

Arquitectura RAID

Administración de Sistemas y Redes

José A. Corrales

ja@uniovi.es

RAID

- Redundant Array of Independent Disks.
- Un grupo de discos físicos es visto por el equipo como un único disco.
- Independiente del sistema operativo.
- Proporciona velocidad, capacidad o tolerancia a fallos dependiendo del tipo.

RAID

- Protegen frente a fallos de discos (excepto RAID 0).
- No protegen frente a errores software.
- Se usan sobre todo con discos reemplazables en caliente.
- Algunas configuraciones permiten reemplazo automático (disco de repuesto).
- Una controladora permite múltiples configuraciones.

RAID

- Software de la controladora RAID habitualmente proporcionado por el fabricante.
- Configuración desde la BIOS o con software específico proporcionado también por el fabricante.
- Existe RAID por software aunque es dependiente del operativo. Menos recomendable.
- Mayores ventajas con discos magnéticos SAS o SATA, menores con SSD (no latencia rotacional, menos margen de mejora).

RAID: configuraciones más habituales

- JBOD (Just a Bunch Of Disks). Discos sueltos, la controladora RAID está en modo transparente. Volumen simple (Microsoft).
- RAID 0. Stripping. Suma de discos.
- RAID 1. Mirror. Espejo de discos.
- RAID 5. Tolerancia a fallos.
- Otras frecuentes: RAID 6, RAID 01, RAID 10, RAID 50

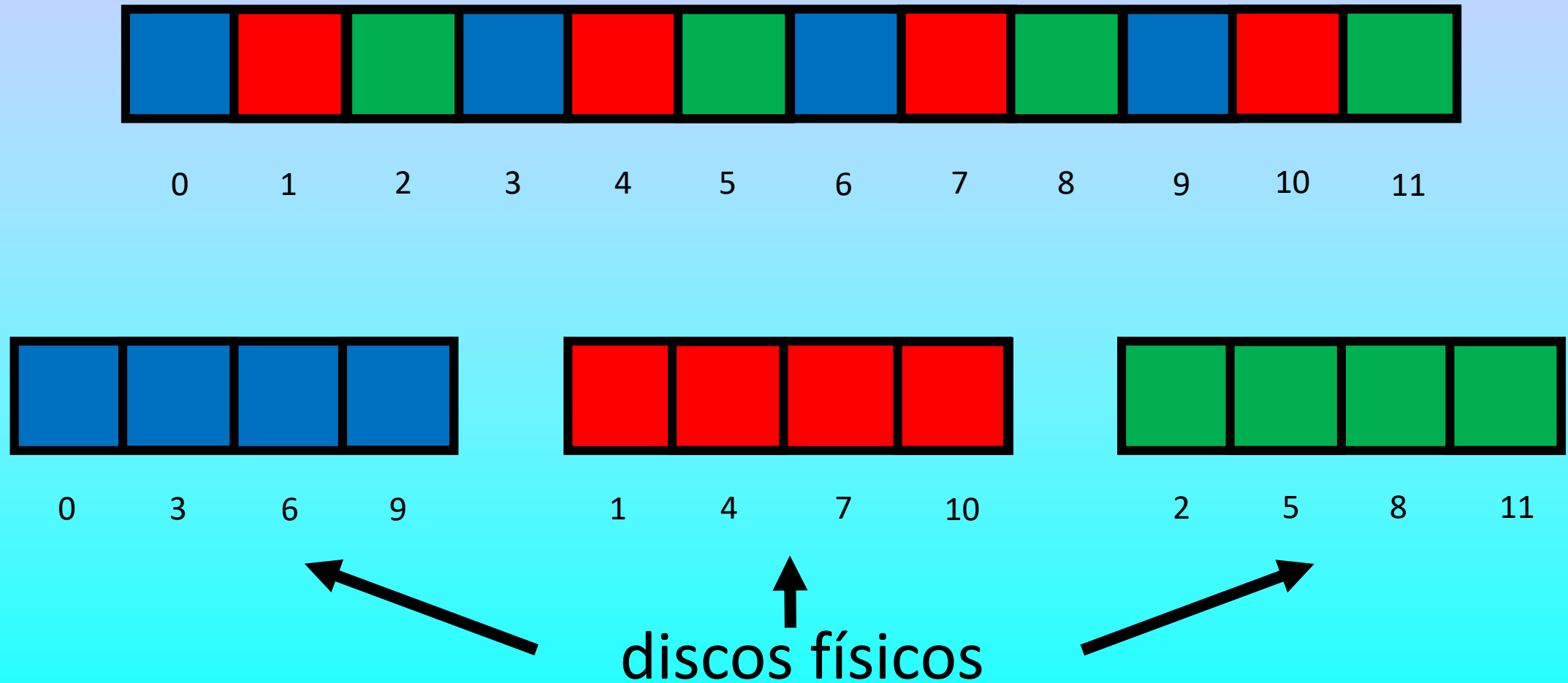
RAID 0 (dividido/stripping)

