

---

03 infra TI

# Ambiente de Armazenamento



---

**Principais componentes de Hosts e Armazenamento; Tipos de conectividade PCI, IDE/ATA, SCSI etc.; Componentes de um drive de disco; Desempenho de drives de disco; Sistemas de arquivos; LVM, Logical Volume Manager**

---

# Principais Componentes do Ambiente de Armazenamento

**Application:** A computer program that provides the logic for computing operations

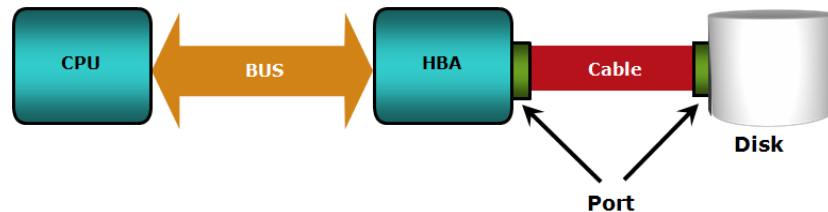
**Database management system (DBMS):** Provides a structured way to store data in logically organized tables that are interrelated

**Host or compute:** A computing platform (hardware, firmware, and software) that runs applications and databases

**Network:** A data path that facilitates communication among various networked devices

**Storage:** A device that stores data persistently for subsequent use.

## Host, Conectividade e Storage



# Hosts: Physical Components

---

Applications runs on hosts that can range from simple laptops to complex server clusters. Physical components of host:

## CPU

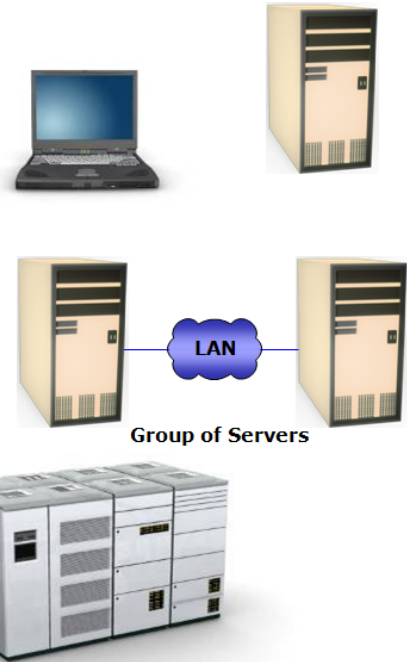
## Storage

Disk device and internal memory

## I/O device

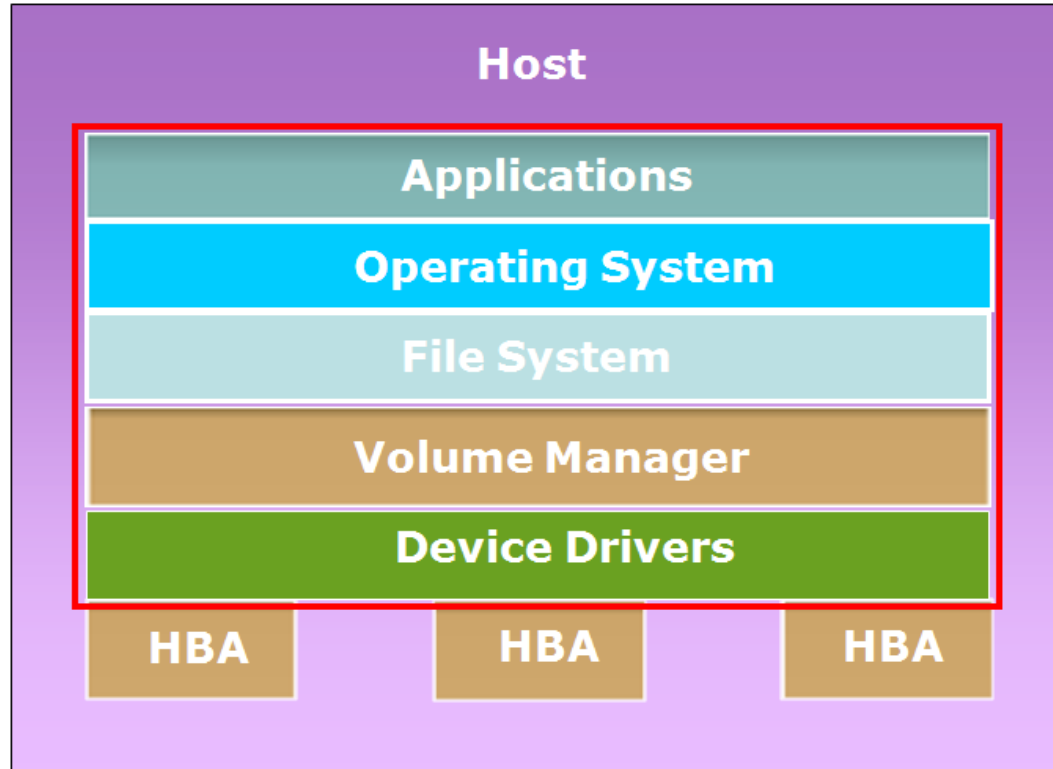
Host to host communications, Network Interface Card (NIC)

