

# Tema 4

---

## Tecnologías de almacenamiento

Firma

# Índice

---

- **Introducción**
- Organización del almacenamiento en el computador
- Discos duros y dispositivos SSD
- RAID

# Almacenamiento permanente

- Concepto

Capacidad de los sistemas informáticos de almacenar información durante un tiempo indefinido, independientemente de que dichos sistemas se encuentren o no conectados a su fuente de alimentación habitual.

- Dispositivos y tecnologías de almacenamiento permanente

Disco duro



Tecnología de almacenamiento

Magnética

CD/DVD



Tecnología de almacenamiento

Óptica

Unidad de estado sólido



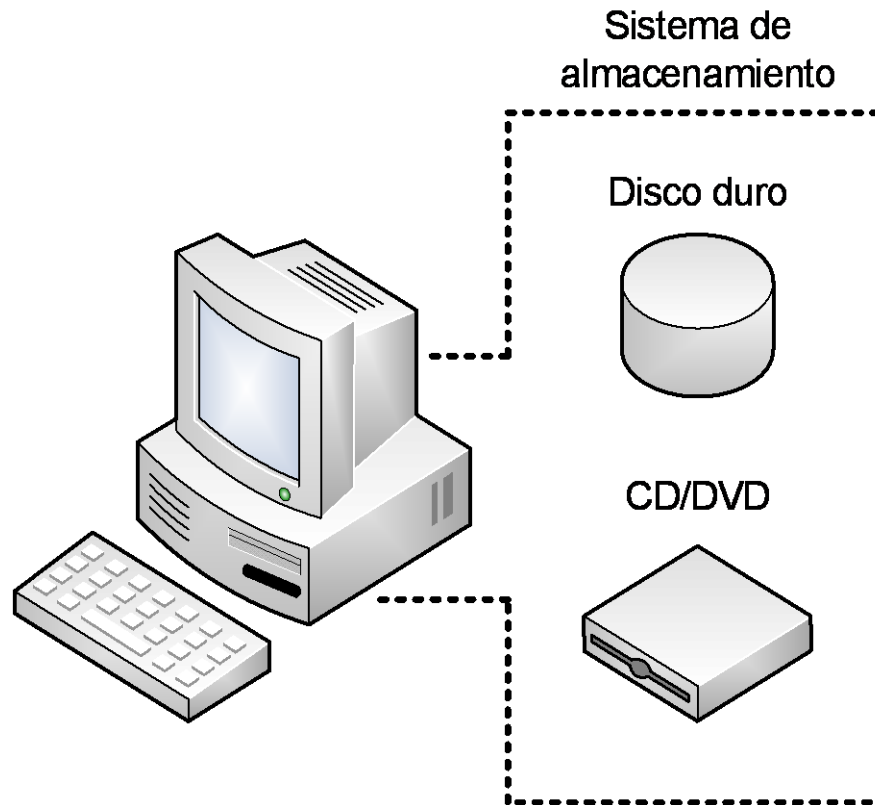
Tecnología de almacenamiento

Memoria flash

# Almacenamiento en el ordenador de escritorio

- Característica básica

Simplicidad: el sistema de almacenamiento está habitualmente formado por un disco duro y una unidad óptica.



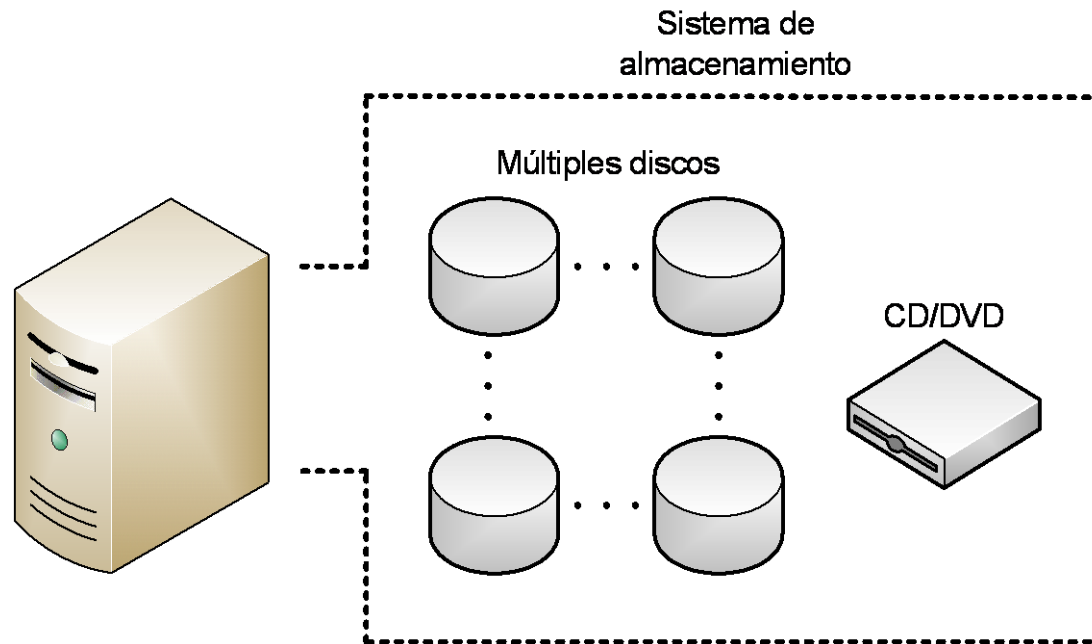
# Almacenamiento en el servidor

- Requisitos habitualmente requeridos

- Elevada capacidad de almacenamiento.
- Elevadas prestaciones: bajo tiempo de acceso y elevada tasa de transferencia.
- Tolerancia a fallos y alta disponibilidad.

- Diseño para el cumplimiento de los requisitos

Basado en múltiples discos habitualmente gestionados mediante controladoras RAID.



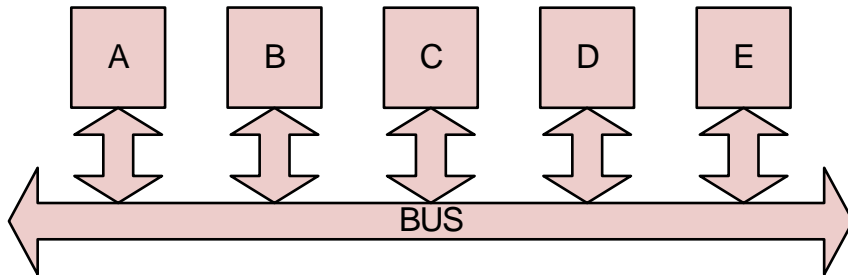
# Índice

---

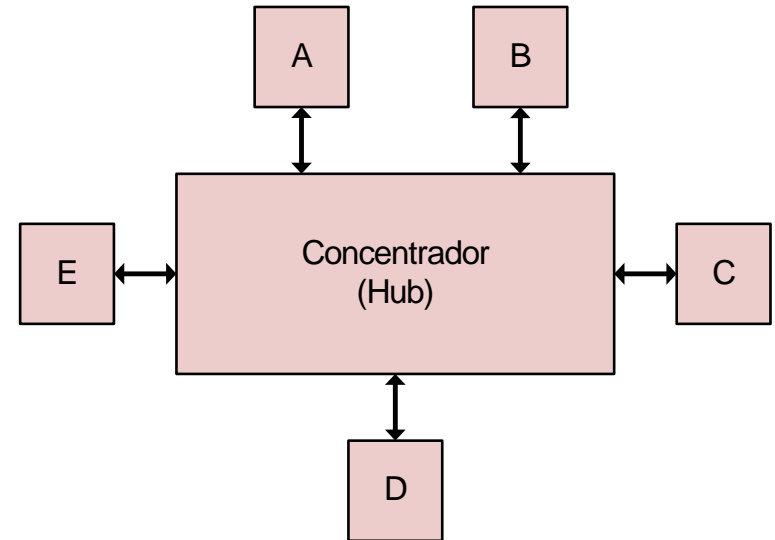
- Introducción
- **Organización del almacenamiento en el computador**
- Discos duros y dispositivos SSD
- RAID

# Repaso: Interconexión de dispositivos en el ordenador

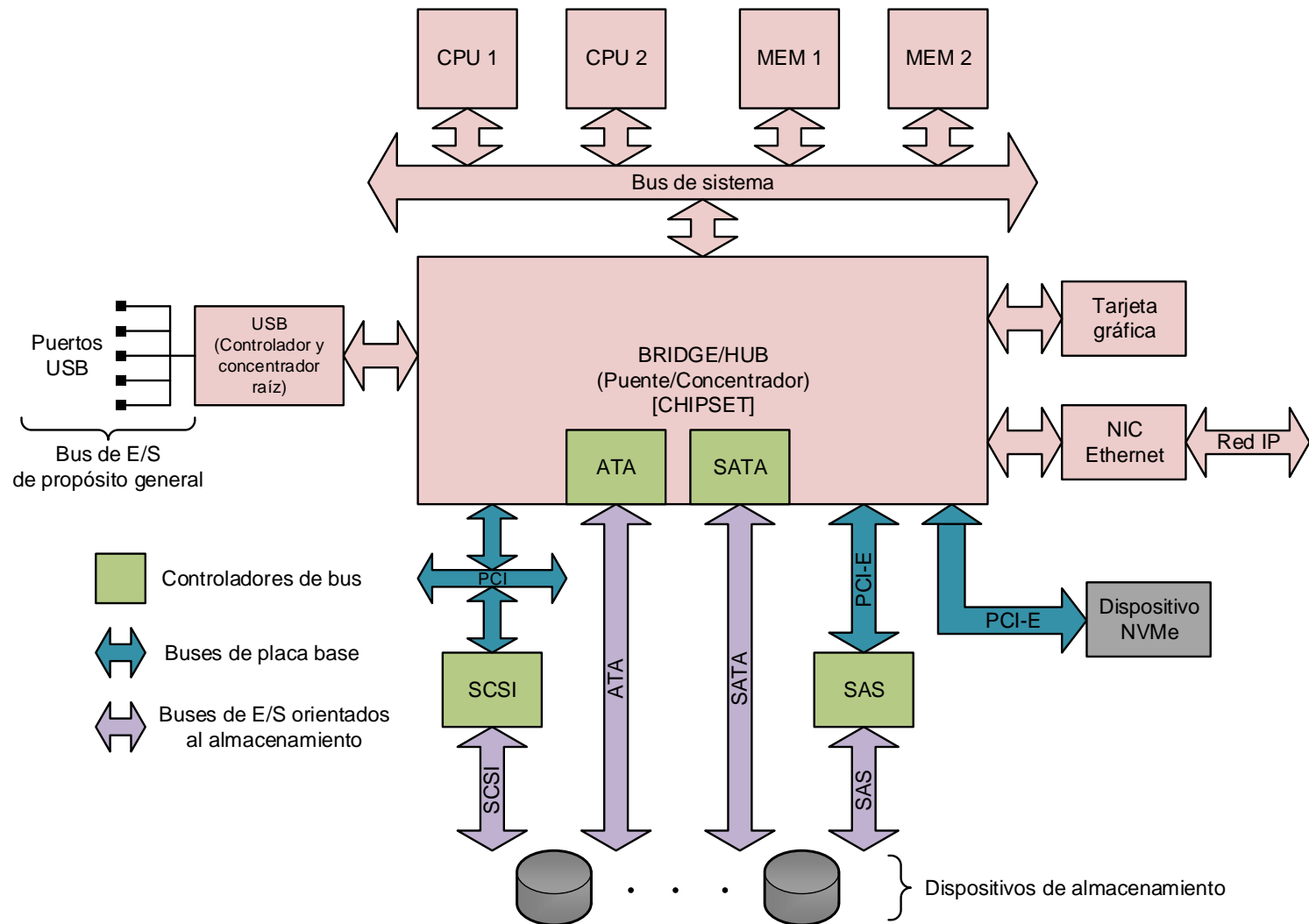
Interconexión en bus



Interconexión mediante concentrador



# Jerarquía de conexiones de un ordenador moderno





## Elementos de la jerarquía de conexiones de un ordenador moderno

- Bus de sistema

Su objetivo es conectar las CPU y la memoria RAM del ordenador. Se trata de un bus especialmente optimizado para proporcionar un elevado ancho de banda de comunicación entre estos dispositivos.

