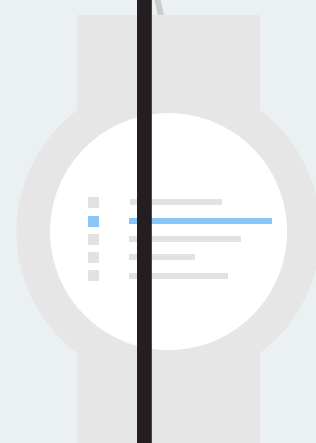
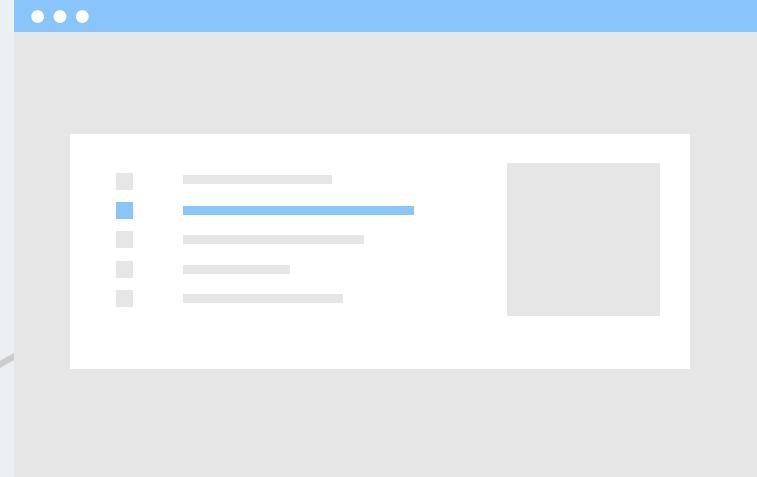
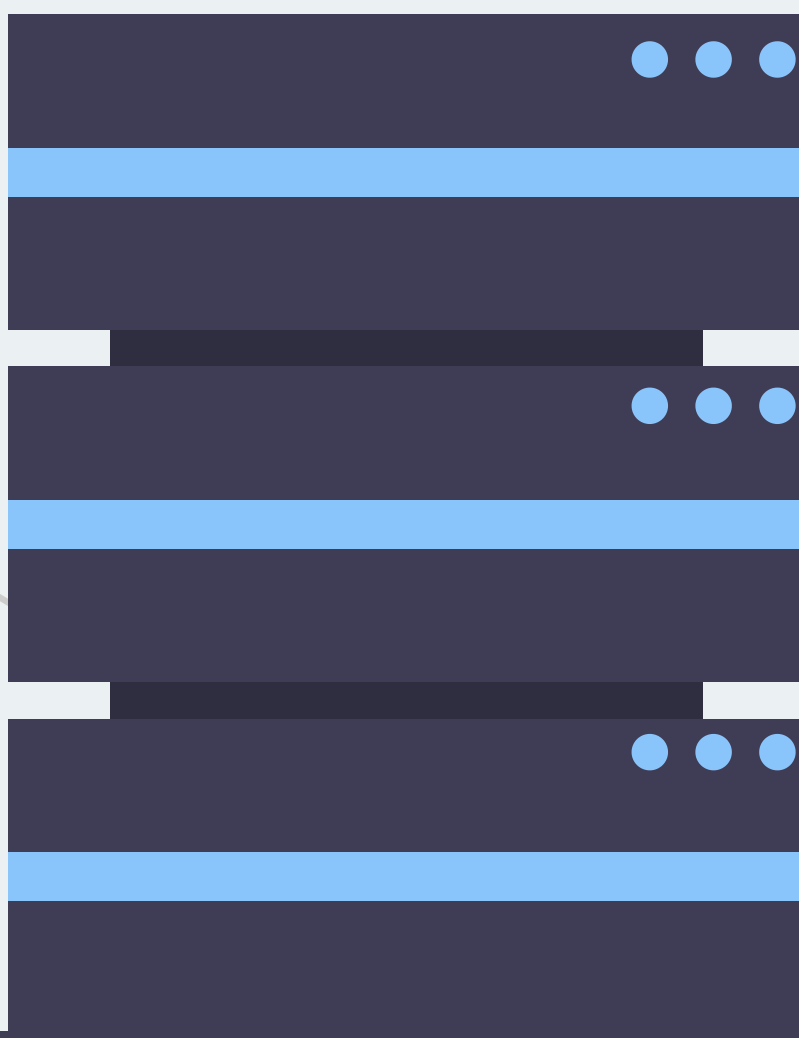