- N discos de capacidad C son vistos como un único disco de capacidad $N \times C$.
- Alta velocidad de lectura/escritura. N operaciones en paralelo.
- Alta capacidad. Mayor que los discos físicos existentes.
- Se aprovecha todo el espacio.

RAID 0 (dividido/stripping)

- Si falla un disco falla todo el RAID.
- Aumenta considerablemente la probabilidad de fallo.
- Discos iguales o bien sólo se usa la capacidad del menor (multiplicada por el número de discos).
- Volumen seccionado (Microsoft).

RAID 0 (dividido/stripping)



RAID 0 (dividido/stripping)

- Sea B el número de bloque del disco RAID 0
- Sea N el número de discos que forman el RAID
- El resto de la división de B entre N devuelve el número de disco físico al que se debe acceder (numerados del 0 al $N-1$)
- El cociente de la división de B entre N devuelve el número de bloque del disco físico al que se debe acceder.

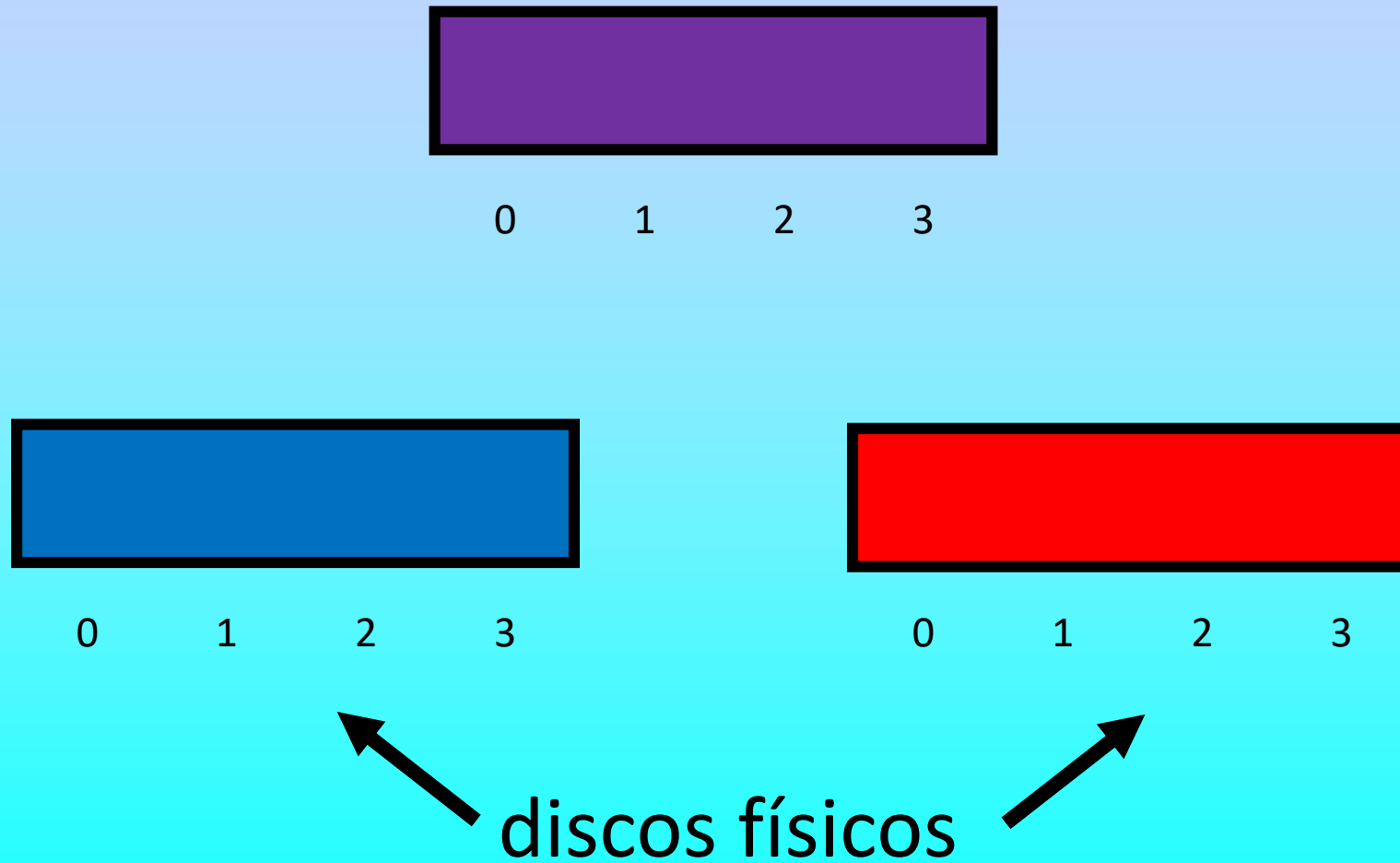
RAID 1 (espejo/mirror)

