Technion

Electronics
Computers
Communications

Information in Storage Devices

049063 – EE Department, Technion

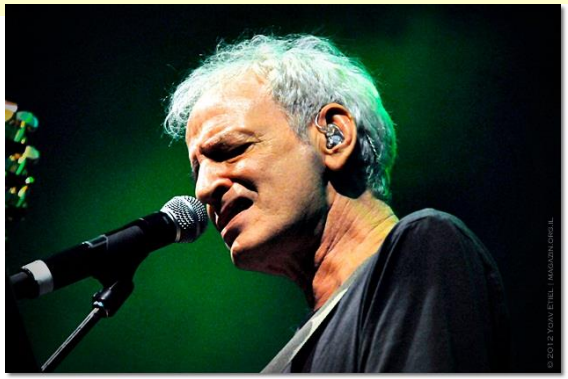
LECTURE 2: HDD AND SSD ACCESS

A Tale of Two Media Stars

- Has been around forever
- Improves, but looks the same
- Predictable performance

- Fast to respond
- Heavily hyped
  - High media exposure
- You know can do wonders
  - But most encounters less exciting

A Tale of Two Media Stars



- Has been around forever
- Improves, but looks the same
- Predictable performance

- Fast to respond
- Heavily hyped
  - High media exposure
- You know can do wonders
  - But most encounters less exciting

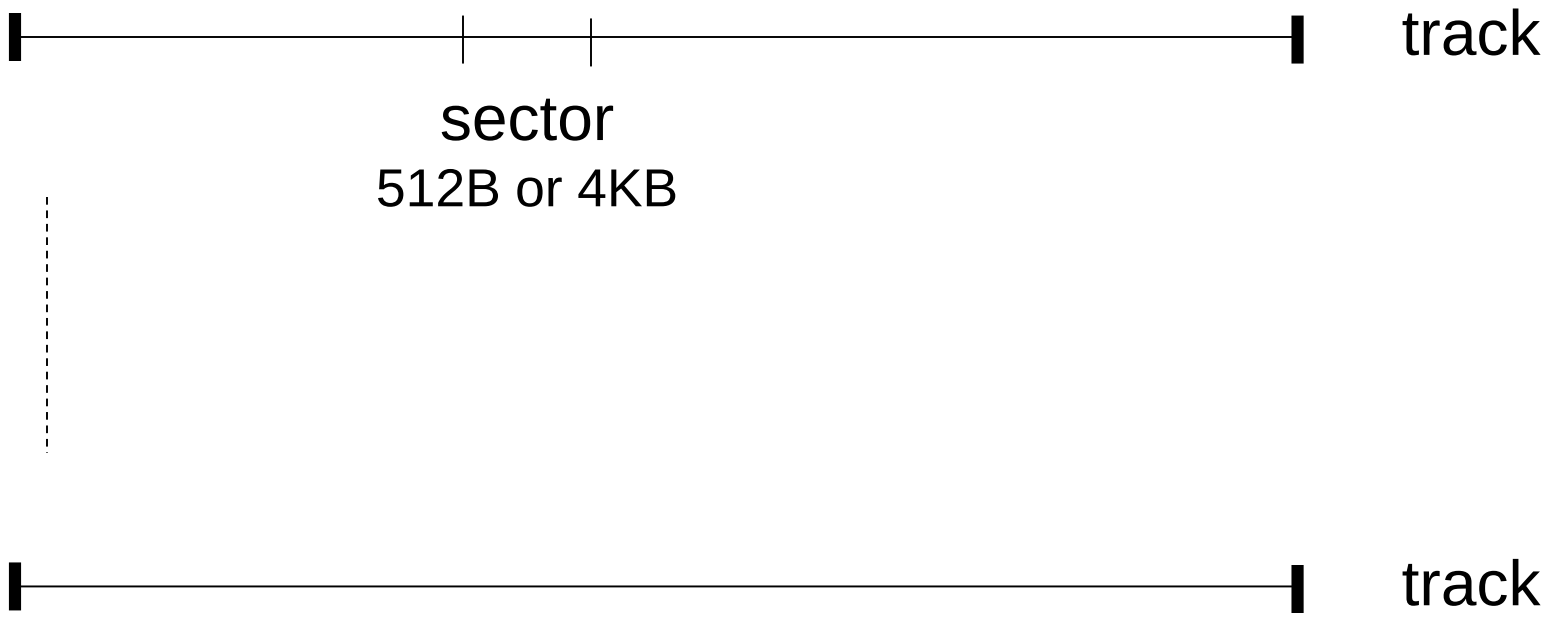
Hard-Disk Drive (HDD)

- Revolving disks, magnetic media
- Invented 1956 (IBM)
  - Size: two refrigerators
  - # disks: 50
  - Capacity: 4MB
- Capacity today: 8TB