Stockage persistent

MiNET

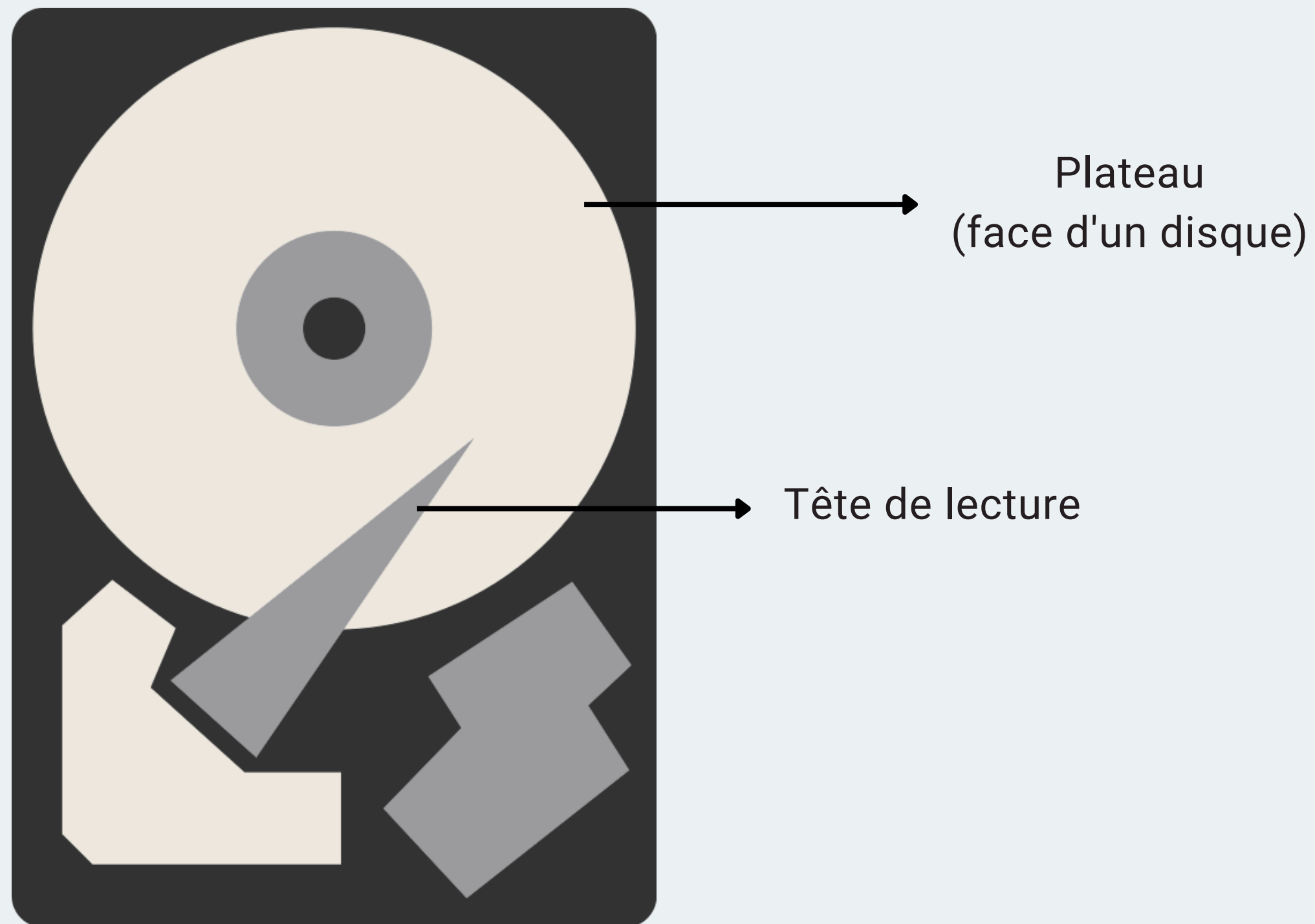


FORMATION



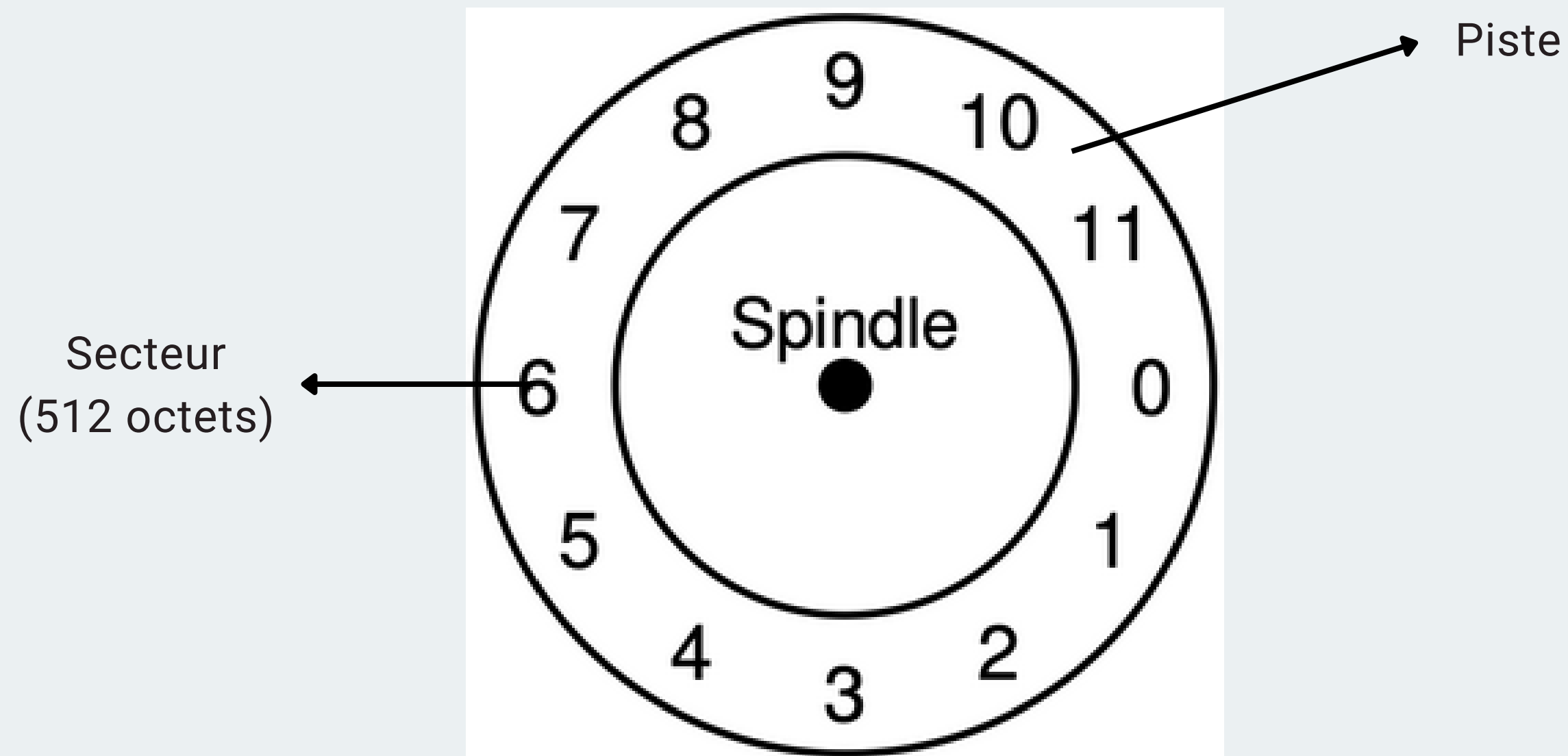
LES DISQUES DURS

GÉOMÉTRIE BASIQUE



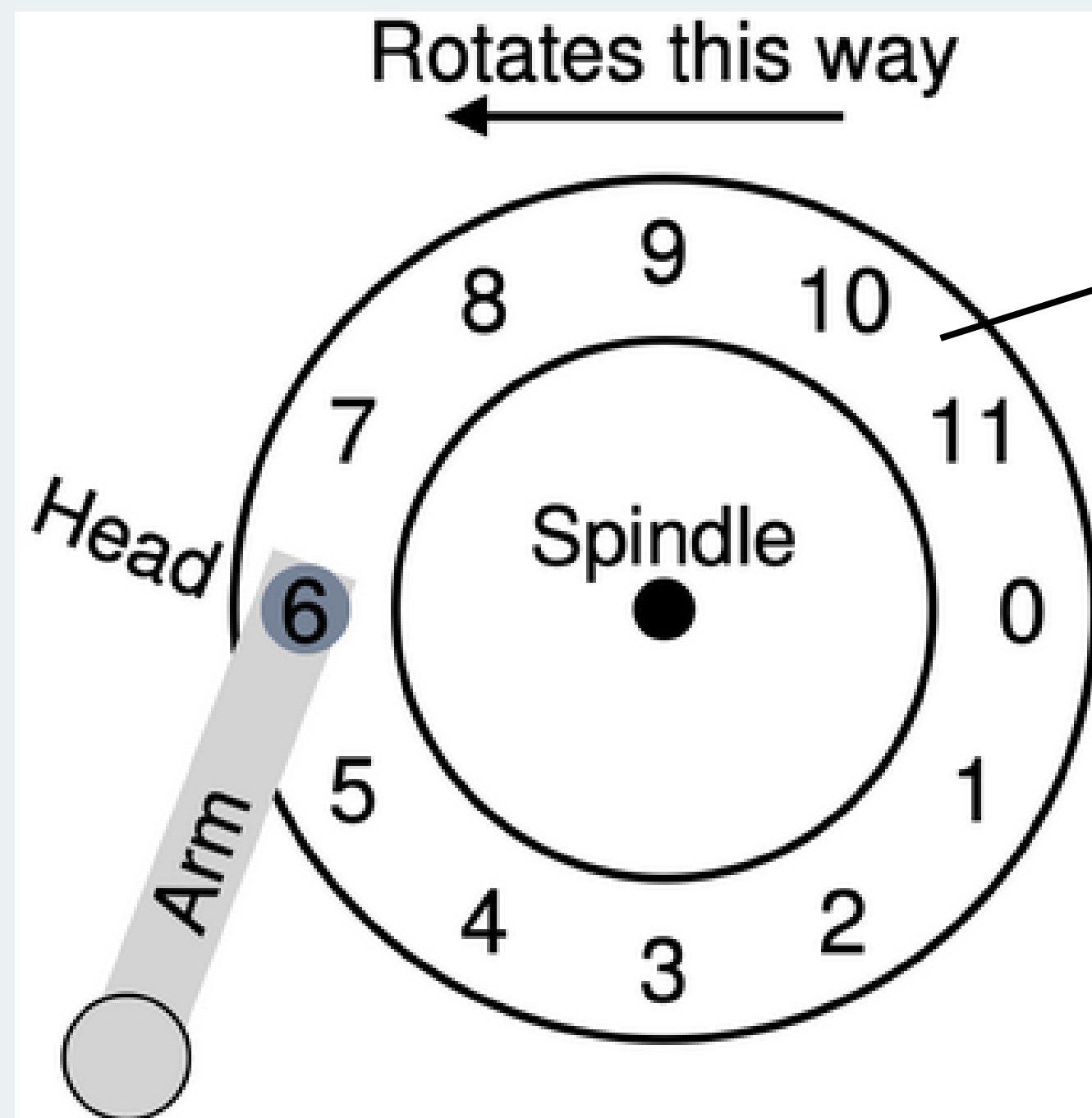
LES DISQUES DURS

MODÈLE SIMPLE - DISQUE MONO-PISTE



LES DISQUES DURS

MODÈLE SIMPLE - DISQUE MONO-PISTE

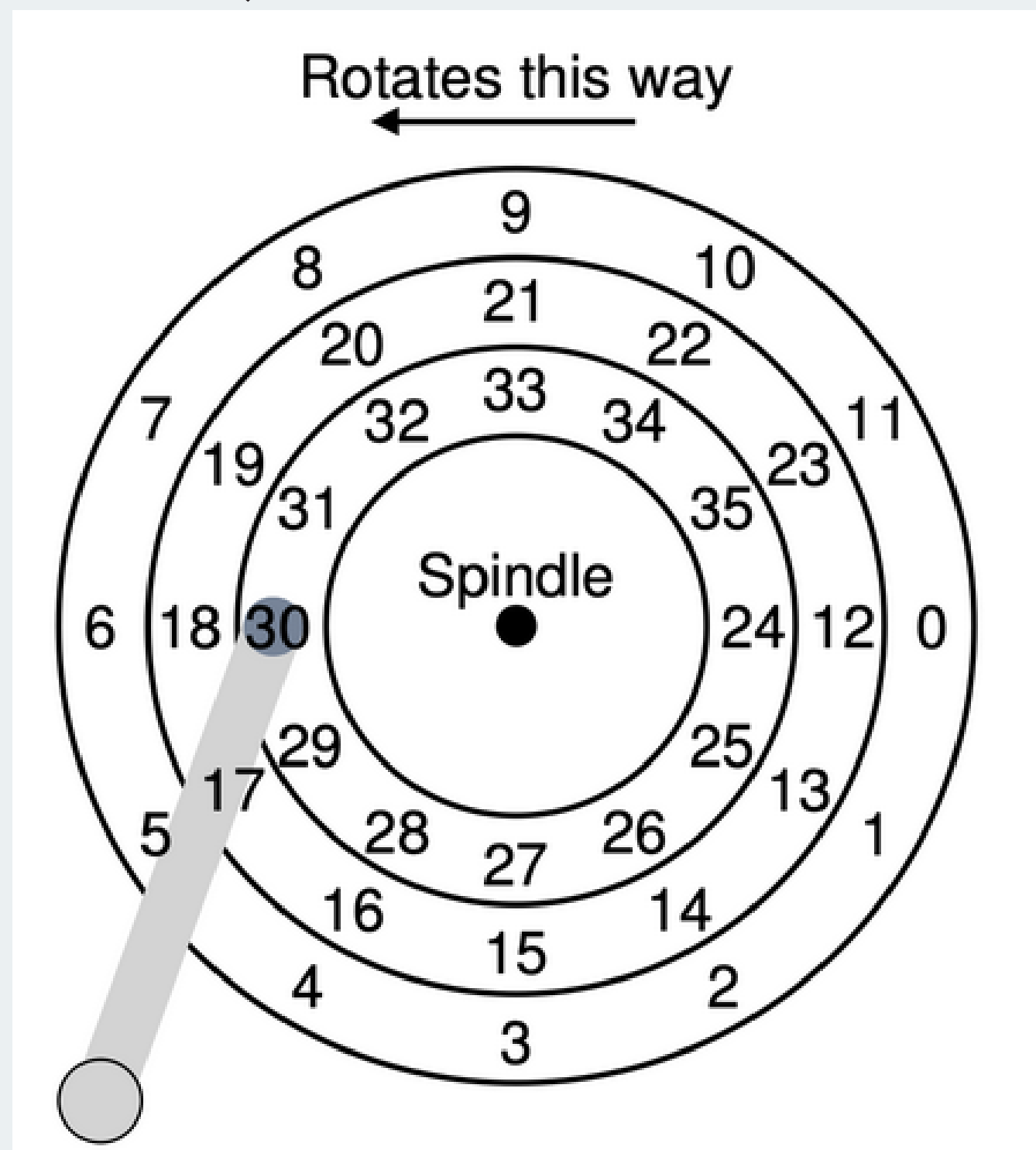


Piste

délai de rotation pour
accéder à un secteur !

