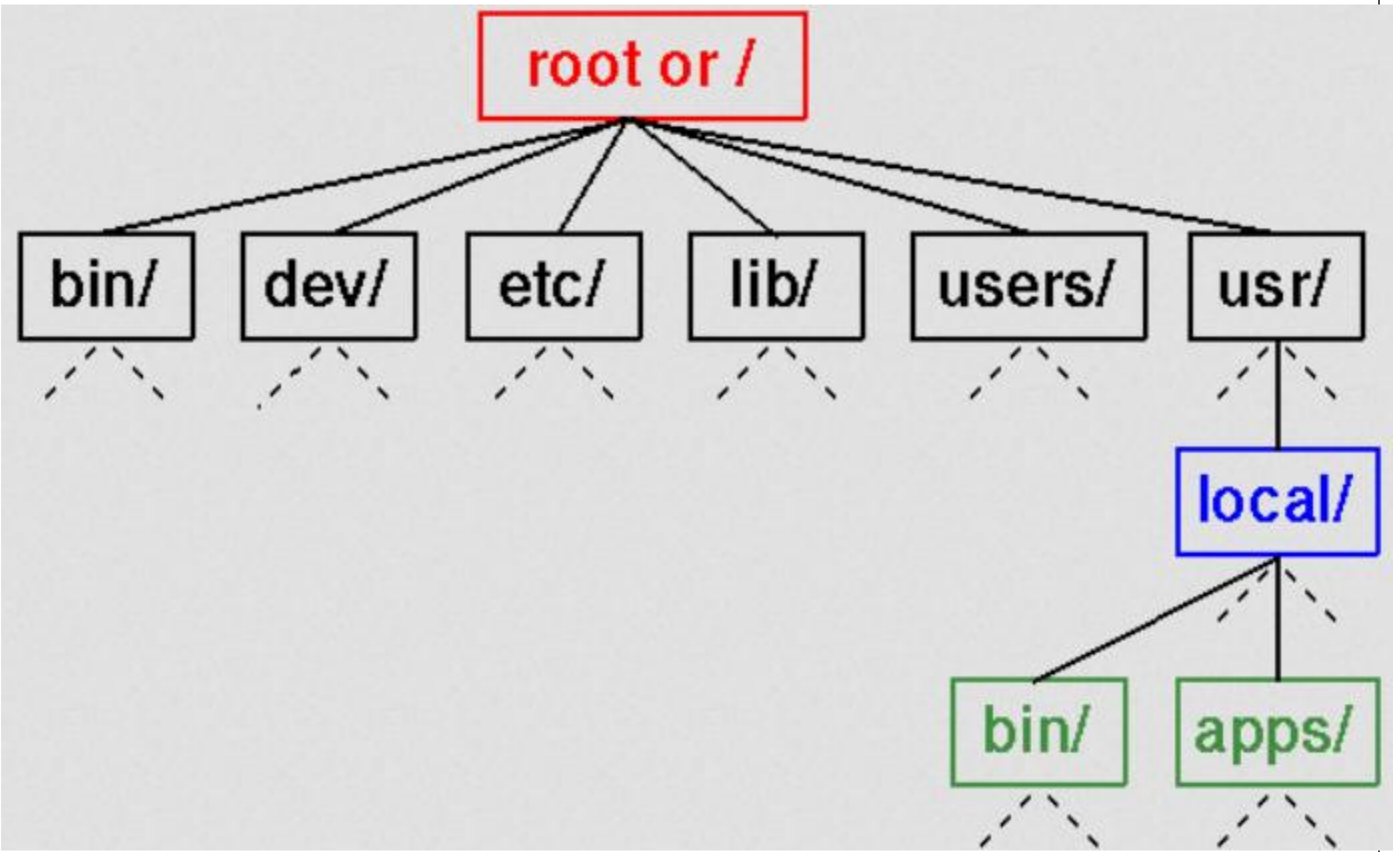
File System Types

- File system responsible for storing, retrieving, and updating a set of files on a disk.
  - Accepts operating system commands to read and write data to the disk.
- Responsible for how the files are named and stored on the disk.
- Responsible for managing access to the file's metadata and the data itself
- Oversees relationships to other files and file attributes.
- Manages available disk space.
- Responsible for reliability of the data on the disk and for organizing that data in an efficient manner.
- Organizes files, directories, and tracks.

