

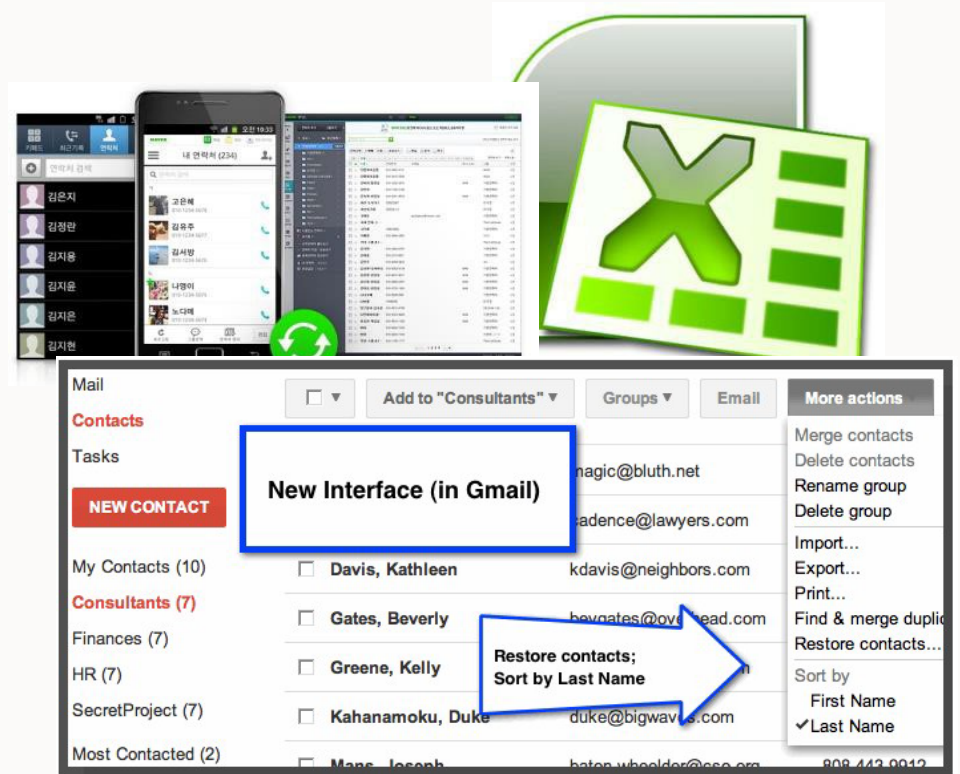
관계형 데이터베이스 개념

경북산업직업전문학교 [1-SEP-2022]
- 무단배포를 금지합니다 -

여딴찰 곳곳 앤뒸겔 핍습깁왓?



가공
및
정리



Data & Information

Data

현실세계에서 관찰이나 측정을 통해
수집한 자료나 값으로,
가공되지 않은 상태를 의미

ex. 명함, 메모, 전화번호 등

재잡, 건상, —곰

ex. 족센렘

HOW?

크룟앤팁야? 뒣숨틀?

억세? 숨립틀팻?

Information

의사결정에 도움을 줄 수 있는
유용한 형태로,
자료를 가공(처리)해서 얻을 수
있는 결과를 의미

ex. 전략고객 리스트 및 전화번호,
청첩장 발송 주소/메일 리스트 등

관리할 Data가 너무 많다면?



많은 Data를 어디다 보관하지?

이 Data가 내가 갖고 있는게 맞나? 이

미지나 문서파일은 어떻게 관리하지?

중요한 자료인데 없어지면 어떡하지?

내가 그 Data를 어디다 뒀더라? 어디서부터
찾아야 되나...

추가항목을 관리하고 싶은데..언제 처음부터
다시 정리하나.....

통계를 내고 싶은데.....

Database의 등장

Concept

다수 사용자의 정보 욕구를 충족시키기 위하여(共用), 서로 관계 있는 데이터를 최소한의 중복으로 統攝해 놓은 데이터의 집합체

(ex. 직원정보/상품목록/고객정보/은행거래내역 등)

Purpose

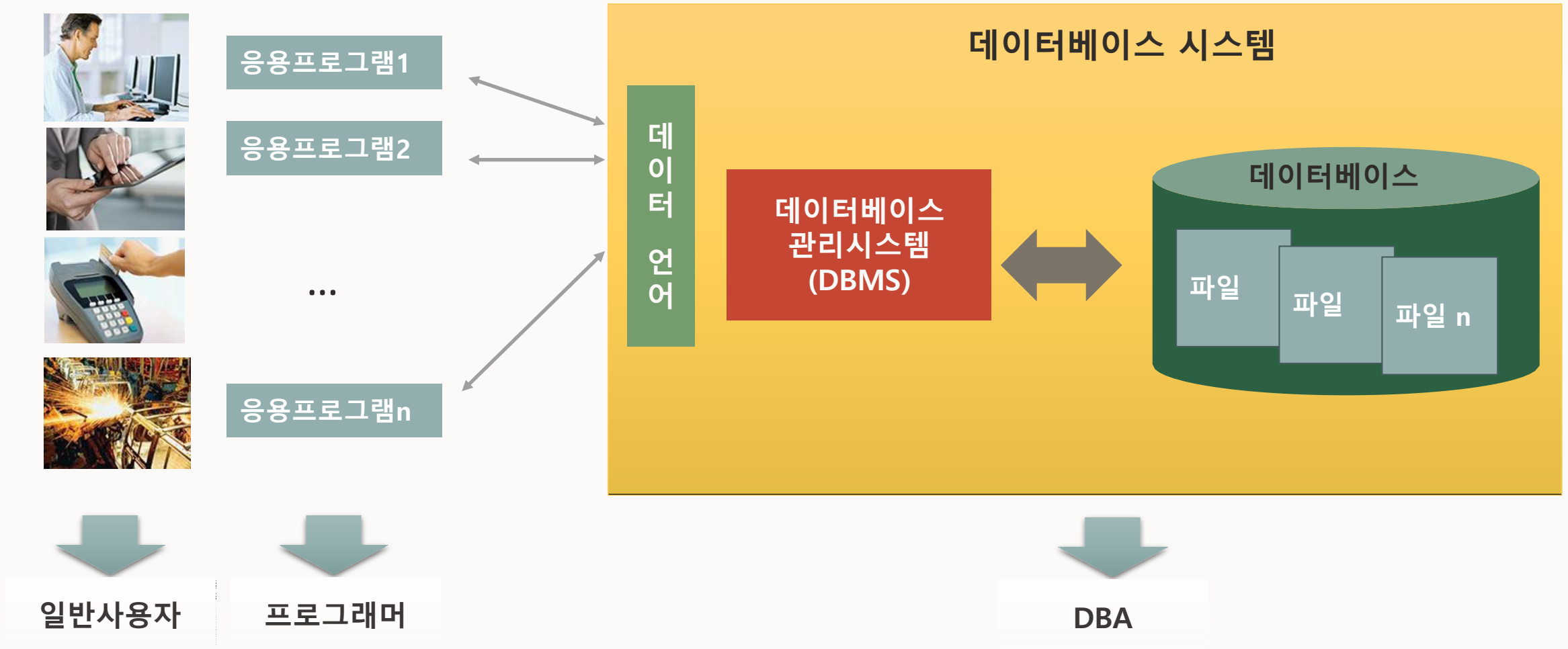
- ✓데이터의 중복 및 비일관성 제거
- ✓접근성(Accessibility)
- ✓체계적 보안과 사용자의 실수를 방지할 수 있는 무결성의 제공
- ✓다양한 데이터 형태 수용/통합