LES DISQUES DURS

MODÈLE SIMPLE - DISQUE MULTI-PISTE

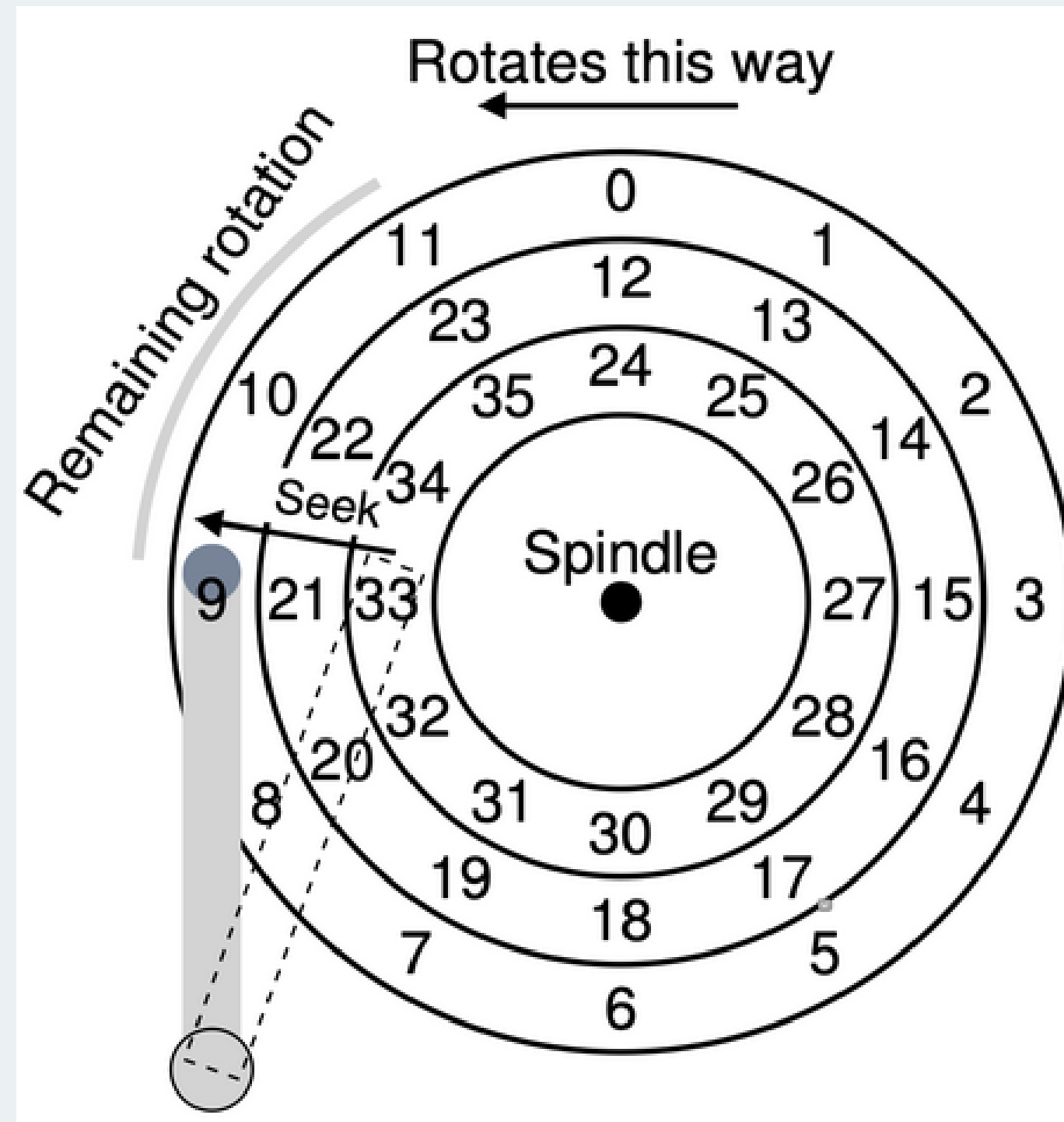


LES DISQUES DURS

MODÈLE SIMPLE - DISQUE MULTI-PISTE

4 phases :

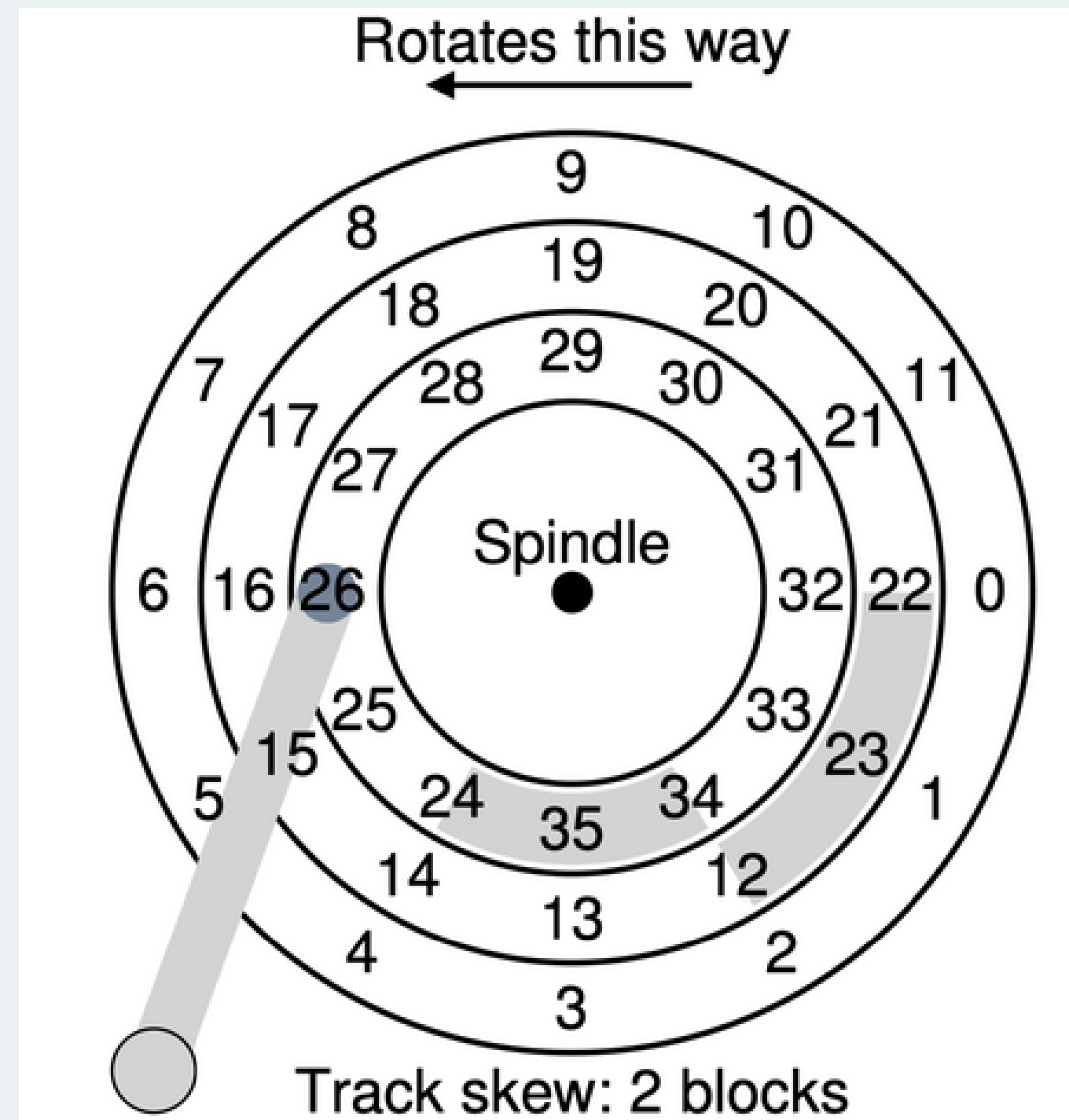
- Accélération
- Roue libre
- Décélération
- Stabilisation



→ délai de recherche d'une piste !

LES DISQUES DURS

MODÈLE PLUS RÉALISTE - DÉCALAGE DES PISTES



Diminue le temps de repositionnement de la tête lors d'un changement de piste

LES DISQUES DURS

TEMPS D'ENTRÉE / SORTIE

	Cheetah 15K.5	Barracuda
Capacity	300 GB	1 TB
RPM	15,000	7,200
Average Seek	4 ms	9 ms
Max Transfer	125 MB/s	105 MB/s
Platters	4	4
Cache	16 MB	16/32 MB
Connects via	SCSI	SATA

LES DISQUES DURS

TEMPS D'ENTRÉE / SORTIE - CHARGES DE TRAVAIL

CHARGE SÉQUENTIELLE

transfert de 100 Mo

Lecture de secteurs contigus sur une piste

Cheetah

Temps d'E/S : 800 ms
Débit d'E/S : 125 Mo/s

Barracuda

Temps d'E/S : 900 ms
Débit d'E/S : 120 Mo/s

CHARGE ALÉATOIRE

Lecture de secteurs répartis aléatoirement sur un plateau

→ Recherche de secteur systématique !

