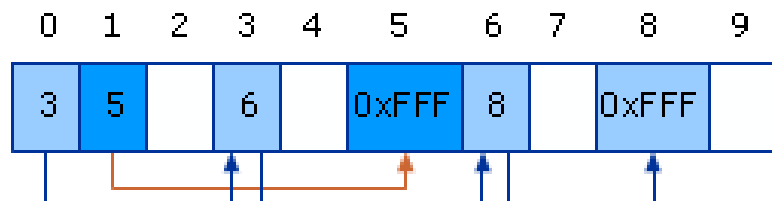
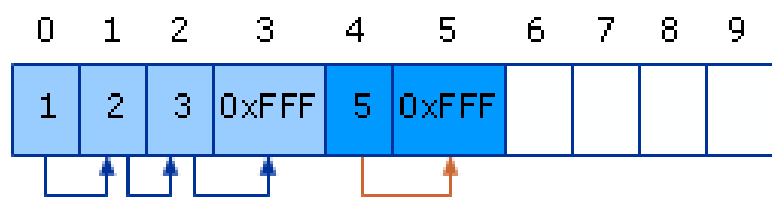


01.1. Проблема фрагментації дискового простору

Volume Before Disk Defragmentation



Volume After Disk Defragmentation



Legend



Фрагментація дискового простору – це нерівномірне розміщення блоків файлів у просторі диску.

Фрагментами називаються «неперервні» кластери блоків.

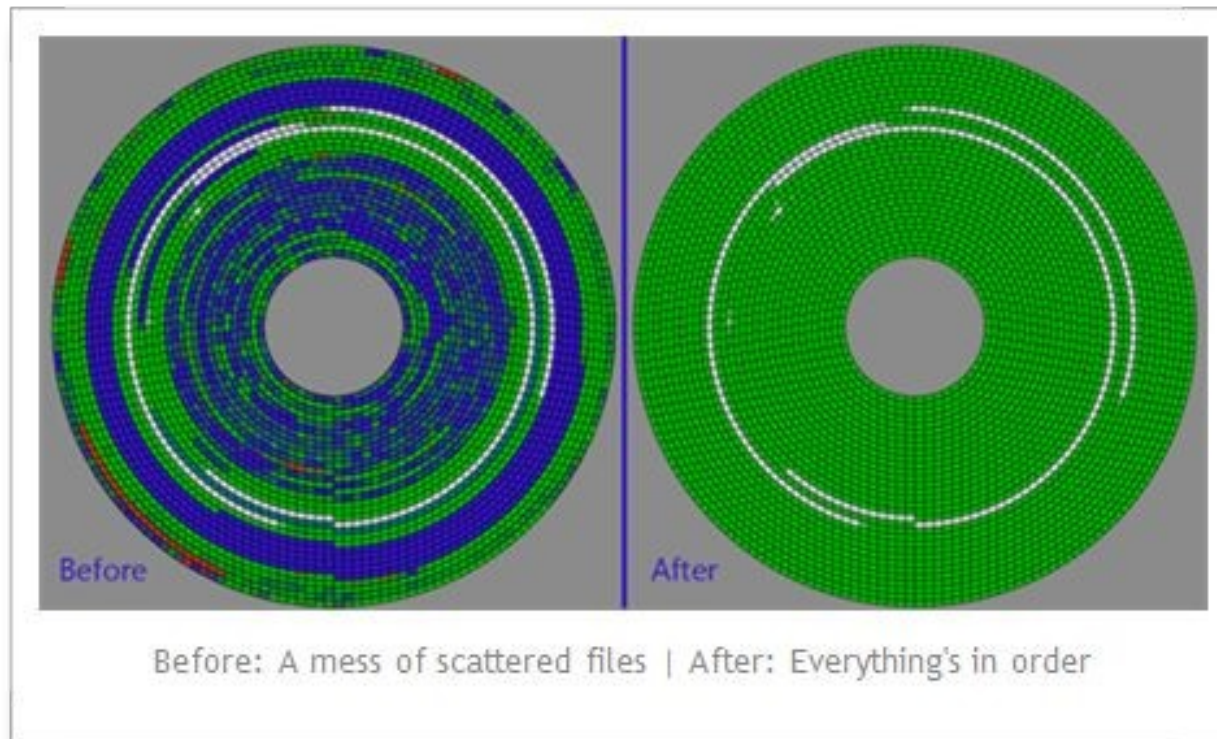
Фрагменти розділені між собою **вільним місцем**.

01.2. Проблема фрагментації дискового простору

Фрагментація є наслідком

- 1) знищення (звільнення) існуючих блоків файлів, внаслідок чого утворюються «розриви» (вільні місця) між фрагментами;
- 2) запису нових блоків файлу у ПЗП, внаслідок чого блоки одного файлу потрапляють в різні місця дискового простору.

01.3. Проблема фрагментації дискового простору



Проблема: file system fragmentation → Чим більша фрагментація, тим нижче продуктивність системи.

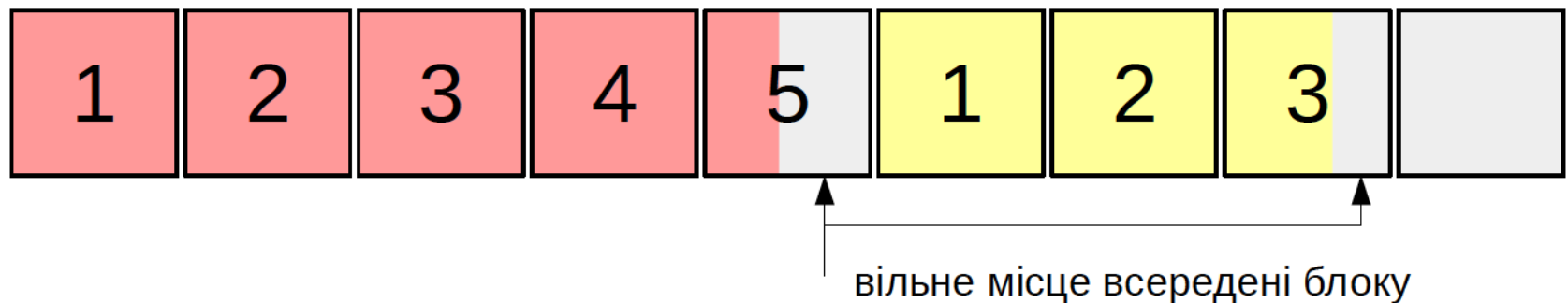
01.4. Проблема фрагментації дискового простору

Види фрагментації:

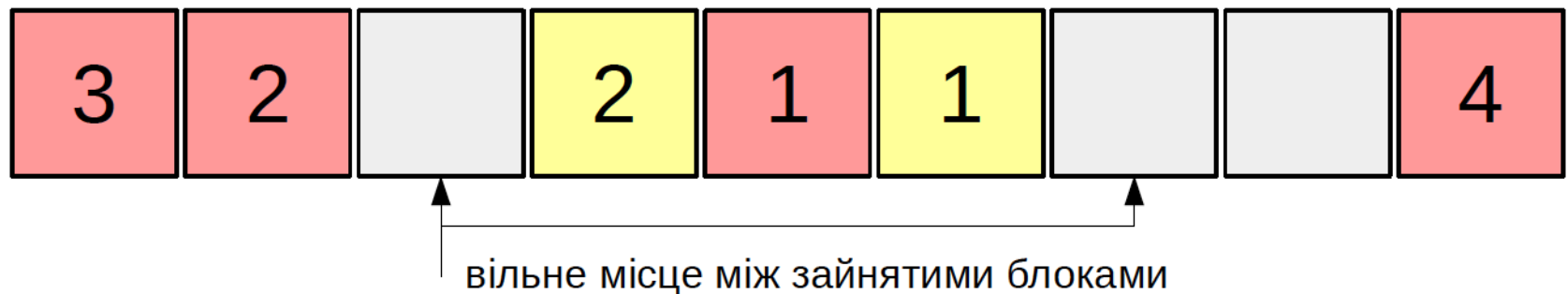
- 1) **внутрішня**: «хвіст» файлу займає не весь блок → втрати вільного місця на диску → чим більше розмір блоку, тим більше внутрішня фрагментація;
- 2) **зовнішня**: рознесення блоків одного файлу у просторі жорсткого диску → це призводить до сповільнення роботи з цим файлом (операції читання/запис) та збільшує кількість рухів зчитуючої головки жорсткого диску.