Host to storage device, Host Bus Adapter (HBA)



# Hosts: Logical Components

---



# Hosts: Logical Components

## Application

Application data access can be classified as:

### Block-level access:

Data stored and retrieved in blocks, specifying the LBA (logical block address)

### File-level access:

Data stored and retrieved by specifying the name and path of files

## Operating system

Resides between the applications and the hardware

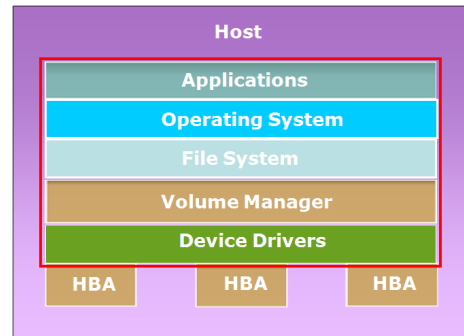
Controls the environment

## File System

File is a collection of related records or data stored as a unit

File system is hierarchical structure of files

Examples: FAT 32, NTFS, UNIX FS, EXT2/3 e **HDFS**



# Hosts: Logical Components

---

## LVM Logical Volume Manager

Responsible for creating and controlling host level logical storage  
Physical view of storage is converted to a logical view by mapping  
Logical data blocks are mapped to physical data blocks

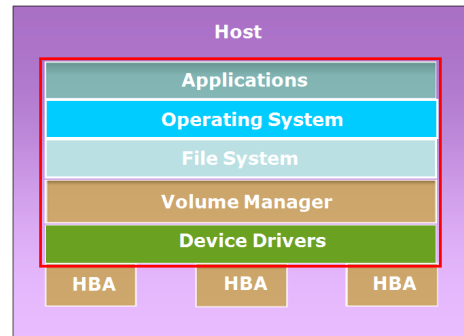
Usually offered as part of the operating system or as third party host software