- Subsistema de concentradores (chipset)

Su objetivo es establecer una vía de comunicación entre los dispositivos conectados al bus de sistema (CPU y memoria) y el resto de dispositivos del computador.

- Bus de placa base

Su objetivo es proporcionar un mecanismo de integración de interfaces en el computador mediante el mecanismo de las ranuras de expansión. Los controladores de buses de E/S pueden conectarse a este bus.

- Controladores de bus

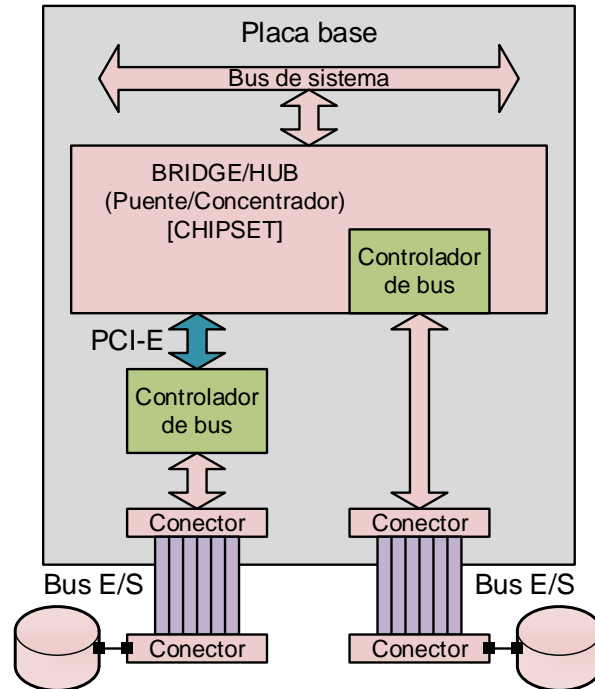
El objetivo de estos dispositivos es controlar buses de E/S. Para ello, gestionan el direccionamiento de los dispositivos en el bus y proporcionan un puente de comunicación entre el bus y otros dispositivos del sistema de interconexiones del computador.

- Buses de E/S

El objetivo de estos buses es proporcionar un mecanismo de conexión de periféricos al computador.

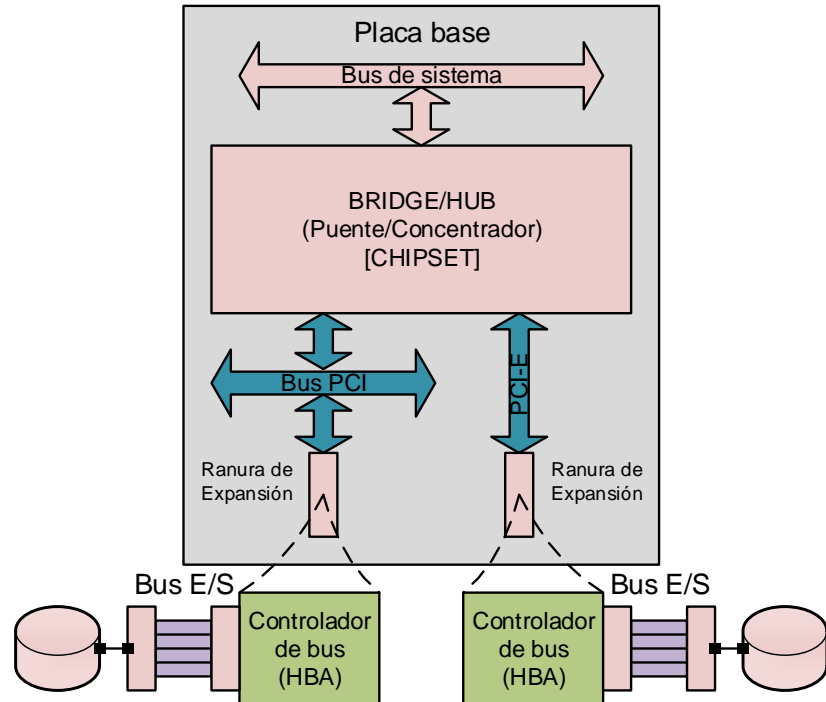
# Alternativas de implementación de los controladores de bus

**Controlador de bus en placa base**



La electrónica del controlador está integrada en la placa base: esta debe proporcionar un conector para conectar el cable que actuará como bus de E/S.

**Controlador de bus en HBA**



La electrónica del controlador se implementa mediante un HBA (Host Bus Adapter), que se integra en la placa base mediante una ranura de expansión perteneciente a un bus de placa base. El HBA proporciona un conector para conectar el cable que actuará de bus E/S.

# Buses de E/S

- El beneficio de la estandarización

Posibilitar la integración entre periféricos y ordenadores de múltiples fabricantes.

- Definición y mantenimiento de los estándares

Los estándares son gestionados, o bien por consorcios de fabricantes del sector, o bien por comités pertenecientes a institutos de estandarización, como el INCIS.

- Aspectos definidos por un estándar de bus

- Características físicas

Define aspectos tales como formatos de cables y conectores, especificaciones eléctricas, frecuencias de trabajo, etc.

- Arquitectura lógica

Define fundamentalmente las líneas que forman el bus así como su funcionalidad.

- Interfaz de programación

Define el conjunto de comandos disponibles para el manejo del bus.

PREGUNTA TÍPICA

# Buses de E/S estándares

- Buses de E/S de propósito general

- USB (Universal Serial Bus)

- ✓ Objetivo

Proporcionar un mecanismo simple para la conexión de periféricos externos a un ordenador.

- ✓ Características más destacadas

- Plug & Play.
    - Capacidad de alimentar a los dispositivos conectados.

- Buses de E/S orientados al almacenamiento

- ATA (Advanced Technology Attachment)

- SCSI (Small Computer System Interface)

- SATA (Serial ATA)

- SAS (Serial Attached SCSI)

# ATA: Aspectos generales

- Nomenclatura alternativa

IDE (Integrated Drive Electronics)

- Arquitectura

- Bus paralelo de ancho 16
- Capacidad de direccionamiento de un controlador: dos dispositivos de almacenamiento.

- Ámbito de uso

Ordenadores personales

- Evoluciones del estándar

7 evoluciones: ATA1 - ATA7

- Prestaciones

Hasta 133MB/s en ATA7.

# ATA: Implementación física

- Controlador

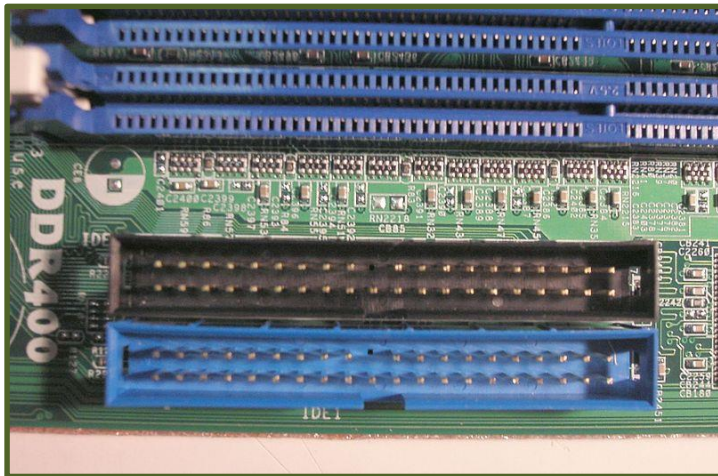
En los últimos años, los controladores ATA se han implementado de forma estándar en placa base, formando parte del chipset.

- Bus

Se implementa mediante un cable de 40 u 80 conductores. En el cable de 80, solo 40 transportan señales, los otros 40 son de aislamiento.

- Conectores

Cada controlador ATA tiene asignado un conector ATA en la placa base, que se utiliza para conectar el cable al controlador.



Conectores ATA



Cable ATA

# SCSI: Aspectos generales

- Arquitectura

- Bus paralelo diseñado con dos tipos de ancho: 8/16
- Capacidad de almacenamiento: 8 dispositivos para ancho 8, 16 para ancho 16.

- Ámbito de uso

Servidores y sistemas de almacenamiento externos (uds. de cinta)

- Evoluciones del estándar

9 evoluciones, desde la SCSI-1 hasta la ultra-320 SCSI

- Prestaciones

320MB/s en la ultra-320 SCSI

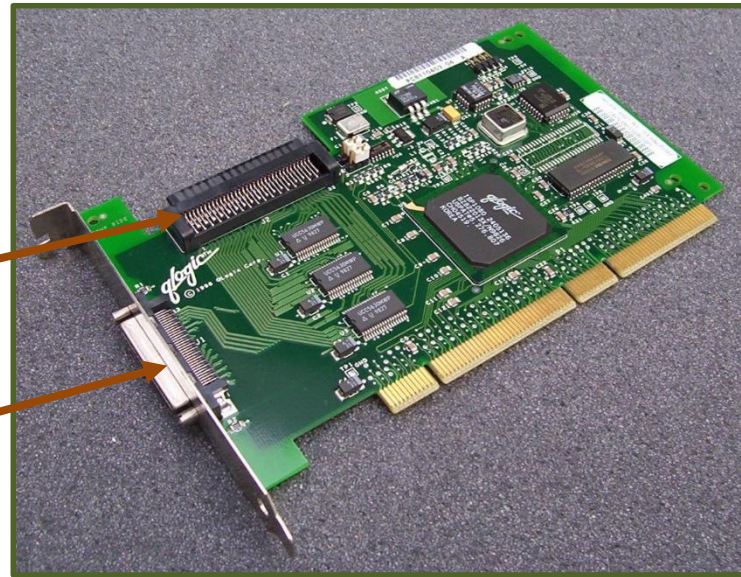
# SCSI: Implementación física

- Controlador

Se implementa con mayor frecuencia mediante HBA

Conector para cable interno

Conector para cable externo



HBA SCSI

- Bus

Se implementa mediante un cable de 68 conductores para los buses de ancho 16



Cable para conexión de dispositivos externos



Cable para conexión de dispositivos internos



# El cambio hacia el bus serie

- Prestaciones (ancho de banda)
  - Factores determinantes en las prestaciones de un bus
    - ✓ Acho (tamaño de los datos transmitidos)
    - ✓ Frecuencia de trabajo
  - Comparativa en prestaciones bus paralelo / bus serie respecto a ancho y frecuencia
    - ✓ Ancho: Bus paralelo ventajoso
    - ✓ Frecuencia: Bus paralelo desventajoso

- Existen un conjunto de problemas de diseño electrónico en los buses paralelos que limitan su frecuencia de trabajo, agravándose dicha limitación en la medida que aumenta la longitud del bus.  
- La longitud requerida para un bus de E/S es significativa.