01.5. Проблема фрагментації дискового простору

1. Внутрішня фрагментація

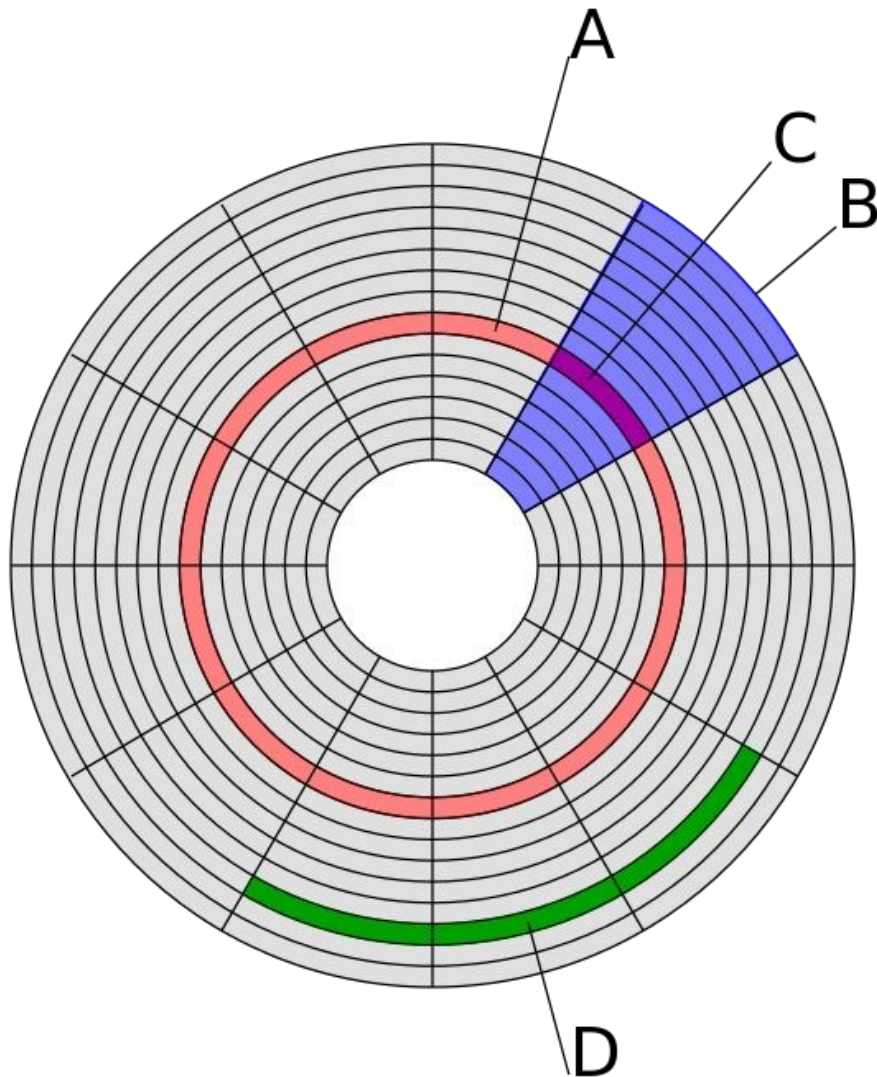


2. Зовнішня фрагментація



Внутрішня та зовнішня фрагментація

01.6. Проблема фрагментації дискового простору



Проблема вибору розміру блоку (data cluster, allocation unit): **на блоки якого розміру розбивати файли?**

Варіанти: Disk sector (C) = 512 байт → 4096 байт (Advanced Format (AF)).

Типові розміри блоку (D) в сучасних системах лежать в межах: від 1 сектора (512 B) до 128 секторів (64 KiB).

01.7. Проблема фрагментації дискового простору

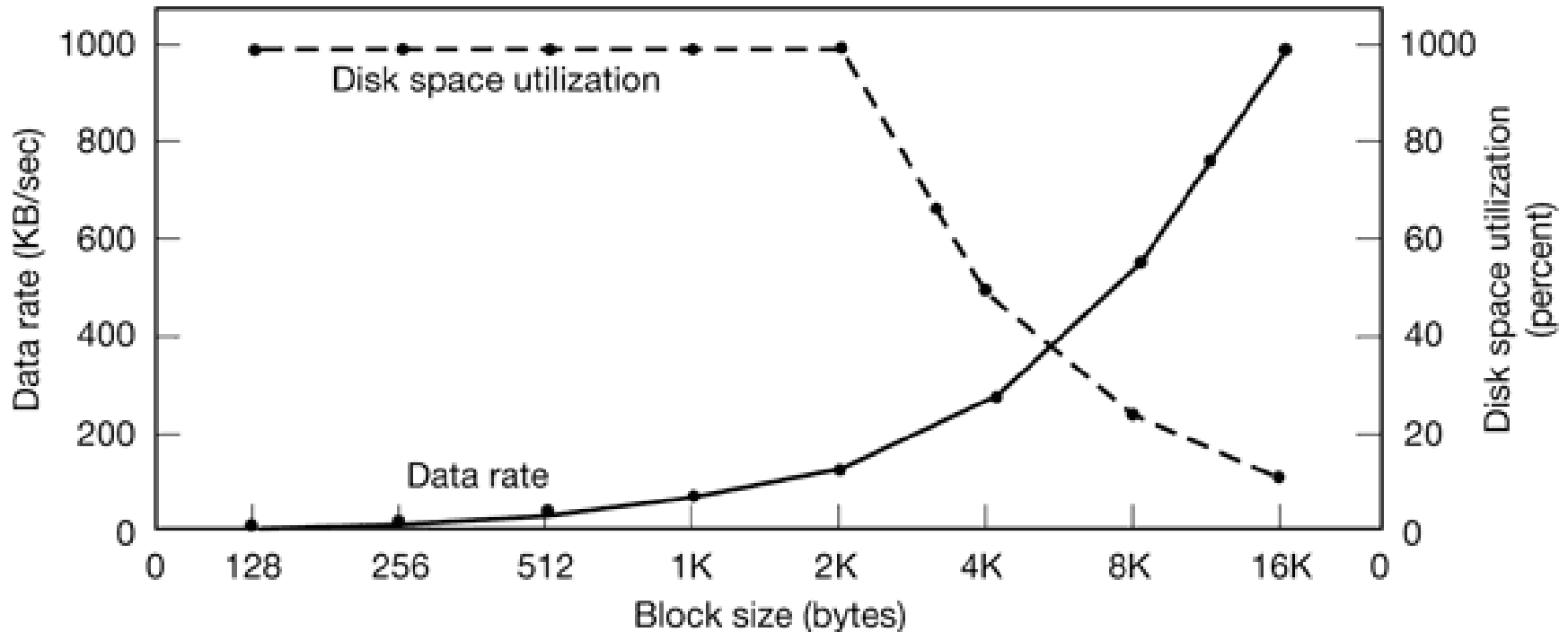
Блок великого розміру:

- більша пропускна здатність операцій вводу/виводу + менший час пошуку даних (seek time);
- більші втрати дискового простору за рахунок внутрішньої фрагментації.

Блок малого розміру:

- менші втрати дискового простору за рахунок зменшення внутрішньої фрагментації;
- більший час пошуку даних на диску (більше блоків → більший час пошуку) + менша пропускна здатність операцій вводу/виводу.