Cheetah

Temps d'E/S : 6 ms
Débit d'E/S : 0,66 Mo/s

Barracuda

Temps d'E/S : 13,2 ms
Débit d'E/S : 0,31 Mo/s

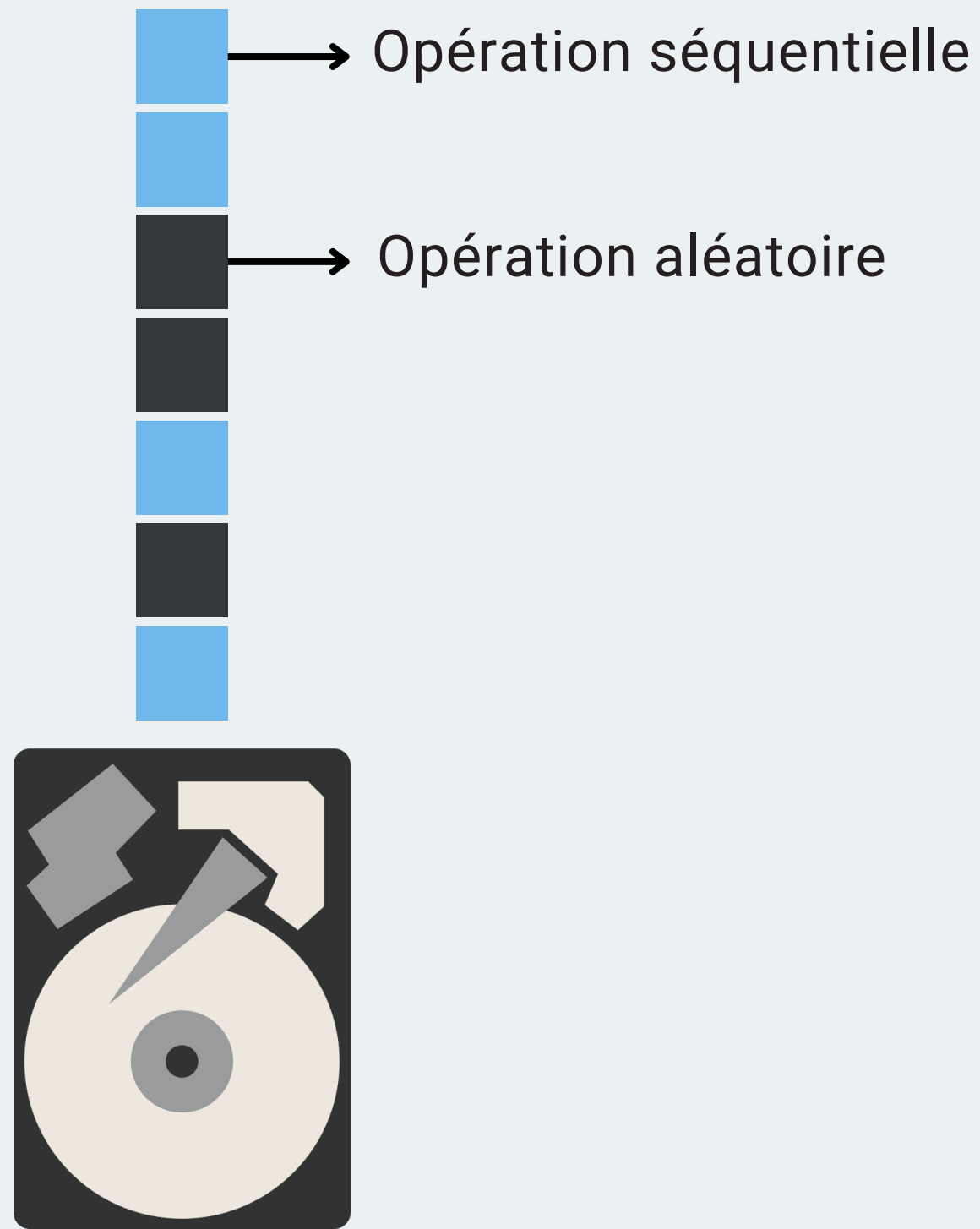
LES DISQUES DURS

TEMPS D'ENTRÉE / SORTIE - BILAN

CHARGE SÉQUENTIELLE >> CHARGE ALÉATOIRE

LES DISQUES DURS

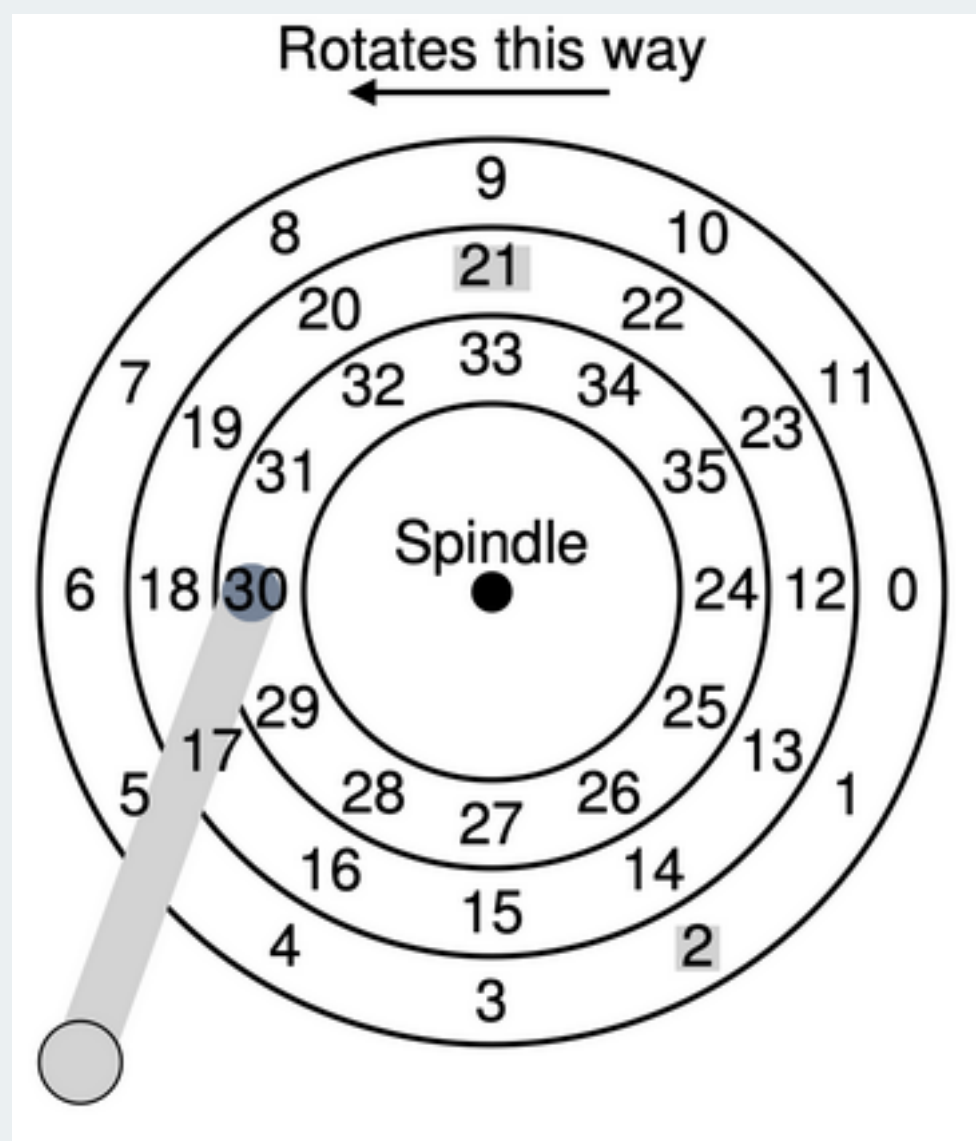
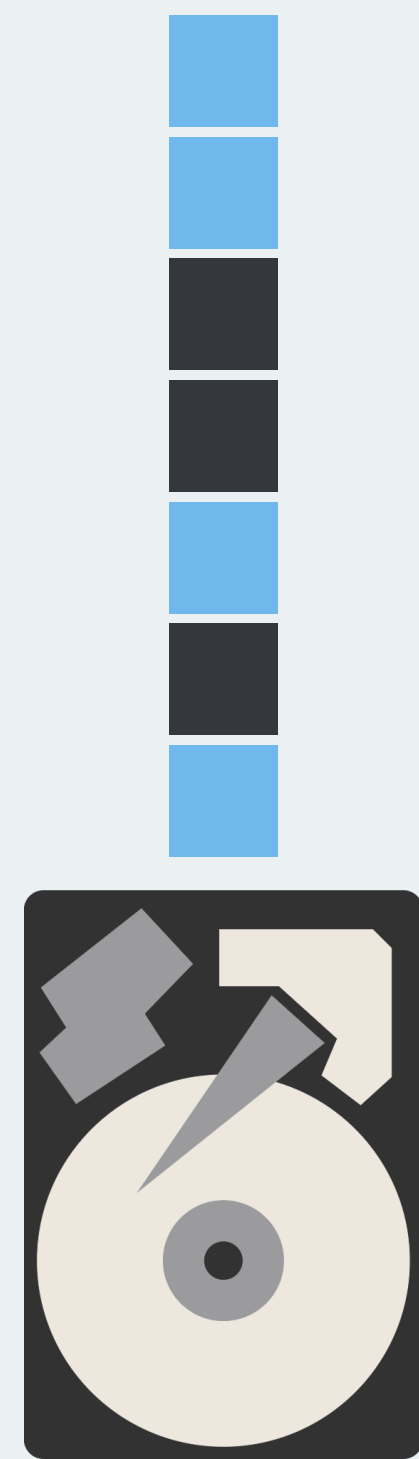
ORDONNANCEMENT DES OPÉRATION D'E/S



Ordonnancer les opérations
différemment peut-il permettre
d'optimiser le temps d'utilisation
global du disque ?