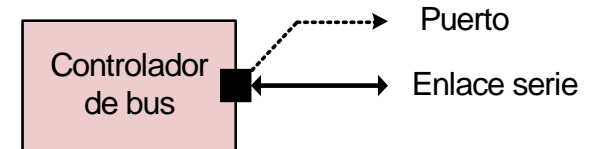
- ✓ Resultado combinado (Acho + frecuencia): Bus serie ventajoso
- Espacio ocupado (*footprint*) por los elementos del bus



- Espacio ocupado: Bus serie ventajoso

# Interconexión de dispositivos mediante bus serie

- Concepto de puerto



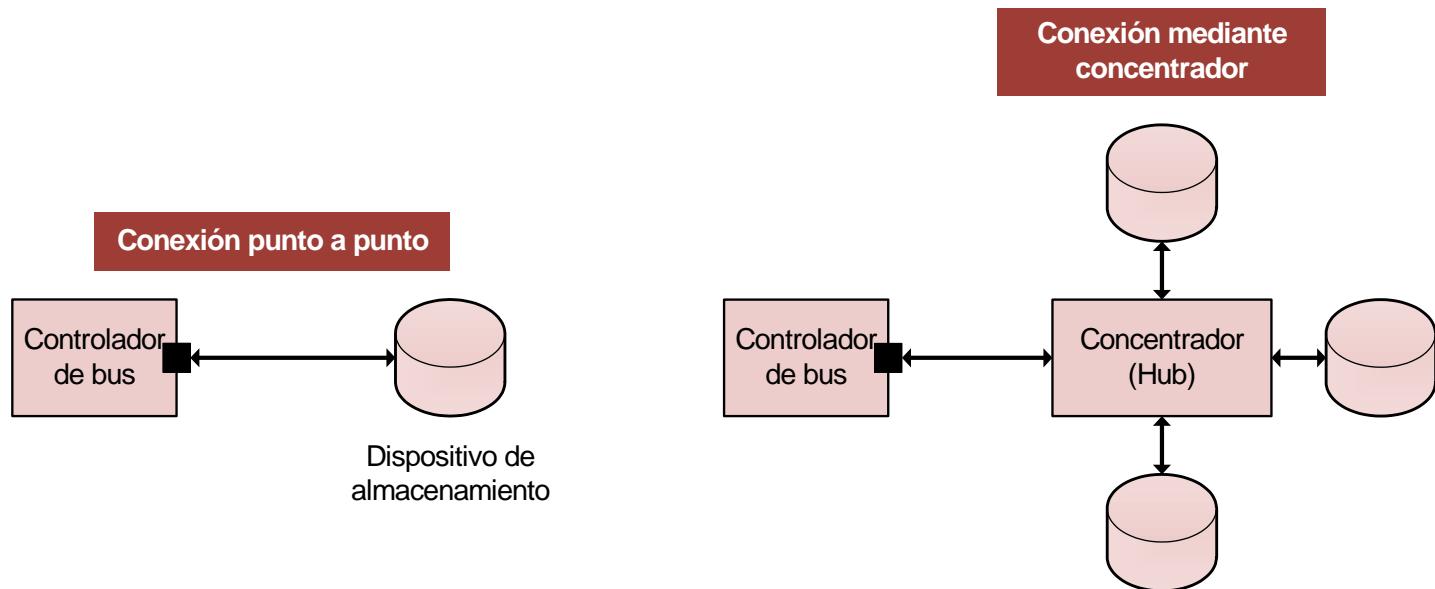
- Modelos de conexión

- Punto a punto

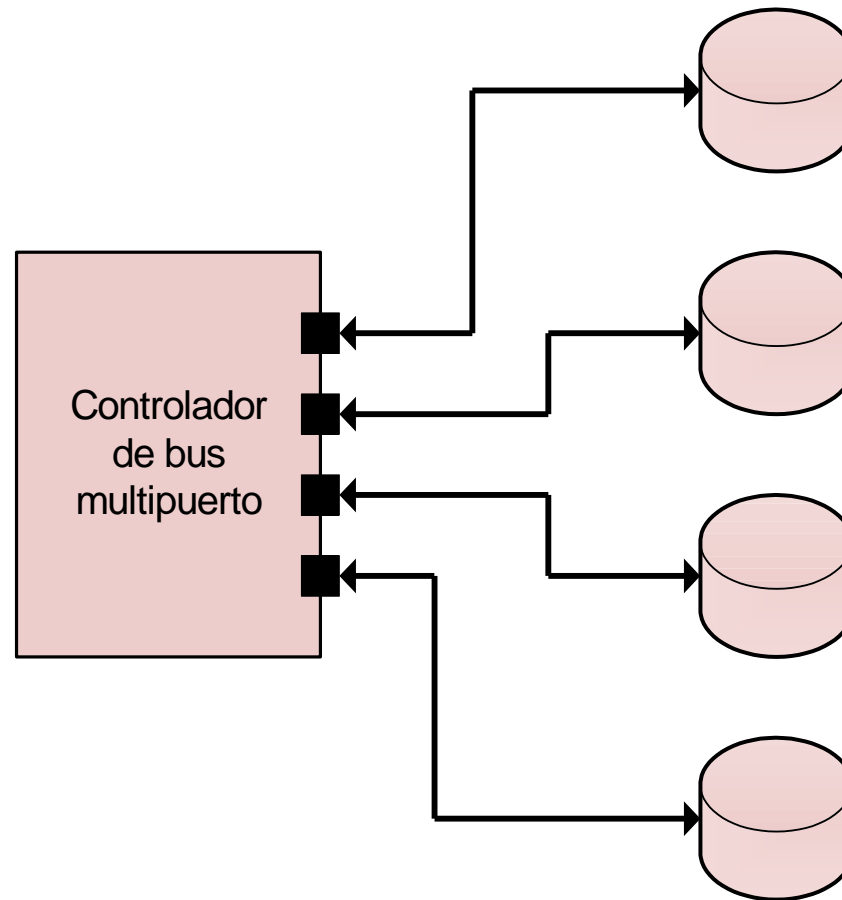
Un puerto, un dispositivo

- Mediante concentrador

Múltiples dispositivos se conectan a un puerto mediante un concentrador



# Controladores multipuerto



# SATA: Aspectos generales

- Nomenclatura

- Compatibilidad con el estándar ATA

- Arquitectura

- Canal de comunicación

- Controlador de bus

- Capacidad de direccionamiento

- Ámbito de uso

# SATA: Estándar

- Evolución del estándar y sus prestaciones

Revisión	Nombre comercial	Frecuencia (GHz)	Ancho de banda (MB/s)
1.0	SATA 1.5 Gbit/s	1,5	150
2.0	SATA 3 Gbit/s	3	300
3.0	SATA 6 Gbit/s	6	600

- Cálculo del ancho de banda (SATA 1.5 Gbit/s)

Nota: Los bytes se codifican con 10 bits con objeto de incluir información de sincronización

- Comparación de prestaciones ATA/SATA
  - ATA-7: 133 MB/s
  - SATA 6 Gbit/s: 600 MB/s

# SATA: Implementación física

- Controlador

Se implementan de forma estándar en placa base, formando parte del chipset.

- Enlace serie

Se implementa mediante un cable de 7 conductores: 4 para los dos canales de transmisión + 3 masas

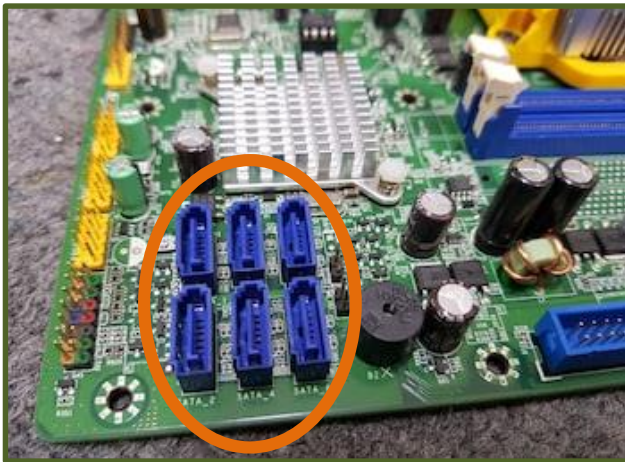
- Cable SATA



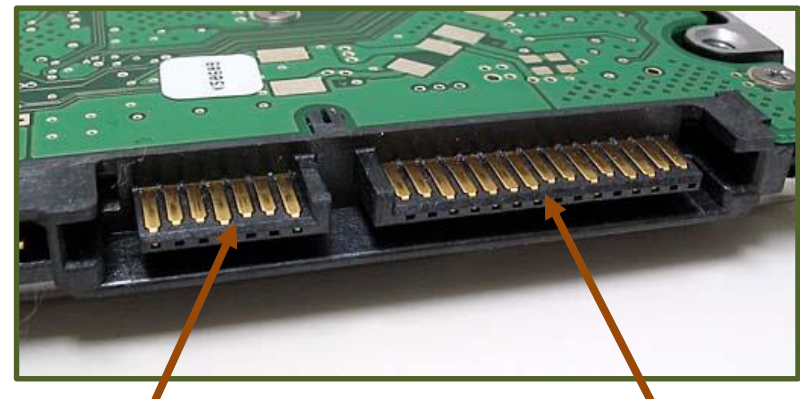
El cable SATA ocupa mucho menos espacio que el cable ATA

- Conectores

Conectores en la placa base



Conectores en el disco



Conector de datos

Conector de alimentación

# SAS: Aspectos generales

- SAS significa Serial Attached SCSI
- Compatibilidad con el estándar SCSI
  - 100% en la interfaz de programación
  - No compatible a nivel físico (cables y conectores)
- Ámbito de uso

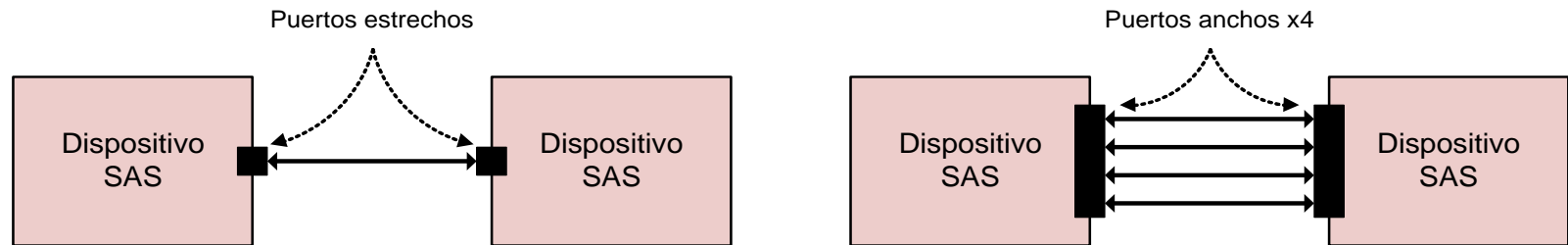
# SAS: Arquitectura (I)

- Canal de comunicación

- Concepto ampliado de puerto

- Puerto estrecho (*narrow port*)

- Puerto ancho (*wide port*)



- Interconexión de dispositivos SAS

- Punto a punto
- A través de concentradores

- Los concentradores SAS se denominan “expanders”
- Se pueden interconectar hasta dos niveles de “expanders”