01.8. Проблема фрагментації дискового простору

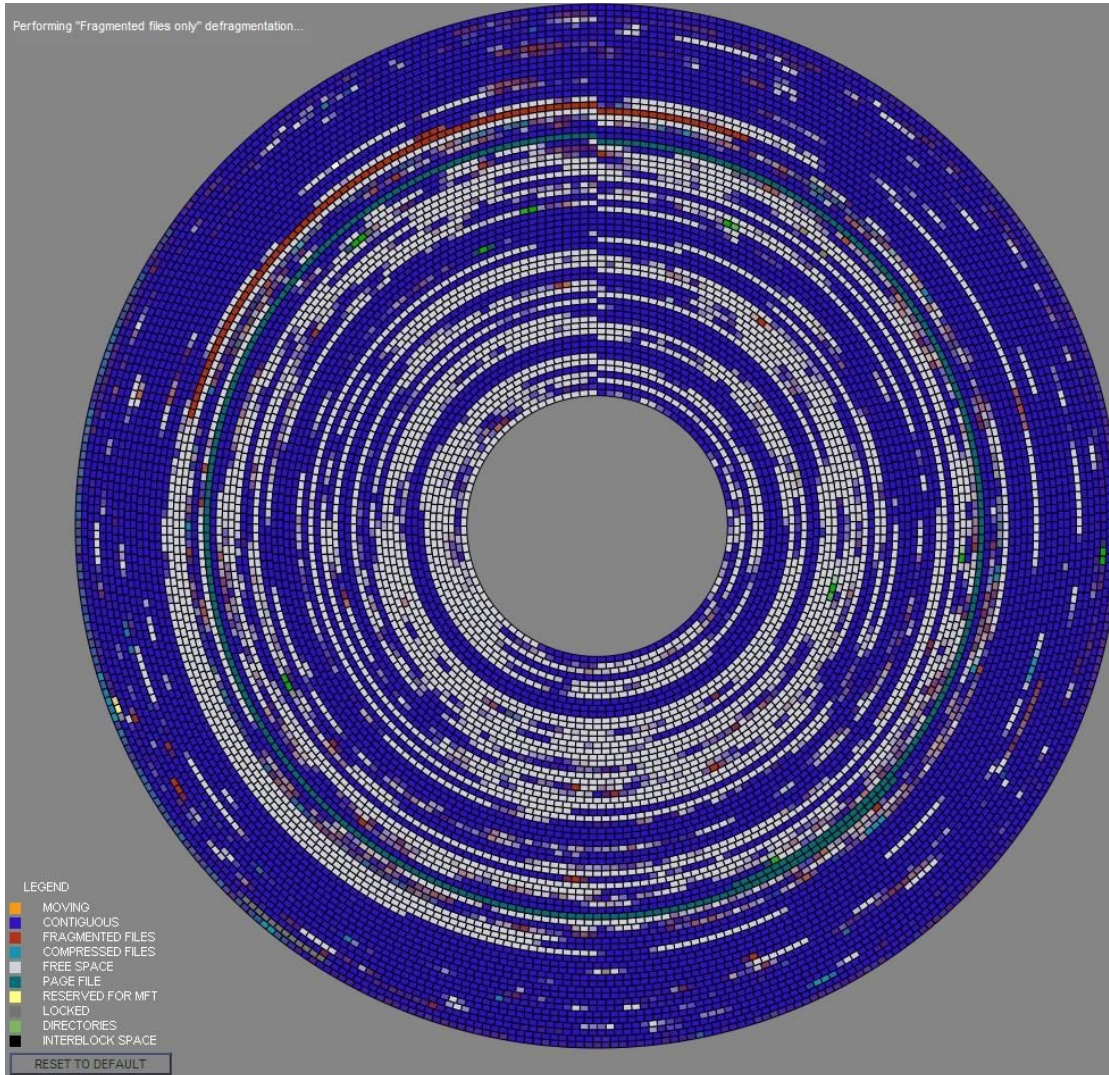


Приклад залежності пропускної здатності та втрат дискового простору від розміру блоку

02.1. Негативні наслідки фрагментації

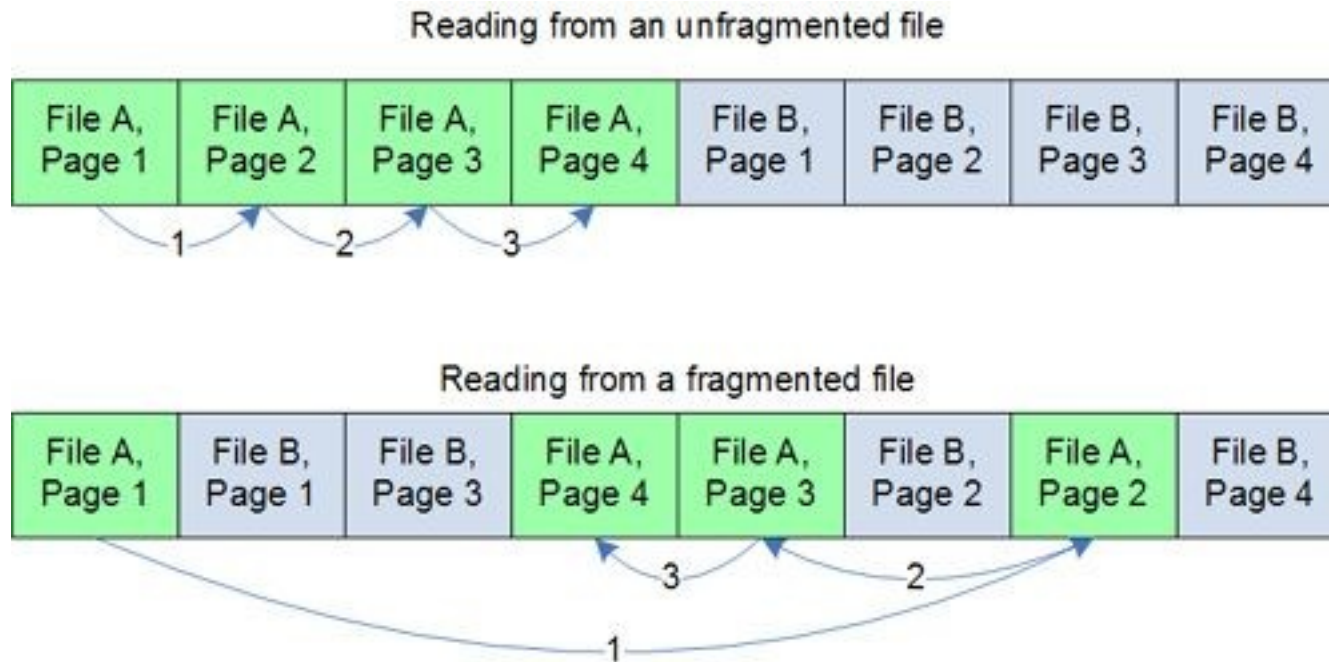
- 1) ускладнення аварійного відновлення даних → в разі сильної фрагментації відновлення даних різко ускладнюється + зростає ймовірність втрати даних;
- 2) зростання кількості фізичних операцій вводу/виводу при роботі з диском → збільшуються втрати енергії + виділяється більше тепла + механічне зношування диска;
- 3) зниження загальної швидкодії системи → наприклад, ускладнення роботи антивіруса, програм резервного копіювання і т.п.

02.2. Негативні наслідки фрагментації



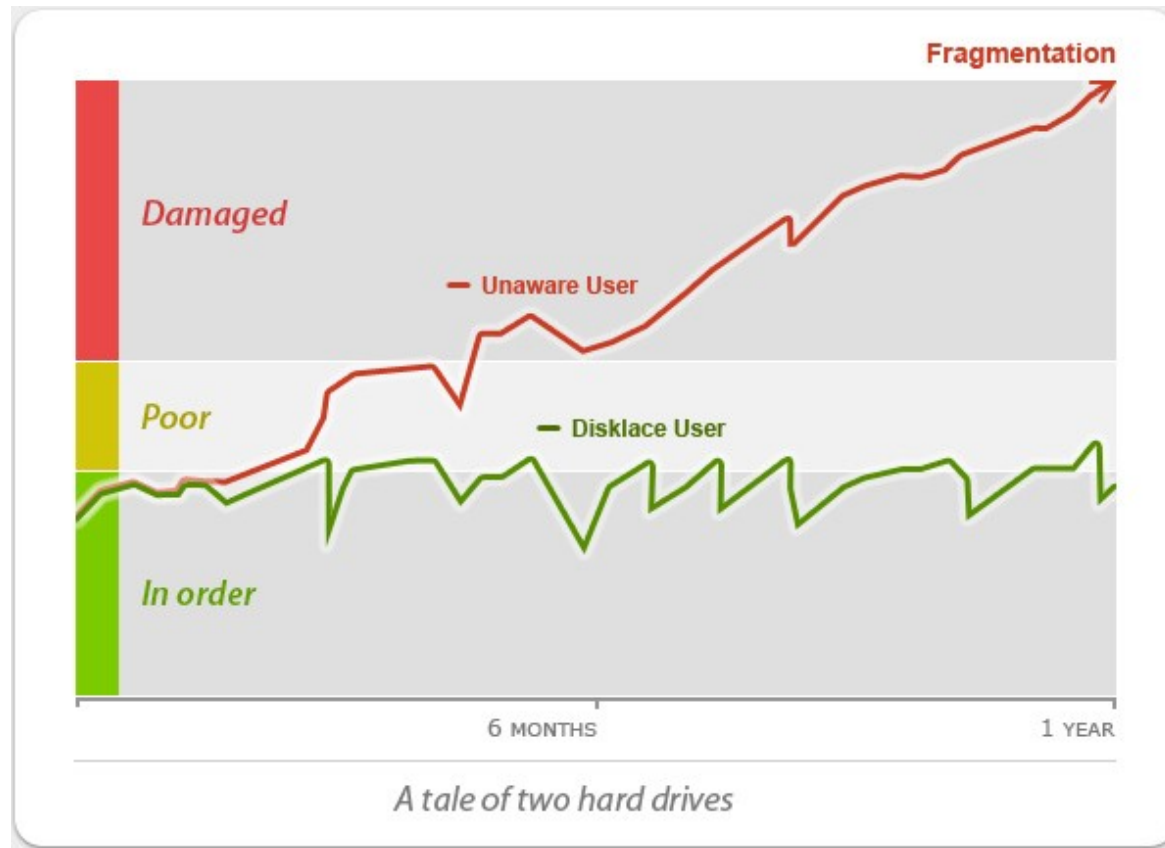
Приклад
фрагментованого
простору диску

02.3. Негативні наслідки фрагментації



Приклад збільшення кількості механічних операцій для читання фрагментованого файлу

