Cloud Computing NETW1009

Lecture 5

Course Instructor: Dr. – Ing. Maggie Mashaly

GUC German University in Cairo

Recall what we discussed so far..

- Cloud Computing
- Cloud Data Centers
- Compute, Storage, Network Systems
- Virtualization
- Systems Virtualization
- Resource Pooling and Provisioning
- Software Defined Infrastructure
- Infrastructure Deployment Models

Lecture 5: ISS - Intelligent Storage Systems

Lecture Outline

- Intelligent Storage Systems
- Intelligent Storage Systems Components
- Hard Disk Drives (HDD)
- Solid State Drives (SSD)
- Types of Intelligent Storage Systems
- Challenges

Intelligent Storage Systems

Intelligent Storage System

A feature-rich RAID (Redundant Array of Independent Disks) array that provides highly optimized I/O processing capabilities

ISS Features

- Supports combination of HDD & SDD
- Service massive amounts of IOPS
- Scale-out Architecture
- Deduplication, Compression, and encryption
- Automated Storage tiering
- Virtual Storage provisioning
- Multi-tenancy
- Supports APIs to integrate with SDDC & Cloud
- Data Protection

Intelligent Storage Systems Components

Controller

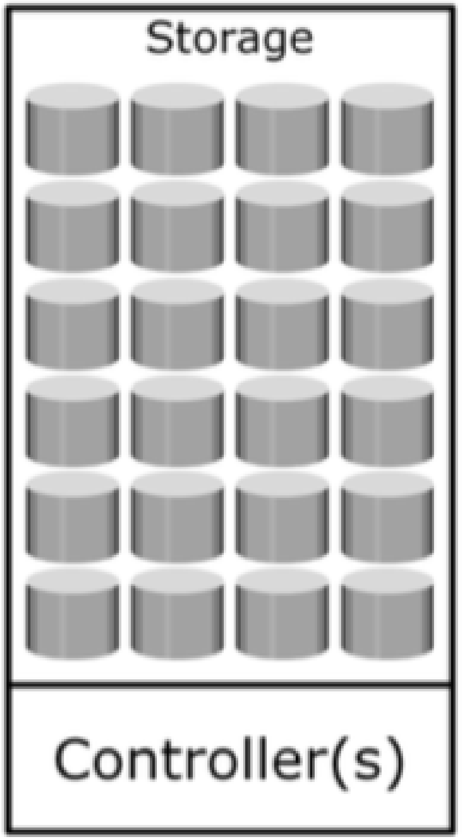
Controls I/O request for reading/writing/data, Could be one of these types:

- Block-based
- File-based
- Object-based
- Unified

Storage

Disks where data is stored, could be one of these types:

- All HDDs
- All SSDs
- Combination of both

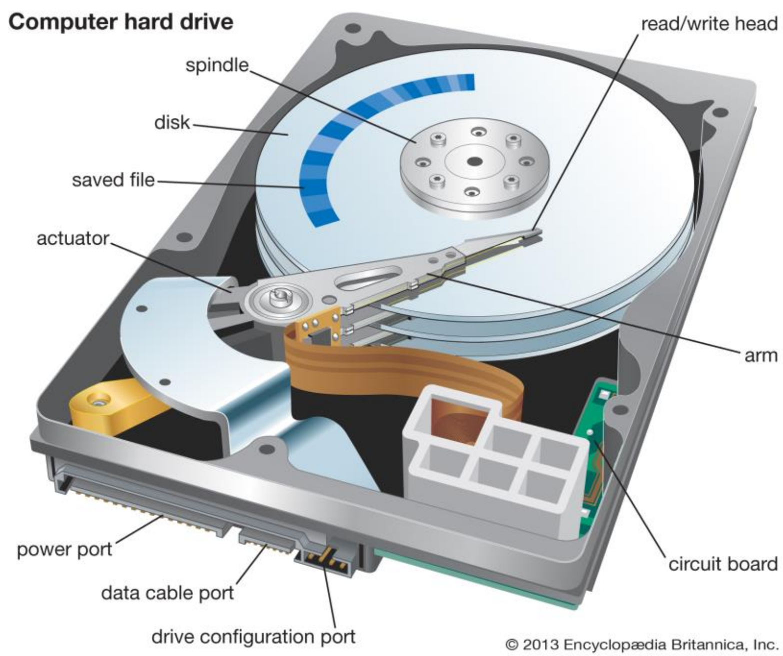


Components of Intelligent Storage Systems I : Hard Disk Drives

Storage – Hard Disk Drives

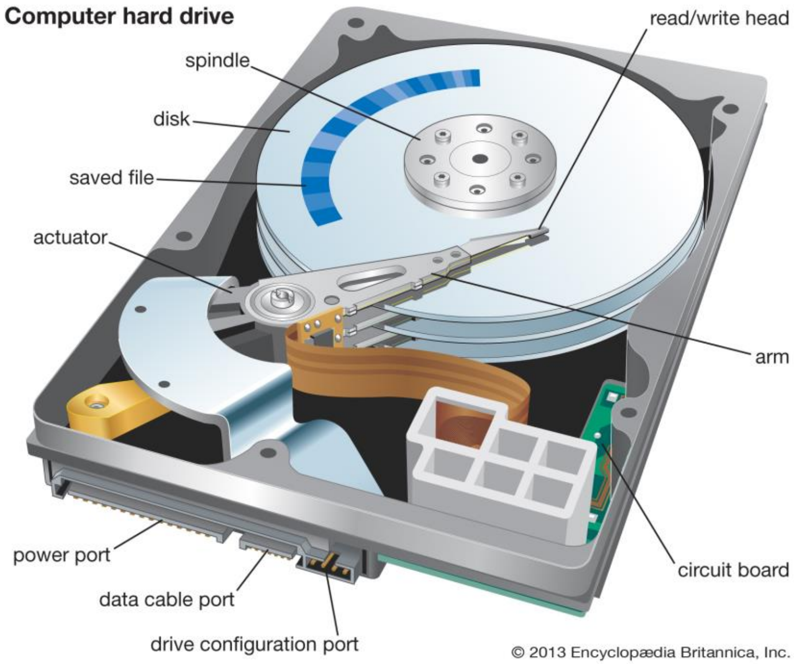
Hard Disk Drives

A hard disk drive is a persistent storage device that stores and retrieves data using rapidly rotating disks (platters) coated with magnetic material



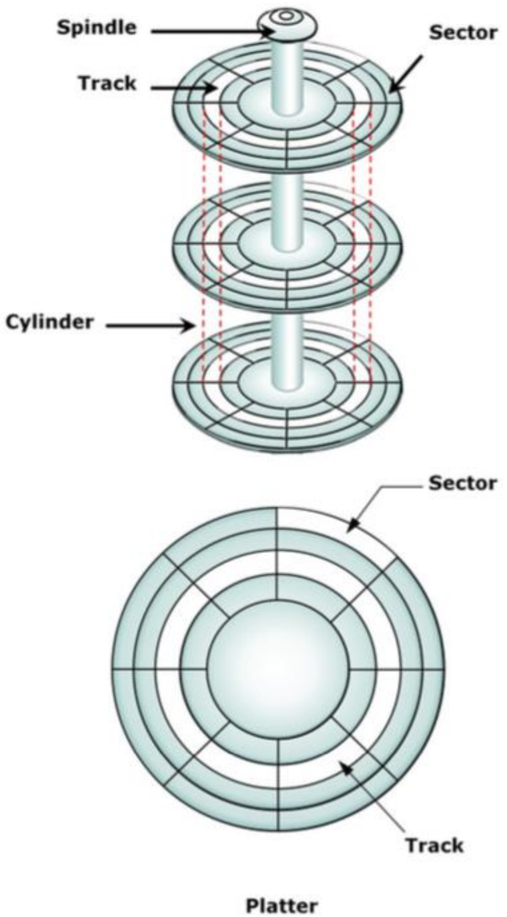
HDD - Components

Platter	• Flat circular disks on which data is recorded in 0s & 1s by polarizing magnetic domains of the disk surface • A typical HDD consists of one or more platters • A set of rotating platters is sealed in a case called Head Disk Assembly (HDA)
Spindle	• Connects all platters and is connected to a motor rotating at a constant speed
Read/Write Head	• Reads/writes data from/to the platters by sensing/changing their magnetic polarization • Drivers have two R/W heads per platter, one for each surface
Actuator Arm Assembly	• R/W heads are mounted on the AAA • It positions R/W head at the location on the platter where data needs to be read/written
Drive Controller Board	• A printed circuit board mounted at the bottom of a disk drive • Consists of a microprocessor, internal memory, circuitry & firmware