LES DISQUES DURS

ORDONNANCEMENT - SHORTEST SEEK TIME FIRST



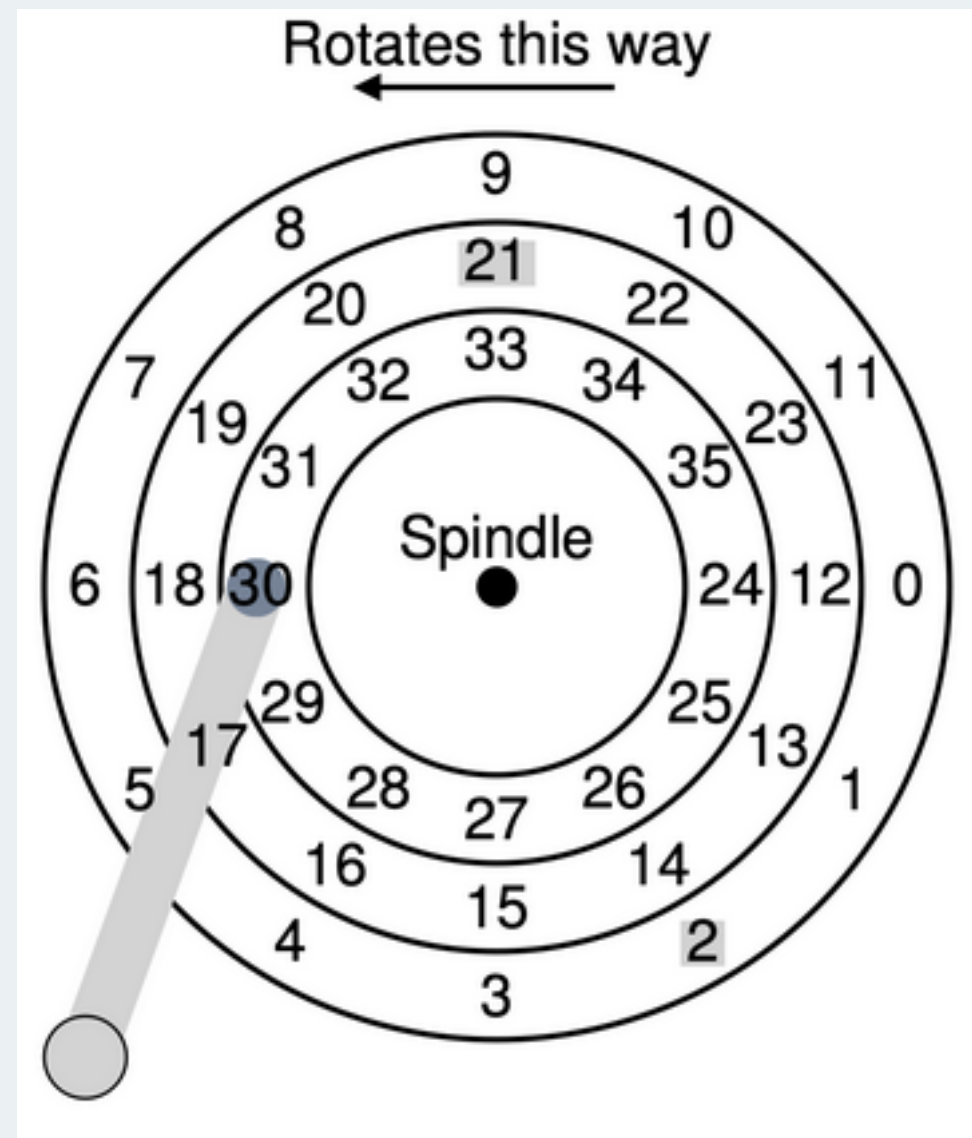
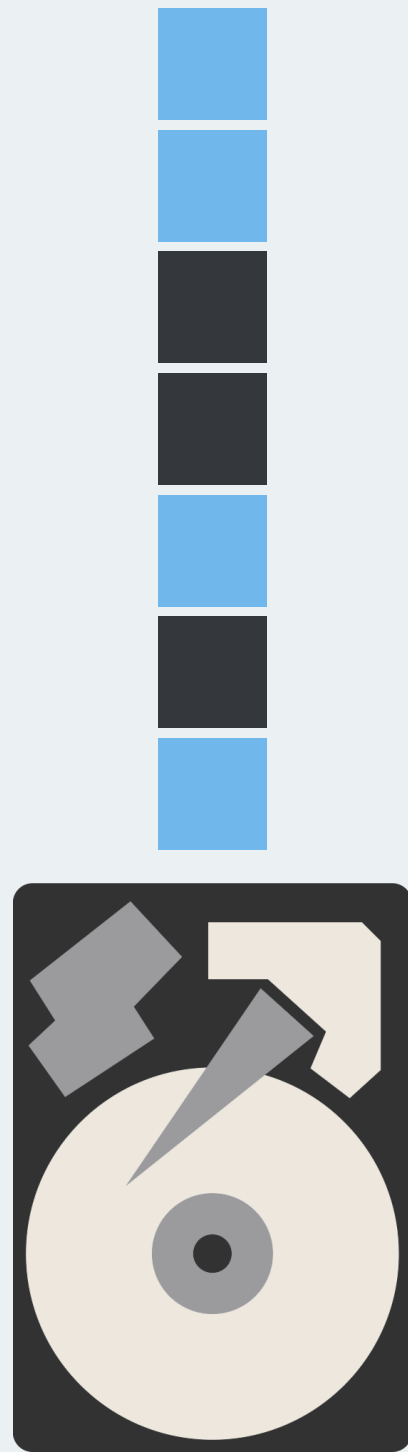
Choisit les opérations sur les pistes les plus proches

Mais :

- Géométrie indisponible pour l'OS
- Famine ✗

LES DISQUES DURS

ORDONNANCEMENT - CIRCULAR SCAN



Choisit les opérations sur
les pistes les plus proches

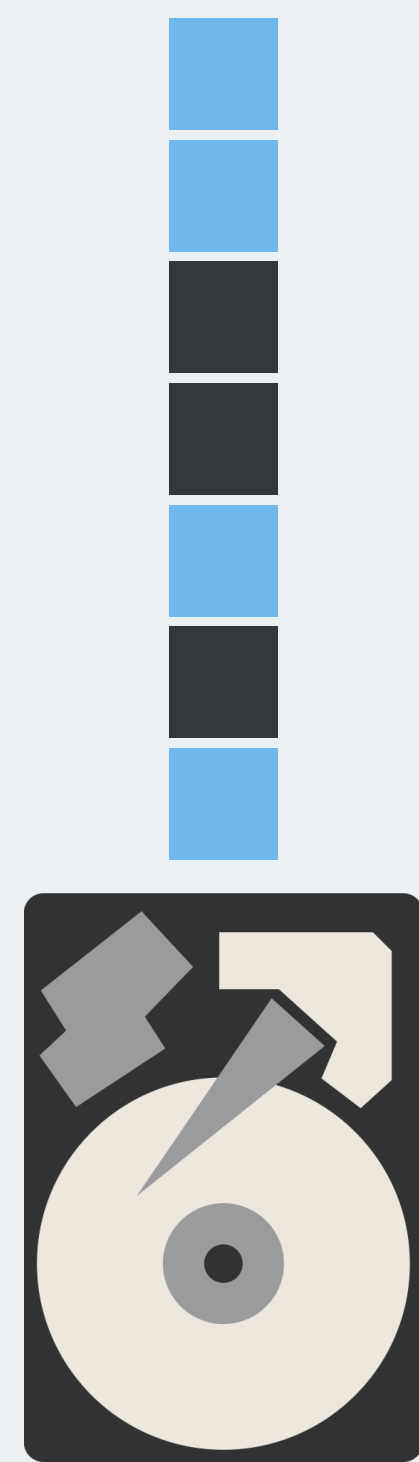
+ Balayage des pistes

Mais :
• ~~Famine~~

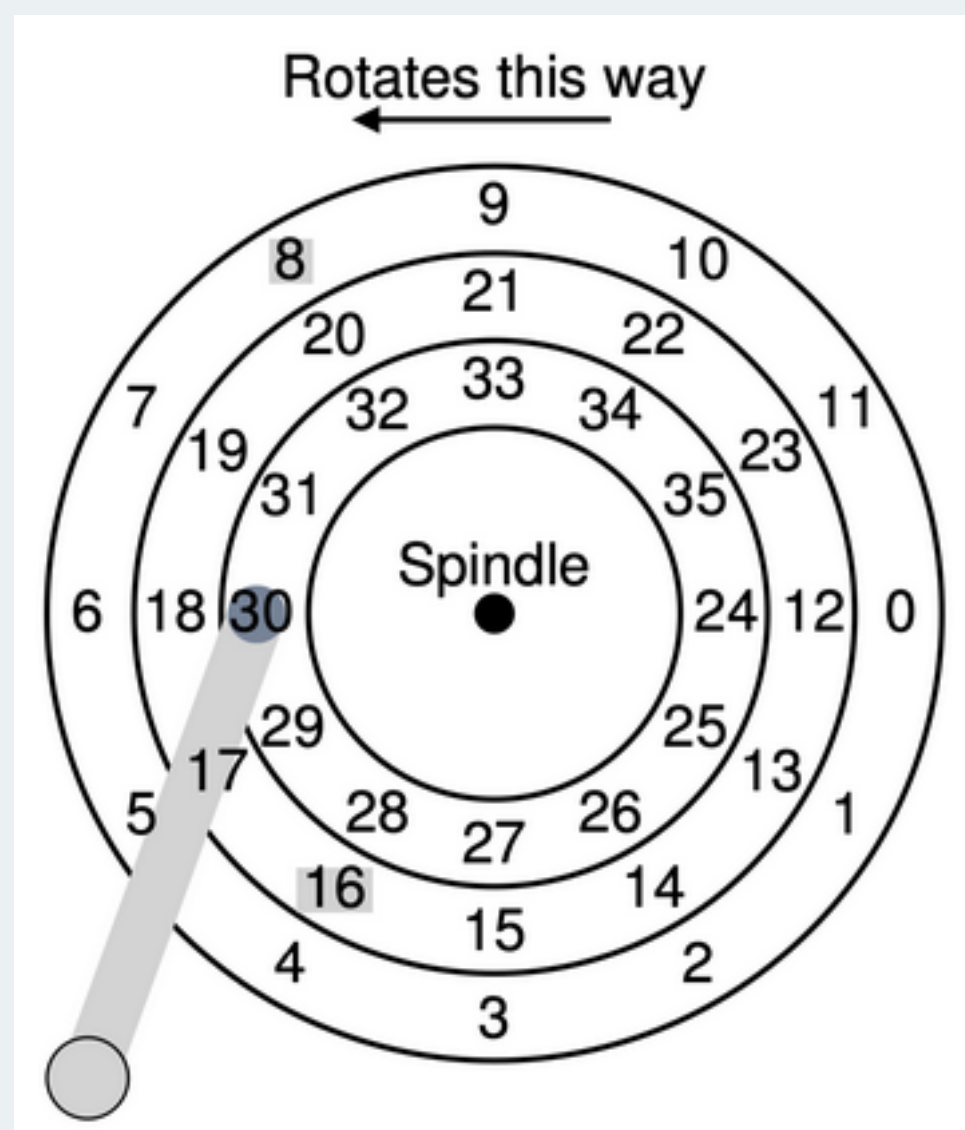
Cependant, ignore la rotation...

LES DISQUES DURS

ORDONNANCEMENT - SHORTEST POSITIONING TIME FIRST



Le problème avec SSTF



Faut-il choisir 8 ou 9 ?