02.4. Негативні наслідки фрагментації



Вплив фрагментації на «час життя» диску

03.1. Вирішення проблеми фрагментації

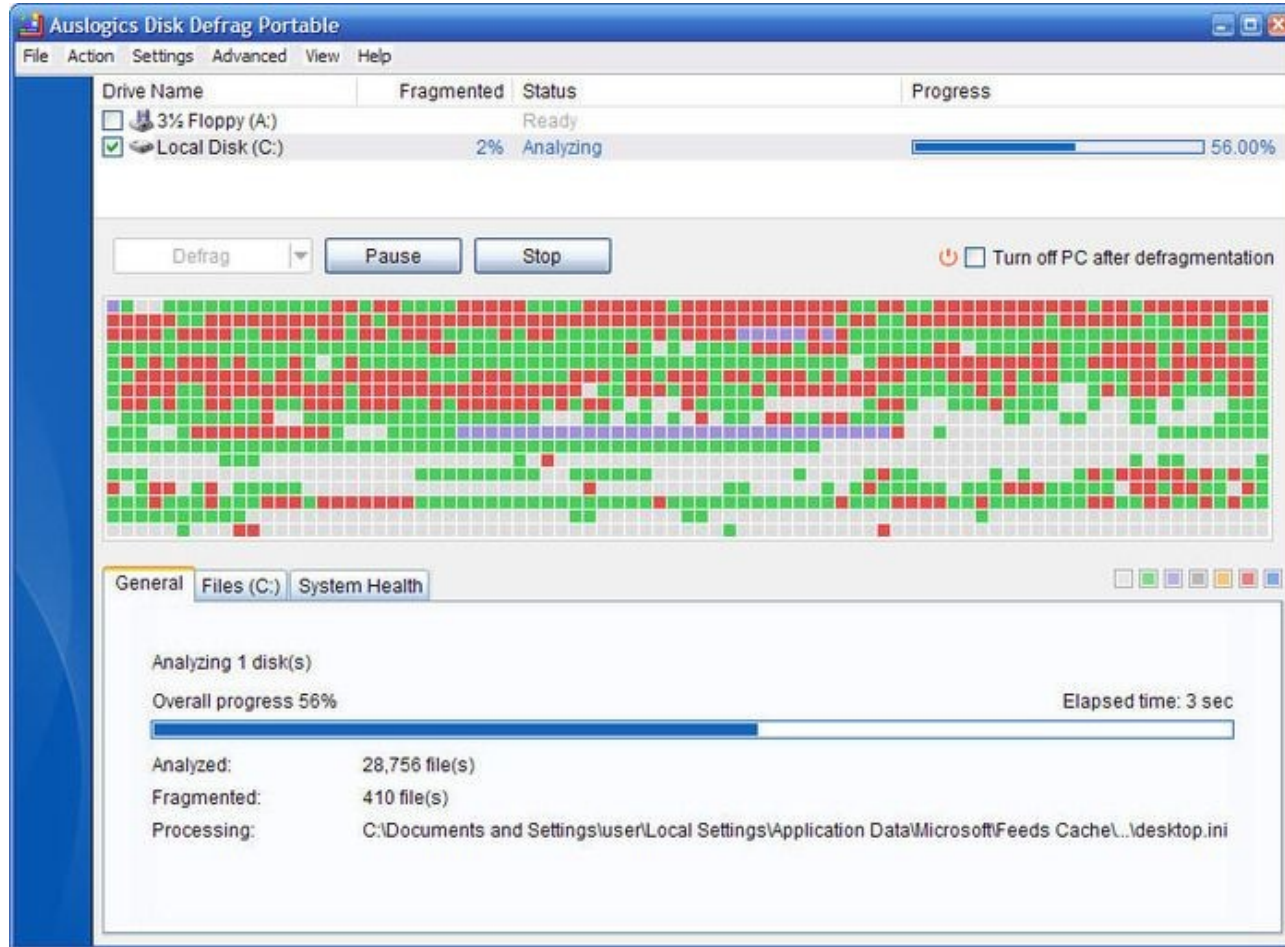
- Дефрагментація: процес оновлення і оптимізації логічної структури розділу диска.
- Програми (утиліти) дефрагментації → вбудовані в ОС + утиліти «третьої сторони».
- Загальна ідея методів дефрагментації: пошук та складання блоків файлів у вільному просторі диску → приклад моделі: узагальнена задача впорядкування набору об'єктів, в який вносяться зміни зовнішнім недетермінованим або частково недетермінованим процесом.

03.2. Вирішення проблеми фрагментації

Два основних етапи:

1. Аналіз файлу → пошук частин/блоків файлу, місця їх розташування на диску та вільного місця для переміщення цих блоків.
2. Здійснення дефрагментації → переміщення блоків файлу в такий спосіб, щоб вони утворювали неперервну послідовність у просторі диску.

03.3. Вирішення проблеми фрагментації



Приклад роботи утиліти дефрагментації

03.4. Вирішення проблеми фрагментації

Дві основні стратегії роботи утиліт/служб дефрагментації:

1. Періодична масова дефрагментація (FAT, NTFS).
2. Поточна дефрагментація → на етапі виконання операцій з файлом (ext2, ext3, ext4).

03.5. Вирішення проблеми фрагментації

Ручна дефрагментація:

- перезапис файлу з одного диску на інший;
- врахування часу «життя» файлу;
- принцип: ближче до початку розділу розміщуються ті файли, які не будуть знищуватись/змінюватись довгий час.

04.1. Методи дефрагментації

1. Проста (часткова) дефрагментація.
2. Повна дефрагментація (дефрагментація вільного простору).
3. Дефрагментація з врахуванням частоти використання/змін даних.
4. Дефрагментація з врахуванням дати останньої зміни файлу.
5. Дефрагментація з врахуванням атрибутів та імен файлів.

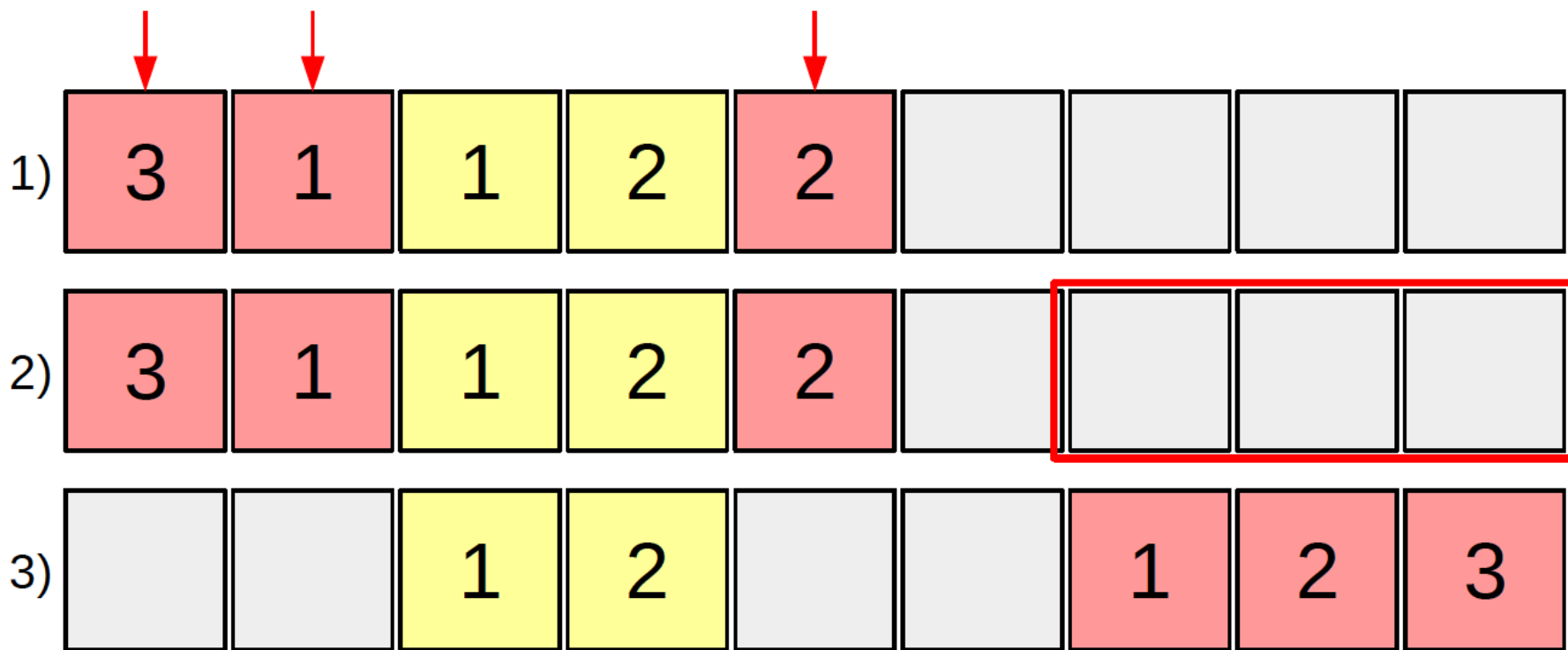
04.2. Методи дефрагментації

1. **Проста (часткова) дефрагментація** — це базовий метод, який використовується в інших методах дефрагментації.