# SAS: Arquitectura (II)

- Direcccionamiento de dispositivos

- Dominio SAS

- Roles de dispositivo en un dominio

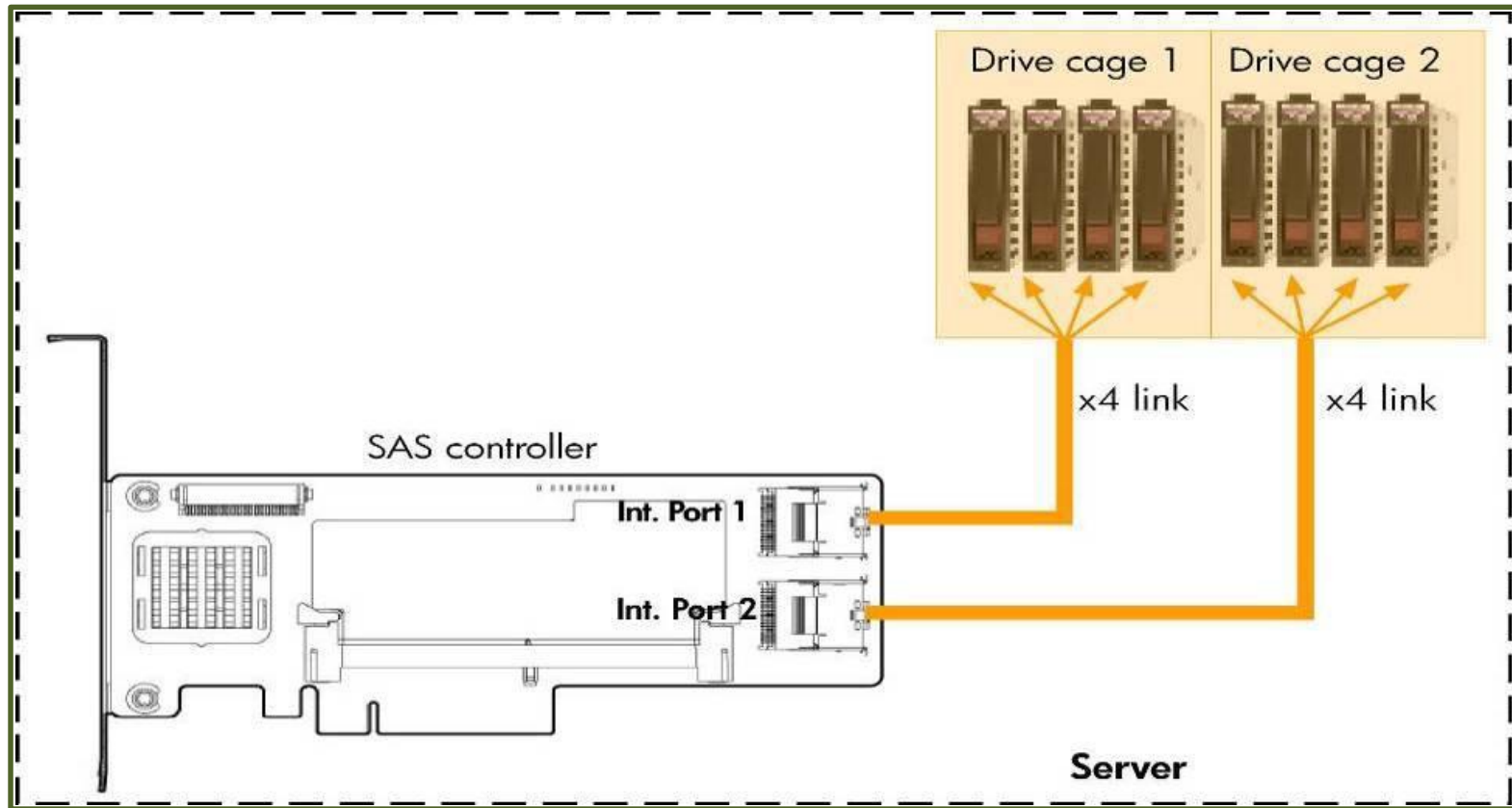
- ✓ *Initiator* (habitualmente, el puerto de un controlador)

- ✓ *Target* (habitualmente, un dispositivo de almacenamiento)

- Capacidad de direccionamiento

# SAS: Ejemplo de topología de interconexión (I)

Topología basada en conexiones punto a punto

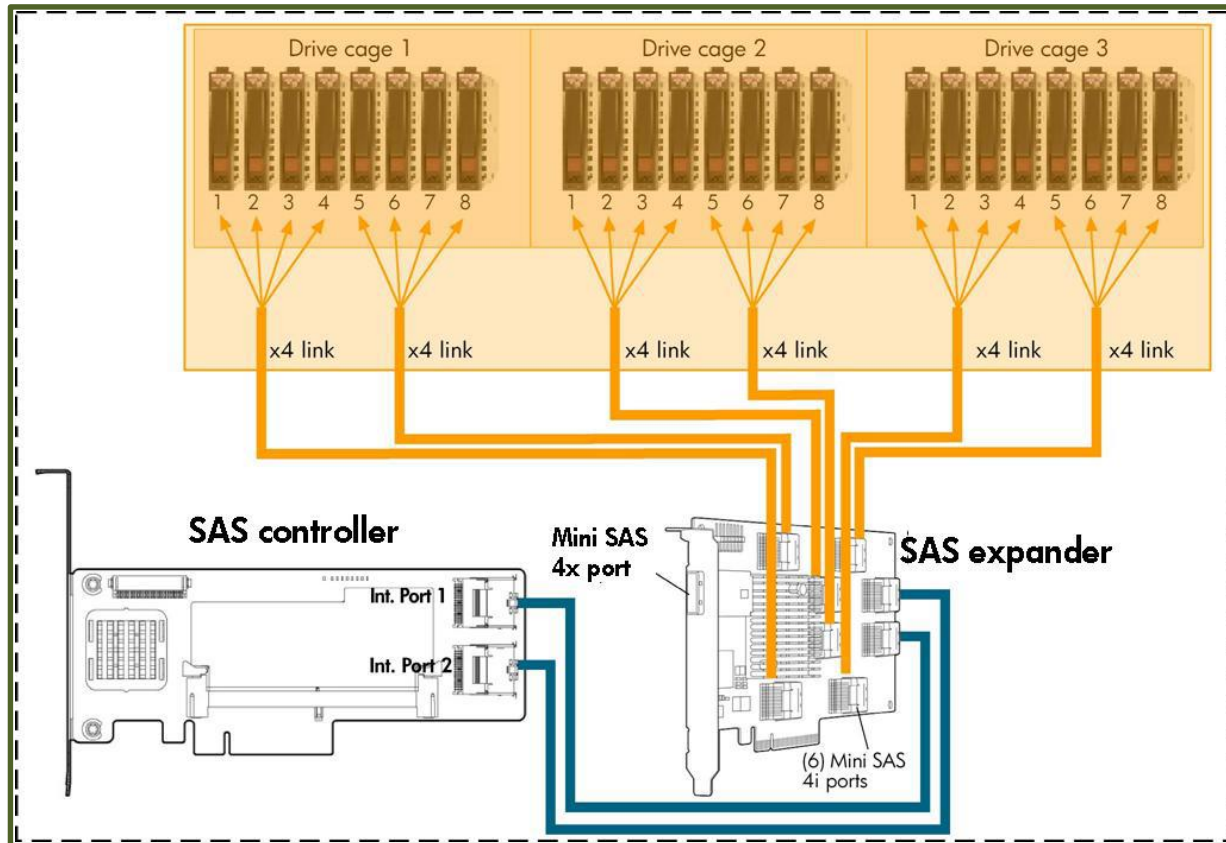


- Topología SAS típica para la organización del almacenamiento interno de un servidor

-

# SAS: Ejemplo de topología de interconexión (II)

Topología basada en concentradores



- Topología SAS para la organización del almacenamiento interno de un servidor de alta capacidad

# SAS: Estándar

- Evolución del estándar y sus prestaciones

Versión	Frecuencia (GHz)	Ancho de banda (MB/s)
SAS-1	3	300
SAS-2	6	600
SAS-3	12	1.200

- Cálculo del ancho de banda (SAS-2)

Nota: como en el caso de SATA, en SAS los bytes se codifican con 10 bits.

- Ancho de banda de un puerto SAS-3 x4

- Comparación de prestaciones SCSI/SAS
  - Ultra-320 SCSI: 320 MB/s
  - SAS-3: 1200 MB/s

# SAS: Implementación física (I)

- Controlador

Se implementa tanto en placa base como mediante HBA



HBA SAS

- Enlace serie

- Se implementa mediante un cable de 7 conductores: 4 para los dos canales de transmisión + 3 masas
- La implementación del enlace SAS es idéntica a la del enlace SATA

# SAS: Implementación física (II)

- SAS Backplane

- Concepto

- Objetivo

- Uso

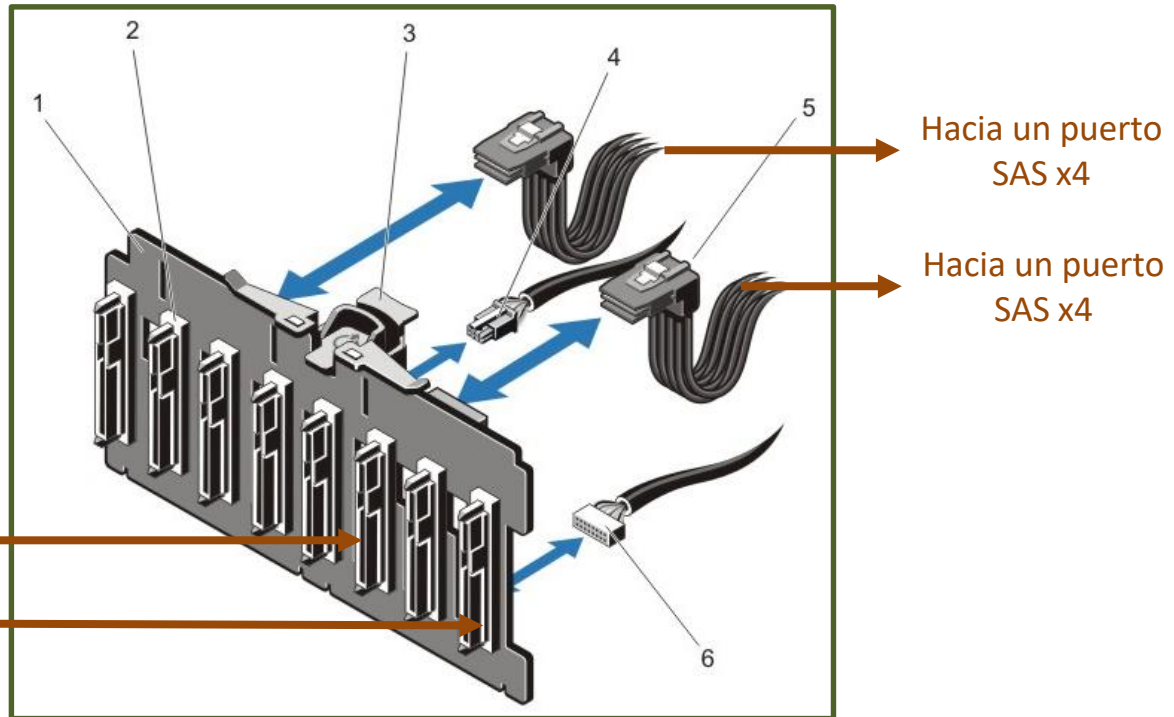
2) Conector SAS para disco

4) Cable de alimentación

5) Cable SAS

Disco SAS

Disco SAS



# SAS: Tecnología de discos con doble puerto

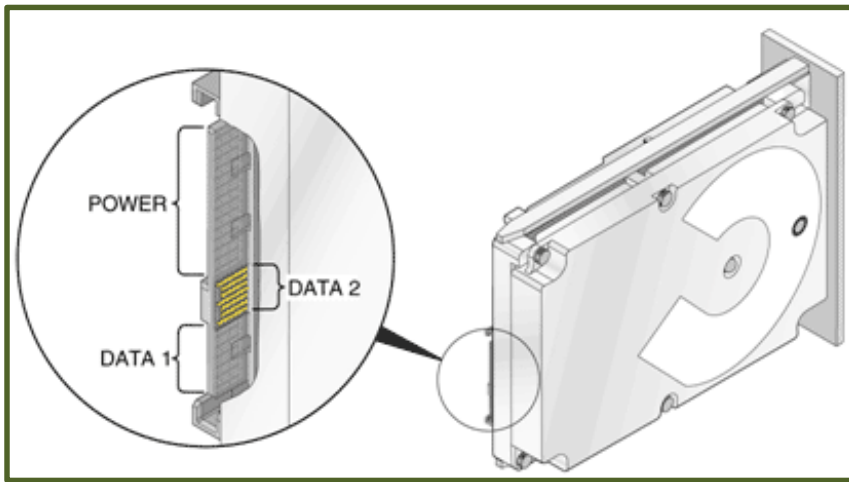
- Concepto

- Objetivos

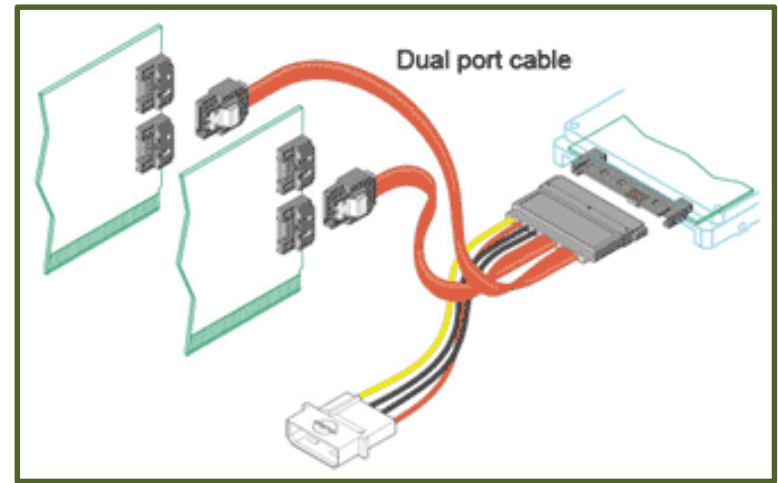
- Permitir la implementación de topologías de conexión de discos con bus redundante