Database **뱃짚뱃**

구분	모델	특징
1세대	파일시스템	파일 시스템은 각각의 파일 단위로 업무와 관련한 데이터를 저장 (SAM 파일)
2세대	계층형 (Hierarchical DBMS)	계층형 데이터베이스 관리시스템은 계층적으로 데이터가 상하 종속적인 관계로 구성
3세대	망형 (Network DBMS)	데이터 구조를 네트워크상의 노드 형태로 논리적으로 표현한 데이터 모델
4세대	관계형 (RDBMS)	이차원의 테이블 즉, 컬럼과 로우로써 이루어진 모델
5세대	객체지향형 (OODBMS)	멀티미디어 데이터의 원활한 처리를 위해 고안된 데이터베이스 이며 객체 지향적 프로그래밍의 개념을 도입
6세대	객체관계형 (ORDBMS)	관계형 데이터베이스의 안정된 성능에 기반하면서 멀티미디어 데이터를 원활히 처리하기 위해 현실적으로 제안된 모델

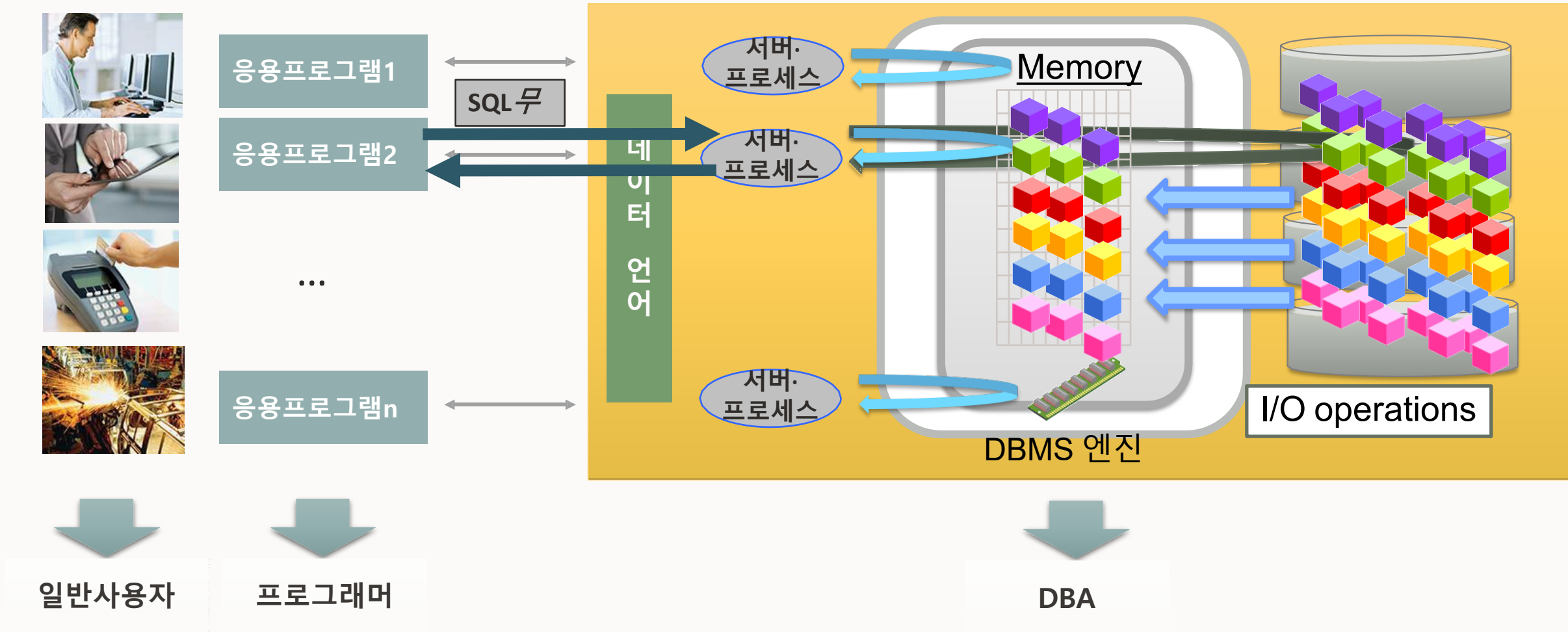
데이터관리 시스템(데이터베이스)