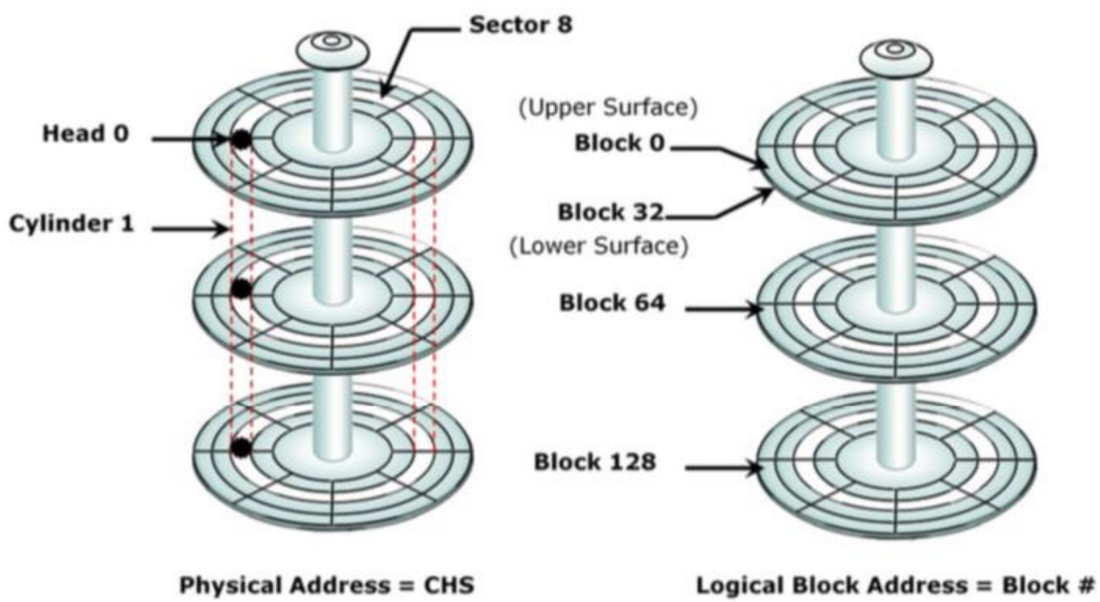


HDD – Physical Disk Structure

Tracks	• Concentric rings on the platter around the spindle • Numbered from zero from the outer platter edge
Sectors	• Each track is divided into smaller units called sectors • Sector is the smallest individually addressable unit of storage • Number of sectors per track varies by drive type • Typically a sector holds 512 bytes of data
Cylinder	• Is a set of identical tracks on both surfaces of each platter



HDD – Logical Block Addressing



In this example:

- 8 sectors per track
- 6 heads
- 4 cylinders
- Total of 8*6*4=192 blocks ranging from Block 0 to Block 191

- Physical Addresses are defined as (CHS): Cylinder, Head & Sector number
- Logical Block Addresses (LBA) has a linear addressing scheme by block number
- Disk Controller translates LBA to CHS

HDD – Performance

Factors affecting performance of a HDD:

1. **Electromechanical device**

   Impacts the overall performance of the storage system

2. **Disk Service Time**

   Time taken by a disk to complete an I/O request, depend on:
   - Seek Time
   - Rotational Latency
   - Data Transfer Rate

Disk Service Time = Seek Time + Rotational Latency + Data Transfer Rate

HDD – Performance

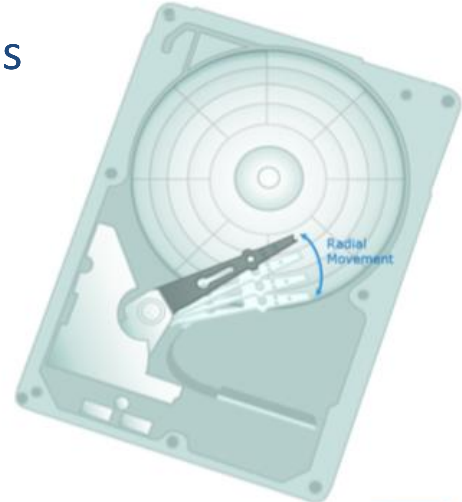
Seek Time

Seek Time

Is the time taken to position the read/write heads across the platter moving along the radius of the platter, thus settling over the correct track

- The lower the seek time, the faster the I/O operation
- Seek time of a disk is specified by its manufacturer in ms
- Seek time specifications include:

Full Stroke	Average	Track-to-track
It is the time taken by the R/W head to move across the entire width of the disk, from innermost to outermost track	It is the average time taken by the R/W head to move from one random track to another, normally listed as 1/3 full stroke	It is the time taken by the R/W head to move between adjacent tracks