- N discos (generalmente 2) de capacidad C son vistos como un único disco de capacidad C.
- Alta velocidad de lectura. N operaciones en paralelo.
- Redundancia a fallos (modo degradado) de N-1 discos.
- Reconstrucción automática.
- Posibilidad de reemplazo automático.
- Posibilidad de usarlo como copia de seguridad.

RAID 1 (espejo/mirror)

- Se pierde el $(N-1)/N$ del espacio. 50% si son dos discos, 66% si son tres y así sucesivamente.
- No mejora la velocidad de escritura.
- Discos iguales o bien sólo se usa la capacidad del menor (multiplicada por el número de discos).
- Volumen reflejado (Microsoft).

RAID 1 (espejo/mirror)



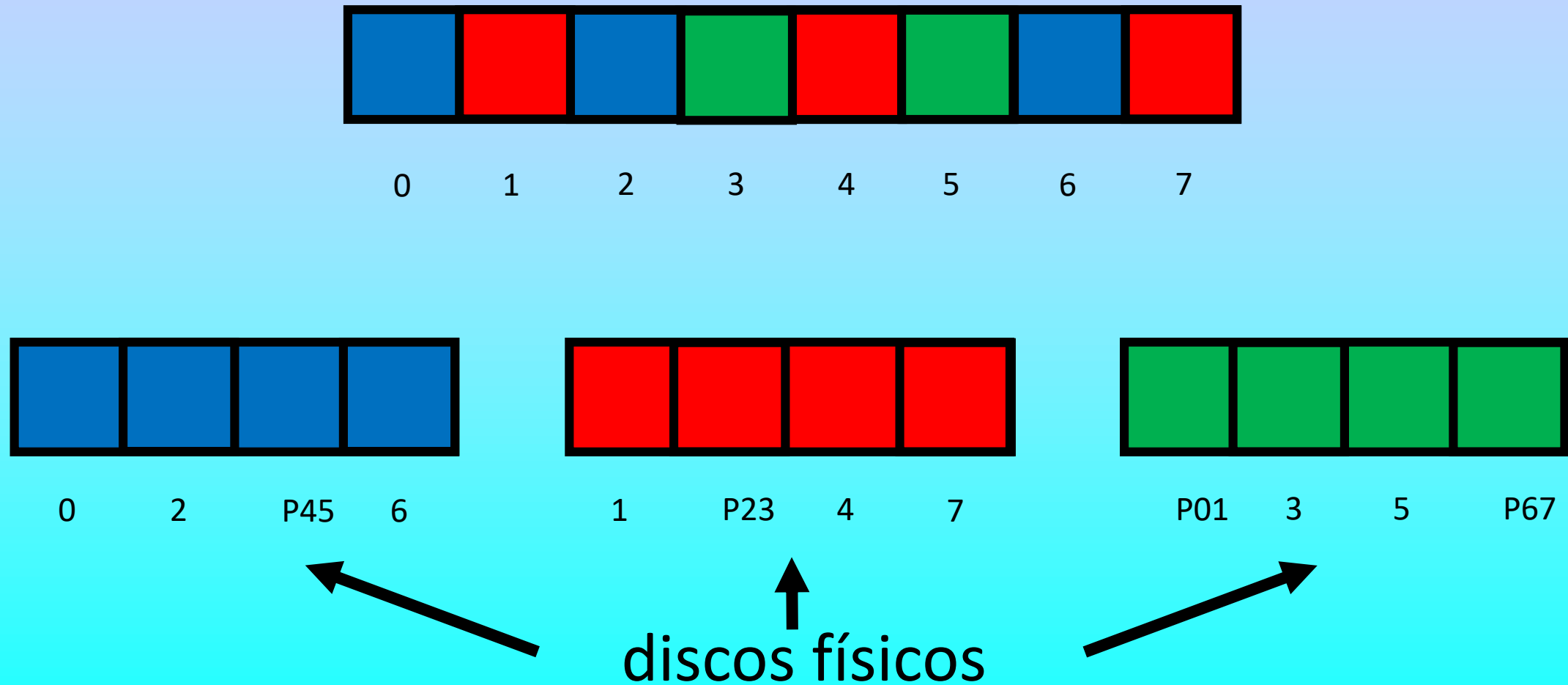
RAID 5 (paridad distribuida/parity)