## Device Drivers

Enables operating system to recognize the device

Provides API to access and control devices

Hardware dependent and operating system specific

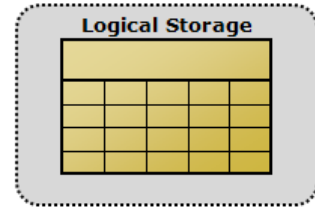


# LVM Logical Volume Manager

---

## LVM Logical Volume Manager

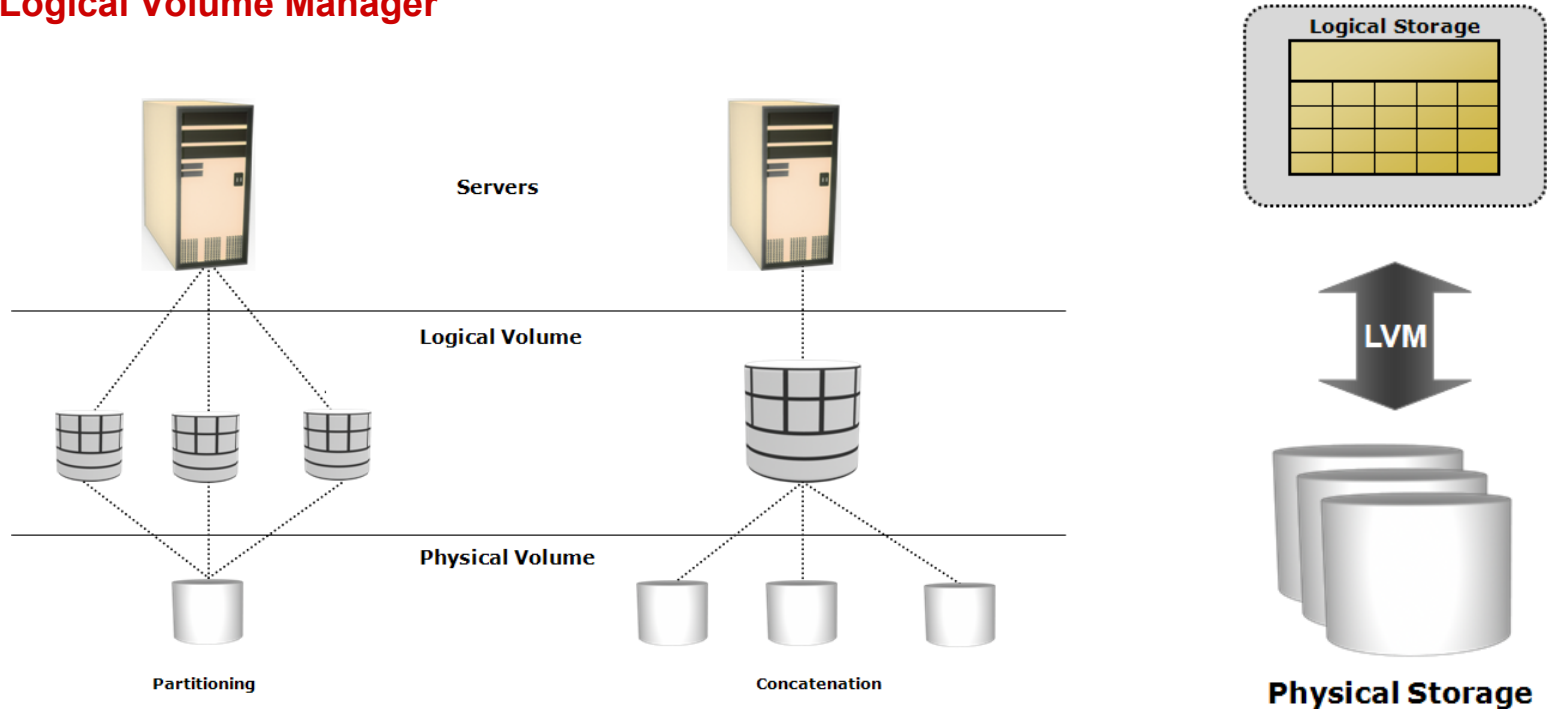
Responsible for creating and controlling host level logical storage  
Physical view of storage is converted to a logical view by mapping