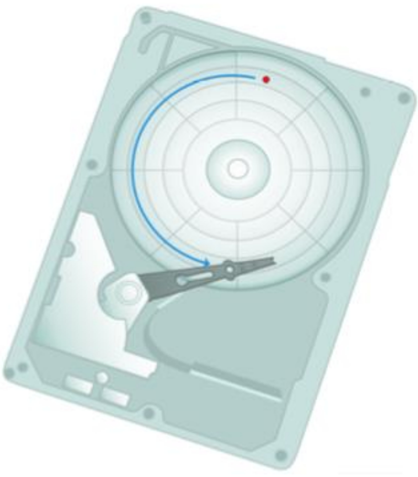


HDD – Performance
Rotational Latency

Rotational Latency

Is the time taken by the platter to rotate and position the data under the R/W head

- Depends on the rotation speed of the spindle
- Measured as one half of the time taken for a full rotation



Example:

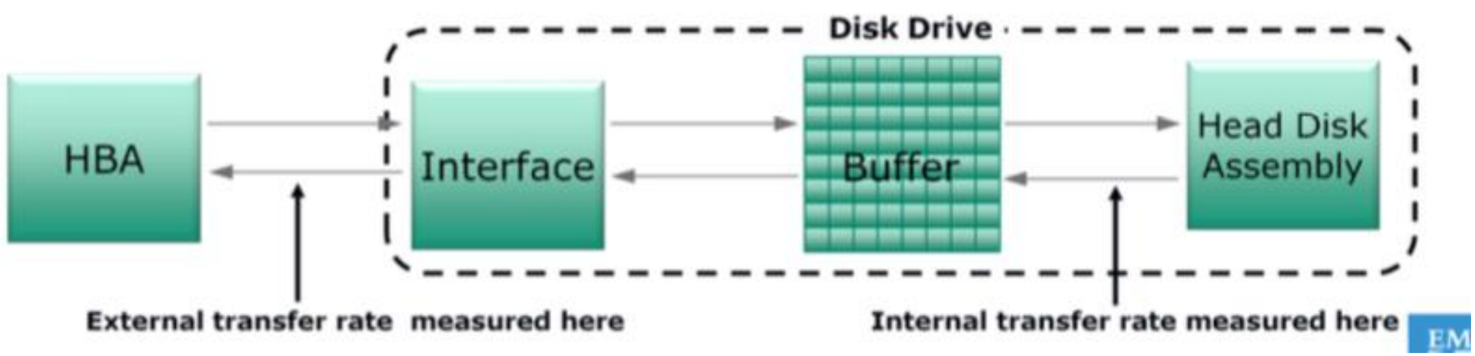
- For ‘X’ rpm, drive latency is calculate in milliseconds as:

$$= \frac{\frac{1}{2} * 1000}{\frac{x}{60}} = \frac{500}{\frac{x}{60}} = \frac{30000}{x}$$

HDD – Performance
Data Transfer Rate

Data Transfer Rate

The average amount of data per unit time that the drive can deliver to the Host Bus Adapter (HBA)



- Internal Transfer Rate: Speed at which data moves from a platter's surface to the internal buffer of the disk
- External Transfer Rate: Rate at which data move through the interface to the HBA

HDD – Performance

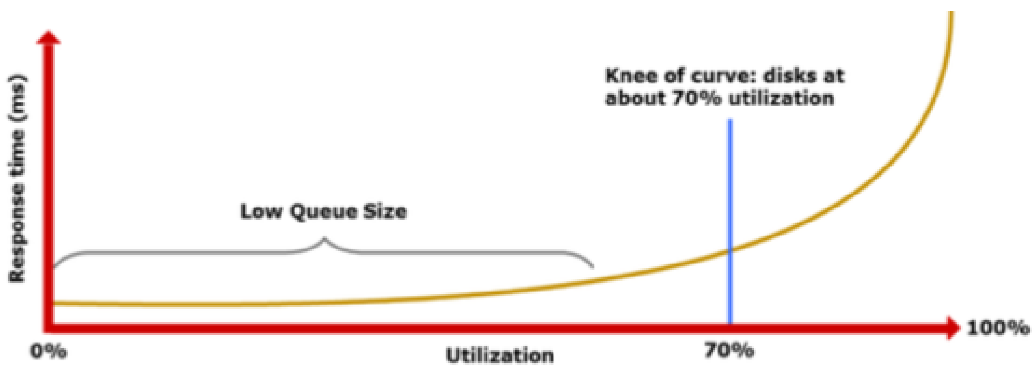
Utilization vs. Response Time

- Based on fundamental laws of disk drive performance:

$$Average\ Response\ Time = \frac{Service\ Time}{1 - Utilization}$$

- Service time is taken by the controller to serve an I/O

- For performance-sensitive applications disks are commonly utilized below **70%** of their I/O serving capability