- N discos de capacidad C son vistos como un único disco de capacidad $(N-1) \times C$.
- Alta velocidad de lectura. N operaciones en paralelo.
- Alta capacidad, mayor que los discos físicos.
- Redundancia a fallos (modo degradado) de un disco.
- Reconstrucción automática.
- Posibilidad de reemplazo automático.

RAID 5 (paridad distribuida/parity)

- Se pierde el $1/N$ del espacio, usado para almacenar la paridad.
- Discos iguales o bien sólo se usa la capacidad del menor (multiplicada por el número de discos).
- No eficiente con muchas escrituras de poco tamaño. Implican lecturas y cálculos de paridad.
- Volumen RAID 5 (Microsoft)

RAID 5 (paridad distribuida/parity)



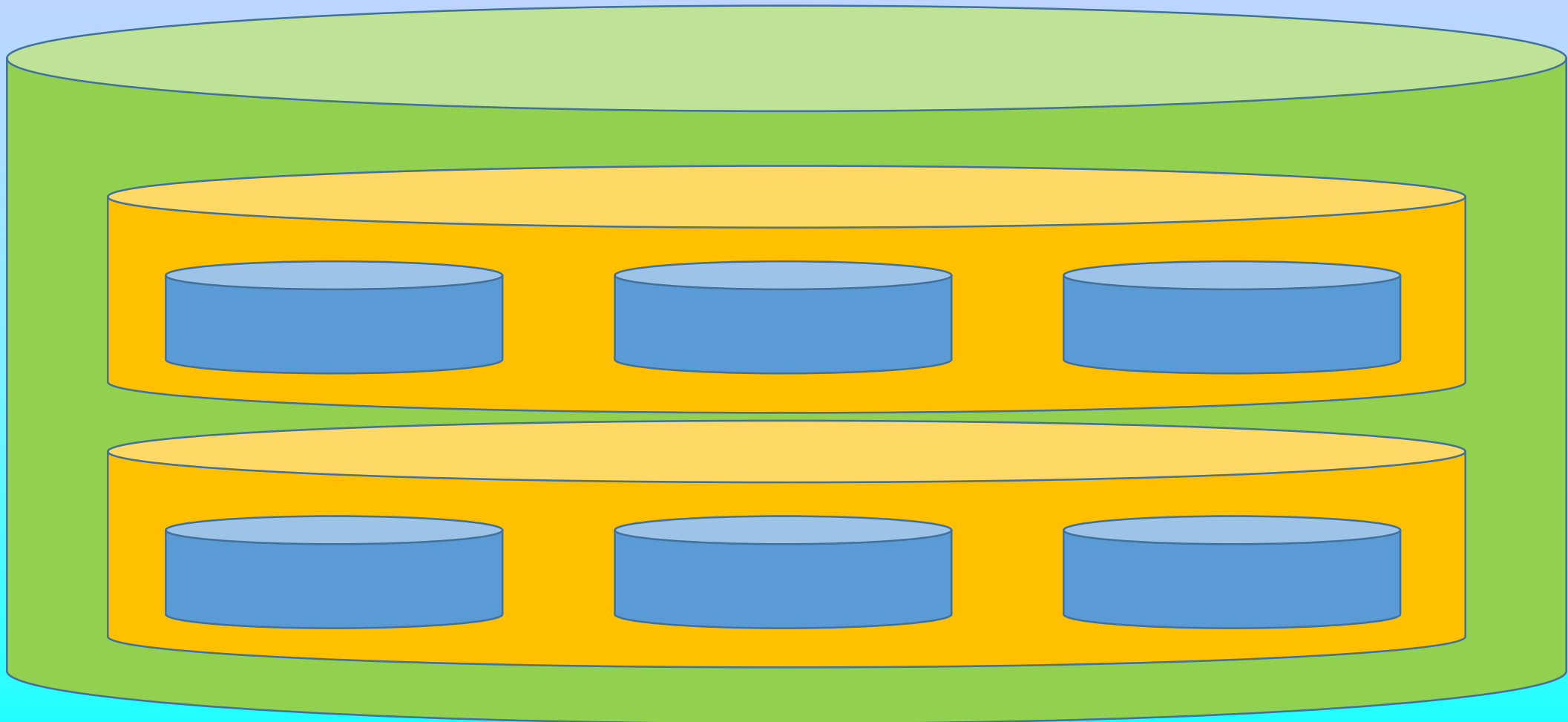
RAID 6 (doble paridad/double parity)