**Physical Storage**

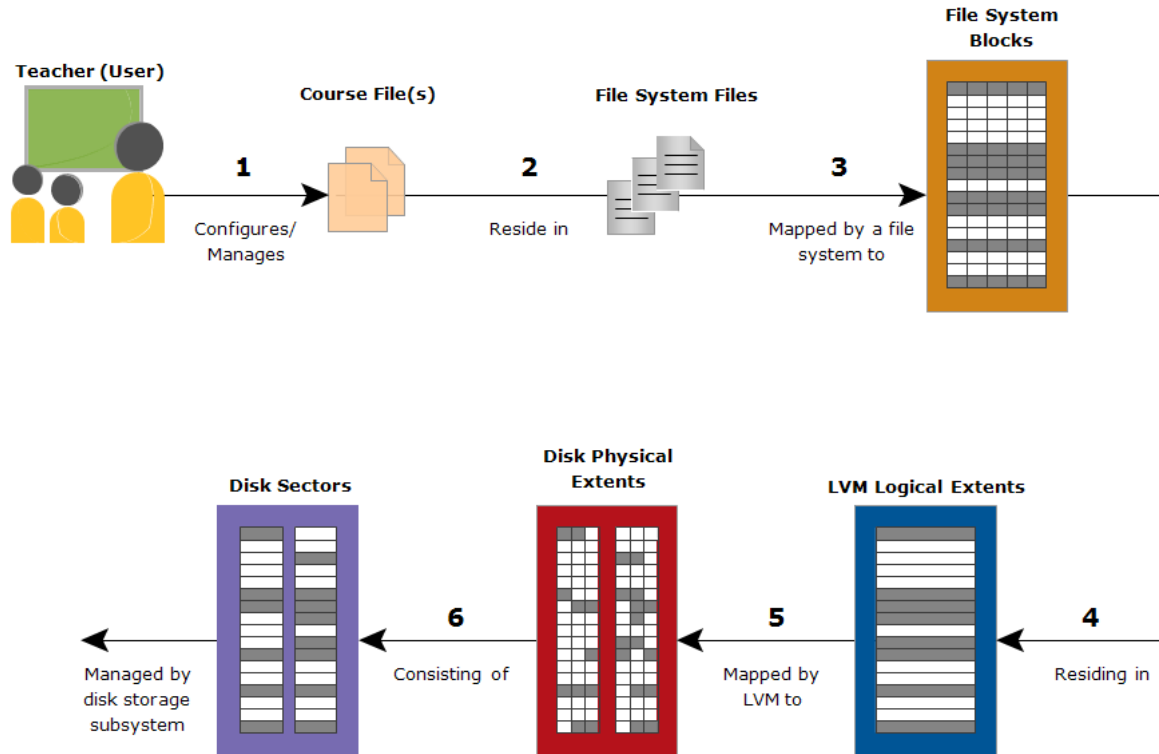
# LVM Partitioning & Concatenation

## LVM Logical Volume Manager





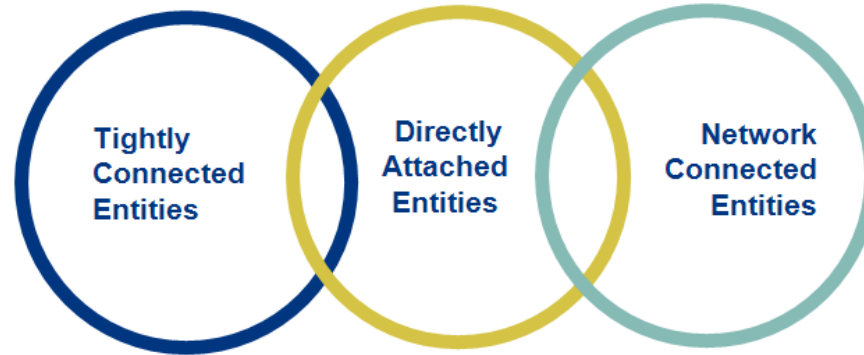
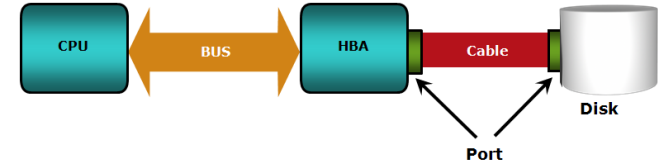
# Files: Storage-User path



# Conectividade

---

**Protocols define a format for communication between sending and receiving devices**



Tightly connected entities such as central processor to RAM, or storage buffers to controllers (example PCI)

Directly attached entities connected at moderate distances such as host to storage (example IDE/ATA)

Network connected entities such as networked hosts, NAS or SAN (example SCSI or FC)