– Scaling with **media** and **head** technologies

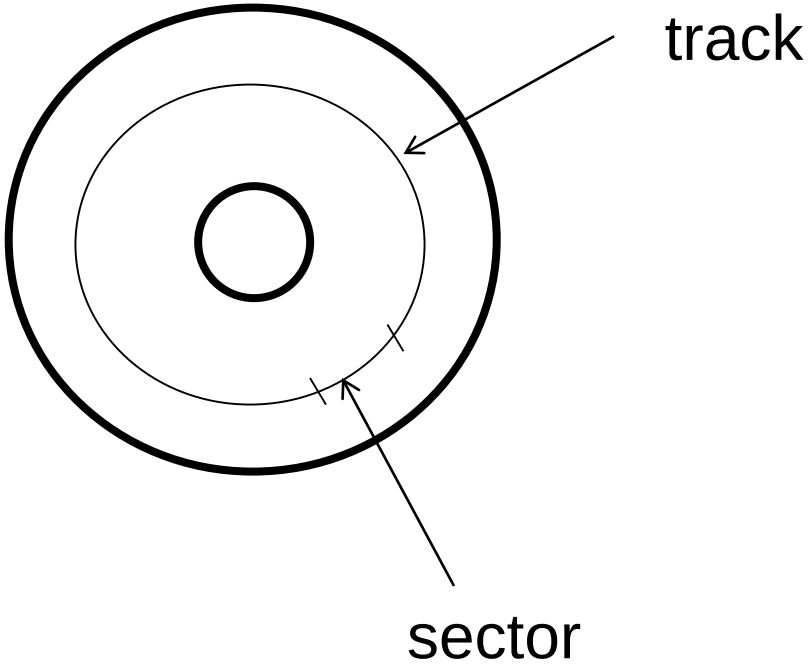
HDD Access

- 1D view



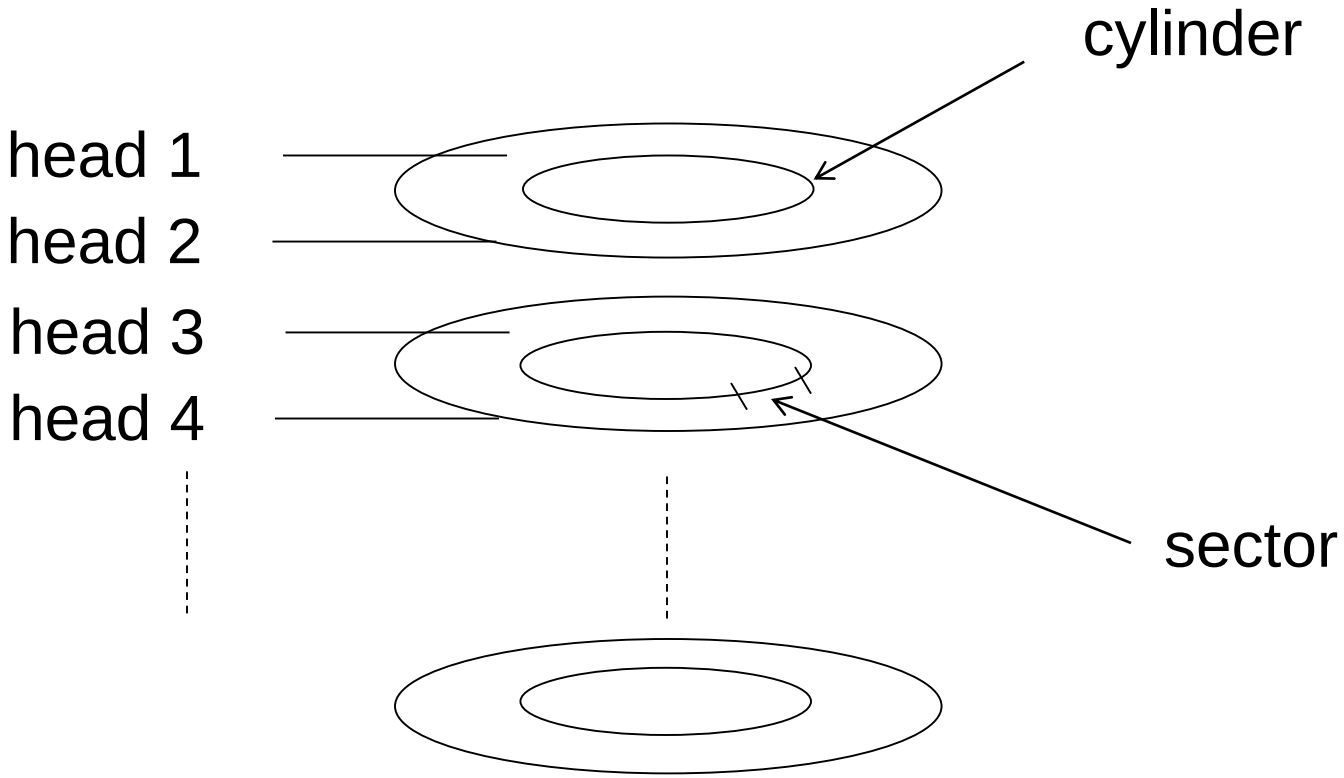
HDD Access

- 2D view



HDD Access

- 3D view



PBA = (Cylinder, Head, Sector) – CHS address

Access Time

T(Read) = T(write) = T(cyl. switch) + T(head switch) + T(rot)

(1) (2) (3)

seek time

Random Access

Define:

Access request Req(L_addr,rw) → Req(L_addr,R) / Req(L_addr,W)

Definition:

A device is called **random access** if any sequence of requests Req_1, Req_2,… → t

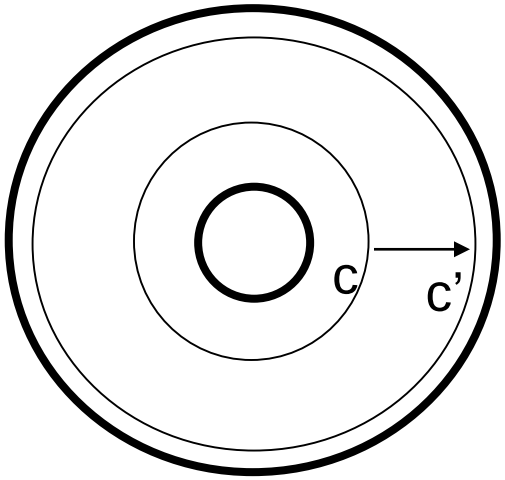
is allowed, and all such sequences exhibit a similar response behavior.

HDD Read/Write ordering

- HDD R/W switch time

Seek Times

$$\mathrm{T}(\mathrm{c} \to c') = \tau \frac{|c - c'|}{\#cyls - 1}$$



Normalized cylinder addresses:

$$\gamma = \frac{c}{\#cyls - 1} \qquad \gamma' = \frac{c'}{\#cyls - 1}$$

$$\mathrm{T}(\mathrm{c} \to c') = \tau\, |\gamma - \gamma'|$$

↑

full-seek time

Seek-Time Distribution

- Max
  - all possible γ' $\max[\mathrm{T}(\gamma)] = \max[\tau\gamma, \tau(1-\gamma)]$
  - all possible γ, γ' $\max[\mathrm{T}] = \tau$

- What is the expected seek time?

$$\mathrm{E}[\mathrm{T}] = ?$$

Expectation given origin γ

- Expectation given γ (uniform γ')

$$\mathrm{E}[\mathrm{T}(\gamma)] = \tau\left[\gamma^2 - \gamma + \frac{1}{2}\right]$$

Overall Expectation

- Expectation (uniform γ, γ')