DBMS는 데이터베이스의 데이터를 저장,수정, 조회,관리할수 있는 소프트웨어

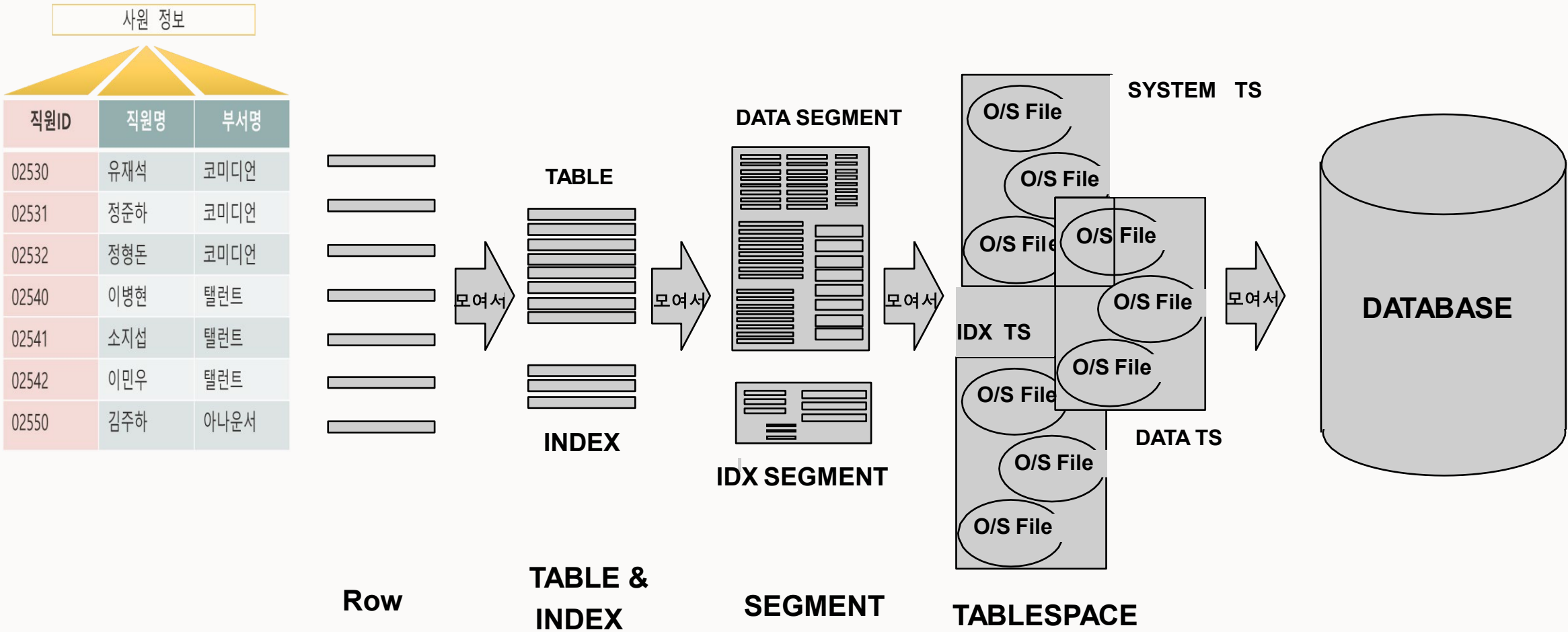


데이터관리 시스템(데이터베이스)