---

# Conectividade

**PCI (Peripheral Component Interconnect) is used for local bus system**

It is an interconnection between microprocessor and attached devices, Has Plug and Play

PCI is 32/64 bit, Throughput is **133 MB/sec**

PCI Express is an enhanced version of PCI bus with higher throughput and clock speed

**Integrated Device Electronics (IDE) / Advanced Technology Attachment (ATA)**

Most popular interface used with modern hard disks

Good performance at low cost, Inexpensive storage interconnect

Used for internal connectivity

**Serial Advanced Technology Attachment (SATA)**

Serial version of the IDE /ATA specification

Hot-pluggable, Enhanced version of bus provides up to **6Gb/s** (revision 3.0)

**Parallel SCSI (Small computer system interface)**

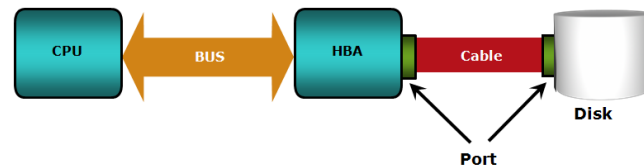
Most popular hard disk interface for servers

Higher cost than IDE/ATA

Supports multiple simultaneous data access

Used primarily in “higher end” environments

Data transfer speeds of **320 MB/s** (SCSI Ultra) to **3 Gb/s** (SAS 300)



**Table 5-1:** Comparison of IDE/ATA with SCSI

FEATURE	IDE/ATA	SCSI
Speed	100, 133, 150 MB/s	320 MB/s
Connectivity	Internal	Internal and external
Cost	Low	Moderate to high
Hot-pluggable	No	Yes
Performance	Moderate to low	High
Ease of configuration	High	Low to moderate
Maximum number of devices supported	2	16

# Storage Medias

## Magnetic Tape

Low cost solution for long term data storage

Limitations

Sequential data access, Single application access at a time, Physical wear and tear and Storage/retrieval overheads

## Optical Disks

Popularly used as distribution medium in small, single-user computing environments