- Duplicar la tasa de transferencia de datos entre el controlador y el disco

Detalle del conector de un disco con doble puerto



Conexión de disco con doble puerto a una topología con bus SAS redundante



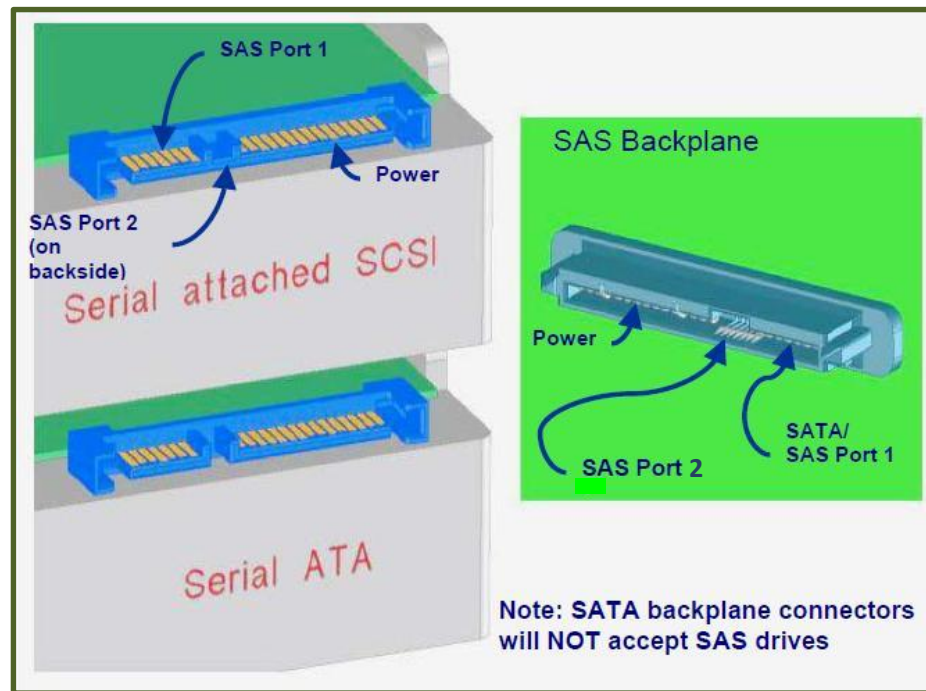
# SAS: Compatibilidad con SATA

- El bus SAS admite dispositivos SATA

- Implicaciones relativas al controlador

- Implicaciones relativas a los conectores

## Comparación conectores SAS/SATA





# NVMe: Aspectos generales

- NVMe significa Non-Volatile Memory Express

- Concepto

- Productos de la implementación del estándar NVMe

- Rediseño del modelo de driver desde cero

- Objetivo

- ¿Qué hace posible el rediseño desde cero?

- Sistemas operativos con soporte nativo para NVMe

Todos los Windows modernos (10 y las versiones de servidor desde 2012 R2) y todas las distribuciones Linux con la versión del kernel 3.3 o superior.

- Diferencia esencial de NVMe respecto a SAS y SATA

SAS y SATA especifican buses físicos (basados en enlaces serie), además de una interfaz de dispositivo lógico. Por el contrario, NVMe solo especifica una interfaz de dispositivo lógico, ya que el bus físico utilizado por los dispositivos basados en esta especificación es el PCIe.

- Ámbito de uso

# Mejoras de NVMe frente a SAS y SATA

- Manejo de múltiples colas de comandos de E/S de mayor profundidad

- Colas de comandos en SATA

Una, de profundidad 32.

- Colas de comandos en SAS

Una, de profundidad 256.

- Colas de comandos en NVMe

Hasta 64 K, de profundidad 64 K.

- Beneficio

- Racionalización y simplificación del código de E/S (usa menos de la mitad de instrucciones)

- Beneficio

# Factores de forma de dispositivos NVMe y formatos de conexión al bus PCIe

NOTA sobre PCIe: Las interfaces PCIe gestionan enlaces serie (*lanes*). El número de enlaces en una interfaz puede ser x1, x2, x4, x8, x16 y x32

- 1)
- 2)
- 3)

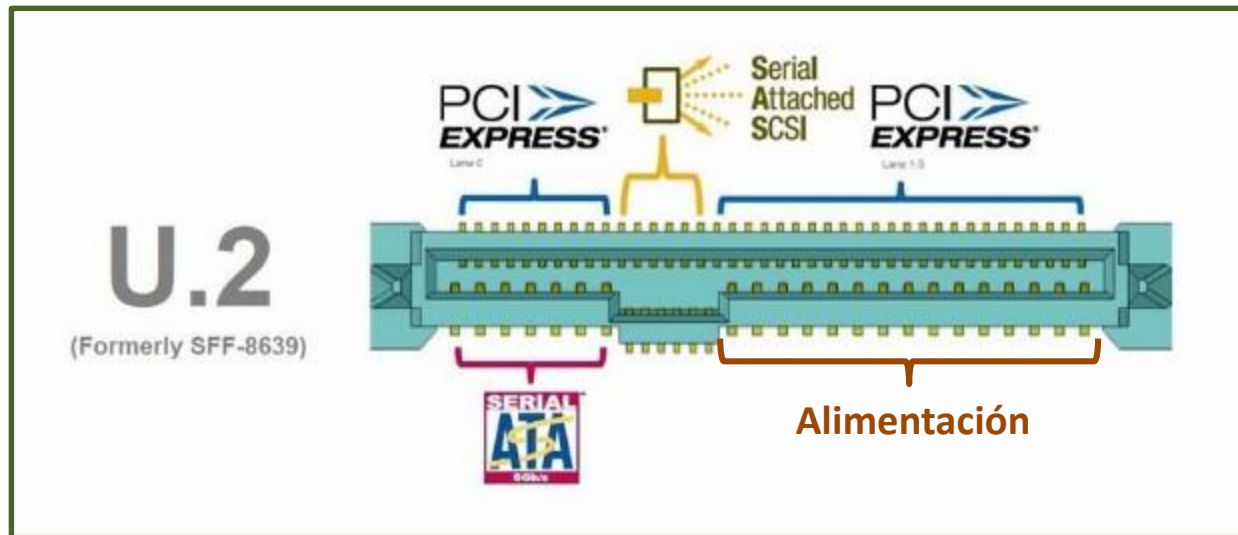
## Anchos de banda PCIe 3.0

x1	984.6 MB/s
x2	1.97 GB/s
x4	3.94 GB/s
x8	7.88 GB/s



# Conector U.2 para backplane

- Diseñado para la conexión en caliente de dispositivos NVMe
- Conexión PCIe x4
- Compatible con dispositivos SAS y SATA



# Índice

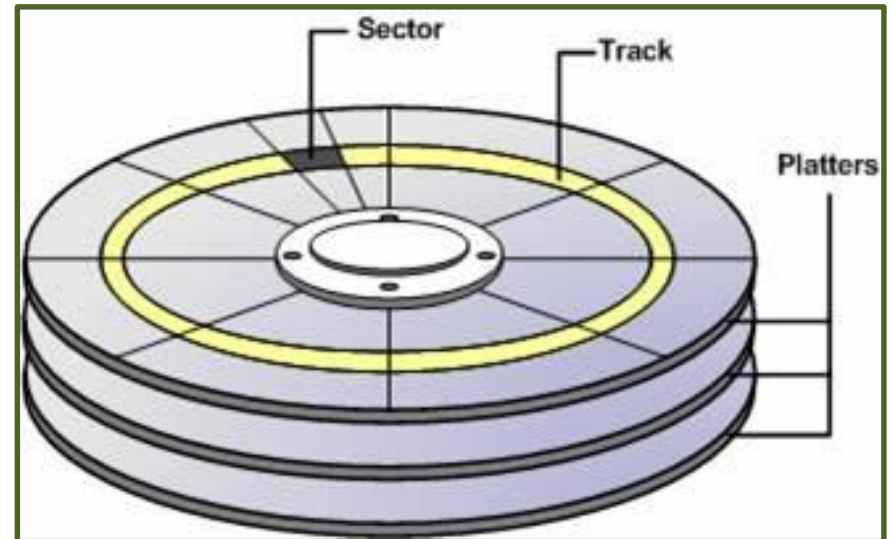
---

- Introducción
- Organización del almacenamiento en el computador
- **Discos duros y dispositivos SSD**
- RAID

# Estructura de un disco duro (HDD)



Organización de la información  
en un disco duro



En un disco duro la información se organiza en superficies, pistas y sectores.

# Características de un disco duro

- Capacidad

- Concepto

- Expresión

- Consumo

- Factor de forma

- Concepto

- Tipos

- ✓ LFF (Large Form Factor):

- ✓ SFF (Small Form Factor):

- Comparación LFF/SFF

Menor footprint	Mayor capacidad	Menor consumo

# Características de un disco duro

- Velocidad de rotación

Valores habituales: \_\_\_\_\_

- Cache

- Interfaz

- Número de puertos de datos

Modelos SATA: \_\_\_\_

Modelos SAS: \_\_\_\_

- Fiabilidad

- Forma de expresión habitual: AFR (Annualized Failure Rate)

- Valores típicos del AFR en discos de clase empresarial



# Características de un disco duro / prestaciones

- Aspectos de las prestaciones
    - Aspecto 1: Velocidad de posicionamiento de las cabezas
- Se expresa mediante el tiempo medio de acceso

Componentes del tiempo medio de acceso

- ✓ Tiempo de búsqueda medio (*average seek time*)

Tiempo medio que tardan las cabezas en moverse desde una pista a otra aleatoriamente elegida.

