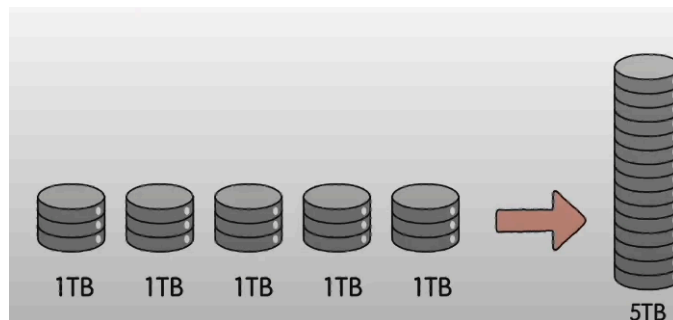


19강. RAID의 정의와 종류

RAID의 정의

RAID (Redundant Array of Independent / Inexpensive Disks)

RAID는 여러 개의 하드 디스크나 SSD를 묶어 마치 하나의 논리적 보조기억장치처럼 사용하는 기술입니다.

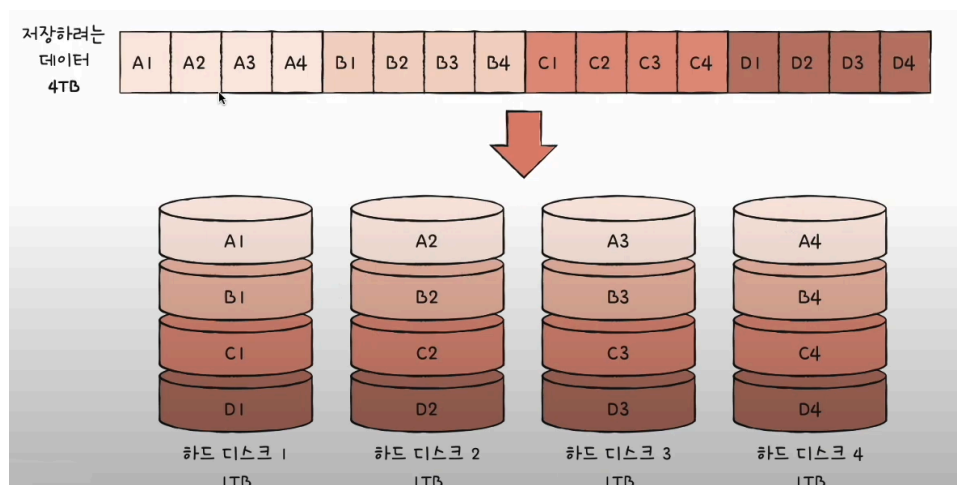


이렇게 묶는 이유는 데이터를 더 안전하게 보관하거나 입출력 속도를 높이기 위함입니다.

RAID는 여러 방식으로 구성할 수 있으며, 이를 각각 RAID 레벨이라고 부릅니다.

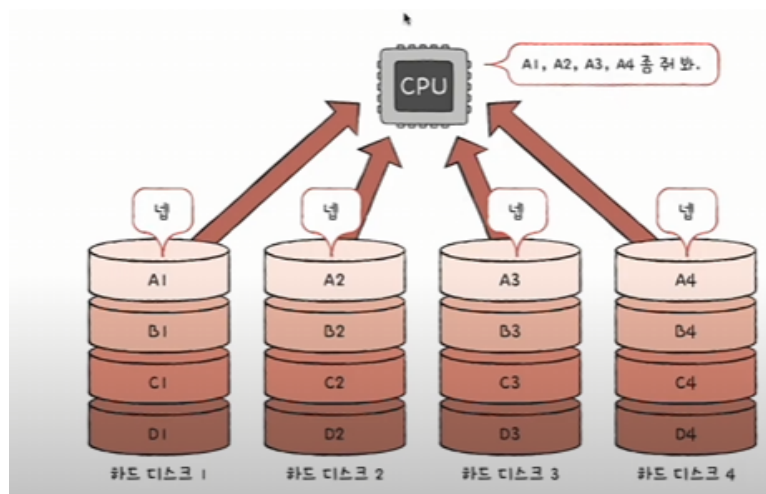
가장 기본적인 형태는 RAID 0부터 RAID 6까지이며, 이외에도 10, 50 처럼 응용된 형태들도 존재합니다. 어떤 RAID 레벨을 선택하느냐에 따라 성능과 안전성 그리고 필요한 디스크 수가 달라집니다.