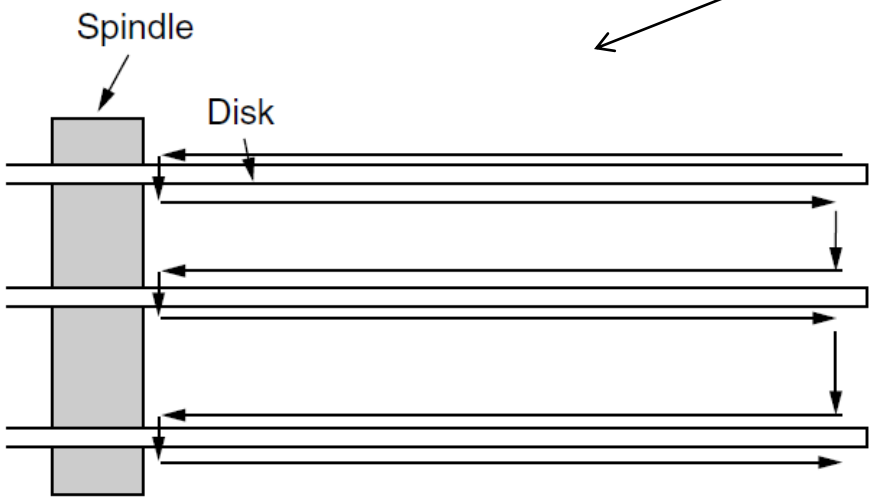
$$\mathrm{E}[\mathrm{T}] = \mathrm{E}\{\mathrm{E}[\mathrm{T}(\gamma)]\} = \frac{\tau}{3}$$

Access Time

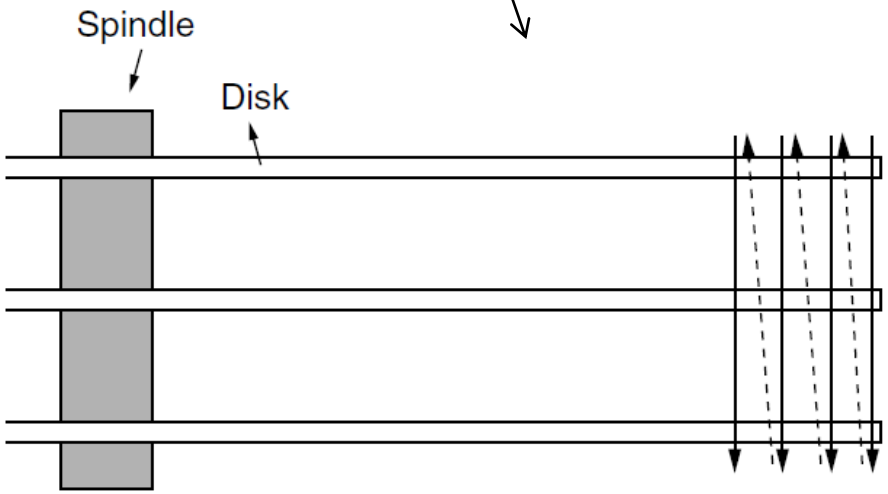
T(Read) = T(write) = T(cyl. switch) + T(head switch) + T(rot)

(1) (2) (3)

or



serpentine mode



cylinder mode

Access Time

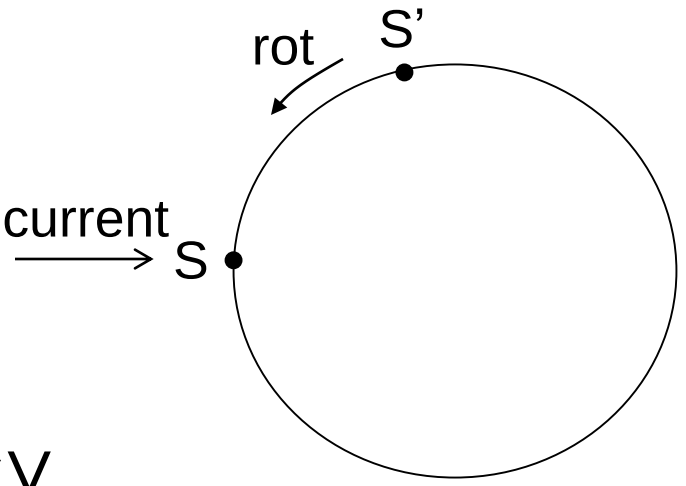
T(Read) = T(write) = T(cyl. switch) + T(head switch) + T(rot)

(1) (2) (3)

rotational latency

Rotational Latency

$$\mathrm{T}(\mathrm{S} \rightarrow S') = T_{rev} \frac{S \dot{-} S'}{\#sectors/rev}$$



- Max rotational latency

$$\max[\mathrm{T}] = T_{rev}$$

- Expectation

$$\mathrm{E}[\mathrm{T}] = \frac{T_{rev}}{2}$$

Command Queueing

- HDD manages command queues
- Allowed out-of-order execution

$S \longrightarrow [S_1, S_2, \ldots, S_N]$

current — reqs in queue

- Optimal choice of next:

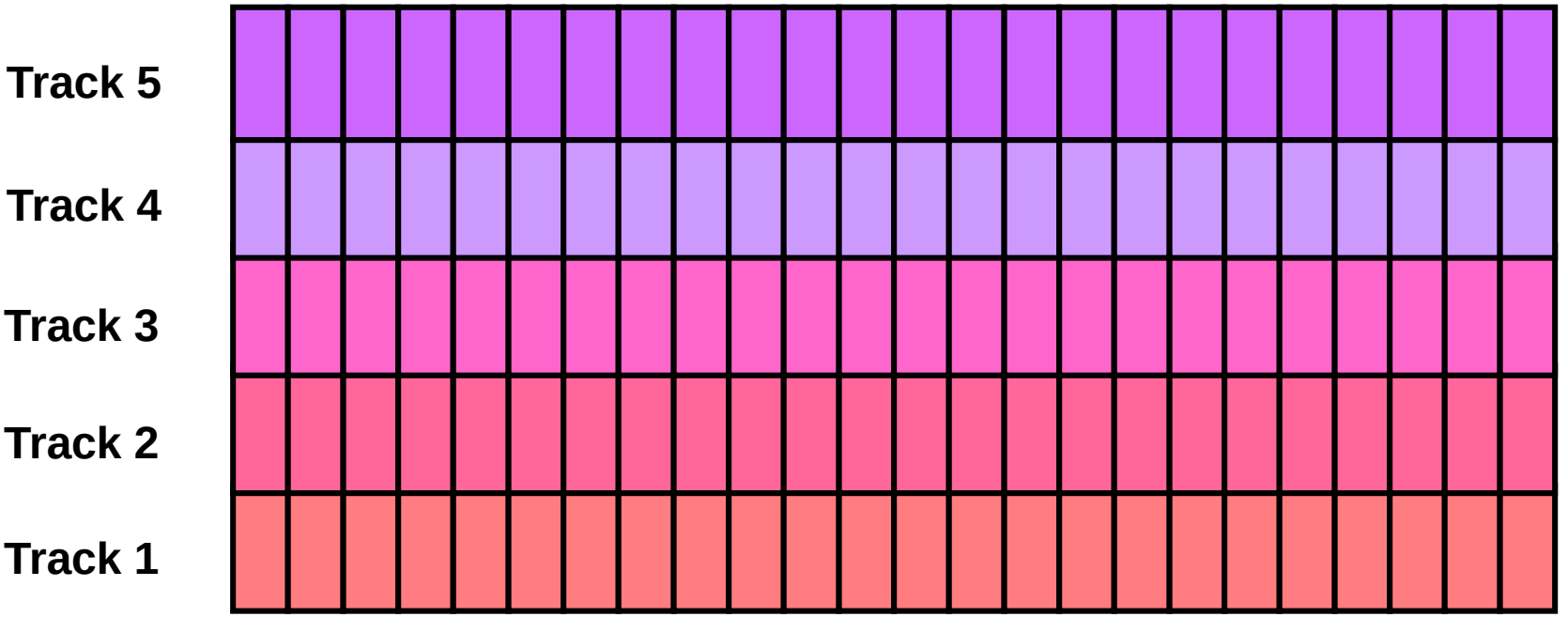
$$S \to S_i : \; i = \arg \min_{j \in 1,\ldots N} T[S \to S_j]$$