- ✓ Latencia (*latency*)

tiempo medio requerido para acceder a un sector, una vez que la cabeza se ha posicionado en la pista que contiene dicho sector. Depende de \_\_\_\_\_

- Aspecto 2: Velocidad de transferencia cabeza/disco

Se expresa como la velocidad de transferencia sostenida (*sustained transfer rate*)

Número de bytes por unidad de tiempo que pueden ser movidos entre el disco y la interfaz que lo controla. Valores máximos actuales: alrededor de 250 MB/seg

# Tipos de acceso a disco y su relación con las prestaciones

- Acceso aleatorio

- Concepto

Acceso a grupos de datos de pequeño tamaño y dispersos

- Ejemplo de escenario de acceso aleatorio

- Factor determinante en la velocidad del acceso aleatorio

- Concepto de IOPS

- Acceso secuencial

- Concepto

Acceso a bloques de datos de gran tamaño ubicados en el disco de forma contigua

- Ejemplo de escenario de acceso secuencial

- Factor determinante en la velocidad de acceso secuencial

# Ejemplos de discos gama empresarial

## Seagate

### Gama Exos orientada a prestaciones

Interfaz	SAS a 12 Gb/s
Factor de forma	SFF
Capacidades (GB)	300 - 2400
Cache (MB)	256
r.p.m.	10.000, 15.000
Velocidad de transferencia sostenida máxima (MB/s)	315
IOP	-
AFR	0,44%

### Gama Exos orientada a capacidad

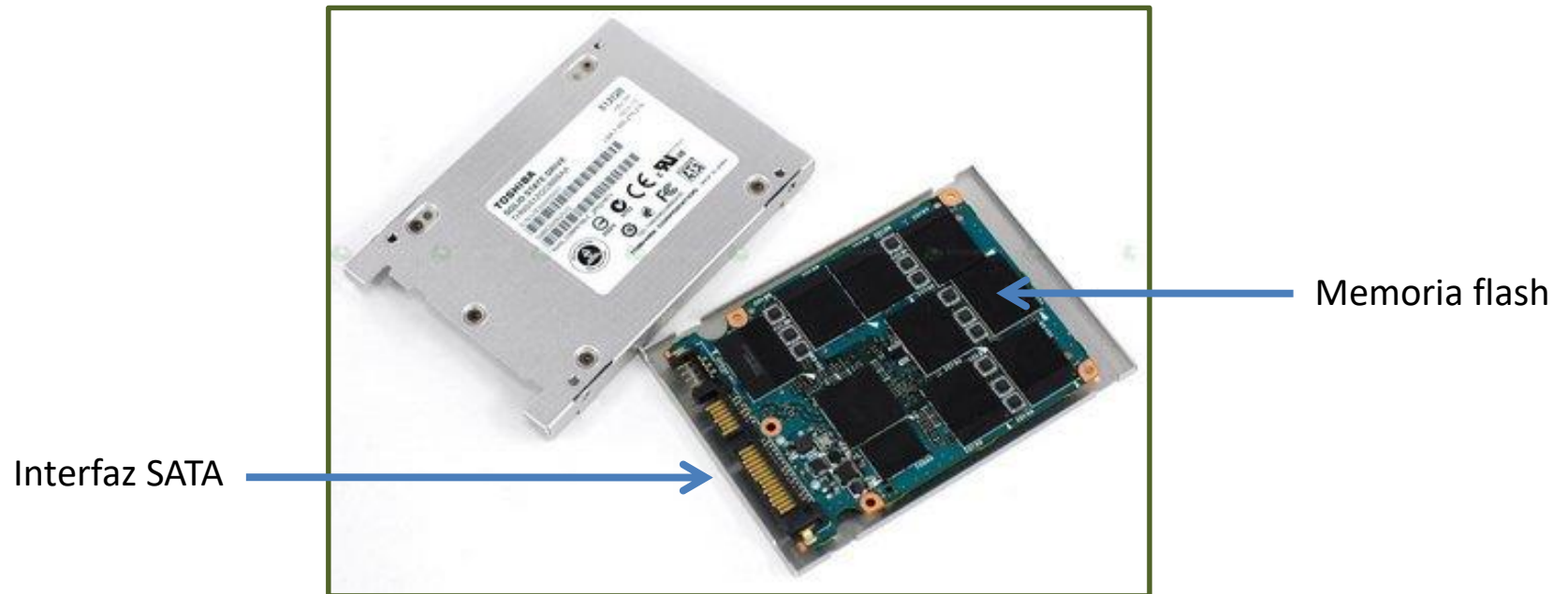
Interfaz	SAS a 12 Gb/s SATA 6 Gb/s
Factor de forma	LFF
Capacidades (TB)	1 - 14
Cache (MB)	256
r.p.m.	7.200
Velocidad de transferencia sostenida máxima (MB/s)	261
IOP	-
AFR	0,35%

# Dispositivos SSD

- SSD significa *Solid State Drive*
- Concepto



Ejemplo de dispositivo SSD con Interfaz SATA



# Factores de forma de dispositivos SSD y sus interfaces de conexión

- 2,5" SFF
  - Interfaces de conexión
    - ✓ SATA
    - ✓ SAS
    - ✓ PCIe x4 (mediante conector U.2)
- M.2
  - Interfaces de conexión
    - ✓ SATA
    - ✓ PCIe x2 o x4
- Tarjeta de expansión PCIe estándar
  - Interfaces de conexión
    - ✓ PCIe x4 o x8

# Factores de forma de dispositivos SSD

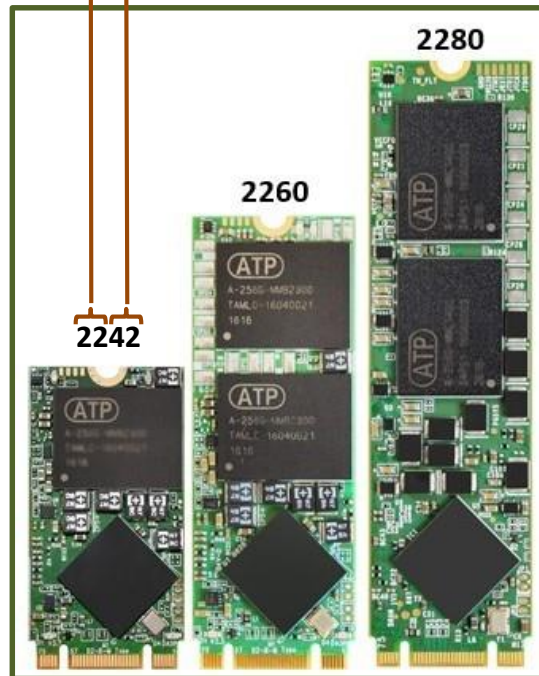
2,5" SFF



M.2

→ Ancho (en milímetros)

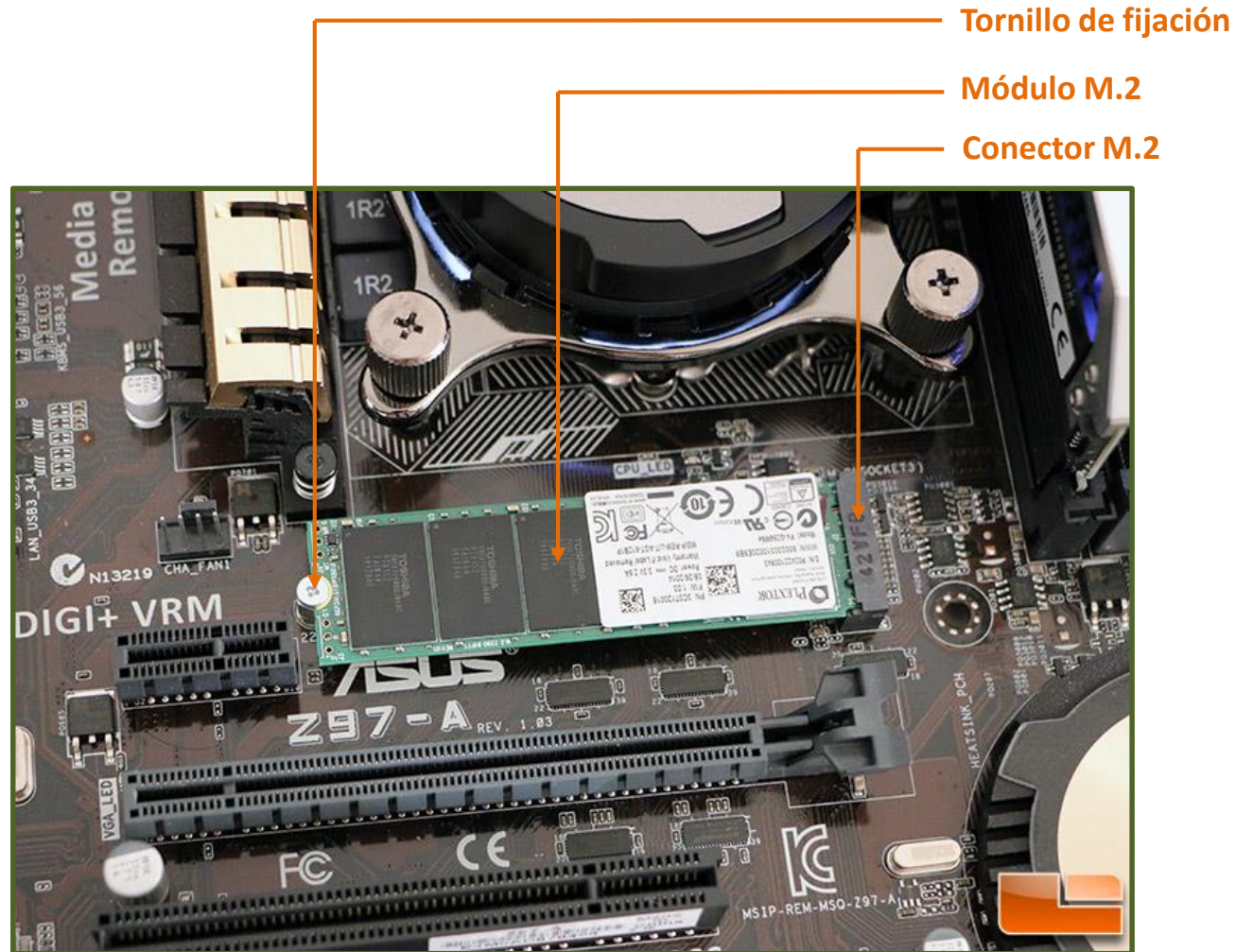
→ Largo (en milímetros)



Tarjeta de expansión PCIe



# Instalación de un dispositivo M.2



# Dispositivos SSD de gama empresarial (Seagate)

Gama Nytro	SAS a 12 Gb/s	SATA a 6 Gb/s
Factor de forma	SFF	SFF
Capacidades	15 TB – 400 GB	3,8 TB – 240 GB
Lectura secuencial sostenida (MB/seg) (1 puerto)	Hasta 1.100	Hasta 564
Lectura secuencial sostenida (MB/seg) (2 puertos)	Hasta 2.100	-
Escritura secuencial sostenida (MB/seg) (1 puerto)	Hasta 970	Hasta 536
Escritura secuencial sostenida (MB/seg) (2 puertos)	Hasta 1.400	-
Lectura aleatoria (IOPS) sostenida (1 puerto)	Hasta 200.000	Hasta 94.000
Lectura aleatoria (IOPS) sostenida (2 puertos)	Hasta 240.000	-
Escritura aleatoria (IOPS) sostenida (1 puerto)	Hasta 80.000	Hasta 61.000
Escritura aleatoria (IOPS) sostenida (2 puertos)	Hasta 80.000	-
AFR	0,35 %	0,5 %



# Comparativa SSD / HDD

Característica	Comparativa
Capacidad máxima	Similares
Prestaciones acceso aleatorio	SSD muy superior (varios órdenes de magnitud)
Prestaciones acceso secuencial	SSD superior
Fiabilidad	Similares
Coste / GB	SSD más caro (x4, aproximadamente)