GUC
German University in Cairo

HDD – Performance
Storage Design: Application Requirements vs. Disk Drive Performance

- Disks required to meet an application's capacity need **(DC)**:

$$DC = \frac{Total\ capacity\ required}{Capacity\ of\ a\ single\ disk}$$

- Disks required to meet an application's performance need **(DP)**:

$$DP = \frac{IOPS\ generated\ by\ an\ application\ at\ peak\ workload}{IOPS\ serviced\ by\ a\ single\ disk}$$

- IOPS serviced by a disk (S) depends upon Disk Service Time (T_s):

$$T_s = Seek\ time + \frac{0.5}{(Disk\ rpm/60)} + \frac{Data\ Block\ Size}{Data\ Transfer\ Rate}$$

- T_s is the time taken for an I/O to cpmplete, therefore IOPS serviced by a disk (S) is equal to ($1/\ T_s$)
  - For performance sensitive applications $S = 0.7 * \frac{1}{T_s}$

$$Disks\ Required\ for\ an\ Application = \max(DC, DP)$$

GUC German University in Cairo

Components of Intelligent Storage Systems II : Solid State Drives

Storage – Solid State Drives

Solid State Drives

A solid state drive is a storage device that uses integrated circuit assemblies & semiconductor technology to store data persistently, typically using flash memory.

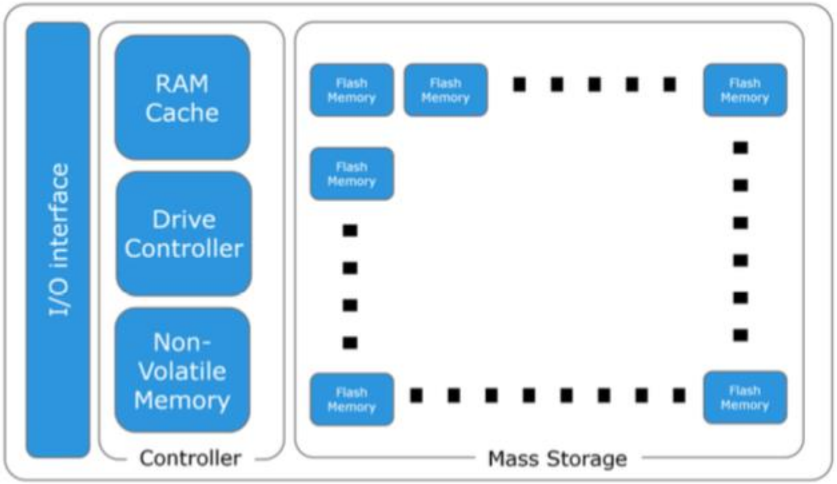
- SSD are superior to mechanical hard disk drives in terms of performance, power use, and availability

- Especially well suited for low-latency applications requiring consistent, low (less than 1 ms) read/write response times.