- Expected latency with N-queue

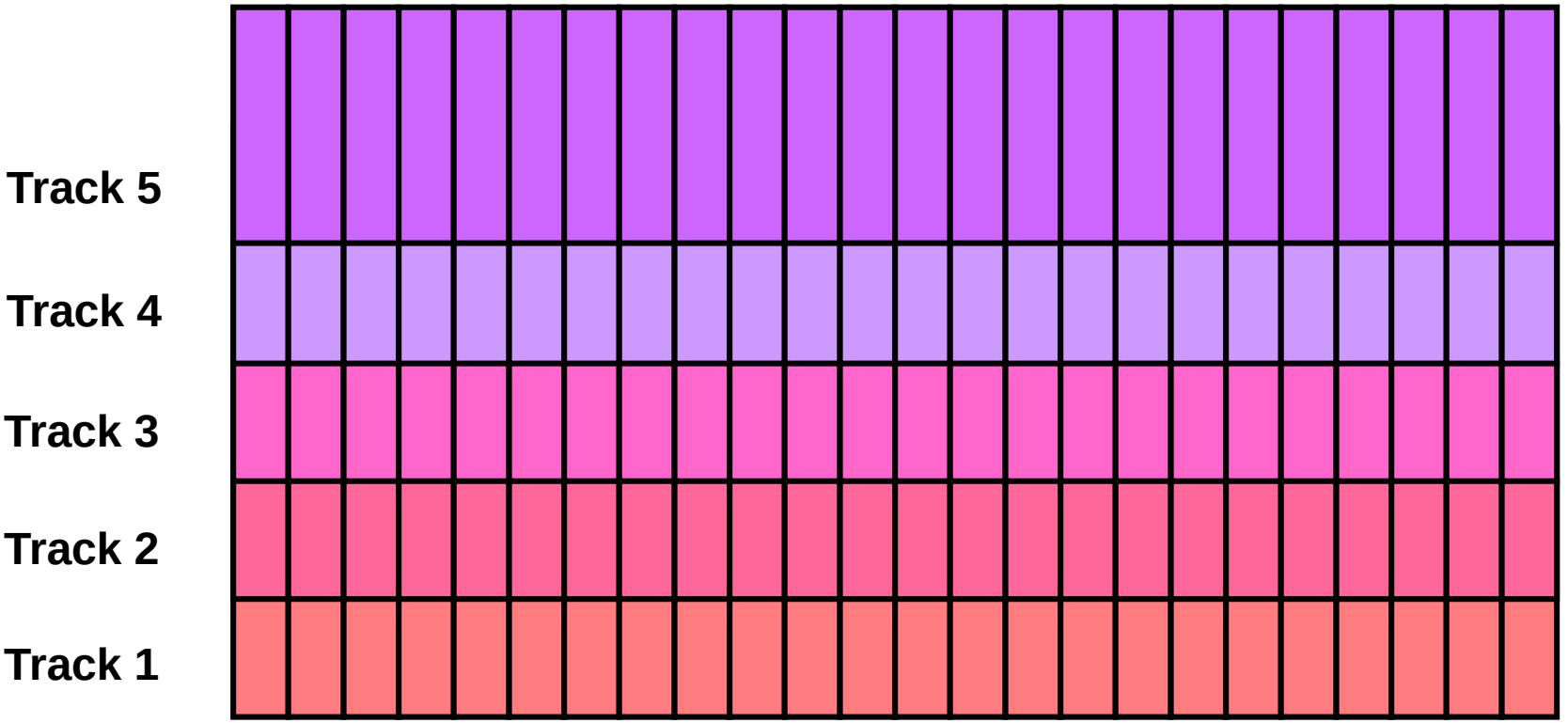
$T[S \to S_j] \sim U[0, T_{rev}]$

$$E[T_N] = E[\min_{j \in 1,\ldots N} T[S \to S_j]] = ?$$

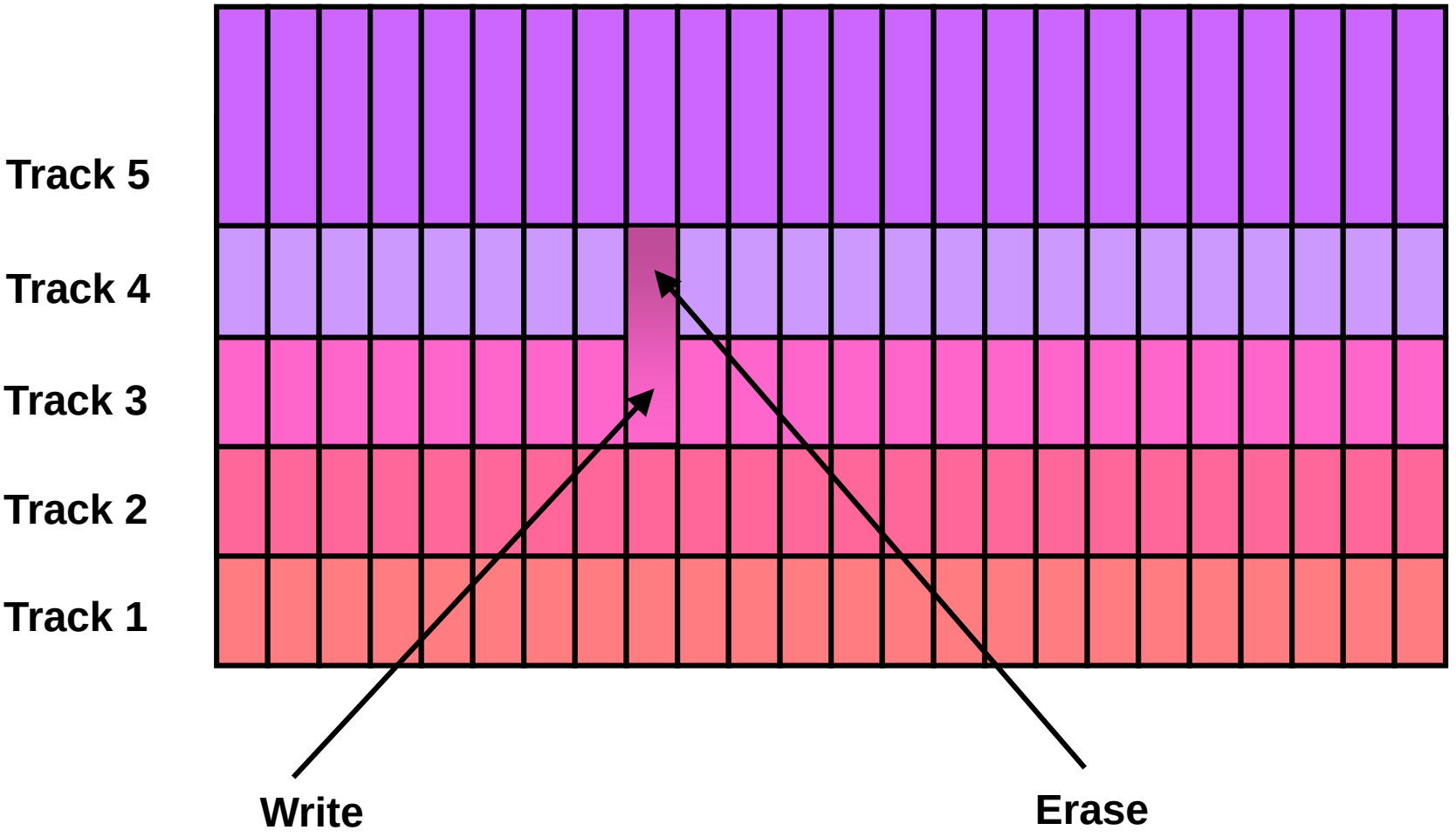
Conventional Recording



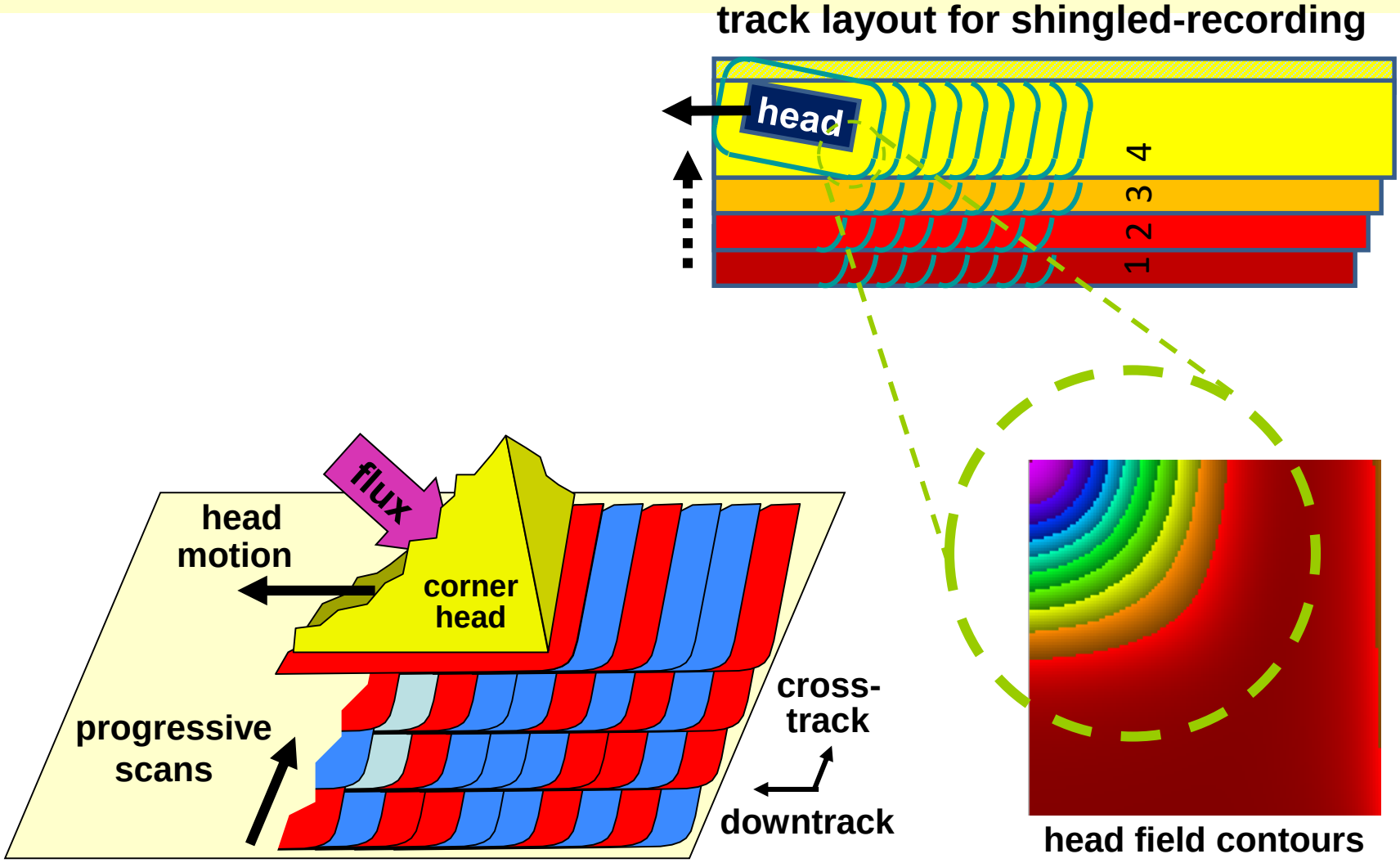