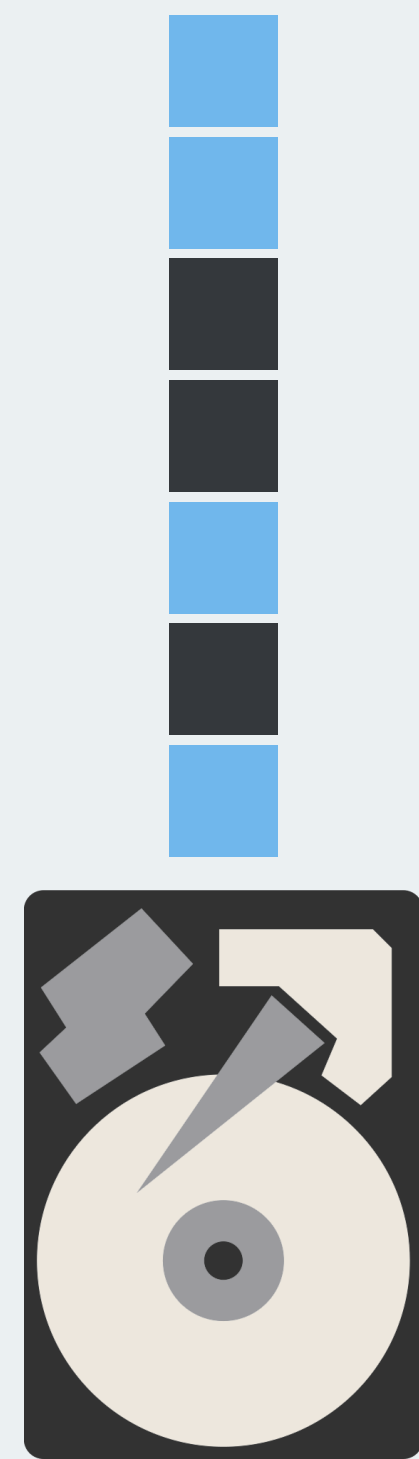
Si temps de recherche >> délai de rotation : SSTF ok ✓

Si temps de recherche << délai de rotation : SSTF pas ok ✗

↓
SPTF

LES DISQUES DURS

ORDONNANCEMENT - AUTRES PROBLÉMATIQUES



Où effectuer l'ordonnancement ?
Lecteur ou OS ?

Attendre avant d'émettre une
opération ?

LES DISQUES DURS

TP & QUIZ

16

TP : <https://gitlabpub.minet.net/xanode/tp-formation-stockage/>

QUIZ :

<https://dashboard.blooket.com/set/63f64d81215002c9838bf276>

REDUNDANT DISK ARRAYS

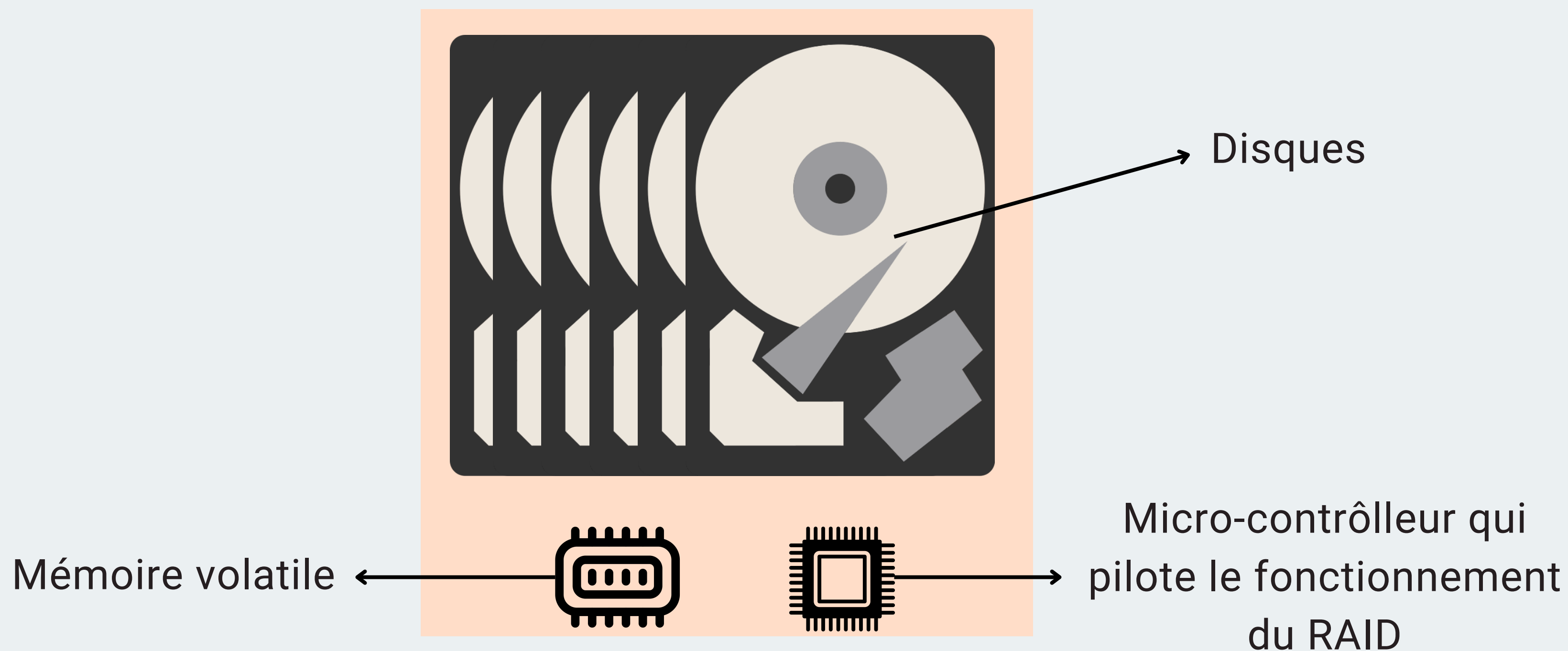
INTRODUCTION



Objectif du RAID : mutualiser les disques

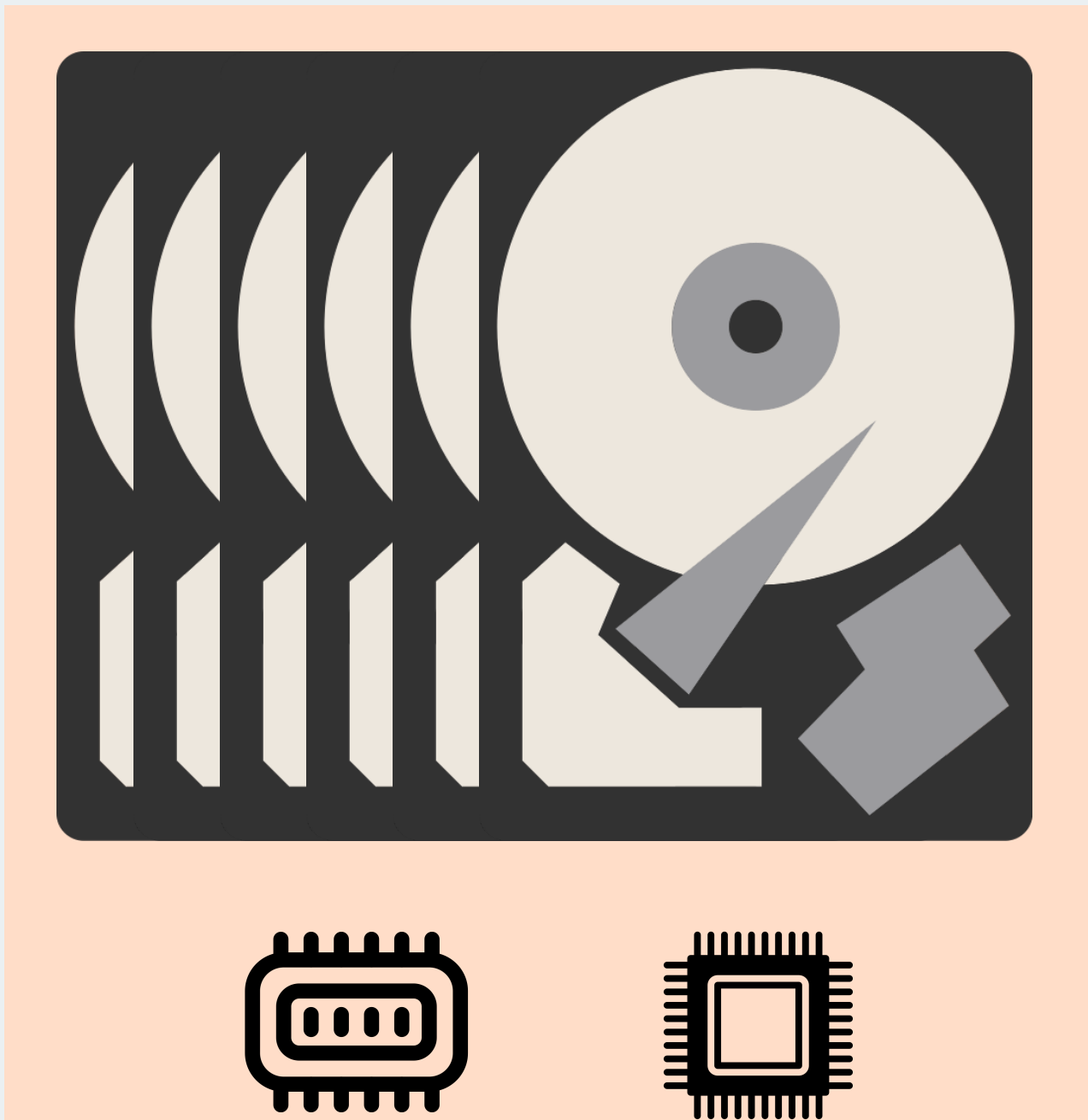
REDUNDANT DISK ARRAYS

INTERFACE & ÉLÉMENTS INTERNES



REDUNDANT DISK ARRAYS

MODÈLE DE DÉFAILLANCE



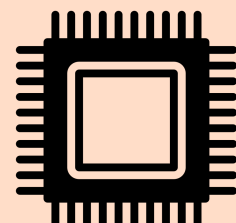
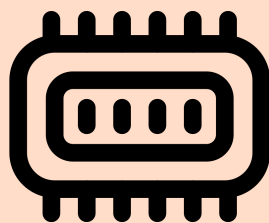
- Une défaillance tolérée
- Arrêt sur panne