- N discos de capacidad C son vistos como un único disco de capacidad $(N-2) \times C$.
- Alta velocidad de lectura. N operaciones en paralelo.
- Alta capacidad, mayor que los discos físicos.
- Redundancia a fallos (modos degradados) de dos discos.
- Reconstrucción automática.
- Posibilidad de reemplazo automático.

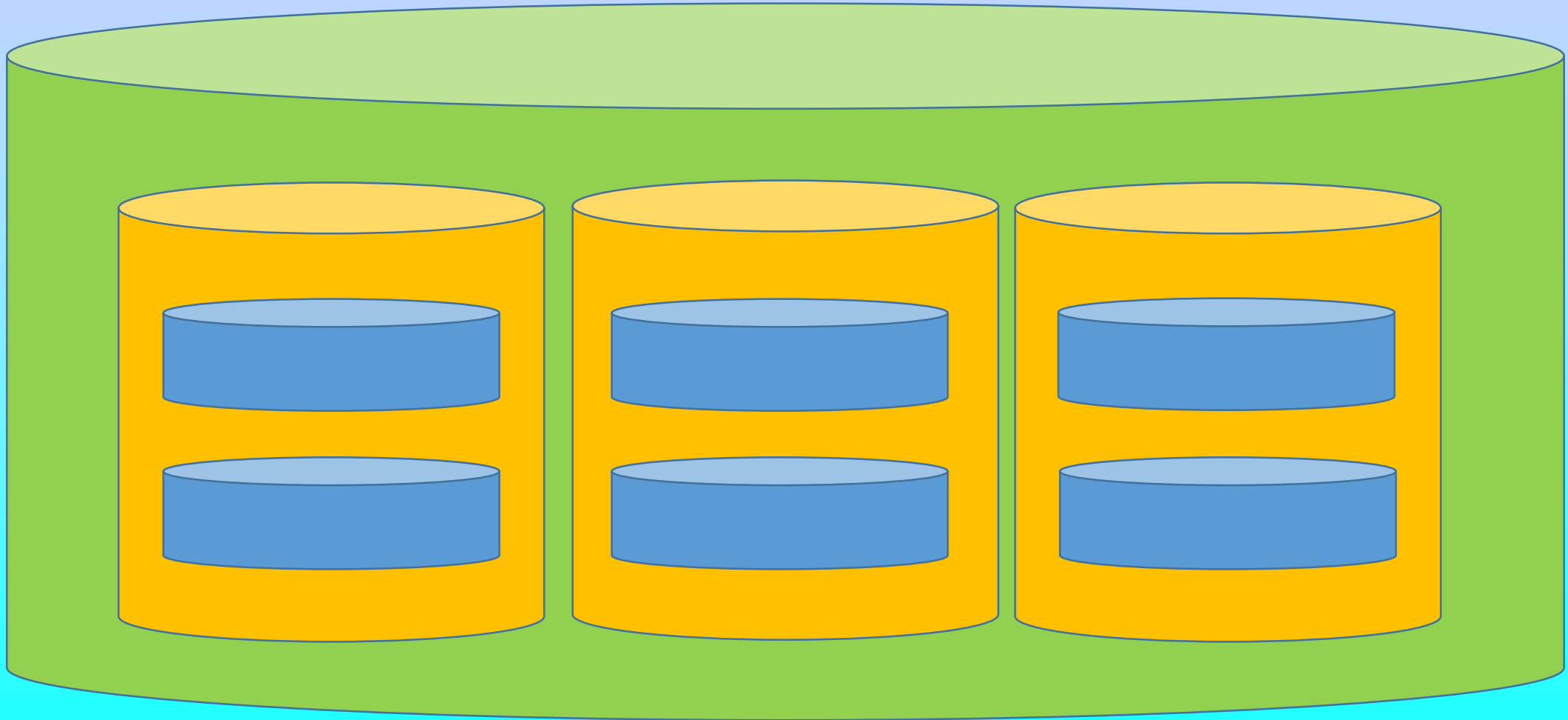
RAID 6 (doble paridad/double parity)

- Se pierde el $2/N$ del espacio, usado para almacenar la paridad. Código de Reed-Solomon.
- Ligeramente más lento en operaciones de escritura.
- Al igual que el RAID 5, no es eficiente con muchas escrituras de poco tamaño.