Алгоритм:

- 1) цикл по частинах файлу → виявлення фрагментів;
- 2) пошук вільного місця → якщо не знайдено вільне місце, то відкласти даний файл і перейти до наступного файлу (п.1);
- 3) якщо вільне місце є, то розмістити у ньому «неперервно» блоки файлу.

04.3. Методи дефрагментації



Приклад виконання простої дефрагментації

04.4. Методи дефрагментації

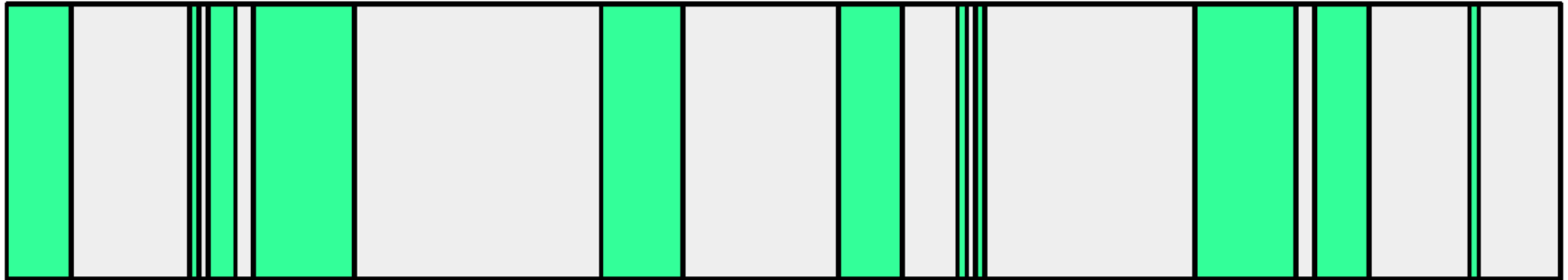
2. **Повна дефрагментація** (дефрагментація вільного простору) → «ущільнення» файлів на початку розділу диску (модель: задача упаковки).

Алгоритм:

- 1) цикл по всіх файлах → проста дефрагментація з тимчасовим записом файлів у наявному вільному місці (наприклад, в кінці розділу диску);
- 2) **розрахунок оптимального перерозміщення** дефрагментованих файлів;
- 3) розміщення дефрагментованих файлів з початку розділу диску згідно розрахунків.

04.5. Методи дефрагментації

1)



2)



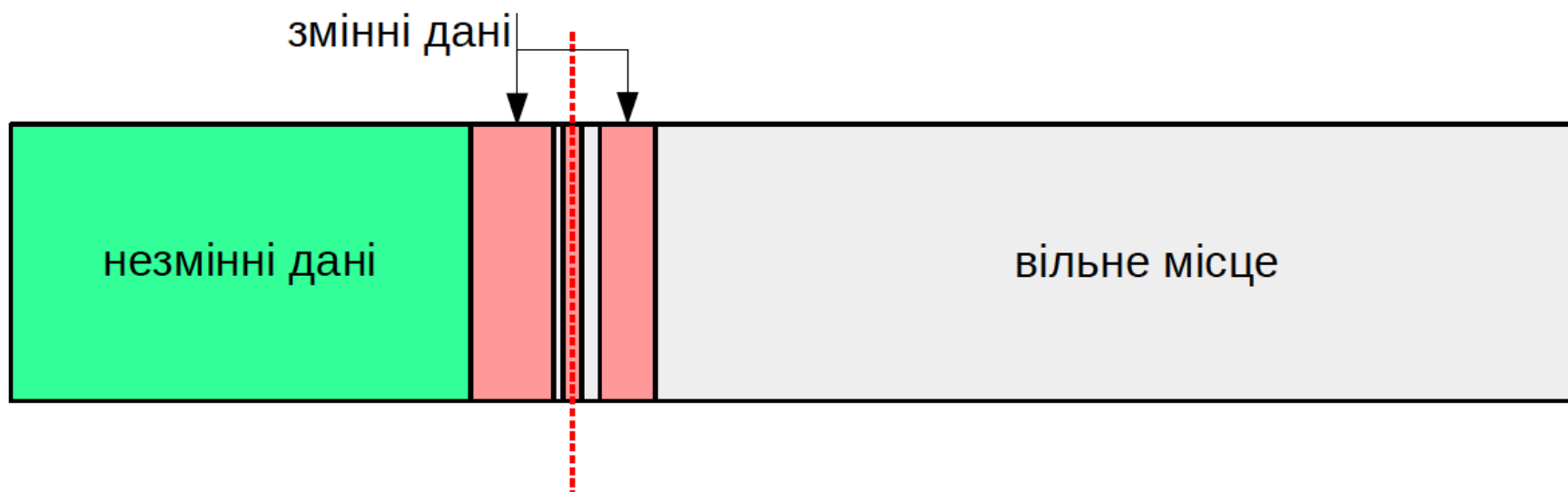
Приклад виконання повної дефрагментації

04.6. Методи дефрагментації

3. Дефрагментація з врахуванням частоти використання/змін даних.

- Евристика: чим більша частота використання файлу, тим більше ймовірність його змін.
- Для кожного файлу визначається частота його використання/змін.
- Чим більше частота, тим ближче до вільного місця розміщуються блоки файлу.
- Блоки файлів, які рідко змінюються, розміщуються ближче до початку розділу диску.

04.7. Методи дефрагментації



Приклад дефрагментації з врахуванням частоти
використання/змін даних

04.8. Методи дефрагментації

4. Дефрагментація з врахуванням дати останньої зміни файлу.

- Евристика: чим давніша дата зміни файлу, тим менше ймовірність його зміни.
- Файли, які не змінювалися тривалий час, розміщуються на початку розділу.
- Файли, які змінювались нещодавно, розміщуються в кінці розділу, що дозволяє, надалі, проводити їх просту дефрагментацію не зачіпаючи інших даних.
- Даний метод вигідно застосовувати для серверних систем, на яких зберігаються бази даних.

04.9. Методи дефрагментації

5. Дефрагментація з врахуванням атрибутів та імен файлів.

- Приклади атрибутів: ім'я файлу, тільки на читання, виконавчий файл і т.п.
- По імені файлу можна визначити його призначення: тимчасовий файл, файл системної бібліотеки і т.п.
- Основний принцип: покласти в основу розміщення блоків файлів на диску деяку логічну структуру, наприклад, розміщувати файли один за одним по алфавіту.

05.1. Переваги та недоліки методів дефрагментації

1. Проста (часткова) дефрагментація.

Переваги:

- 1) висока швидкість;
- 2) споживає мало обчислювальних ресурсів.

Недоліки:

- 1) не дефрагментує вільне місце;
- 2) при відсутності вільної ділянки потрібного розміру, не дефрагментує файл.

05.2. Переваги та недоліки методів дефрагментації

2. Повна дефрагментація.

Переваги:

- 1) різко підвищує продуктивність роботи системи з сильно фрагментованими дисками;
- 2) за рахунок дефрагментації вільного місця запобігає подальшій дефрагментації.

Недоліки:

- 1) займає тривалий час;
- 2) створює велике навантаження на диск;
- 3) не підходить для роботи у фоновому режимі.

05.3. Переваги та недоліки методів дефрагментації

3. Дефрагментація з врахуванням частоти використання/змін даних.

Переваги:

- 1) підвищує поточну продуктивність роботи системи (це важливо, наприклад, для серверів);
- 2) дозволяє об'єднати більшість вільного місця на диску.

Недоліки:

- 1) потребує багато вільного місця для початку роботи;
- 2) споживає додаткові обчислювальні ресурси (на визначення частоти використання/змін даних та збереження цієї інформації).