Write once and read many (WORM): CD-ROM, DVD-ROM

Limited in capacity and speed

## Disk Drive

Most popular storage medium with large storage capacity

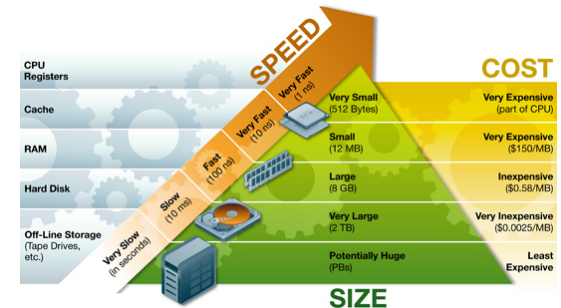
Random read/write access

Ideal for performance intensive online application

## Solid State Media or FLASH DRIVES

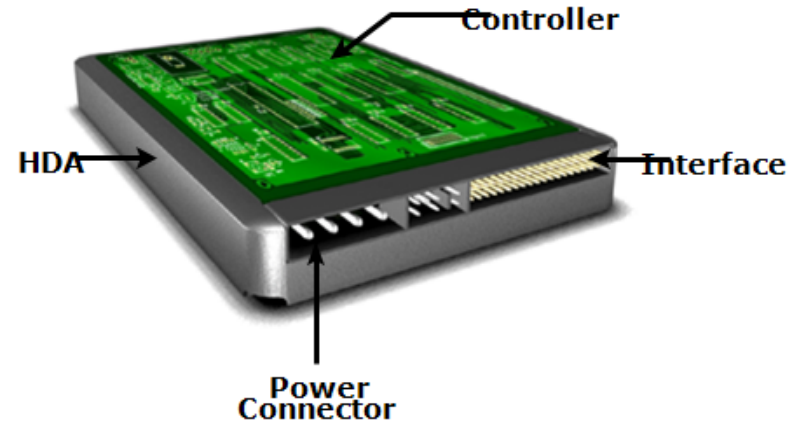
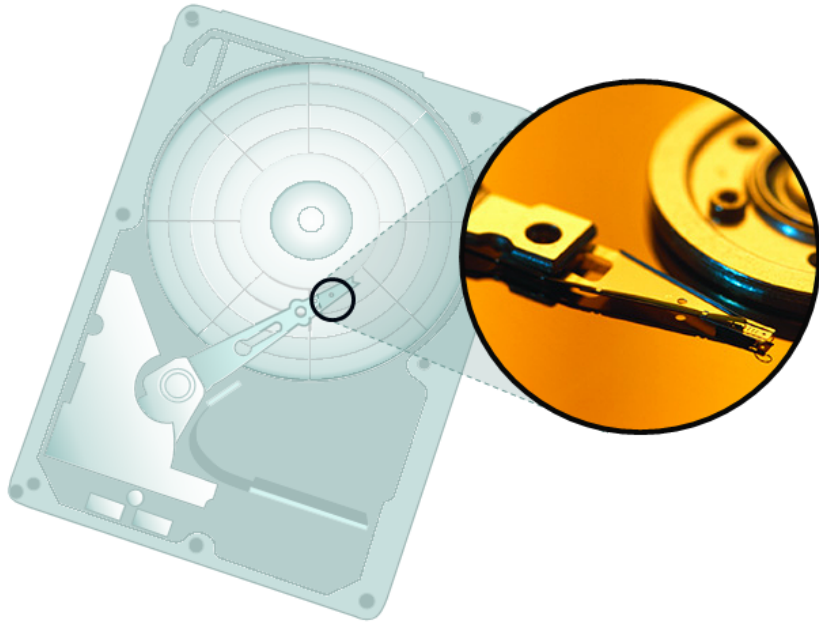
Expensive

Sem partes móveis, como circuitos integrados e placas-mãe em computadores



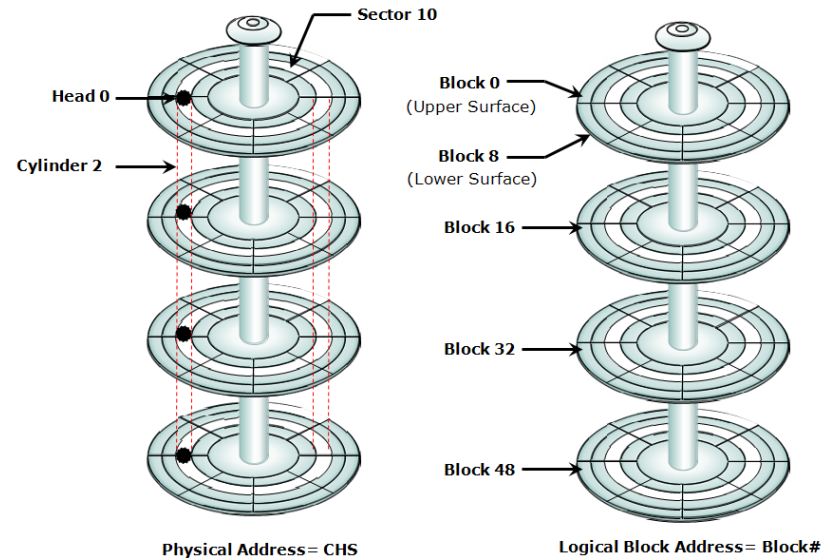
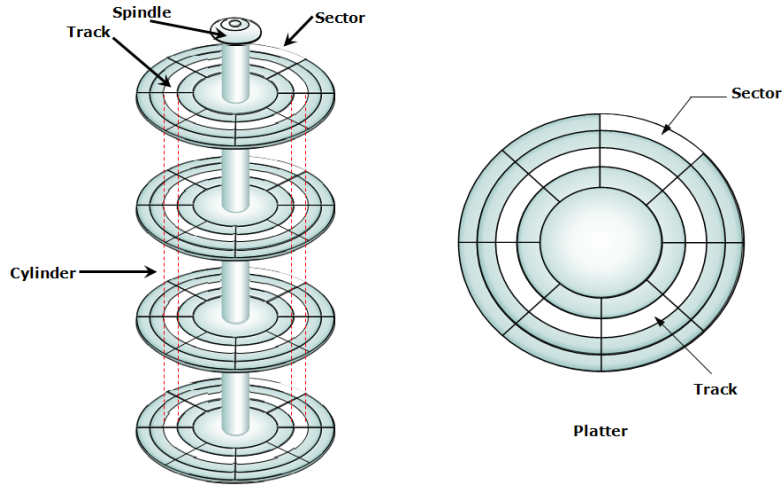
# Disk Drive Components