DBMS는 데이터베이스의 데이터를 저장,수정, 조회,관리할수 있는 소프트웨어

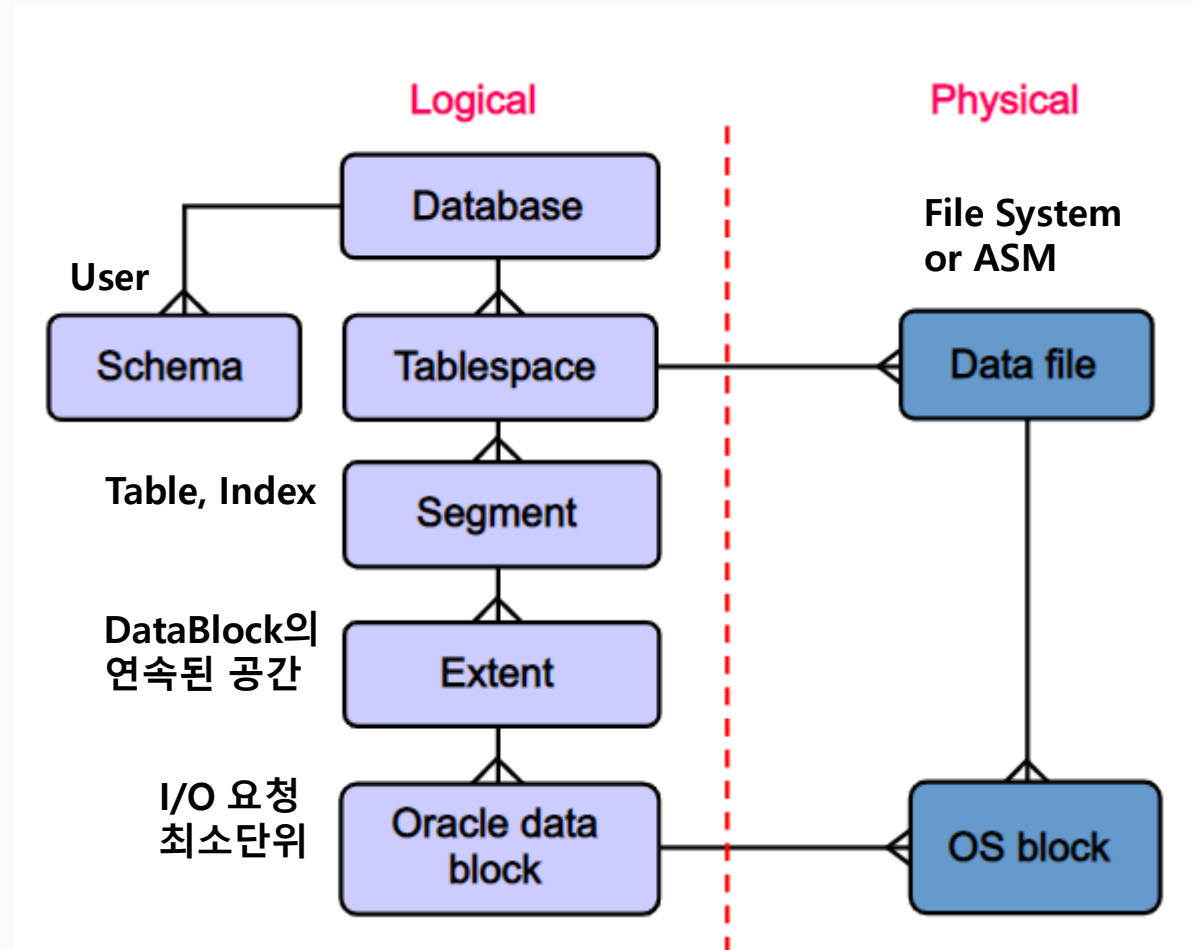
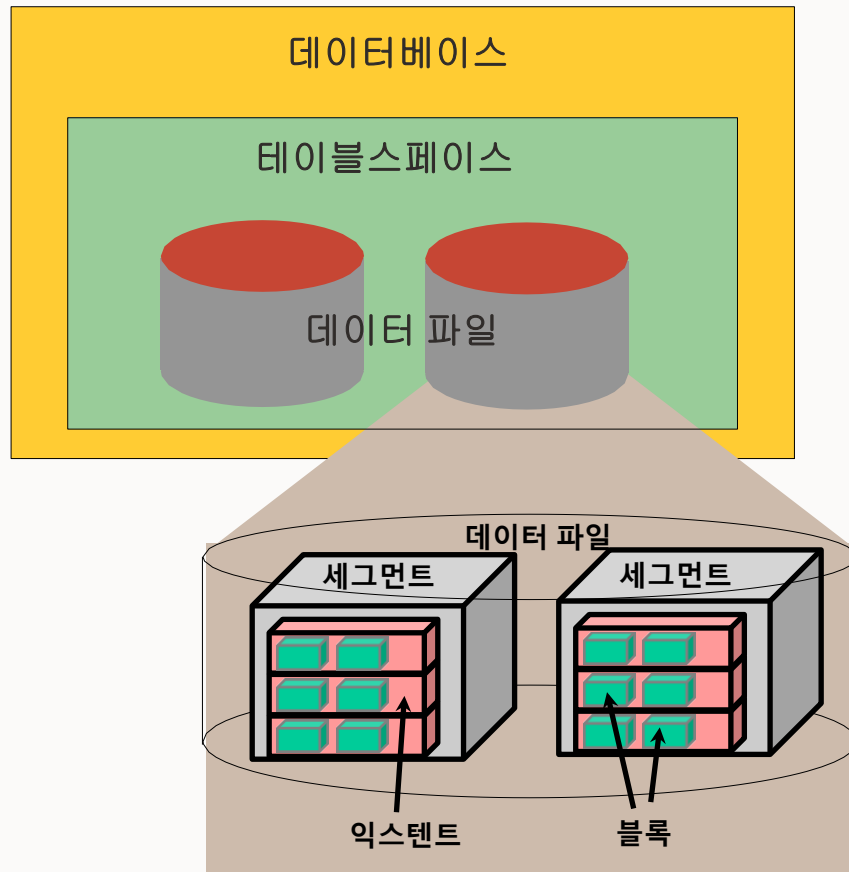