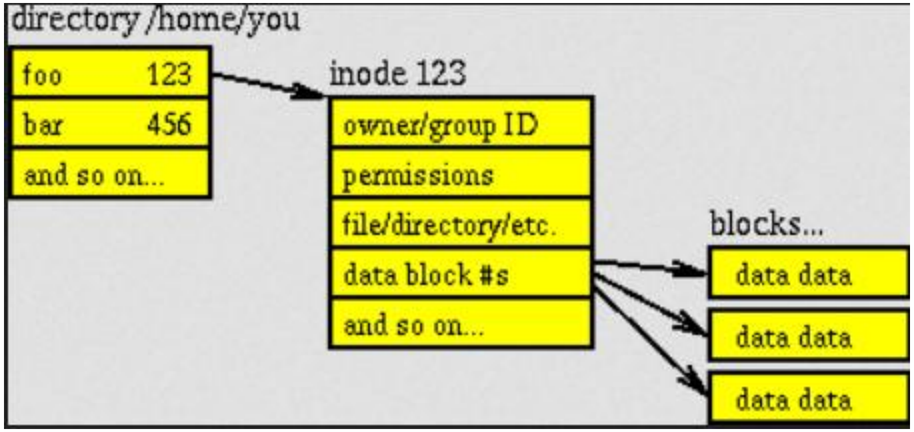
Unix File System

- The Unix file system (UFS) primary file system for Unix and Unix-based operating systems.
- Uses a hierarchical file system structure where the highest level of the directory is called the root
  - All other directories span from that root.
  - Under the root directory, files are organized into subdirectories.
- All files on a Unix system are related to one another in a parent-child relationship.
- All share a common parental link to the top of the hierarchy.

UFS



root or /
bin/
dev/
etc/
lib/
users/
usr/
local/
bin/
apps/



directory /home/you
foo 123
bar 456
and so on...
inode 123
owner/group ID
permissions
file/directory/etc.
data block #s
and so on...
blocks...
data data
data data
data data

Extended File System (EXT)

- First file system created specifically for Linux.
- Metadata and file structure based on the Unix file system.
- Default file system for most Linux distributions.
- EXT is currently on version 4, or EXT4,
  - Introduced in 2008
  - Supports a larger file and file system size.
- EXT4 backward compatible with EXT3 and EXT2
  - Allows for mounting an EXT3 and EXT2 partition as an EXT4 partition.

File Allocation Table (FAT)

- Legacy file system provides good performance
  - Does not deliver the same reliability and scalability as newer file systems.
- If a user has a drive running a FAT32 file system partition, they can connect it to a computer running Windows 7 and retrieve the data from that drive
  - Windows 7 still supports the FAT32 file system.
- FAT is used by a variety of removable media
  - floppy disks, solid state memory cards, flash memory cards, and portable devices.
- FAT does not support the advanced features like encryption, VSS, and compression.

New Technology File System (NTFS)

- Proprietary file system developed by Microsoft.
  - First available with Windows NT 3.1.
  - FAT file system replacement.
- Improved performance and reliability, larger partition sizes, and enhanced security.
  - Starting with version 1.2, NTFS added support for file compression.
  - NTFS version 3.0 added support for volume shadow copy service (VSS), which keeps a historical version of files and folders on an NTFS volume.
  - Shadow copies allow you to restore a file to a previous state without the need for backup software.
- Best practice to create a shadow copy volume on a separate disk to store the files.

Encrypting File System (EfS)

- Provides an encryption method for any file or folder on an NTFS partition
  - Transparent to user.
- Encrypts a file by using a file encryption key (FEK)
  - Which is associated with a public key that is tied to the user.
  - Encrypted data is stored on an alternate location from the encrypted file.
- To decrypt file, EFS uses the private key of the user to decrypt the public key that is stored in the file header.
- If the user loses access to their key, a recovery agent can still access the files.
- NTFS does not support encrypting and compressing the same file.

Disk Quotas

- Allow an administrator to set user disk space thresholds.
- Gives an administrator the ability to track and limit amount of disk space for each user.
- Administrator can set a warning threshold and a deny threshold and deny access to the user once they reach

Virtual Machine File System

- Virtual machine file system (VMFS) is VMware's cluster file system.
  - Used with VMware ESX server and vSphere.
  - Created to store virtual machine disk images, including virtual machine snapshots.
- Allows for multiple servers to read and write to the file system simultaneously, while keeping individual virtual machine files locked.
- VMFS volumes can be logically increased by spanning multiple VMFS volumes together.

Z File System (ZFS)

- A combined file system and logical volume manager designed by Sun Microsystems.
- Provides protection against data corruption and support for high storage capacities.
  - Also provides volume management, snapshots, and continuous integrity checking with automatic repair.
- Was created with data integrity as its primary focus.
- Designed to protect the user's data against corruption.
- 128-bit file system.
- Uses a pooled storage method, which allows space to be used only as it is needed for data storage.

File Systems Compared

File System	Maximum File Size	Maximum Volume Size	Encryption	Resizable Volumes
Unix file system (UFS)	32 PB	1 YB	No	Offline but cannot be shrunk
File allocation table (FAT32)	4 GB	2 TB	No	Offline
New technology file system (NTFS)	16 TB	256 TB	Yes	Online
Virtual machine file system (VMFS)	2 TB	64 TB	No	Offline but cannot be shrunk*
Z file system (ZFS)	16 EB	16 EB	Yes	Online but cannot be shrunk

* Newest version of VMFS allows dynamic resizing but must be supported by the OS for it to be utilized without a reboot or additional sizing tools.

Questions?