Un disque peut se trouver dans 2 états :

- Fonctionnel et en marche
- Défaillant et définitivement perdu

REDUNDANT DISK ARRAYS

ÉVALUATION D'UN RAID

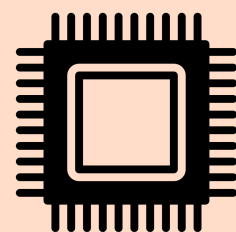
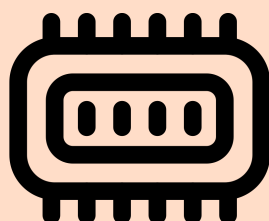


3 critères d'évaluation :

- Capacité
- Fiabilité
- Performance

REDUNDANT DISK ARRAYS

RAID 0 : VOLUME AGRÉGÉ PAR BANDES



Disque 0	Disque 1	Disque 2	Disque 3
0	1	2	3
4	5	6	7
8	9	10	11
12	13	14	15

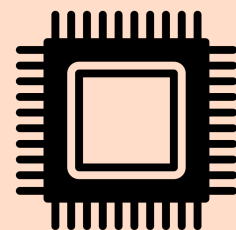
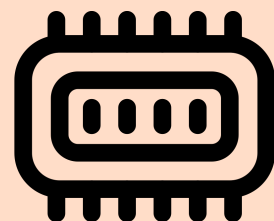
Capacité : tip top

Fiabilité : pas tip top

Performance : RÀS

REDUNDANT DISK ARRAYS

RAID 1 : DISQUES EN MIRROIR



Disque 0	Disque 1	Disque 2	Disque 3
0	0	1	1
2	2	3	3
4	4	5	5
6	6	7	7

Capacité : pas tip top

Fiabilité : tip top

Performance : celle du moins bon disque

REDUNDANT DISK ARRAYS

RAID 4 - AGRÉGATION PAR BANDES À PARITÉ



Disque 0	Disque 1	Disque 2	Disque 3	Disque 4
0	1	2	3	P0
4	5	6	7	P1
8	9	10	11	P2
12	13	14	15	P3

Capacité : un disque sans données

Fiabilité : défaillance d'un disque tolérée

Performance : celle du moins bon disque

REDUNDANT DISK ARRAYS

RAID 5 - AGRÉGATION PAR BANDES À PARITÉ RÉPARTIE



Disque 0	Disque 1	Disque 2	Disque 3	Disque 4
0	1	2	3	P0
5	6	7	P1	4
10	11	P2	8	9
15	P3	12	13	14
P4	16	17	18	19

Capacité : pareil que RAID-4

Fiabilité : pareil que RAID-4

Performance : meilleure que RAID-4 dans certaines situations

REDUNDANT DISK ARRAYS

CONCLUSION

Choisir le RAID dépend du besoin :

Performance

RAID-0

**VOLUME AGRÉGÉ
PAR BANDES**

Performance + Fiabilité

RAID-1

DISQUES EN MIRROIR

Capacité + Fiabilité

RAID-5

**VOLUME AGRÉGÉ
PAR BANDES À
PARITÉ RÉPARTIE**

REDUNDANT DISK ARRAYS

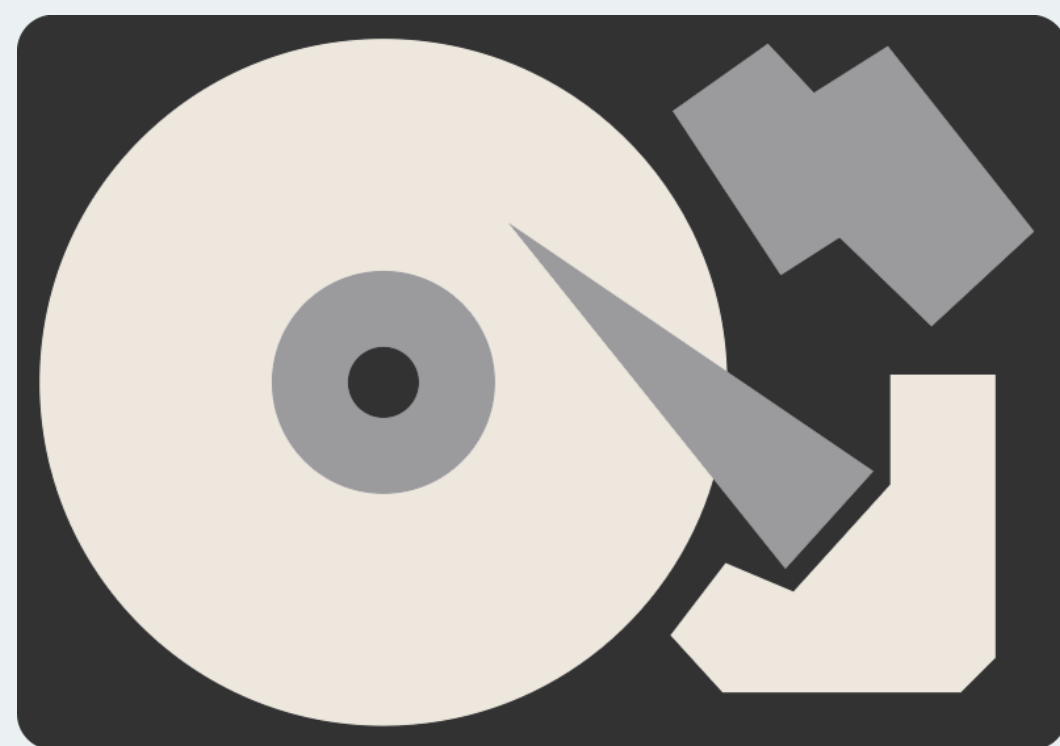
TP & QUIZ

TP : <https://gitlabpub.minet.net/xanode/tp-formation-stockage/>

QUIZ :
<https://dashboard.blooket.com/set/63f652eb04efbdc117ae63ef>

SYSTÈME DE FICHIERS

INTRODUCTION



Disque dur
Matériel



Système de fichier
Logiciel

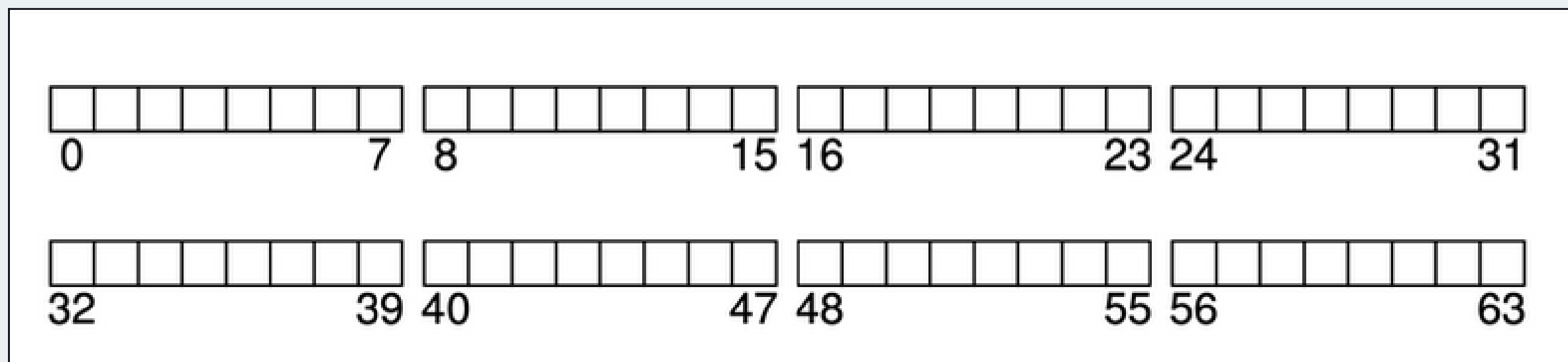


Utilisateur

SYSTÈME DE FICHIERS

ORGANISATION GLOBALE - BLOCS

Division d'un disque en **bloc**

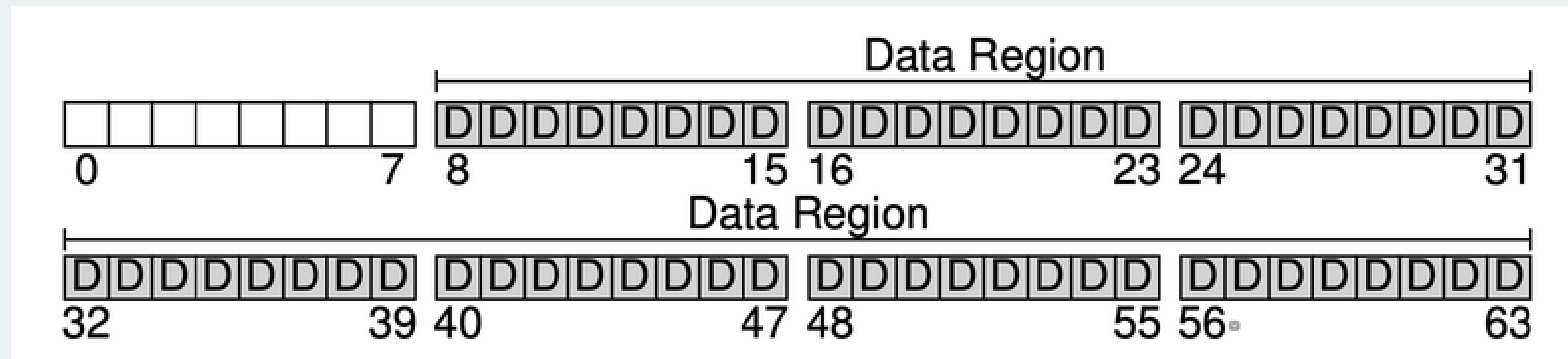


Découpage en 64 blocs de 8 secteurs = 4 Ko

SYSTÈME DE FICHIERS

ORGANISATION GLOBALE - RÉGION DE DONNÉES

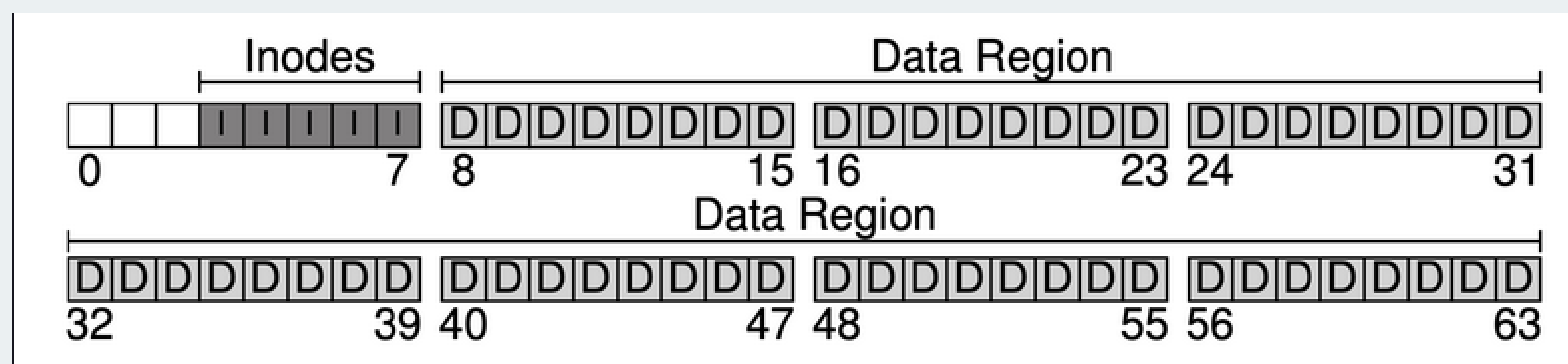
On réserve la majeure partie du disque pour les données de l'utilisateur :