Database 데이터 저장구조



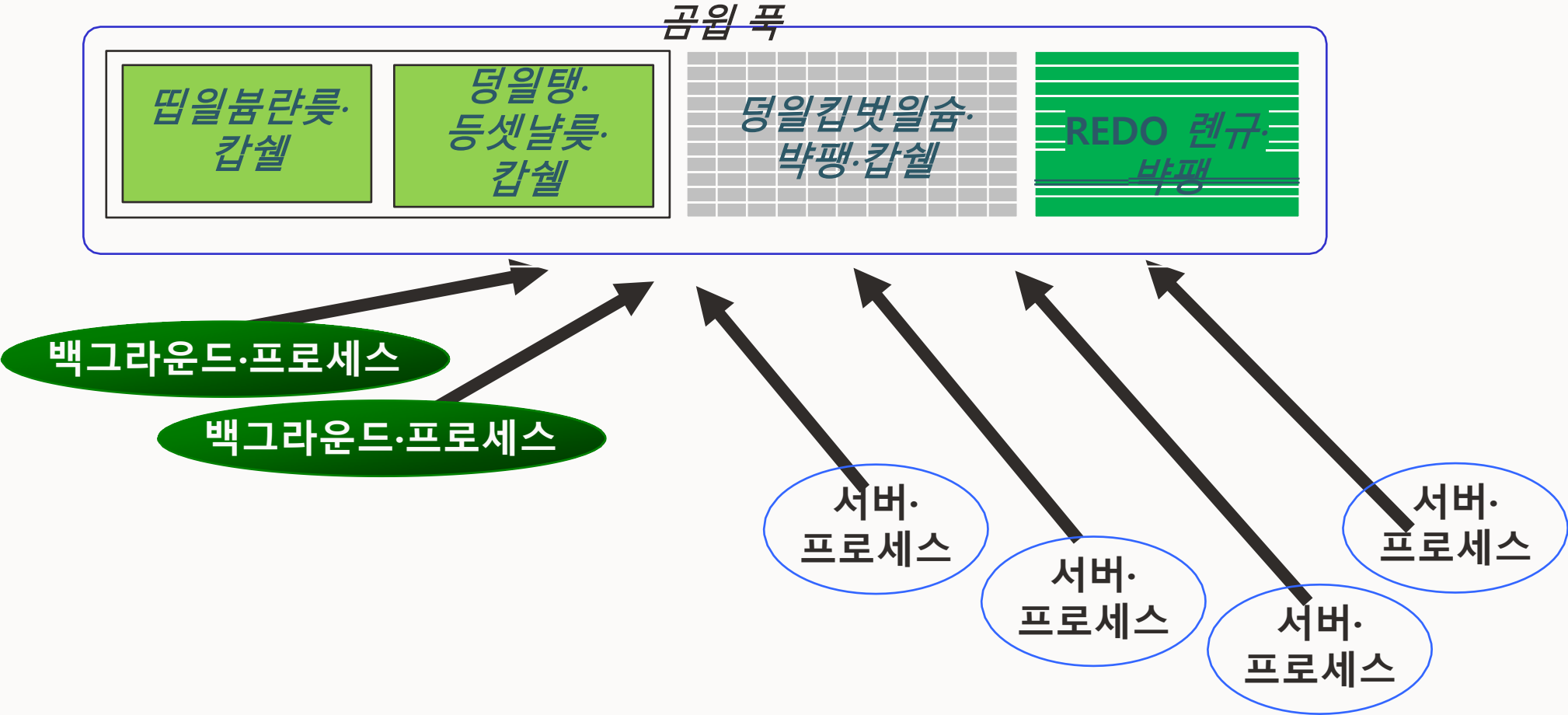
Oracle Database 데이터 저장구조

Database Storage Architecture



Database 관리 프로세스

복수의 프로세스가 동시에 캐쉬(메모리)에 대해 액세스



데이터 언어는?

SQL(Structured Query Language)

- 관계형 데이터베이스에 접근(생성, 변경, 삭제, 조회)하기 위한 **비절차적 표준언어**

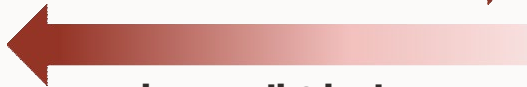


사용자

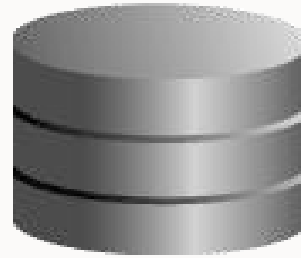
SQL (명령)



자료 (데이터)



	건물 평면적(m²) [단순 설치가능 면적]	BIPV 발전 가능 전력 (0.7kWp/9.9m²)
주거용	2,109,637,400	153,408,705
농수산용	98,445,850	6,960,818
광공업용	167,178,500	11,820,702
상업용	496,920,750	35,135,811
공공용	27,626,900	1,963,417
문교사회용	138,952,150	9,824,899
기타	97,116,650	6,866,834
합계	3,195,878,200	225,971,186



데이터베이스

◆ SQL 언어의 특징

- ✓ 이해하기 쉬운 형태로 표현
- ✓ 대화식 질의어로 사용가능
- ✓ 데이터 정의, 데이터 조작, 제어기능 제공
- ✓ 레코드 집합단위로 처리
- ✓ 비절차적 언어

데이터 언어는?

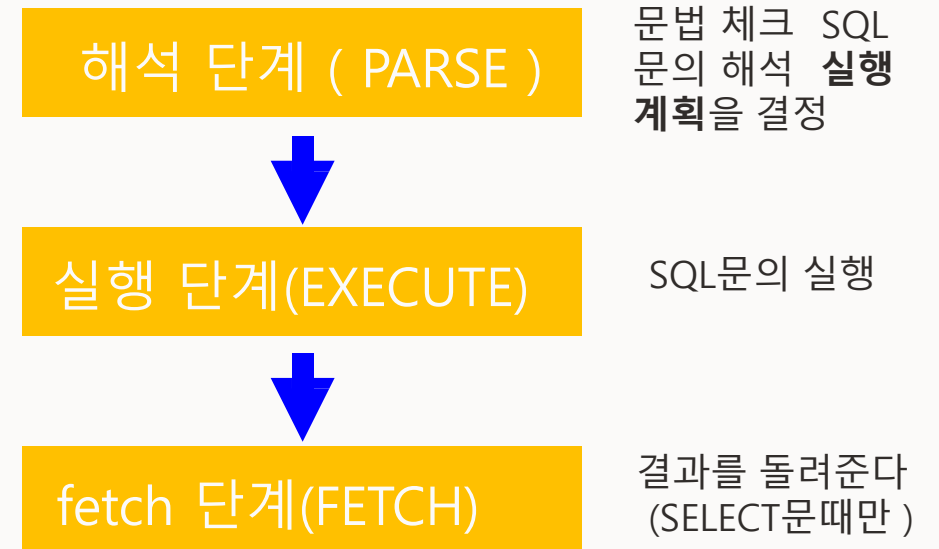
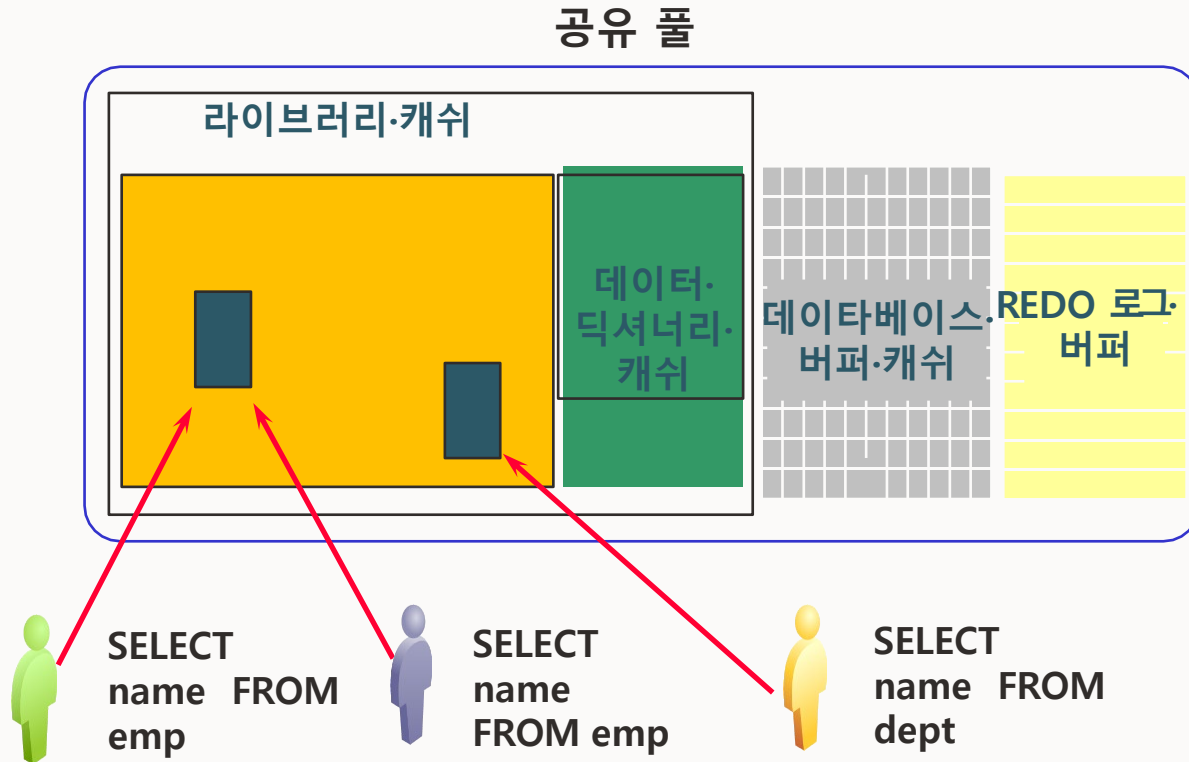
SQL문의 유형