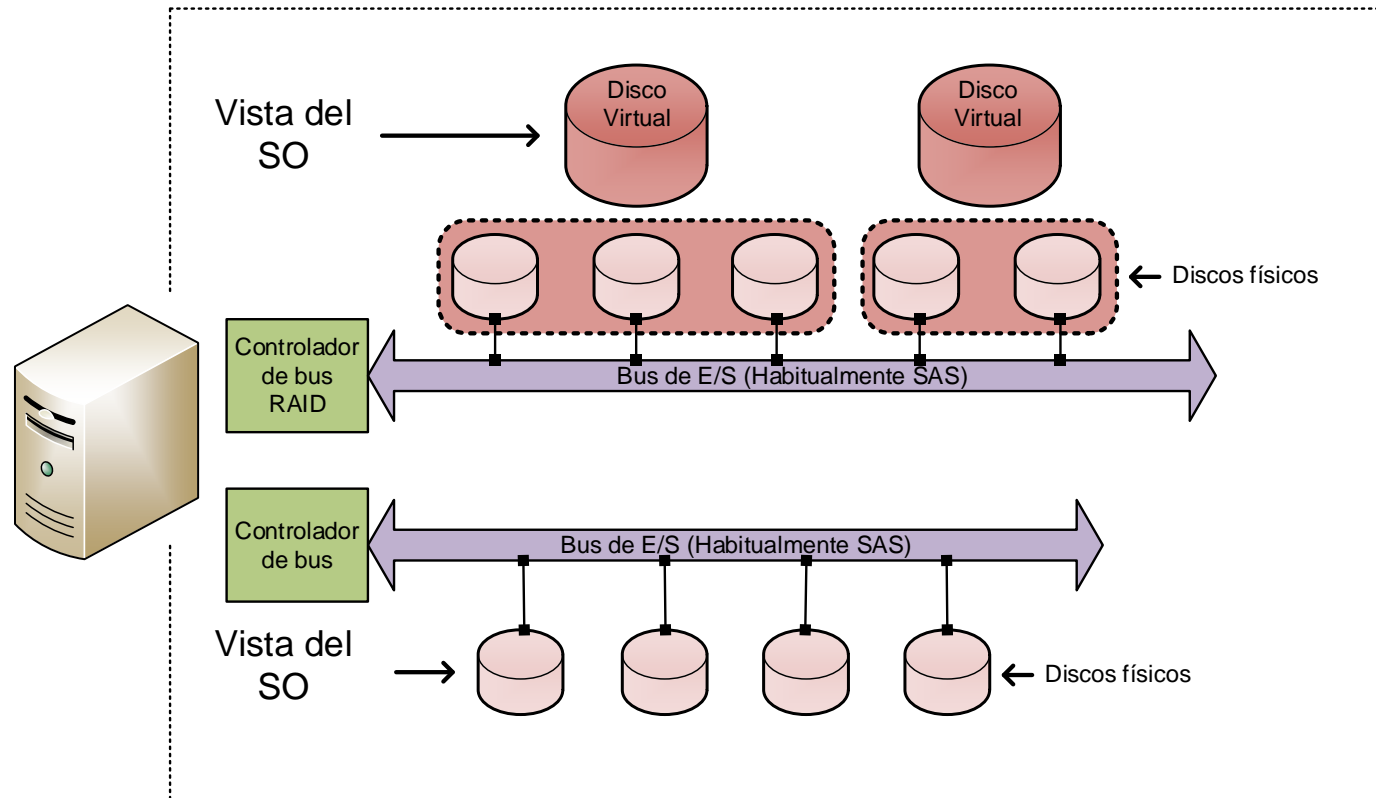
# Índice

---

- Introducción
- Organización del almacenamiento en el computador
- Discos duros y dispositivos SSD
- **RAID**

# Concepto

- RAID significa Redundant Array of Independent Disks
- Concepto



# Aspectos generales

- Objetivos

- Incremento de prestaciones

Se consigue mediante fragmentación de datos (*data striping*). \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

- Tolerancia a fallos

Se consigue mediante redundancia en el almacenamiento de la información. \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

- Niveles RAID

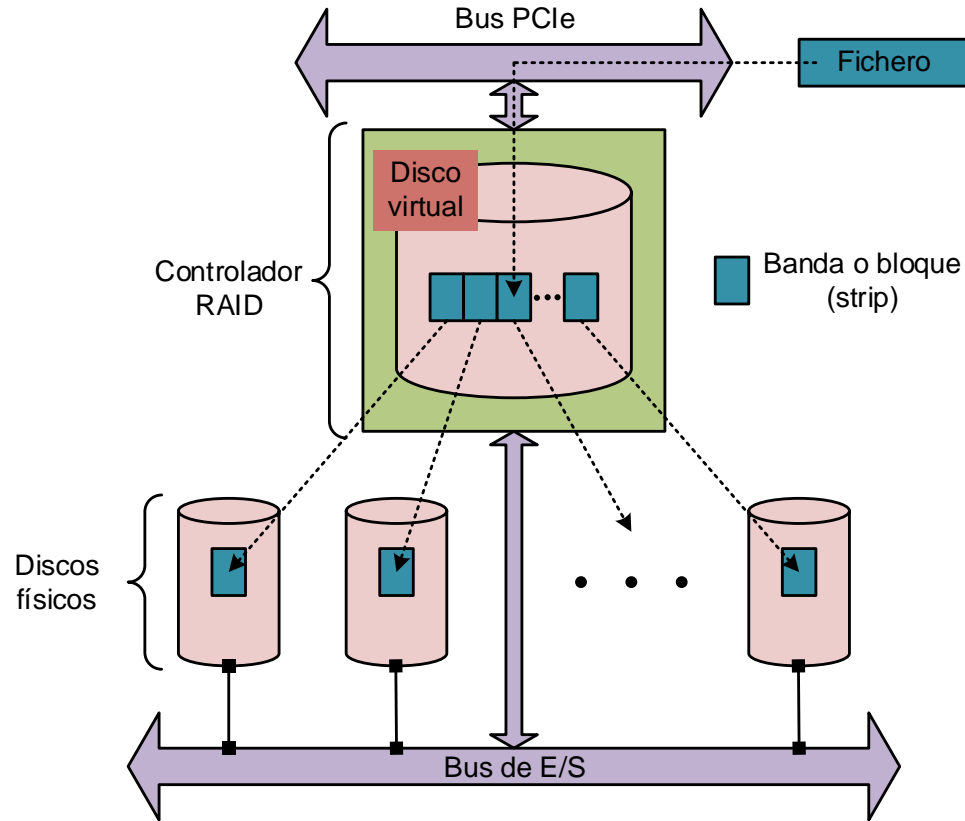
- Concepto

\_\_\_\_\_

- Niveles RAID más comunes

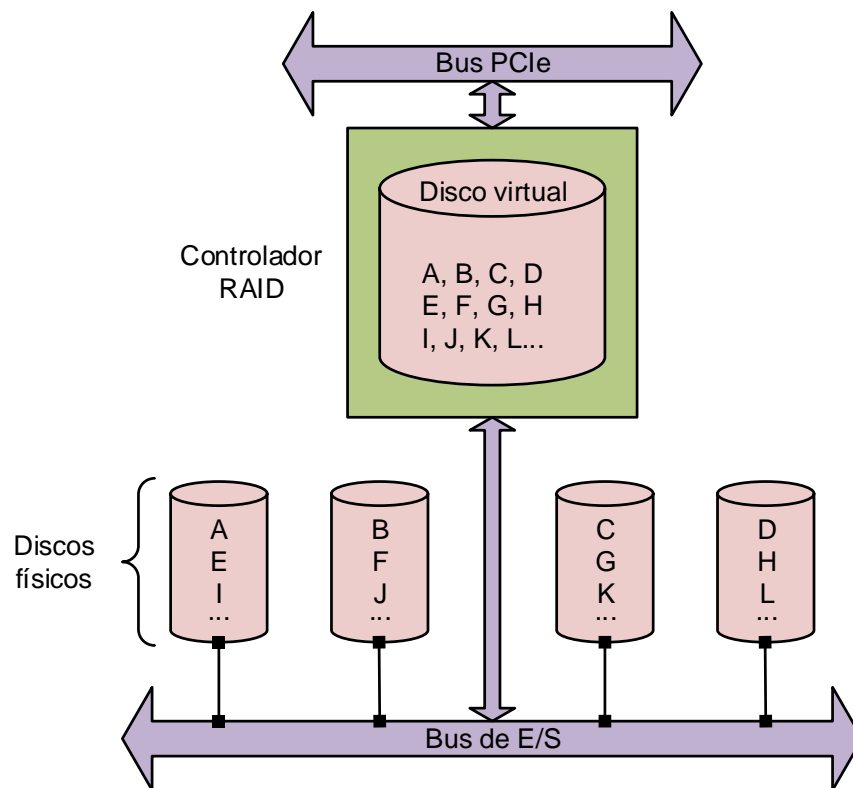
0, 1, 10, 5 y 6.

# Concepto de banda o bloque (*strip*)



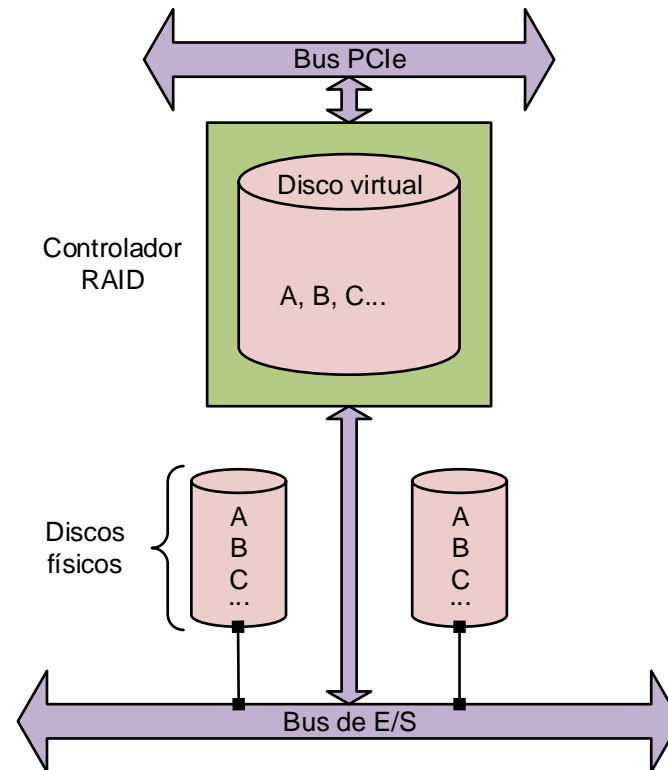
- Definición: \_\_\_\_\_
- Tamaño: Mínimo alrededor de 2KB, máximo alrededor de 2MB, y típico alrededor de 64KB.
- Aplicación: \_\_\_\_\_

# RAID 0: Fragmentación (*striping*)



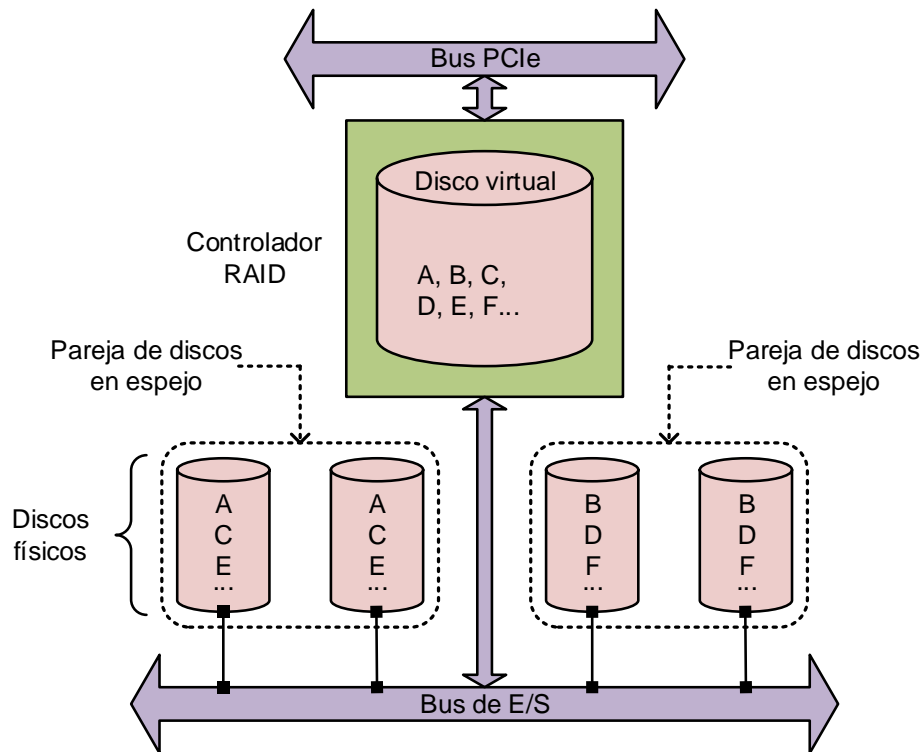
- Funcionamiento: Los bloques a escribir se almacenan en los discos siguiendo una distribución secuencial
- Redundancia: \_\_\_\_
- Prestaciones en operaciones de lectura: \_\_\_\_\_
- Prestaciones en operaciones de escritura: \_\_\_\_\_