Shingled Recording



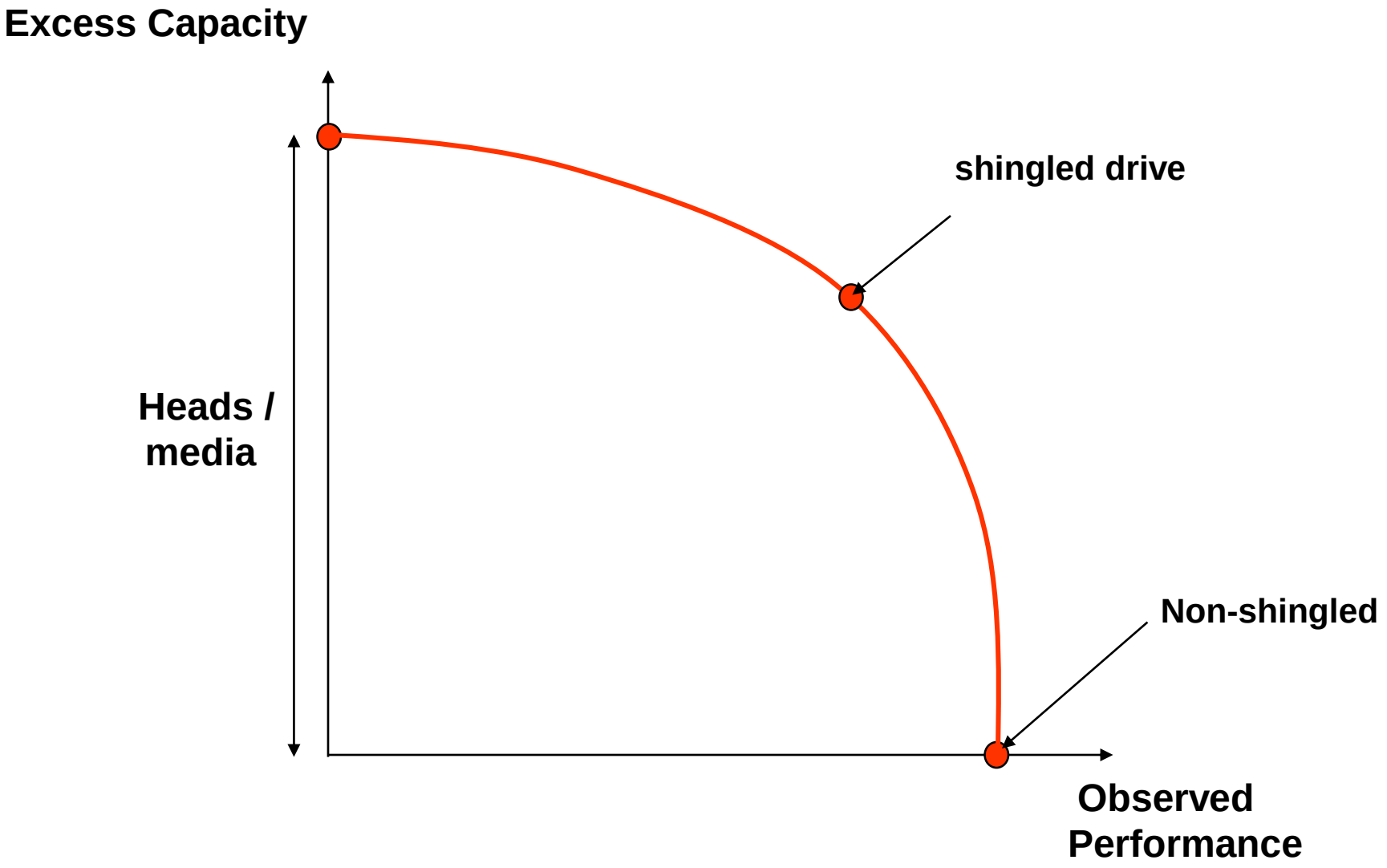
Shingled Recording – No Random Write



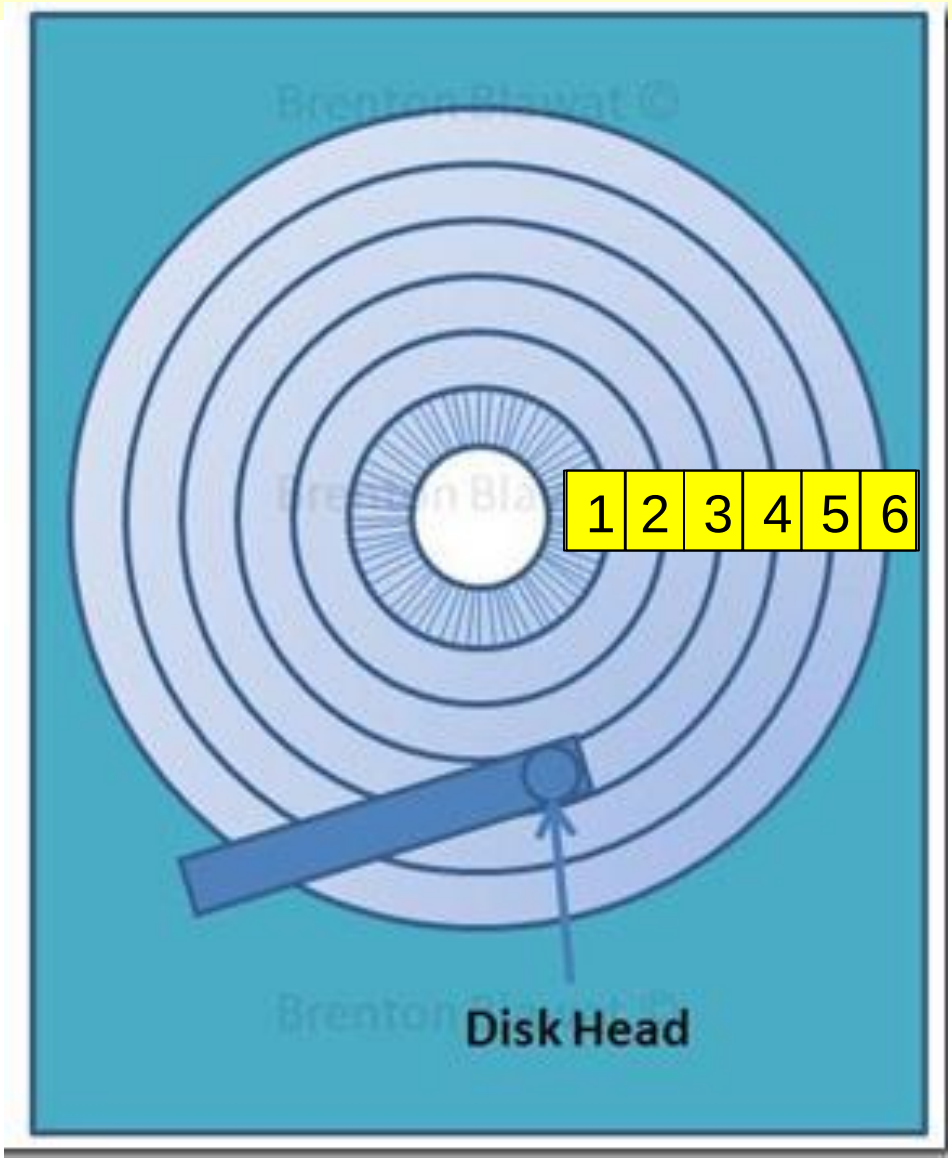
Shingled Recording Magnetics



Shingled Drive Tradeoff



Performance with Shingling

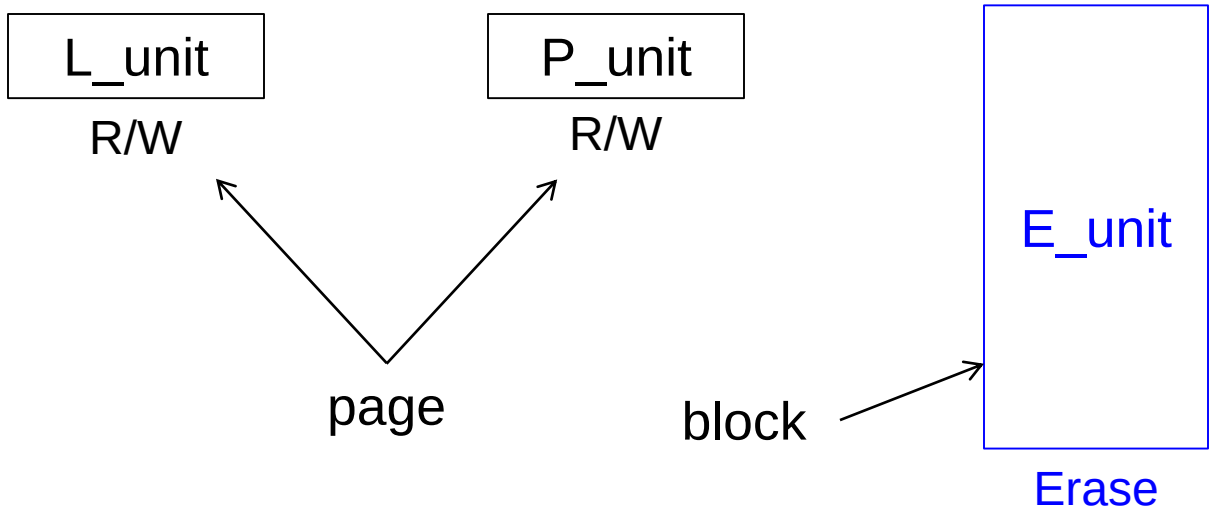


Solid-State Drive (SSD)

- Silicon-based array of memory cells
  - with standard interface (HDD replacement)
- Invented 1995 (M-Systems, Israel)
  - Uses NAND flash for maximal density
- More expensive, but much faster
- Capacity scales by “Moore’s law”