RAID 0

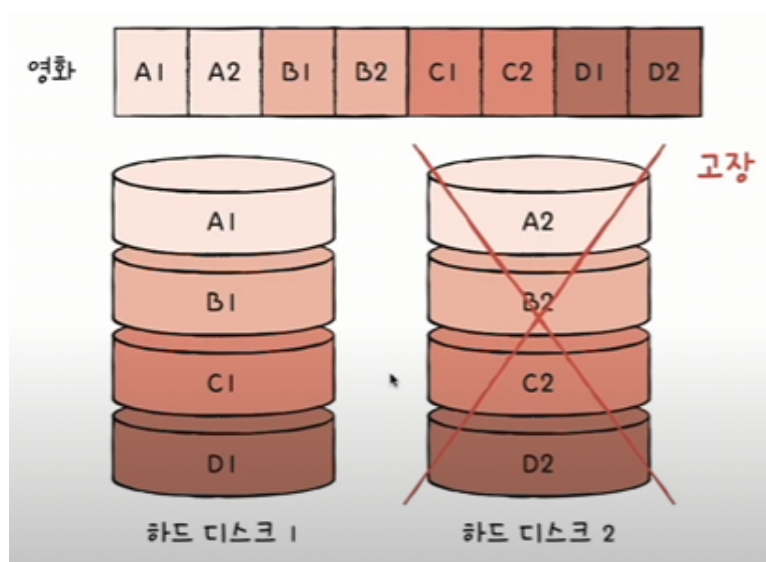


먼저 RAID 0은 데이터를 단순히 나누어 여러 디스크에 분산 저장하는 방식입니다. 데이터를 스트라이프(stripe) 단위로 나누어 디스크에 번갈아 기록하는데, 이를 스트라이핑(striping)이라고 합니다.

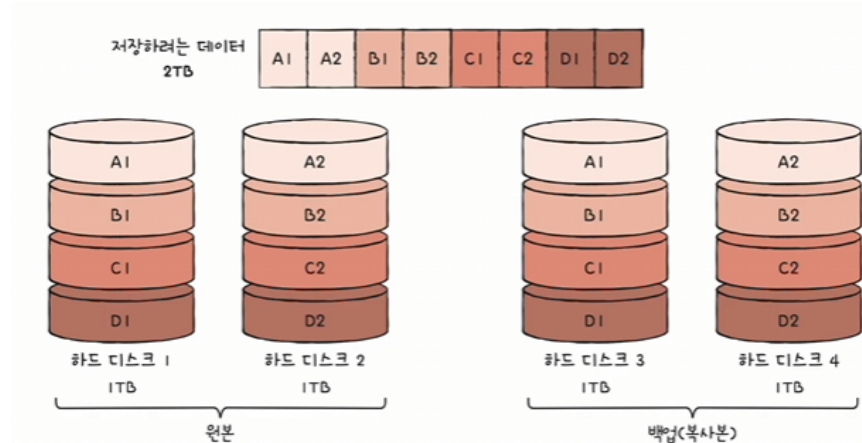
RAID 0의 가장 큰 장점은 입출력 속도가 크게 향상된다는 점입니다. 여러 디스크가 동시에 데이터를 읽고 쓰기 때문에 성능이 높아집니다.



단점으로는 데이터의 안전성이 전혀 보장되지 않는다는 점입니다. 어느 한 디스크라도 고장이 나면 전체 데이터가 손실됩니다.