RAID 01 (0+1)



RAID 10 (1+0)



RAID 01 versus 10

- En el ejemplo anterior 01 hay 6 discos. Posibles fallos de dos discos:

$$C_2^6 = \binom{6}{2} = \frac{6!}{2!4!} = 15$$

- De las cuales son destructivas uno de cada RAID 0:

$$3^2 = 9$$

- Se estropea el RAID en un $9/15=0,6=60\%$ de los casos

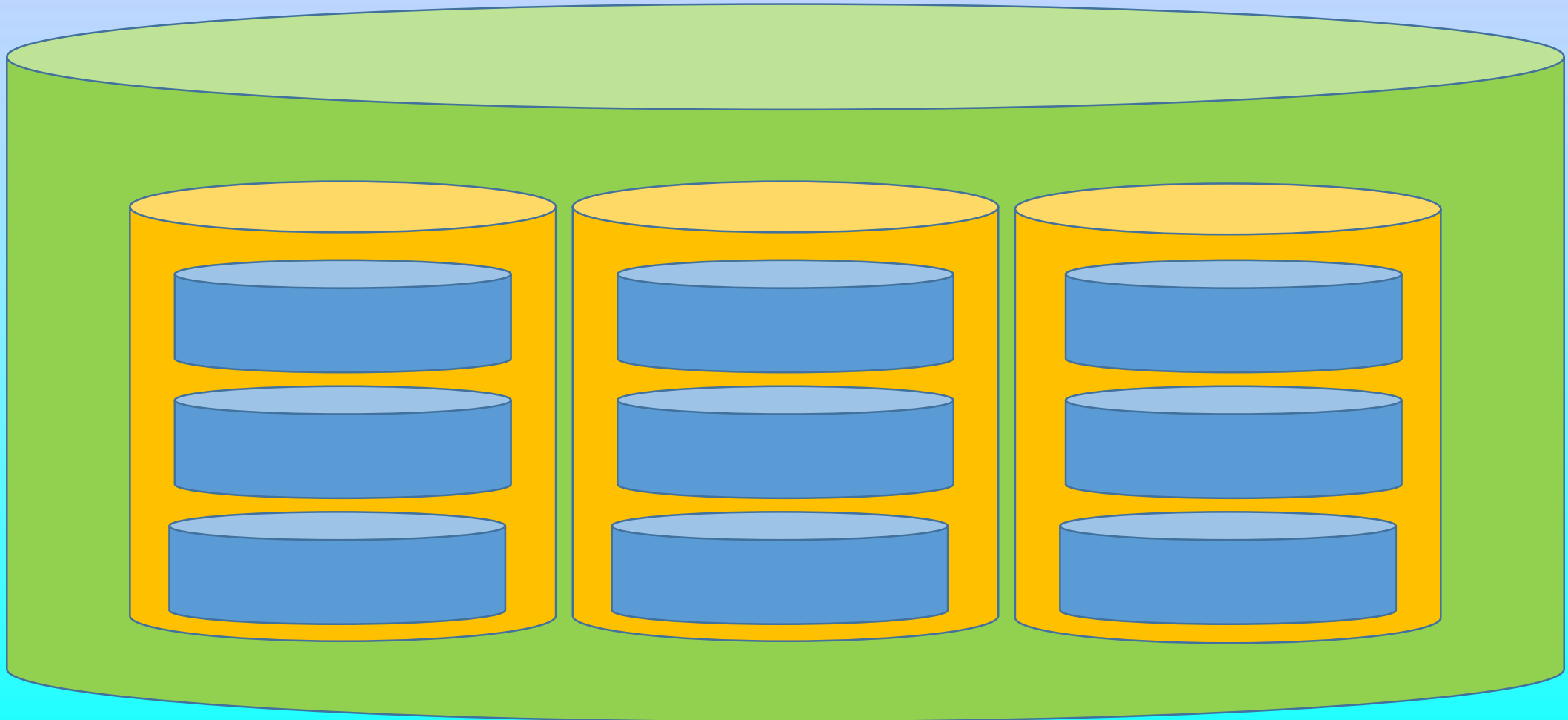
RAID 01 versus 10

- En el ejemplo anterior 10 hay 6 discos. Posibles fallos de dos discos:

$$C_2^6 = \binom{6}{2} = \frac{6!}{2!4!} = 15$$

- De las cuales son destructivas dos de cada RAID 1, es decir sólo hay tres combinaciones destructivas.
- Se estropea el RAID en un $3/15=0,2=20\%$ de los casos
- Conclusión, mejor el RAID 10 y si hay poca escritura mejor el RAID 6 que no llegaría a estroperse.