명령어의 종류	명령어	설명
QUERY (데이터 검색, 조회)	SELECT	데이터를 조회하거나 검색하기 위한 명령어
DML (데이터 조작어)	INSERT UPDATE DELETE	테이블의 데이터를 조작(새로운 데이터 입력, 수정, 삭제)하는 명령어 : Data Manipulation Language
DDL (데이터 정의어)	CREATE, DROP ALTER, RENAME TRUNCATE, COMMENT	테이블과 같은 데이터의 구조를 정의(생성, 변경, 제거 등)하는데 사용되는 명령어 : Data Definition Language
TCL (트랜잭션 제어어)	COMMIT, ROLLBACK SAVEPOINT	논리적인 작업의 단위를 묶어서 DML에 의해 조작된 결과를 작업단위별로 제어하는 명령어 : Transaction Control Language
DCL (데이터 제어어)	GRANT REVOKE	Object들을 사용하도록 권한을 주고 받는 명령어 : Data Control Language

데이터 언어는?

SQL(Structured Query Language)

- SQL문이나 PL /SQL 블록을 해석해 , 실행 계획을 공유 SQL 영역에 저장
사용자가 동일 SQL를 실행했을 경우 , 같은 공유 SQL 영역을 이용

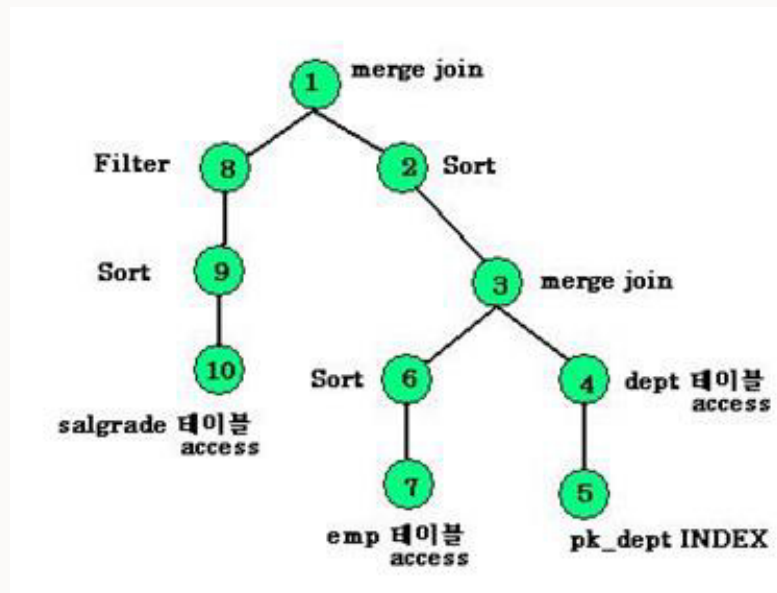


Optimizer

SQL문에 대하여 이를 해석하고 최적의 실행계획을 수립하는 역할

데이터 차제의 통계정보를 기반으로 SQL실행 시 소요되는 실질적인 비용을 계산하여 최적의 방식을 선택하여 수행 (Cost-Based Optimizer)

데이터베이스 성능의 핵심적인 요인(오라클의 핵심 기술)



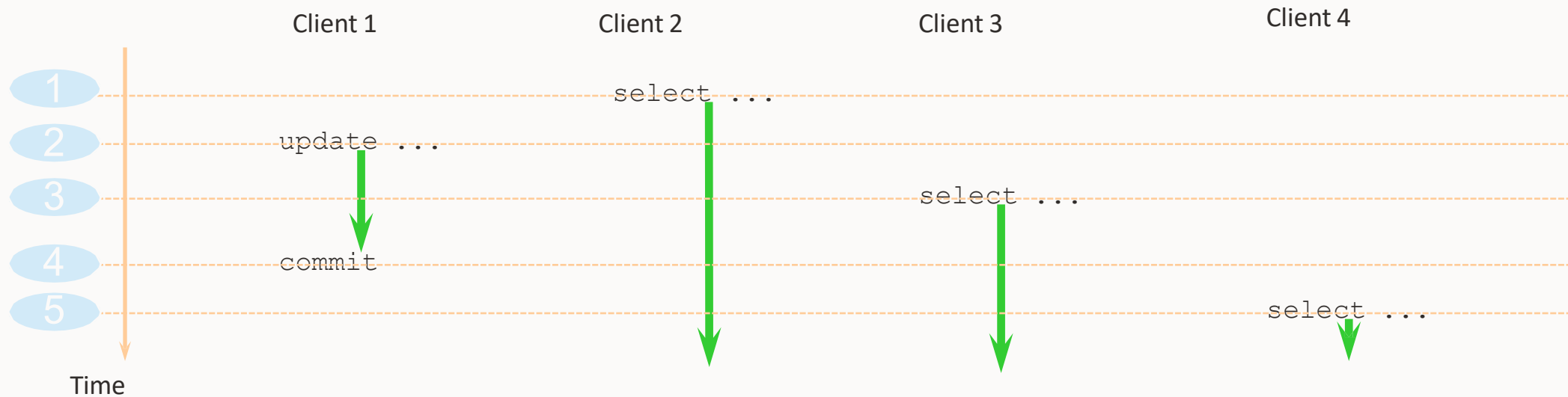
- 저장하는 데이터에 대한 특성을
통계정보를 기반으로 실질적인 비용을 계산하여 수행
(Cost-Based Optimizer = CBO)
 - ✓ 데이터베이스 시스템 환경 및 DB Parameter 정보
 - ✓ 인덱스 유무 및 테이블내 데이터 분포(통계 정보)
 - ✓ SQL내 사용컬럼 및 조건, 테이블의 Join 순서, Join 방법