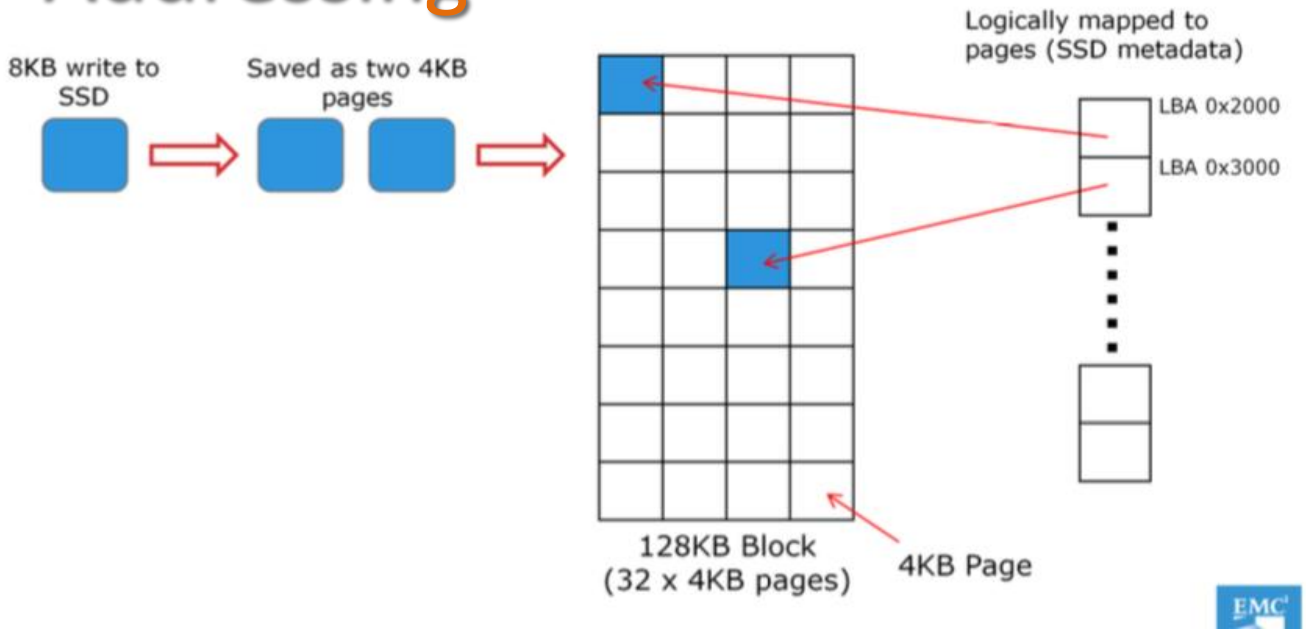
GUC German University in Cairo

SSD - Components

I/O Interface	• Enables connecting power and data connections to SSD Drive
RAM Cache	• Used in the management of data being read & written from the SSD as a cache • Used for SSD’s operational programs & mapping tables • Enhances the overall performance of the SSD
Drive Controller	• Manages all drive functions, e.g. (encryption/decryption, write coalescing)
Non-Volatile Memory	• Stores The SSD’s operational software & data • Not all SSD models have it, some store their programs and data to the driver's mass storage
Flash Memory	• Number and capacity of flash memory chips vary directly with the SSD’s capacity • Consume less power than HDDs • Require much less cooling as they don’t have moving parts as HDDs • Have multiple parallel I/O channels to the controller, thus higher internal bandwidth

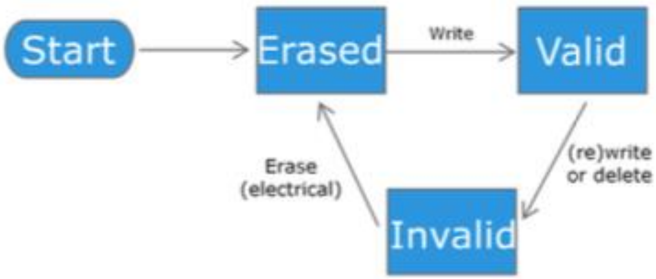


SSD –Addressing

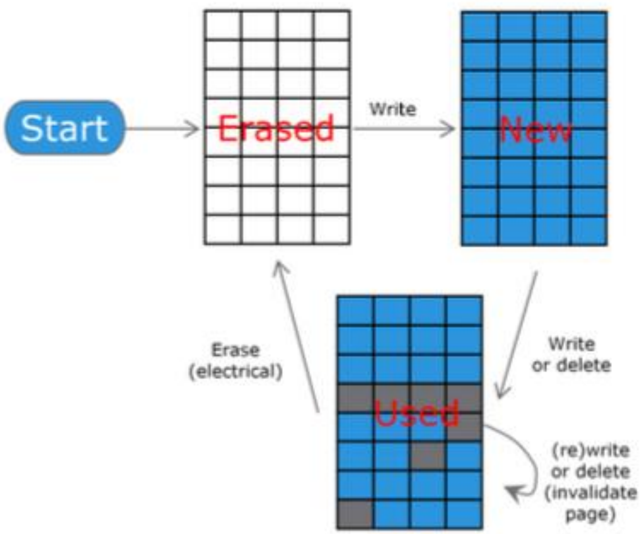


- A Page is the smallest object that can be read/written on a solid state drive
- Page's capacity is not standard like sectors in HDD, they depend on architecture of the solid state memory chip. Typical page capacities are 4 KB, 8 KB and 16 KB
- A block is made up of pages, and only entire blocks can be written/erased. Pages are assembled into full blocks in the cache RAM then written to the block storage object