RAID 1

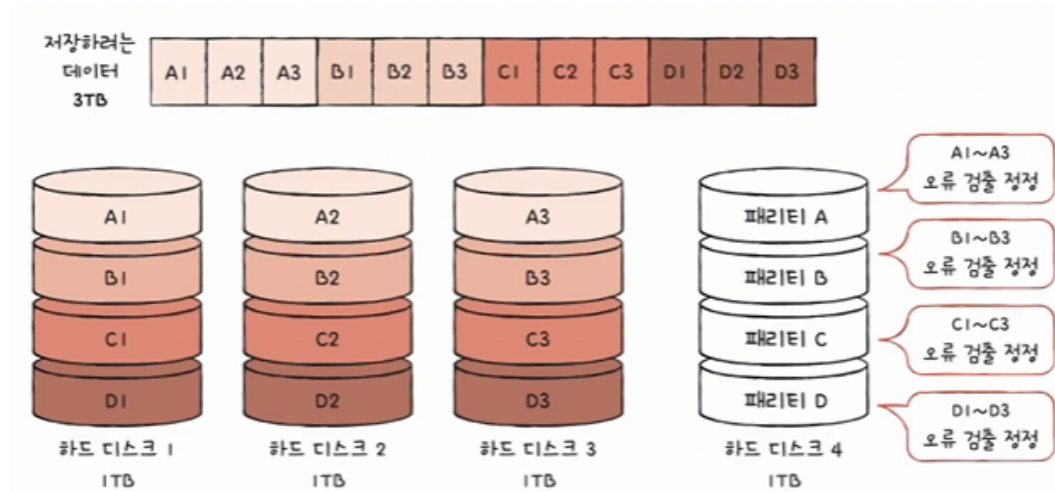


RAID 1은 미러링(mirroring) 방식으로 동작합니다. 데이터를 저장할 때 원본과 동일한 복사본을 동시에 다른 디스크에 기록합니다. 이 방식은 안전성이 매우 높아 한쪽 디스크가 손상 되더라도 다른 디스크에서 데이터를 복구할 수 있습니다.

- 미러링(mirroring) : 복사본을 만드는 방식

하지만 저장 효율이 낮아 실제로 사용할 수 있는 용량은 전체 디스크 용량의 절반에 불과합니다. 또한 쓰기 속도는 두 군데에 동시에 기록해야 하기 때문에 느려질 수 있으며, 많은 디스크가 필요해 비용이 높아지는 단점이 있습니다.

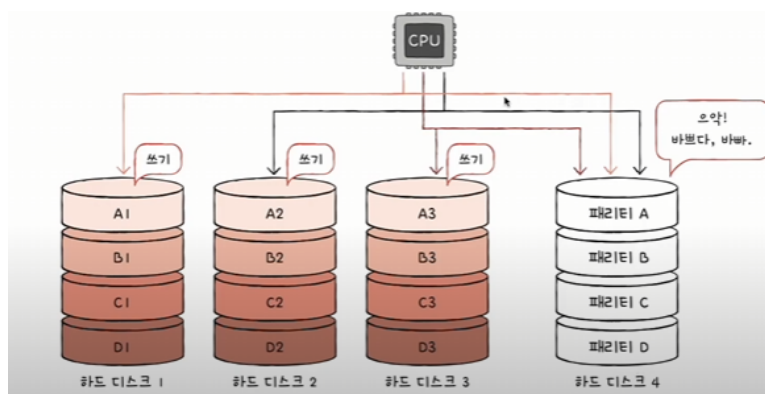
RAID 4



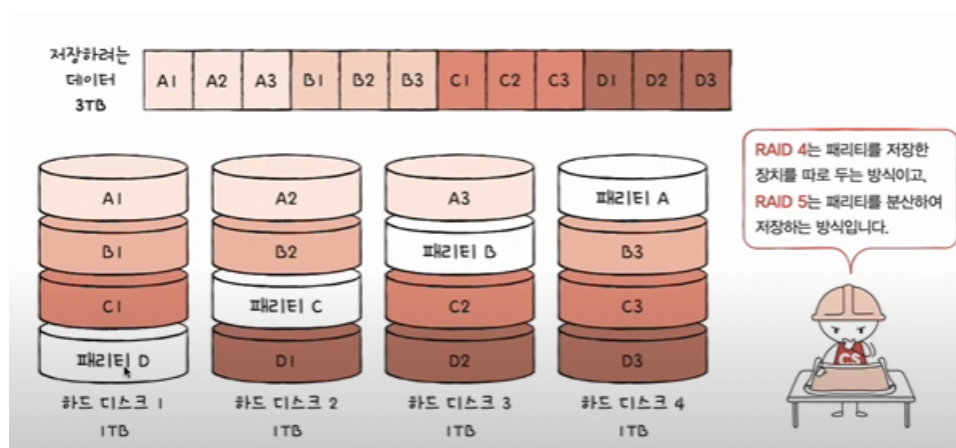
RAID 4는 RAID 1처럼 모든 데이터를 복사하지 않고, 대신 패리티(parity) 정보를 별도의 디스크에 저장하는 방식입니다. 패리티는 오류 검출을 위한 추가 데이터로, 이를 이용하면 특정 디스크에서 오류가 발생했을 때 문제를 확인할 수 있습니다.