---



# Disk Drive

## Estrutura física e Endereçamento Lógico



# Disk Drive Performance

---

## Disk Service Time

Time taken by a disk to complete an I/O request

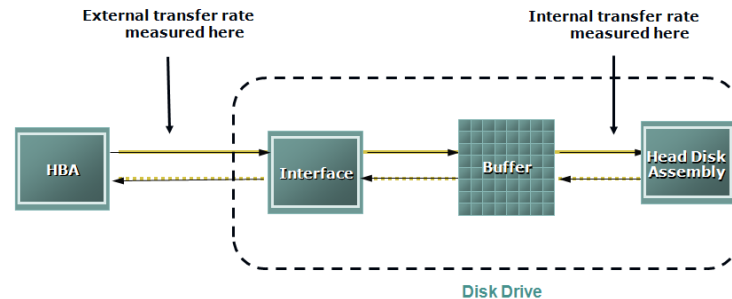
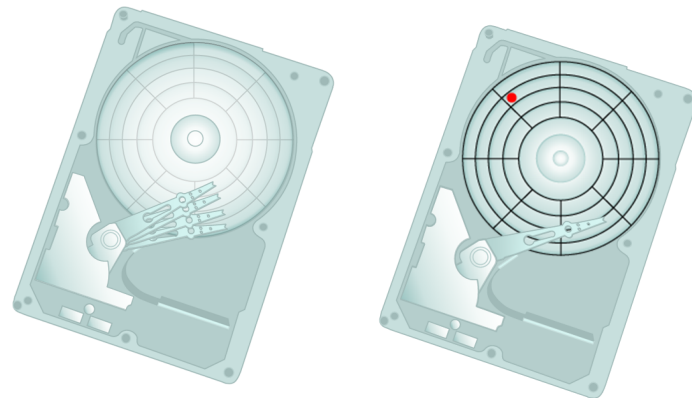
Seek Time

Rotational Latency

Appx. 5.5 ms for 5400-rpm drive, 2.0 ms for 15000-rpm drive

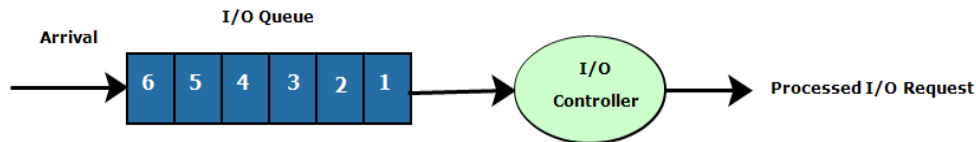
Data Transfer Rate

Qual é maior ?