# RAID 1: Espejo (*mirroring*)



- **Funcionamiento:** Los bloques a escribir se copian en todos los discos del RAID. Este tipo de RAID se configura con dos discos.
- **Redundancia:** \_\_\_\_
- **Eficiencia en el almacenamiento:** \_\_\_\_
- **Prestaciones en operaciones de lectura:** \_\_\_\_
- **Prestaciones en operaciones de escritura:** \_\_\_\_

# RAID 10: Espejo (*mirroring*)

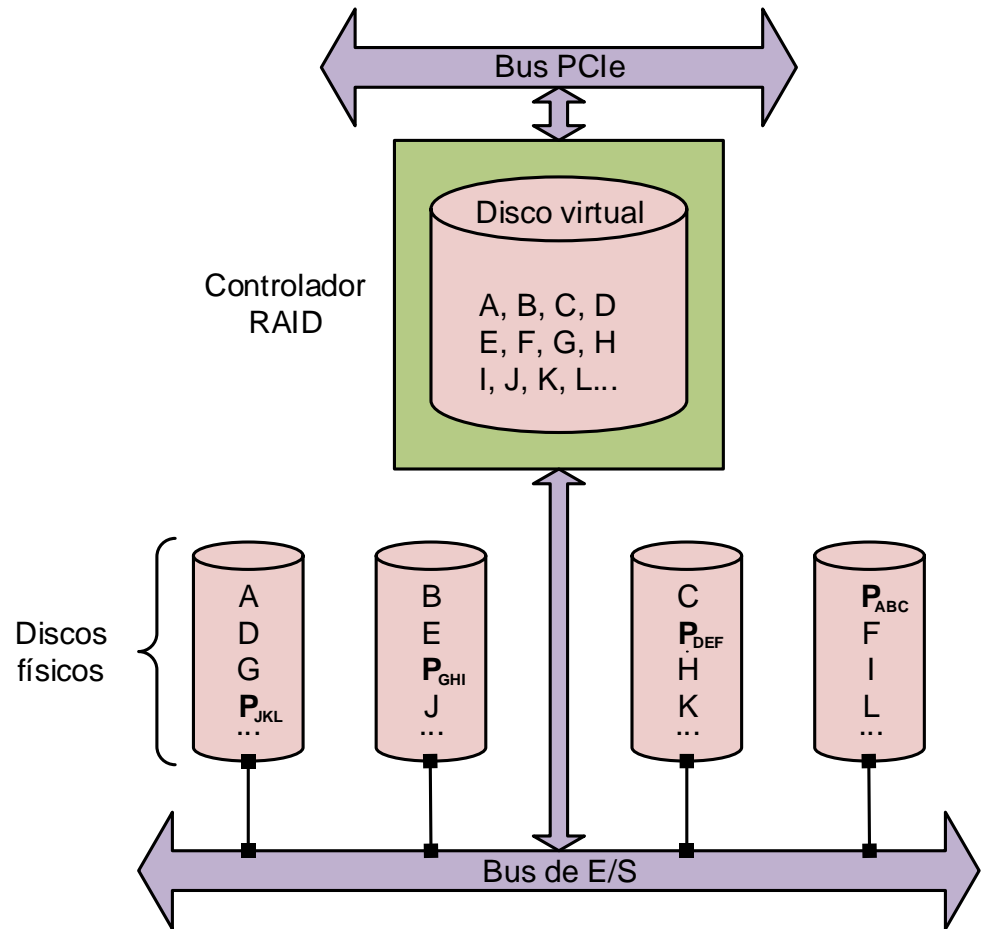


- **Funcionamiento:** Los discos se organizan en parejas que funcionan en espejo. Los bloques se almacenan en las parejas de discos siguiendo una distribución secuencial.
- **Redundancia:**
- **Eficiencia en el almacenamiento:**
- **Prestaciones en operaciones de lectura:**
- **Prestaciones en operaciones de escritura:**



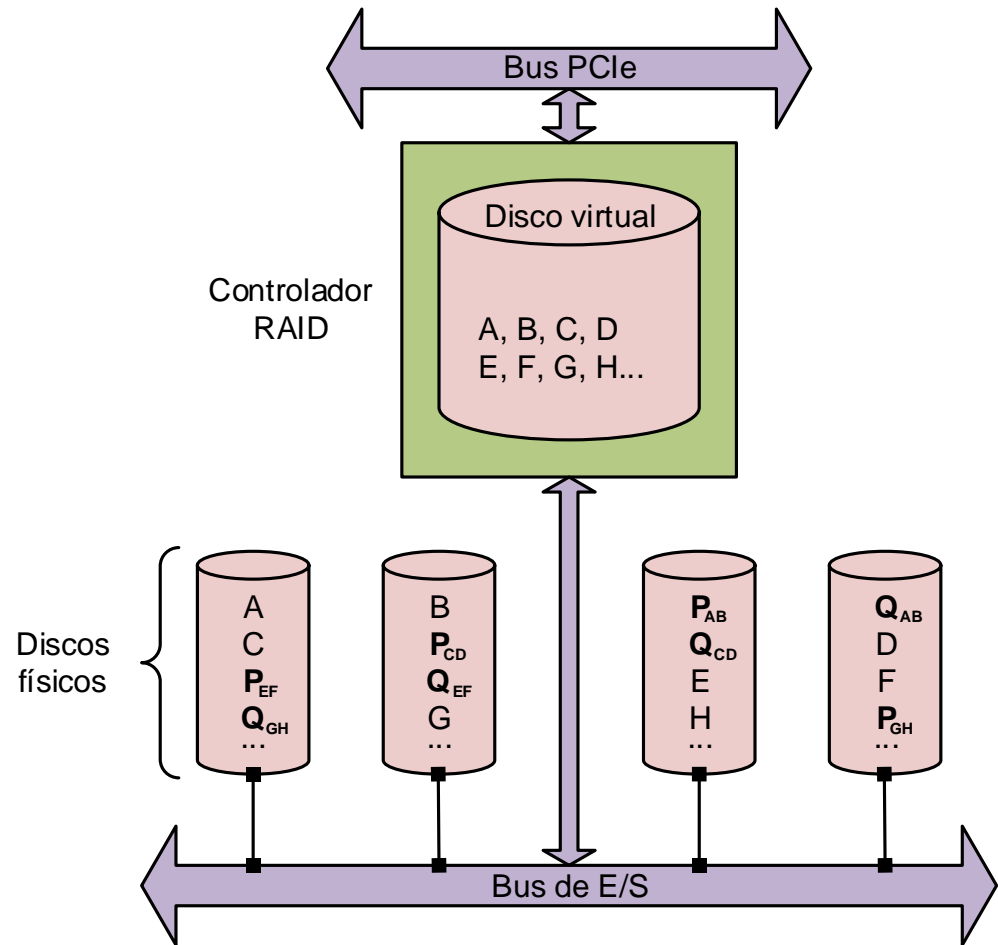
# RAID 5: Paridad (*parity*)

- **Funcionamiento:** Dado un RAID de N discos, para cada (N-1) bloques de datos a escribir se genera un bloque de paridad.
- El bloque de paridad se calcula realizando la operación XOR sobre los bloques de datos correspondientes.
- Los bloques de datos se distribuyen en los discos de manera secuencial.
- Los bloques de paridad se intercalan entre los bloques de datos, de manera que queden uniformemente distribuidos entre todos los discos.
- Ante la pérdida de cualquier bloque de datos (debido al fallo de un disco), la información de dicho bloque puede calcularse a partir de los bloques de datos restantes del grupo, junto con el bloque de paridad.
- **Redundancia:** \_\_\_\_
- **Eficiencia en el almacenamiento:** \_\_\_\_
- **Prestaciones en op. de lectura:** \_\_\_\_
- **Prestaciones en op. de escritura:** \_\_\_\_



# RAID 6: Paridad doble (*parity*)

- **Funcionamiento:** Dado un RAID de N discos, para cada (N-2) bloques de datos a escribir se generan dos bloques de paridad.
- Los bloques de paridad se calculan realizando la operación XOR sobre los bloques de datos correspondientes.
- Los bloques de datos se distribuyen en los discos de manera secuencial.
- Los bloques de paridad se intercalan entre los bloques de datos, de manera que queden uniformemente distribuidos entre todos los discos.
- El RAID es tolerante al fallo de dos discos.
- **Redundancia:** \_\_\_\_\_
- **Eficiencia en el almacenamiento:** \_\_\_\_\_
- **Prestaciones en op. de lectura:** \_\_\_\_\_
- **Prestaciones en op. de escritura:** \_\_\_\_\_



# Cache de escritura

- ¿Qué algoritmo de escritura se utiliza?

- ¿Cómo funciona?

- ¿Qué ocurre si se interrumpe la energía en el servidor?

- Denominación de la tecnología utilizada

Flash-backed write cache (FBWC)

# Conveniencia de la cache en función del tipo de dispositivos del RAID

- RAID de dispositivos HDD

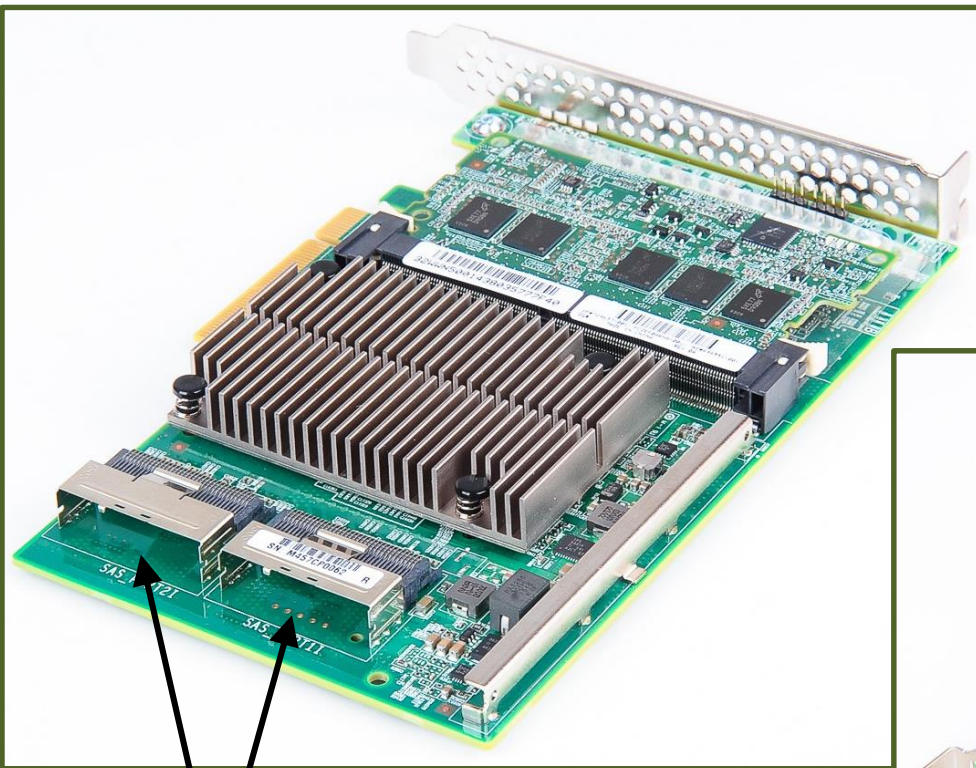
La cache es esencial para obtener unas prestaciones razonables, especialmente si el porcentaje de escrituras en la carga de trabajo es significativo.

- RAID de dispositivos SSD

La cache no es necesaria

# Ejemplo de controladora RAID de clase empresarial

## HPE Smart Array P830



Puertos SAS x8

Condensador



FBWC

# HPE Smart Array P830

Propiedades esenciales	
Interfaz de placa base	PCIe
Interfaz de disco	SAS 12 Gb/s
Puertos físicos	2 x 8
Niveles de RAID soportados	0, 1, 10, 5, 50, 6, 60, 10 ADM
Cache	4 GB

Ejercicio: cálculo del tiempo medio de acceso de un disco 10K, sabiendo que el tiempo de búsqueda medio es de 3ms

---

## Ejercicio: ejemplo de cálculo de paridad en RAID 5

---