RAID 50 (5+0)



LVM (volumen lógico/logical volume)

- Varios discos físicos forman un grupo de volumen.
- El grupo de volumen se reparte en discos lógicos del tamaño que se desee.
- Permite redistribuir el tamaño de los discos lógicos.
- Volumen distribuido (Microsoft).

LVM (volumen lógico/logical volume)

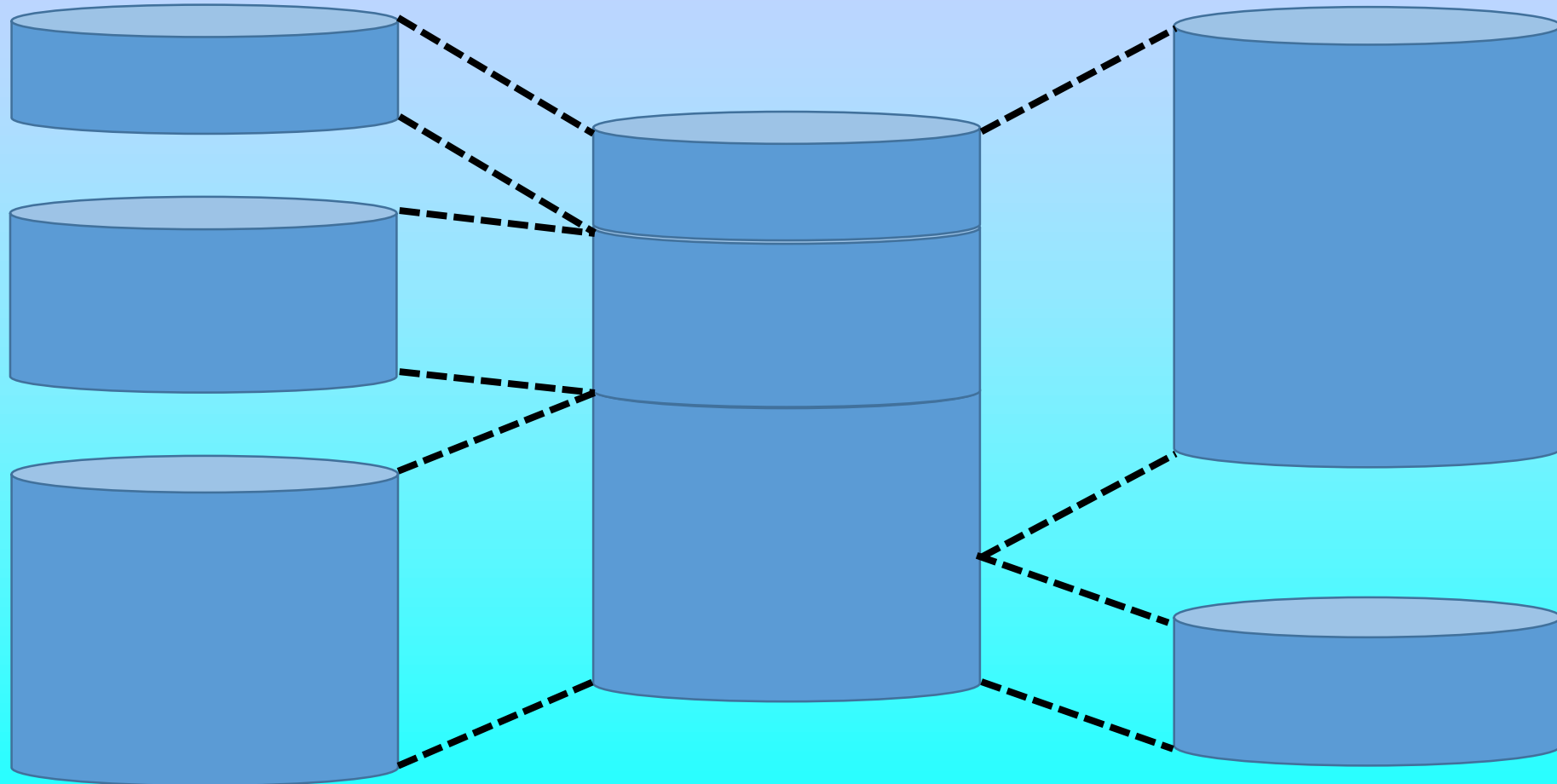
- Exclusivamente por software y dependiente del sistema operativo.
- El fallo de un disco afecta a todo el volumen lógico.
- Es conveniente usarlo sobre algún RAID redundante.