SSD – Page & Block States



Flash memory page states



Block state diagram

SSD – Performance

1. **Access Type**
   - SSD performs random reads the best
   - SSDs use all internal I/O channels in parallel for multi-threaded large block I/Os

2. **Drive State**
   - New SSD or SSD with substantial unused capacity offers best performance

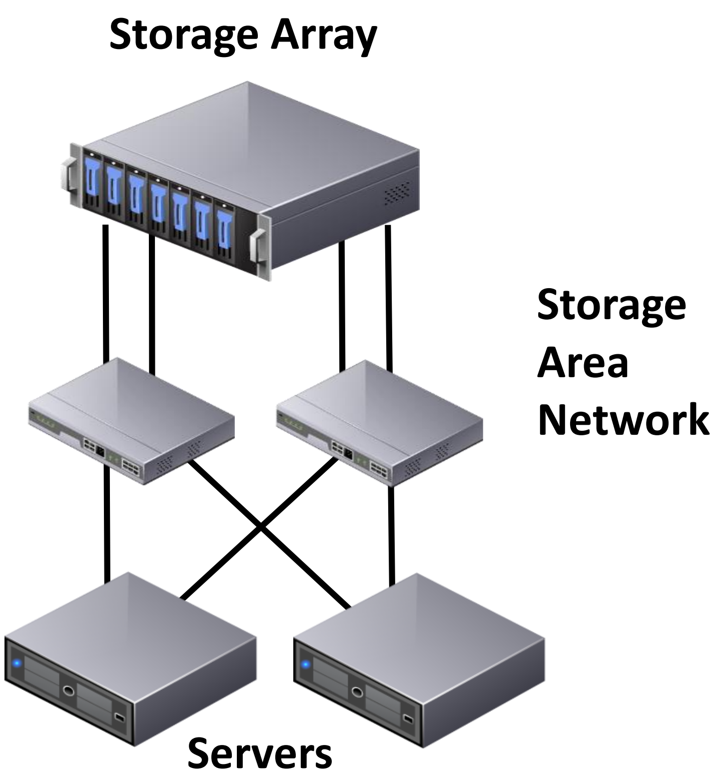
3. **Workload Duration**
   - SSDs are best for workloads with short bursts of activity

HDD vs. SSD - A detailed view

Intelligent Storage Systems

Storage Hardware: Challenges

- Configuring very large amounts of storage to a server is complex, large number of drives may be required
- Having many drives leads to an increased risk of drive failures
- RAID provides a way to aggregate and manage multiple drives
- SAN & NAS RAID arrays provide network based, high capacity storage solutions and deliver incredibly high performance



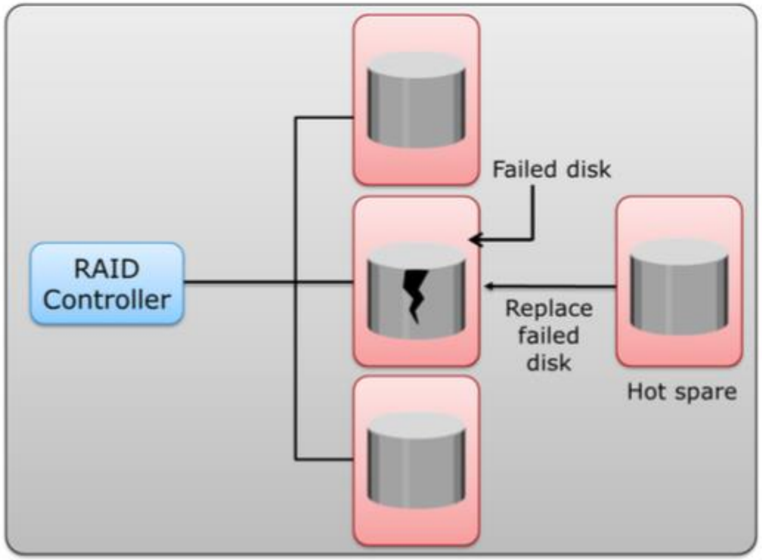
Storage Hardware: Challenges

Dynamic Disk Sparing (Hot Sparing)