06.1. Згладжування показників зношення блоків у твердотільних накопичувачах

- Твердотільний накопичувач (SSD, solid-state drive) — це ПЗП на мікросхемах флеш-пам'яті, якими керує спеціалізований контролер.
- На даний час для побудови SSD найбільш часто використовують технологію **NAND flash** (Toshiba, 1989).
- Пристрій пам'яті на основі NAND flash складається з блоків (**blocks**), кожний з яких в свою чергу містить деяку кількість сторінок (**pages**).

06.2. Згладжування показників зношення блоків у твердотільних накопичувачах

- Для кожної сторінки додатково виділяється декілько байтів ($1/32$ від розміру сторінки) для збереження контрольних сум (корекція помилок).
- Типові розміри сторінок: 512; 2.048; 4.096 bytes.
- Типові розміри блоків:
16KiB (32x512); 128KiB (64x2k);
256KiB (64x4k); 512KiB (128x4k).

06.3. Згладжування показників зношення блоків у твердотільних накопичувачах

- У пристроях пам'яті на основі NAND flash читання та запис нових даних відбувається **по-сторінково**, а стирання даних (та перезапис старих даних) можна здійснити лише **по-блоково**.
- Переважна більшість сучасних пристроїв пам'яті на основі NAND flash поступає у продаж з деякою незначною кількістю пошкоджених блоків (**bad blocks**), що є основним фактором здешевлення цих пристроїв.

06.4. Згладжування показників зношення блоків у твердотільних накопичувачах

Основні завдання контролера:

- 1) корекція помилок читання/запису;
- 2) відслідковування (mapping) пошкоджених блоків (bad blocks);
- 3) згладжування показників зношення (wear leveling);
- 4) кешування даних;
- 5) виконання операцій перезапису старих даних з врахуванням позначок їх «актуальності» (Garbage collection).

06.5. Згладжування показників зношення блоків у твердотільних накопичувачах

Переваги SSD в порівнянні з HDD:

- 1) відсутність рухомих частин (більша стійкість до механічних пошкоджень, менша вага, відсутність шуму);
- 2) менше споживання енергії (приблизно вдвічі) і відтак менше виділення тепла (26-30°C vs. 33-44°C);
- 3) більша швидкість читання/запису даних (bandwidth: 550MB/s vs. 200MB/s) та менший час відгуку (response time: 0,05ms vs. 6ms).

06.6. Згладжування показників зношення блоків у твердотільних накопичувачах

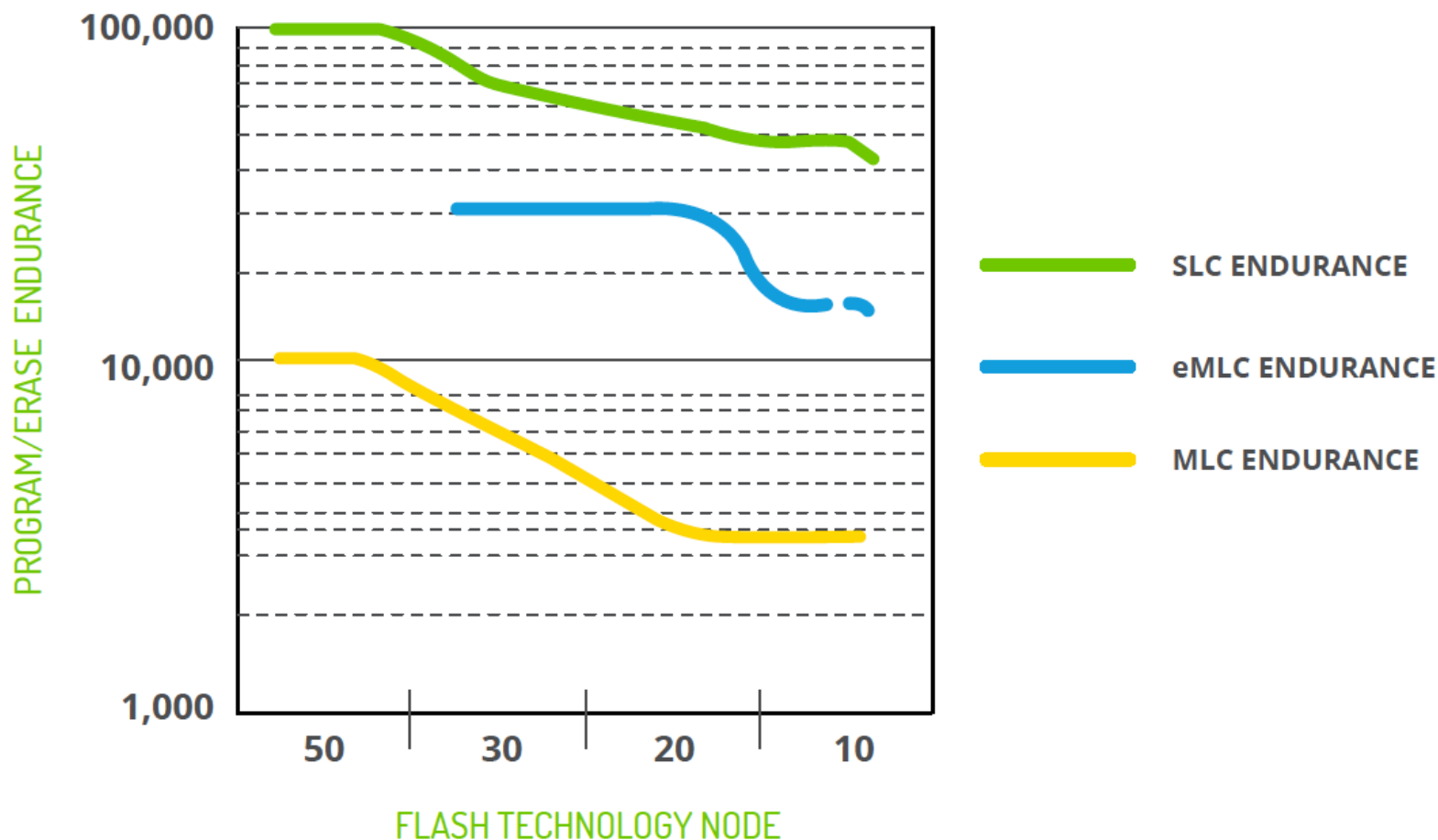
Недоліки SSD в порівнянні з HDD:

- 1) більша вартість (що найменше вдвічі більша);
- 2) відносно менший об'єм пам'яті при тих самих фізичних розмірах пристрою (цей недолік поступово зникає з розвитком технології);
- 3) наявність ефекту зношення (менший «час життя»).

06.7. Згладжування показників зношення блоків у твердотільних накопичувачах

- **Ефект зношення (wear)**: кількість операцій запису (стирання-запису) комірок флеш-пам'яті обмежена внаслідок незворотних змін у структурі р-n переходів.
- Типові показники зношення лежать в межах від 1000 до 100.000 операцій запису для різних типів флеш-пам'яті.
- Типи флеш пам'яті: Single-level cell (SLC) flash, multi-level cell (MLC) flash та ін.