# Disk Drive Performance Calc

---



I/O arrival rate,  $a$

Average inter-arrival time,  $R_a = 1 / a$

Utilization,  $U = R_s / R_a$

Average response time,  $R = R_s / (1 - U)$

Average queue size =  $U^2 / (1 - U)$

Time spent by a quest in queue =  $U \times R$



# Disk Drive Performance Calc

---

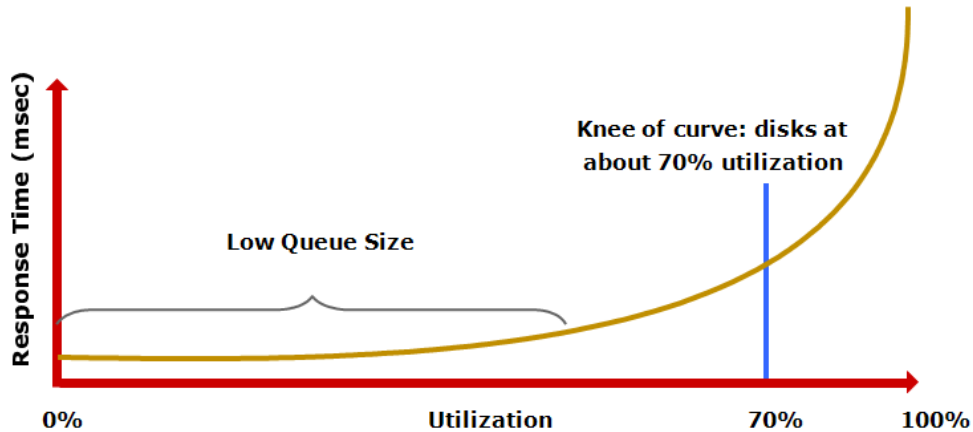
Consider a disk I/O system in which an I/O request arrives at a rate of 100 I/Os per second. The service time,  $R_s$ , is 8 ms.

I/O arrival rate, $a$	100 IOPS
Average inter-arrival time, $R_a = 1 / a$	10 ms
Utilization, $U = R_s / R_a$	8 ms / 10 ms = 0,8 = 80%
Average response time, $R = R_s / (1 - U)$	8 ms / (1-0,8) = 40 ms
Average queue size = $U^2 / (1 - U)$	3,2
Time spent by a quest in queue = $U \times R$	32 ms

---

# Utilização x Performance

---



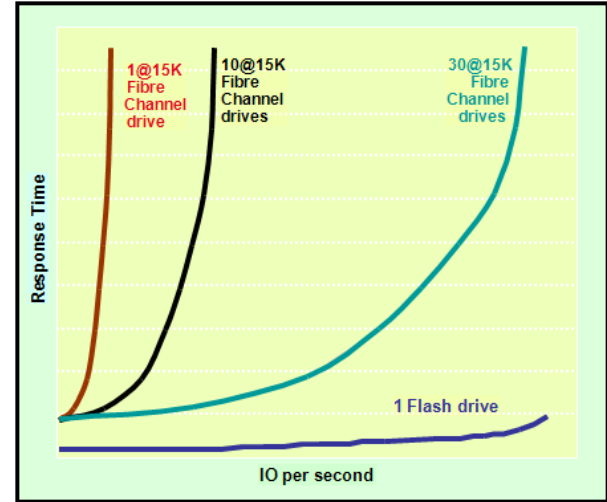
Consider a disk I/O system in which an I/O request arrives at a rate of 100 I/Os per second. The service time,  $R_s$ , is 4 ms.

Utilization of I/O controller ( $U = a \times R_s$ )

Total response time ( $R = R_s / (1 - U)$ )

Calculate the same with service time is doubled

# Flash Disk Drives



# Discussão e exercícios

---

Dê exemplos de conexões PCI e SCSI.

Um banco de dados requer um disco de 2TB. Mas os disk drives disponíveis são somente de 500GB. Que componente lógico do sistema pode ser utilizado para solução desse problema e como ?

Um disco com 500GB tem mesmo 500GB úteis?

Um sistema emprega 10 discos de 500GB e vem apresentando problemas de performance no I/O (alto tempo de resposta). Tendo disponível apenas mais volumes de disco como você resolveria esse problema?

Altere o exemplo de cálculo de performance de discos para 3000 IOPS. Qual o tempo de resposta e tamanho de fila obtidos?

---

# Leitura recomendada

---

## Capítulo 2

**Information Storage and Management Storing, Managing, and Protecting Digital Information in Classic, Virtualized, and Cloud Environments**

2nd Edition Edited by Somasundaram Gnanasundaram, Alok Shrivastava

---