LVM (volumen lógico/logical volume)

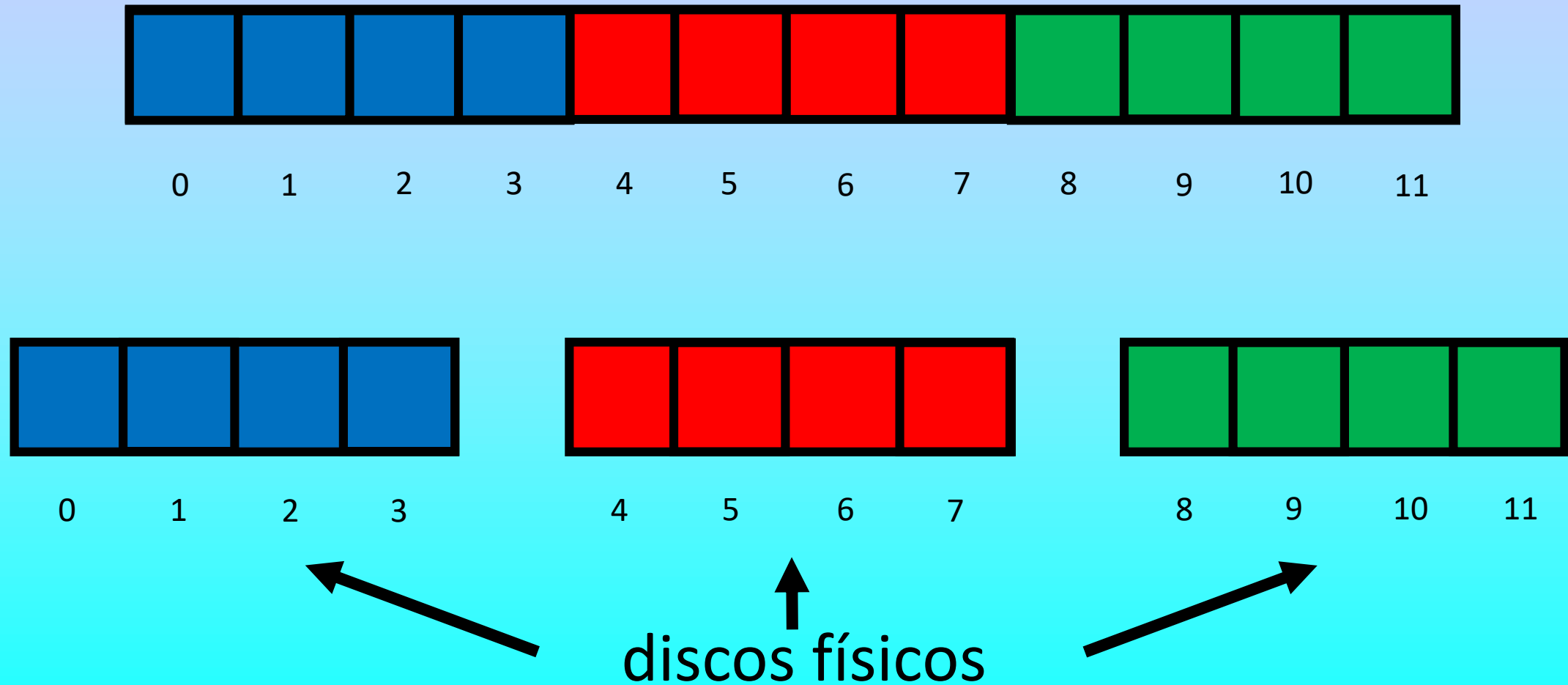
discos físicos o RAID > 0

grupo volumen

discos (volúmenes) lógicos



LVM (volumen lógico/logical volume)



LVM (volumen lógico)

Herramientas (Linux):

- pvcreate
- vgcreate, vgdisplay, vgrename
- lvcreate, lvdisplay, lvextend, lreduce