06.8. Згладжування показників зношення блоків у твердотільних накопичувачах

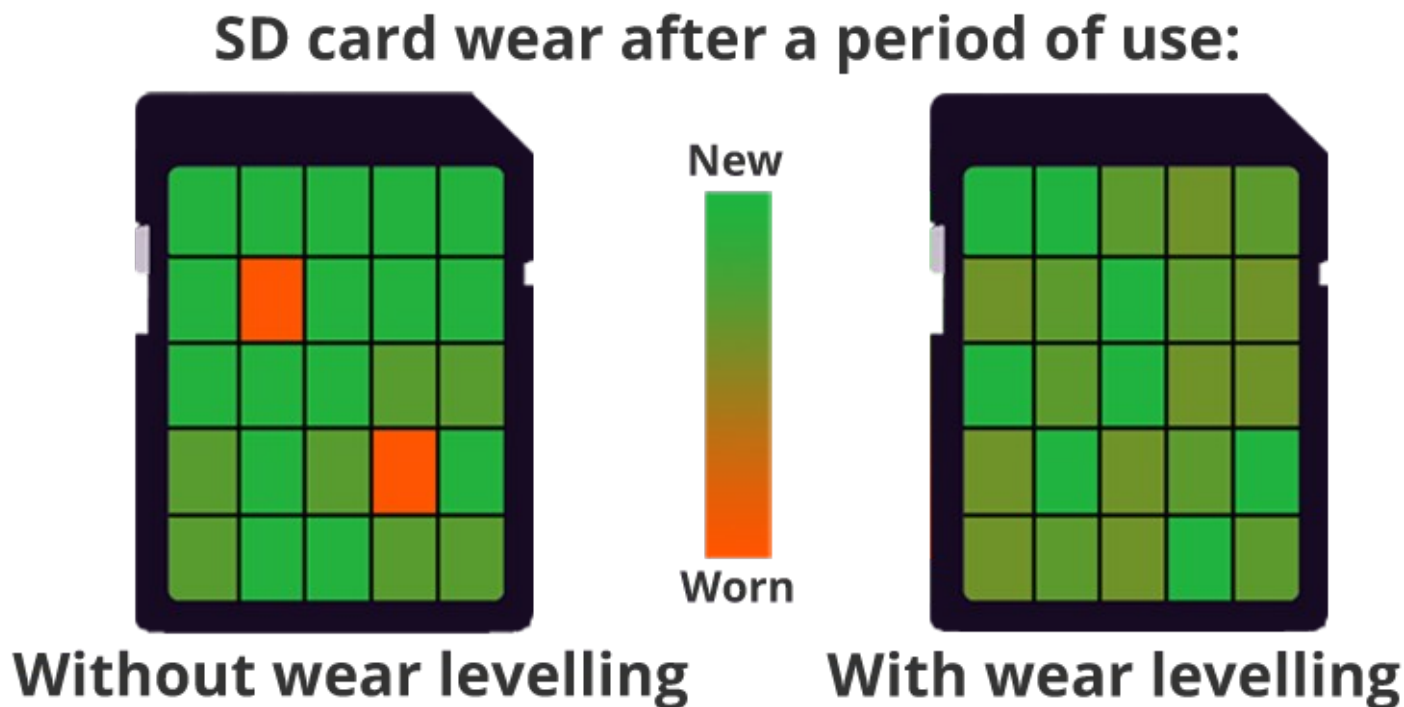


Типові показники максимальної кількості циклів запису для різних типів флеш-пам'яті в залежності від тех.процесу (nm)

06.9. Згладжування показників зношення блоків у твердотільних накопичувачах

- «Витривалість» на запис (**write endurance**) визначається у вигляді рейтингу «витривалості» (**endurance rating**) і дорівнює максимальній кількості стирань на блок, яку гарантує виробник.
- За рахунок використання алгоритмів згладжування показників зношення (wear leveling) у сучасних SSD вдається значно зменшити наслідки ефекту зношення.
- Типові значення write endurance rating для сучасних SSD складають від 1 до 5 млн.

06.10. Згладжування показників зношення блоків у твердотільних накопичувачах



Результат роботи алгоритму згладжування показників зношення (wear leveling)

06.11. Згладжування показників зношення блоків у твердотільних накопичувачах

Мета: забезпечити **рівномірний кількісний розподіл** операцій запису між усіма блоками флеш-пам'яті.

Результат:

- 1) рейтинг «витривалості» (endurance rating) ПЗП на основі флеш-пам'яті збільшується пропорційно до кількості блоків;
- 2) виключається передчасне зношення і відмова окремих блоків (тобто приблизно **усі** блоки використовують **весь** свій ресурс циклів запису).

06.12. Згладжування показників зношення блоків у твердотільних накопичувачах



Задача згладжування показників зношення (wear leveling)

06.13. Згладжування показників зношення блоків у твердотільних накопичувачах

- LBA (Logical Block Address) → адреса блоку даних, яку використовує ОС в операціях читання/запису з ПЗП на основі флеш-пам'яті.
- PBA (Physical Block Address) → фіксована фізична адреса блоку даних у ПЗП на основі флеш-пам'яті.
- Відповідність LBA та PBA підтримується за допомогою таблиці (lookup table).

06.14. Згладжування показників зношення блоків у твердотільних накопичувачах

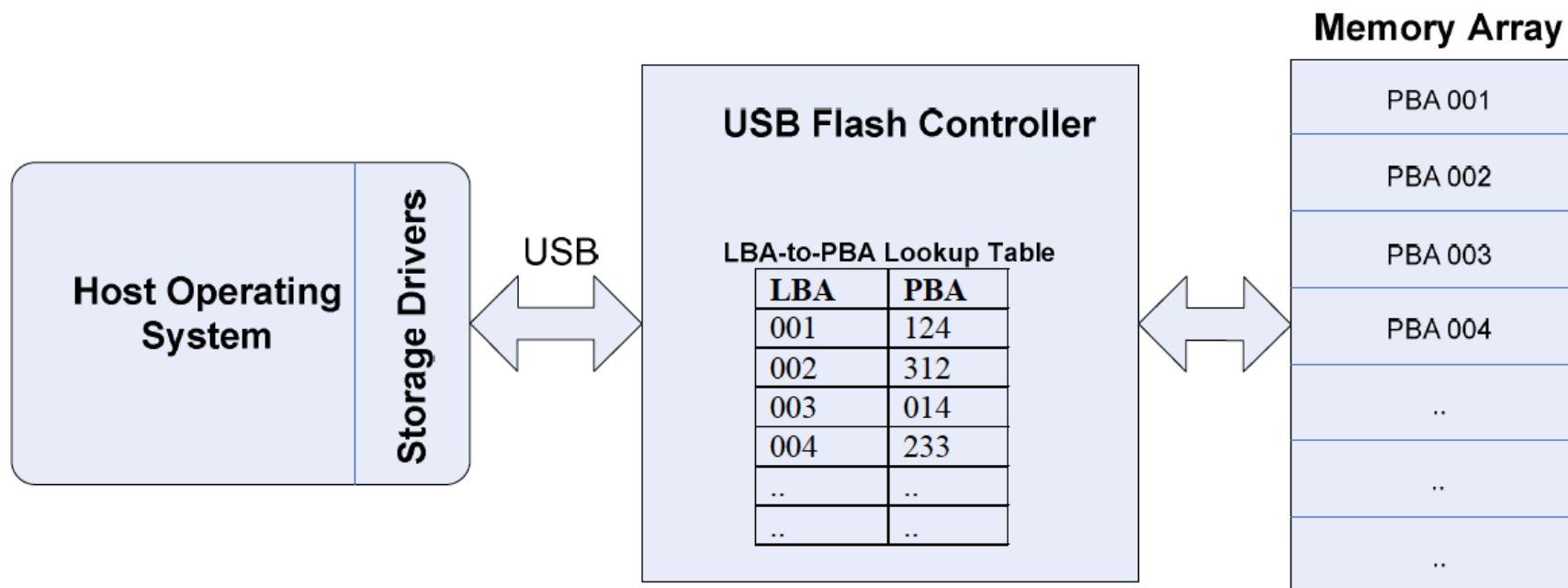
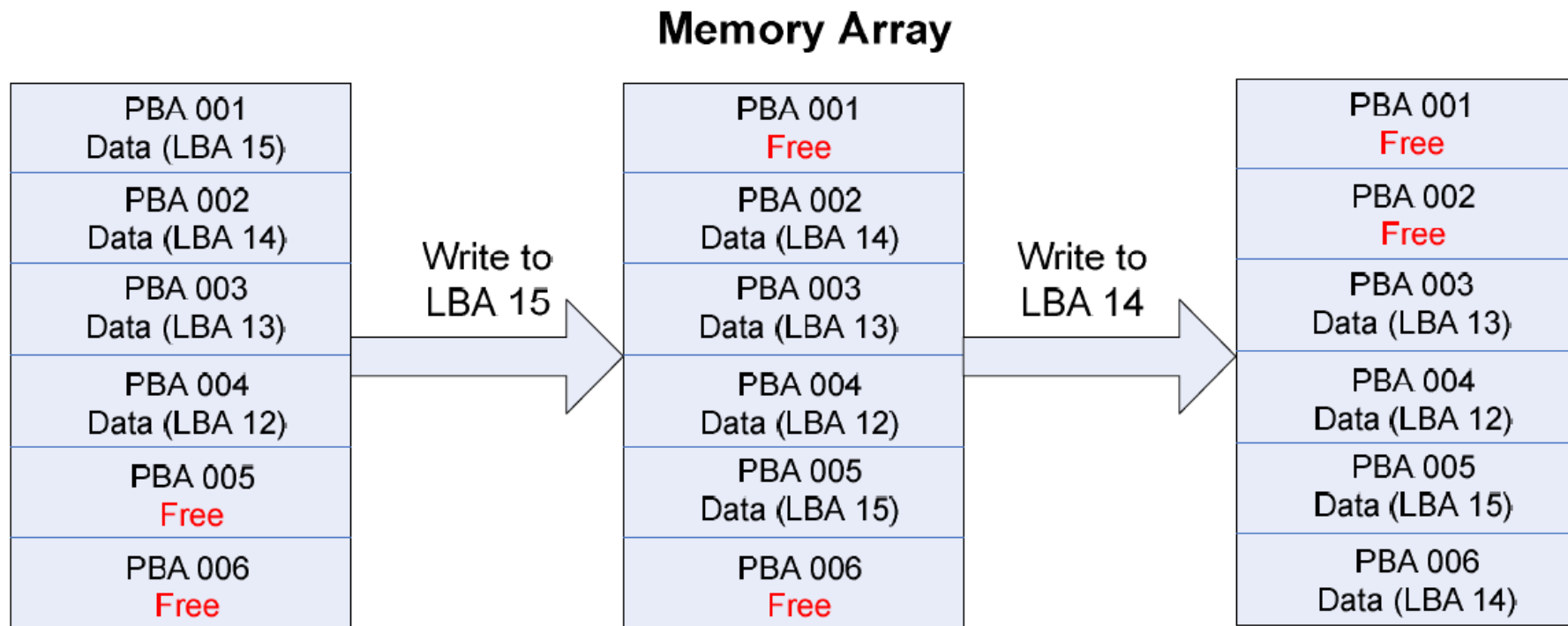