ACID



ACID : DBMS(RDBMS)에서 제공해야 하는 기본 개념/기능

- Atomicity : 원자성, 트랜잭션은 ALL or Nothing으로 처리되어야 함.
- Consistency : 일관성, 트랜잭션이 성공하면 해당 상태가 반드시 유지되어야 함.
- Isolation : 고립성, 트랜잭션 수행중 다른 트랜잭션의 연산에 영향을 받으면 안됨.
- Durability : 지속성, 성공적인 트랜잭션은 영원히 반영되어야 함.



TRANSACTION

데이터베이스 시스템에서 사용되는 쪼갤 수 없는 업무처리의 단위(Atomicity = All or Nothing)

100만원 계좌이체 할 경우,



COMMIT

입력, 수정 및 삭제한 자료가 문제가 없다고 판단되었을 경우, 변경사항을 데이터베이스에 반영하는 것
(Consistency, Isolation)

세션 1

① SELECT * FROM 직원; [데이터 건수 10건]

② INSERT INTO 직원 VALUES('10', 송혜교);
[데이터 추가 1건]

③ SELECT * FROM 직원; [데이터 건수 11건]

⑤ COMMIT;

세션 2

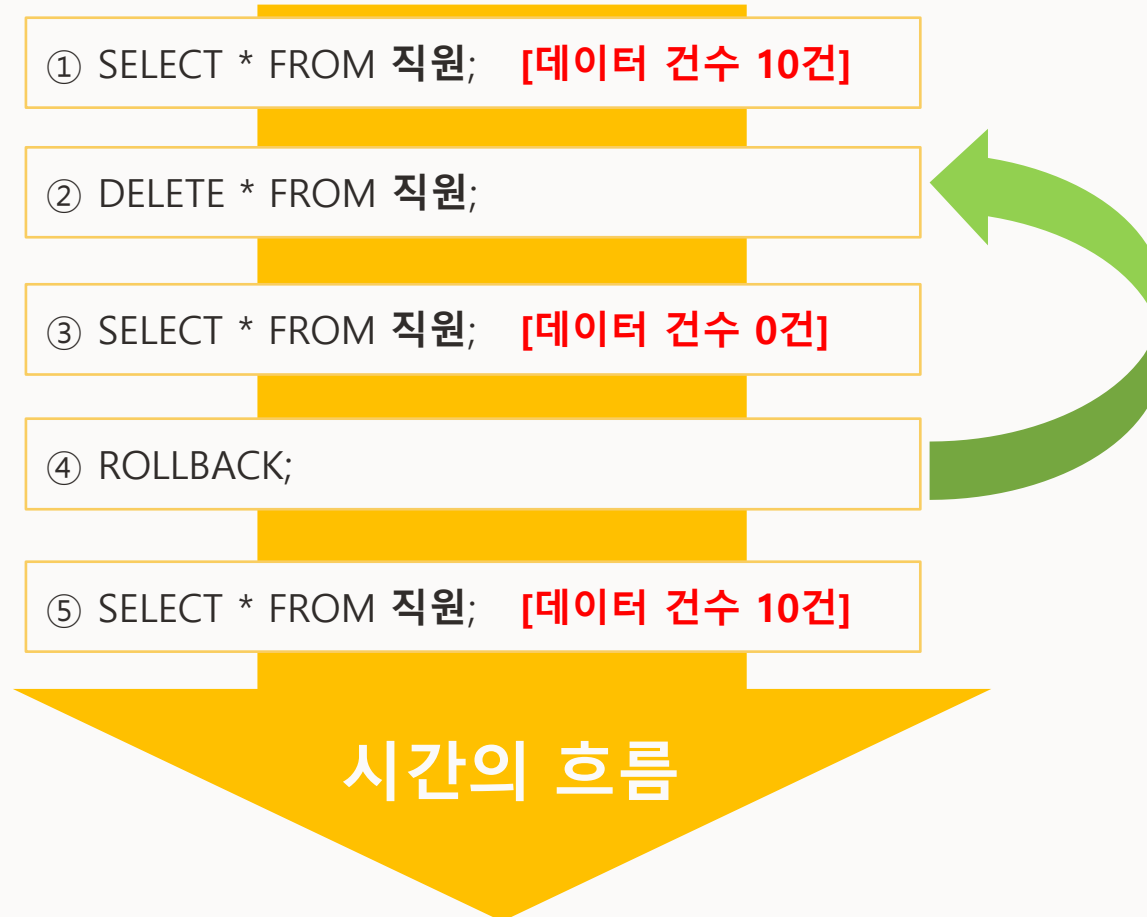
④ SELECT * FROM 주문; [당일 탕 갺샙 10 갺]

⑥ SELECT * FROM 주문; [당일 탕 갺샙 11 갺]