Flash: No Random-Access Erase

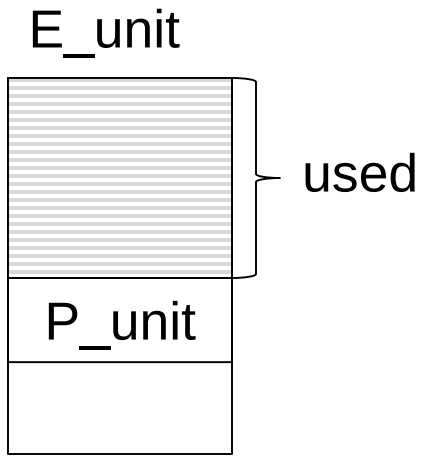
- New: Erase unit



- Physical erase of cells: only full blocks
- Write → program/erase

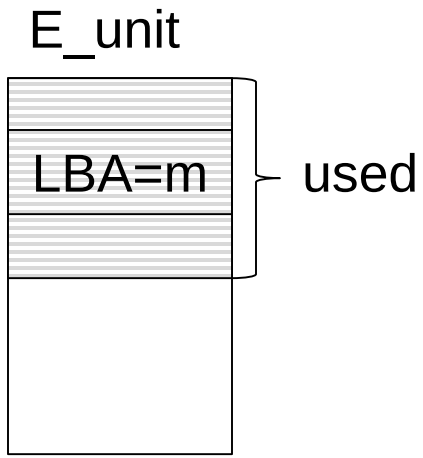
No In-Place Updates

W L_write: L_unit P_write

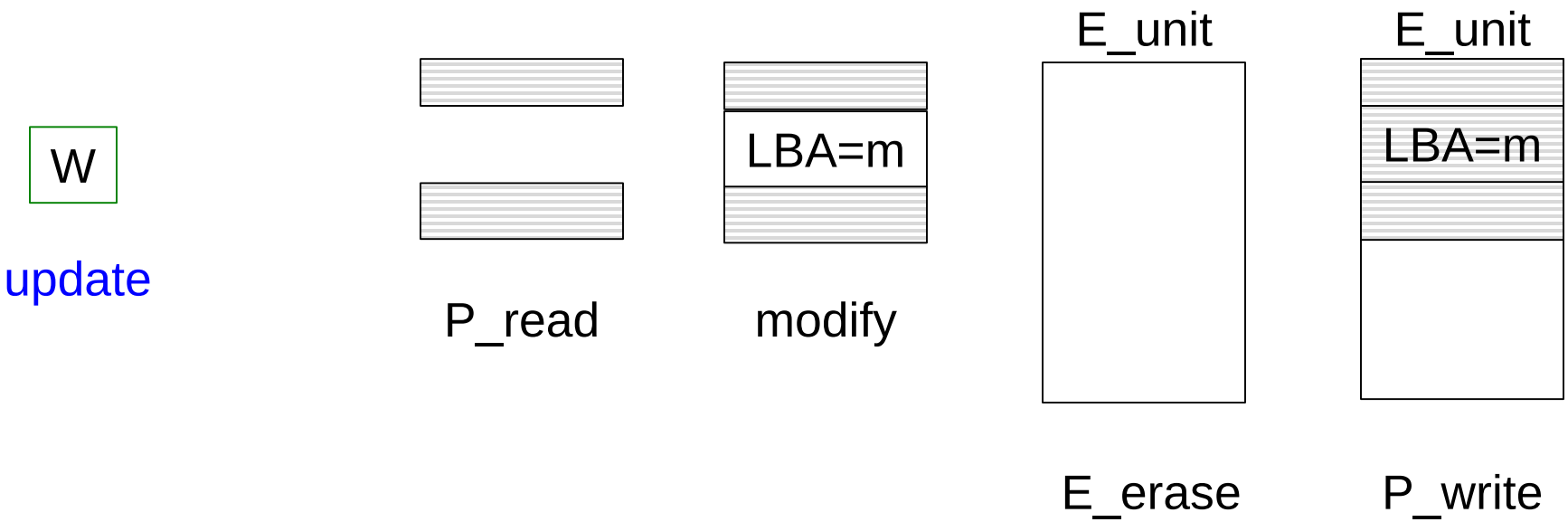


W L_write: LBA=m P_write?

update



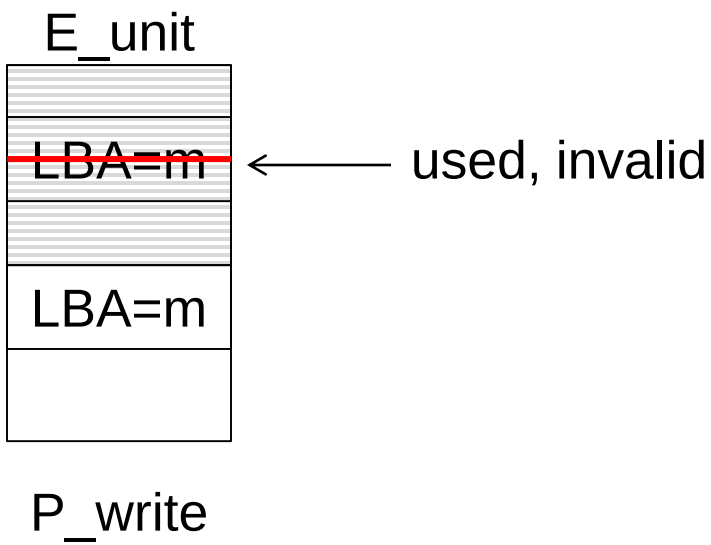
Option 1: RMEW



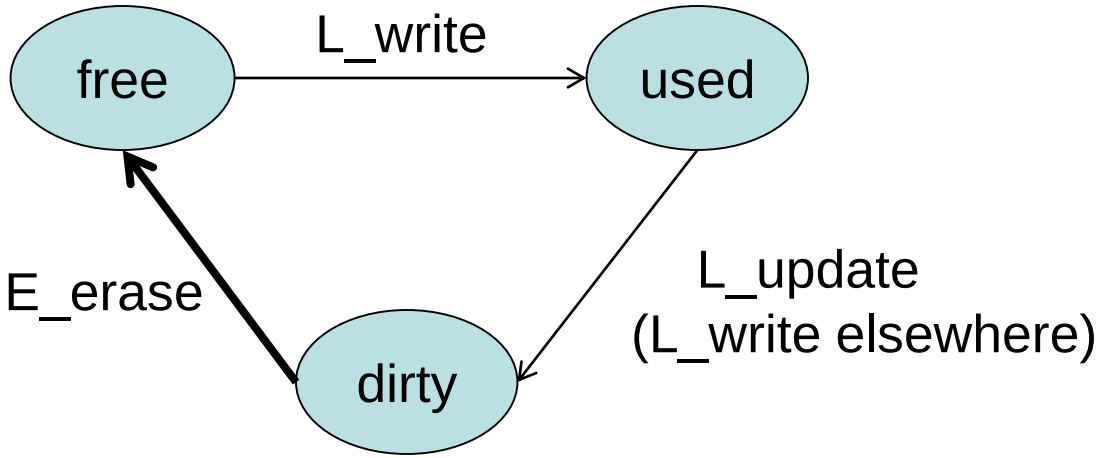
Option 2: Invalidation

W

update



Flash State Diagram



Issues

Option 1: RMEW

- Time
- Wear

Option 2: Invalidation

- Over-provisioning
- Indirection