Схема використання lookup table

06.15. Згладжування показників зношення блоків у твердотільних накопичувачах

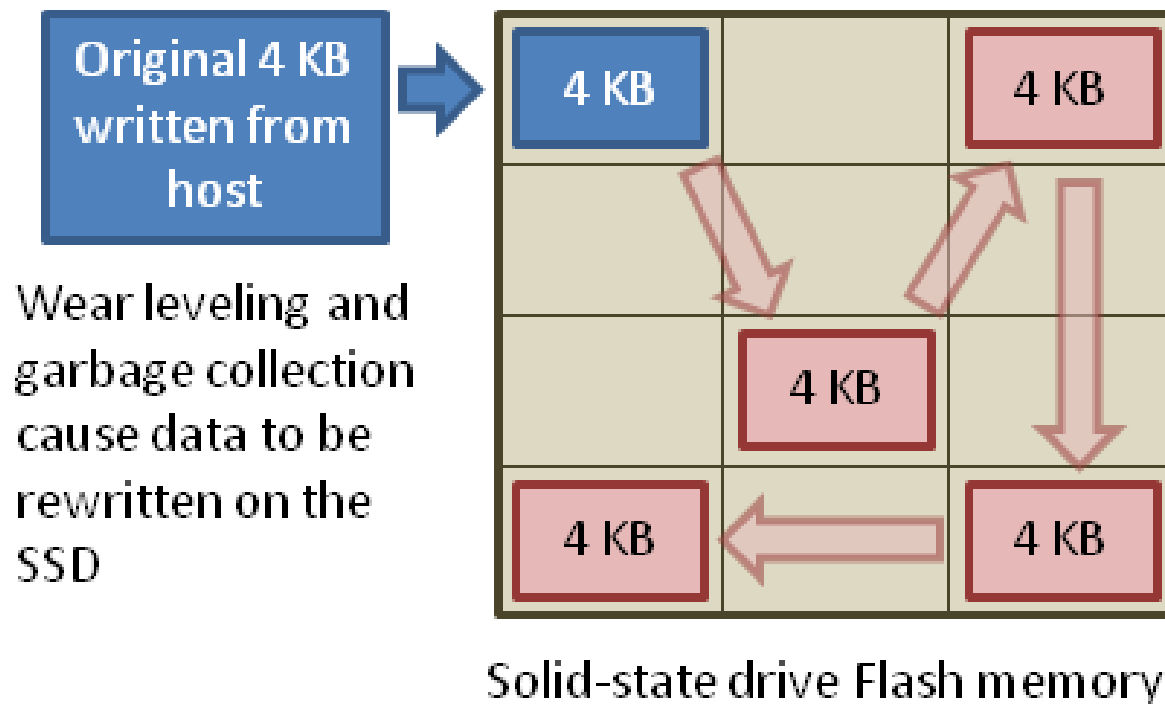
- Для кожного РВ зберігається значення кількості здійснених операцій запису $N(b)$.
- Принцип роботи: при надходженні нових даних, які потрібно [пере]записати (**hot data**), для них обирається РВ з найменшим значенням $N(b)$.
- Якщо в цьому РВ зберігаються потрібні дані (**cold data**), то вони в свою чергу перезаписуються у інший РВ за тим самим принципом.

06.16. Згладжування показників зношення блоків у твердотільних накопичувачах



Приклад переміщення даних між PBs

06.17. Згладжування показників зношення блоків у твердотільних накопичувачах



Явище надлишкового запису (write amplification)

06.18. Згладжування показників зношення блоків у твердотільних накопичувачах

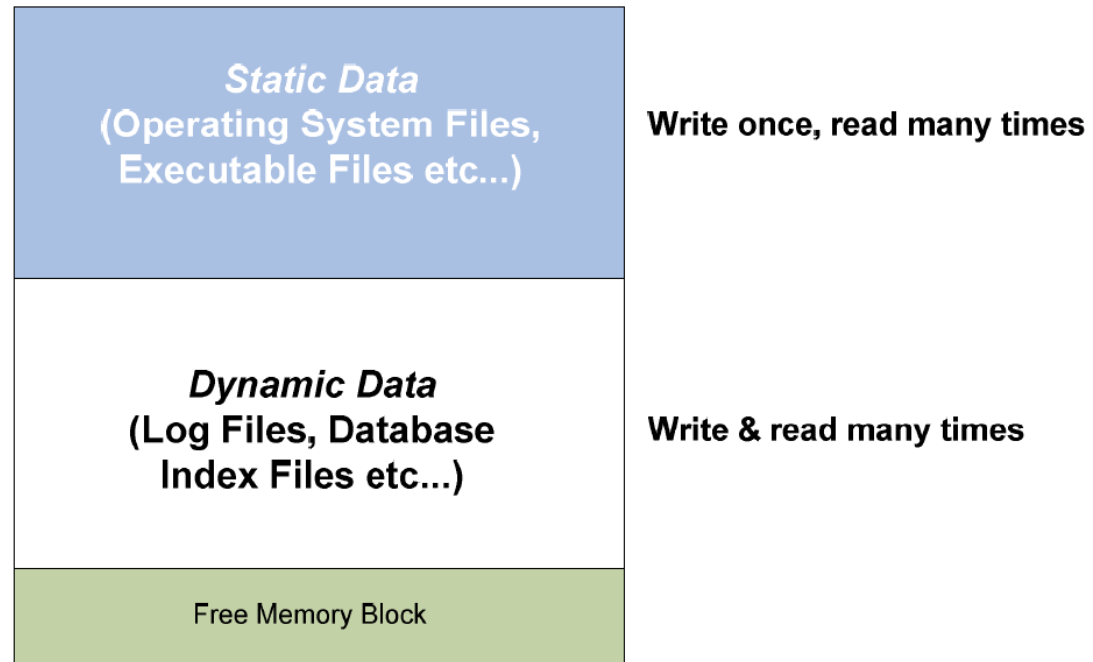
- Явище надлишкового запису (**write amplification**): кількість фізичних операцій запису у твердотільному ПЗП перевищує кількість логічних операцій запису даних.
- Ще одною причиною надлишкового запису є «асиметрія» операцій читання/запису нових даних (по-сторінкове) та перезапису старих даних + стирання (по-блокове), яка також призводить до додаткових операцій перерозміщення блоків (Garbage collection).

06.19. Згладжування показників зношення блоків у твердотільних накопичувачах

- При цьому надлишкові операції по перезапису блоків з потрібними даними (cold data) в інші блоки в свою чергу збільшують зношення (=«розплата» за забезпечення рівномірності).
- Відтак в алгоритмах згладжування показників зношення (wear leveling) потрібно знаходити **оптимальне співвідношення** між перерозміщенням даних, які змінюються (hot data), та даними, які не змінюються (cold data).

06.20. Згладжування показників зношення блоків у твердотільних накопичувачах

Two types of data stored in
Flash memory



Підхід до побудови алгоритмів згладжування: hot data апроксимуються як динамічні дані (**dynamic data**), а cold data апроксимуються як статичні дані (**static data**).

06.21. Згладжування показників зношення блоків у твердотільних накопичувачах

Стратегії роботи алгоритмів згладжування показників зношення (wear leveling):

1) Динамічне згладжування (**dynamic wear leveling**): перерозміщення блоків відбувається тільки для даних, які позначені динамічними, та вільних блоків. Мета: зменшити обчислювальні витрати та час роботи алгоритма.

2) Статичне згладжування (**static wear leveling**): перерозміщення блоків відбувається для всіх даних (в тому числі статичних) та вільних блоків. Мета: забезпечити максимально рівномірне згладжування.

06.22. Згладжування показників зношення блоків у твердотільних накопичувачах

Параметр	Статичне згладжування	Динамічне згладжування
«Витривалість» (endurance)	Більший очікуваний «час життя»	Менший очікуваний «час життя»
Швидкість роботи	Менша	Більша
Складність реалізації	Більша	Менша
Типове використання	SSD	USB Flash Drive

Порівняння статичного та динамічного згладжування