SYSTÈME DE FICHIERS

ORGANISATION GLOBALE - INODES

Suivit des informations de chaque fichier
(emplacement, propriétaire, droits, etc.)



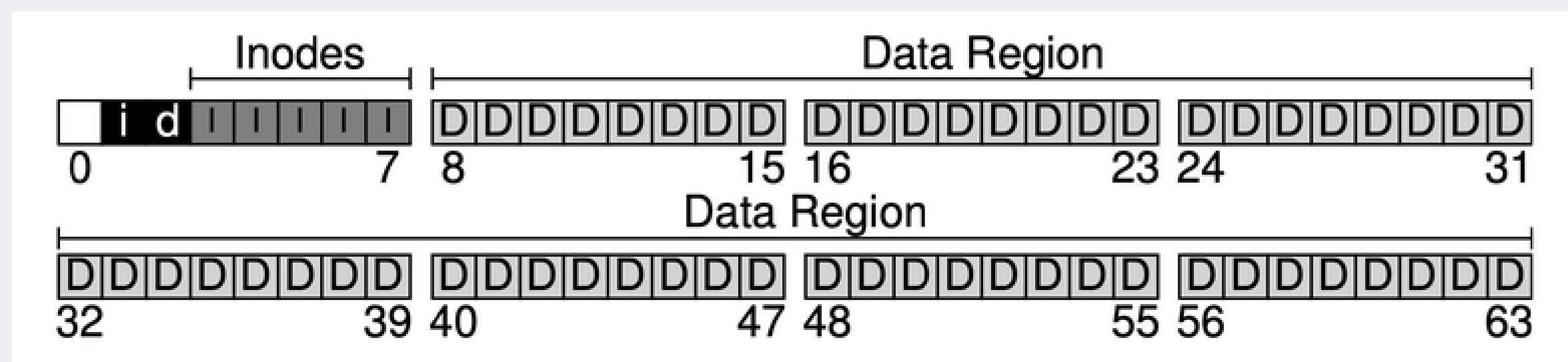
Réservation de 5 blocs pour les inodes

→ **Limite le nombre de fichiers !**

SYSTÈME DE FICHIERS

ORGANISATION GLOBALE - STRUCTURES D'ALLOCATION

Besoin de savoir si un inode ou un bloc de donnée est alloué ou non

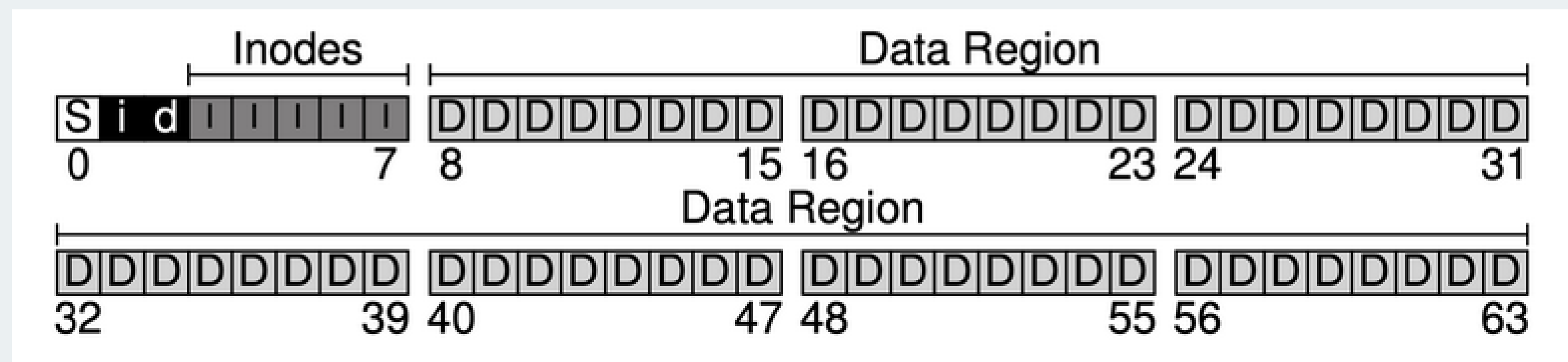


*On fait correspondre des bits aux blocs,
leurs états indiquent la disponibilité des blocs*

SYSTÈME DE FICHIERS

ORGANISATION GLOBALE - SUPERBLOC

Conserve les informations primaires du système de fichier
(nombre d'inodes, position du premier bloc d'inodes, etc.)

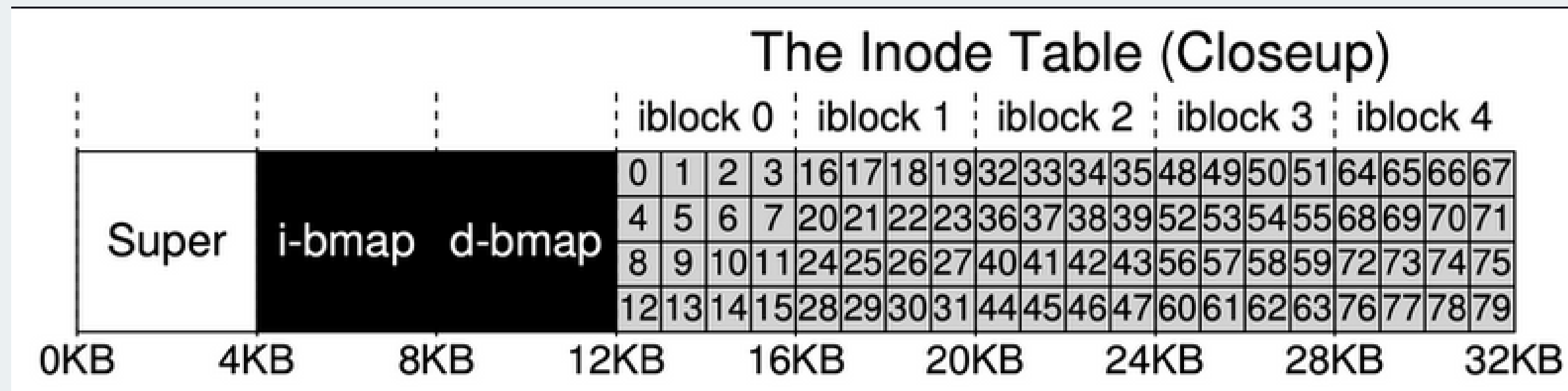


Lorsque qu'on monte la partition, le système lit le superbloc pour initialiser les paramètres du volume

SYSTÈME DE FICHIERS

ORGANISATION DES FICHIERS : NŒUDS D'INDEX

Structure qui contient les métadonnées d'un fichier donné



*Un inode contient tout ce qu'il faut savoir au sujet d'un fichier
(type, taille, nombre de blocs, permissions, etc.)*

SYSTÈME DE FICHIERS

ORGANISATION DES FICHIERS : NŒUDS D'INDEX

Exemple d'inode d'ext2 :

Size	Name	What is this inode field for?
2	mode	can this file be read/written/executed?
2	uid	who owns this file?
4	size	how many bytes are in this file?
4	time	what time was this file last accessed?
4	ctime	what time was this file created?
4	mtime	what time was this file last modified?
4	dtime	what time was this inode deleted?
2	gid	which group does this file belong to?
2	links_count	how many hard links are there to this file?
4	blocks	how many blocks have been allocated to this file?
4	flags	how should ext2 use this inode?
4	osd1	an OS-dependent field
60	block	a set of disk pointers (15 total)
4	generation	file version (used by NFS)
4	file_acl	a new permissions model beyond mode bits
4	dir_acl	called access control lists

SYSTÈME DE FICHIERS

ORGANISATION DES DOSSIERS

Beaucoup plus simple : contient une liste d'inodes qui correspondent aux fichiers contenus par le dossier

SYSTÈME DE FICHIERS

CHEMINS D'ACCÈS - LECTURE

Flux d'opération :

	data bitmap	inode bitmap	root inode	foo inode	bar inode	root data	foo data	bar data [0]	bar data [1]	bar data [2]
open(bar)			read	read	read	read	read			
read()					read		read			
read()					write read				read	
read()					write read					read

Temps

SYSTÈME DE FICHIERS

CHEMINS D'ACCÈS - ÉCRITURE

Flux d'opération :

	data bitmap	inode bitmap	root inode	foo inode	bar inode	root data	foo data	bar data [0]	bar data [1]	bar data [2]
create (/foo/bar)			read			read				
				read			read			
		read write			read write		write			
				write						
write()	read write				read			write		
					write read					
write()	read write							write		
					write read					
write()	read write									write
					write					

Temps

SYSTÈME DE FICHIERS

TP & QUIZ

38

TP : <https://gitlabpub.minet.net/xanode/tp-formation-stockage/>