RAID 4는 RAID 1보다 디스크 효율이 높지만, 모든 패리티 정보가 하나의 디스크에 집중되기 때문에 이 디스크가 병목 지점이 되어 성능 저하가 발생할 수 있습니다. 또한 패리티는 오

류 검출에는 유용하지만 복구 기능은 제한적입니다.



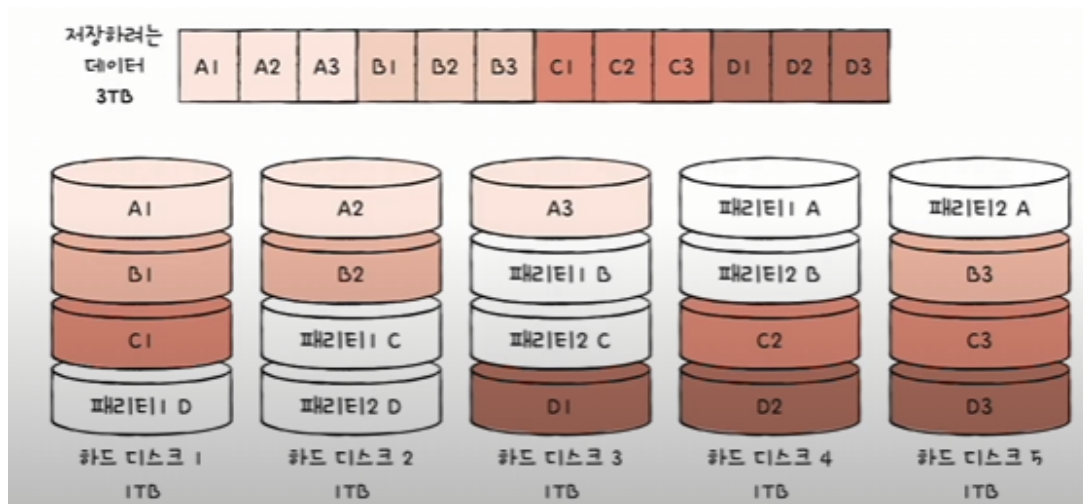
RAID 5



RAID 5는 RAID 4의 문제를 개선한 방식으로, 패리티 정보를 한 디스크에 몰아 저장하지 않고 여러 디스크에 분산하여 기록합니다. 덕분에 RAID 4에서 발생하던 병목 현상이 줄어들며, 하나의 디스크가 고장 나더라도 패리티 정보를 통해 데이터를 복구할 수 있습니다.

RAID 5는 성능과 안전성의 균형이 좋아 실무에서 널리 사용됩니다.

RAID 6



RAID 6은 RAID 5의 확장판으로, 패리티를 두 종류 저장합니다. 이중 패리티 덕분에 동시에 두 개의 디스크가 손상되더라도 데이터를 복구할 수 있어 안전성이 더 높습니다.

하지만 패리티를 두 배로 계산하고 기록해야 하기 때문에 쓰기 성능은 RAID 5보다 떨어집니다.

RAID 정리

이처럼 RAID는 레벨마다 각각의 장단점이 있습니다.

어떤 상황에서는 성능이 우선일 수도 있고, 어떤 상황에서는 데이터 안전성이 더 중요할 수도 있습니다.

따라서 RAID를 설계할 때는 “무엇을 최우선으로 할 것인가”를 기준으로 최적의 레벨을 선택하는 것이 중요합니다.