Hot Sparing

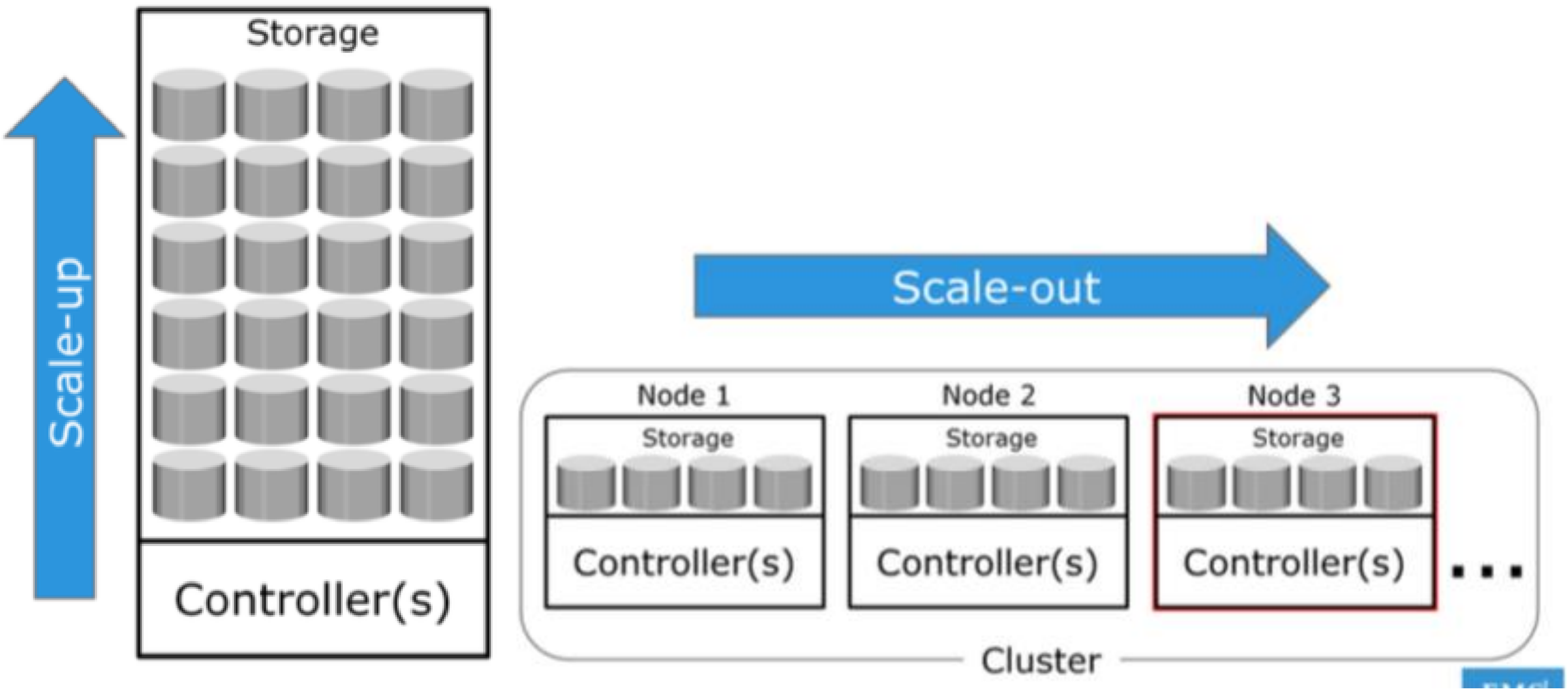
Refers to the process that temporarily replaces a failed disk drive with a spare drive in a RAID array by taking the identity of the failed disk drive.

- After the dynamic disk sparing process; when a new disk is added to the system, data from the hot spare is copied to it
- Hot spare then returns to its idle state, ready to replace the next failed drive
- Some systems implement multiple hot spares to improve data availability



GUC German University in Cairo

ISS Architectures
Scale-up vs. Scale-out



ISS Architectures

Scale-up vs. Scale-out

Scale-up Architecture

- Involves upgrading or adding controllers & storage
- Provides the capability to scale the capacity and performance of a single storage system
- Has a fixed capacity ceiling and limited scalability
- Performance starts degrading when reaching the capacity limit

Scale-out Architecture

- Involves upgrading and adding controllers & storage
- Provides the capability to maximize the capacity by adding nodes without causing any downtimes
- Pools the resources in the cluster and distributes the workload across all nodes
- Performance improves linearly as more nodes are added to the system

References

- “Cloud Infrastructures and Services - CIS” Course by Dell Technologies
- “Information Storage and Management – ISM” Course by Dell Technologies
- “IT Solutions for Digital Businesses - Virtualization and the Journey to the Modern Digital Workspace” Course by Vmware

For any inquiries e-mail me on:
maggie.ezzat@guc.edu.eg