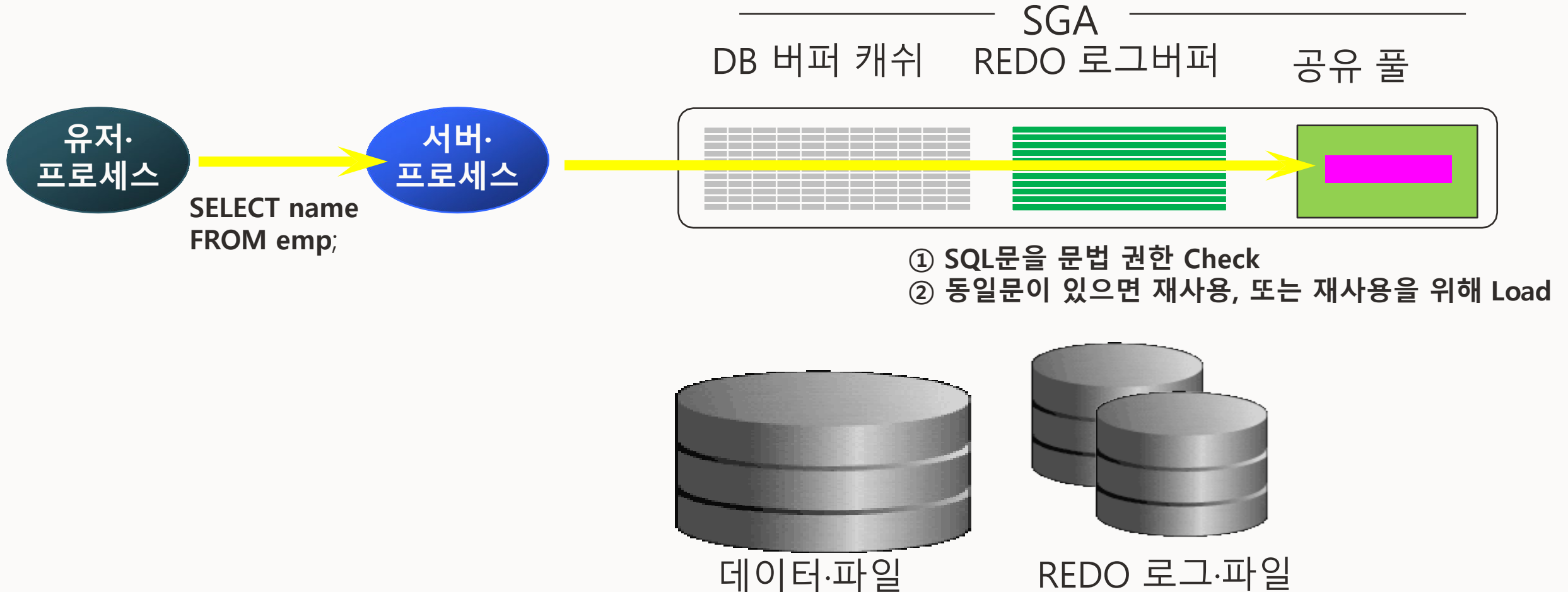
습 / 을 흙룩

ROLLBACK

데이터 변경사항이 취소되어 데이터를 이전 상태로 복구하는 것 (Consistency)

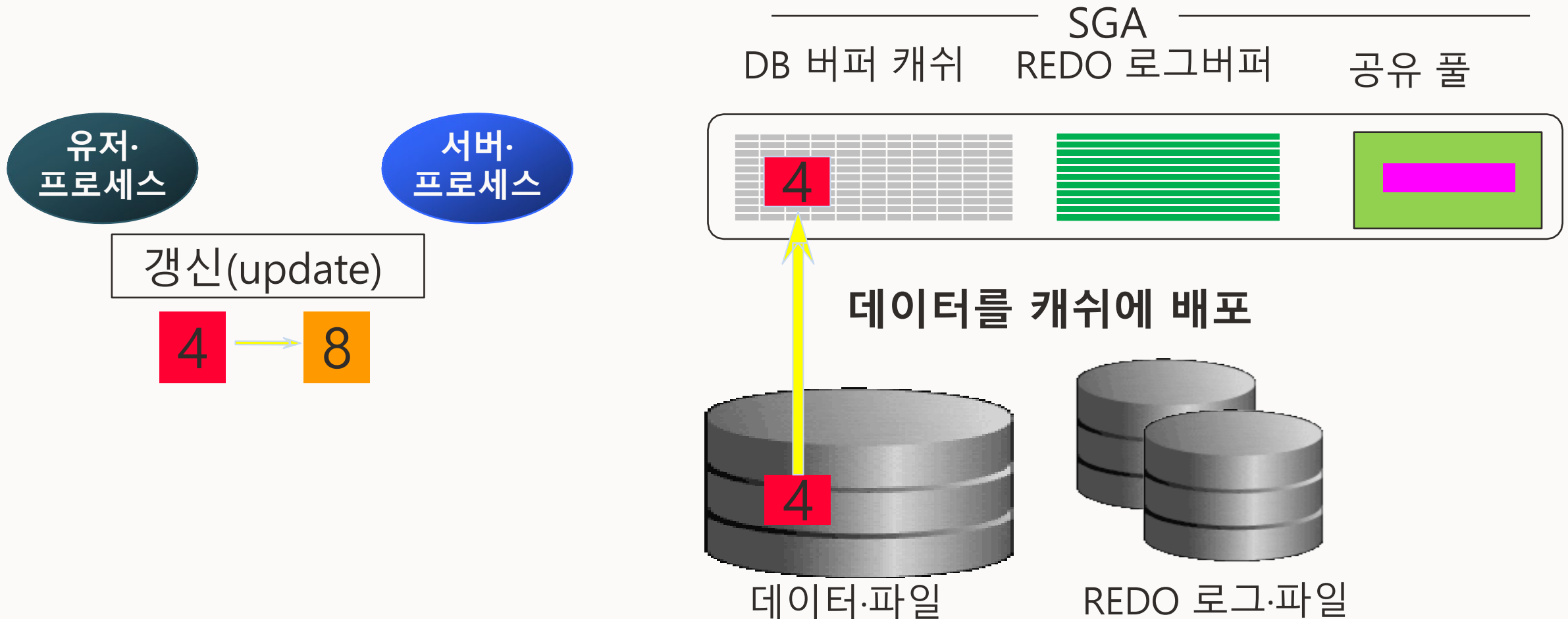


데이터 변경의 지속성(Durability) 보장 방안



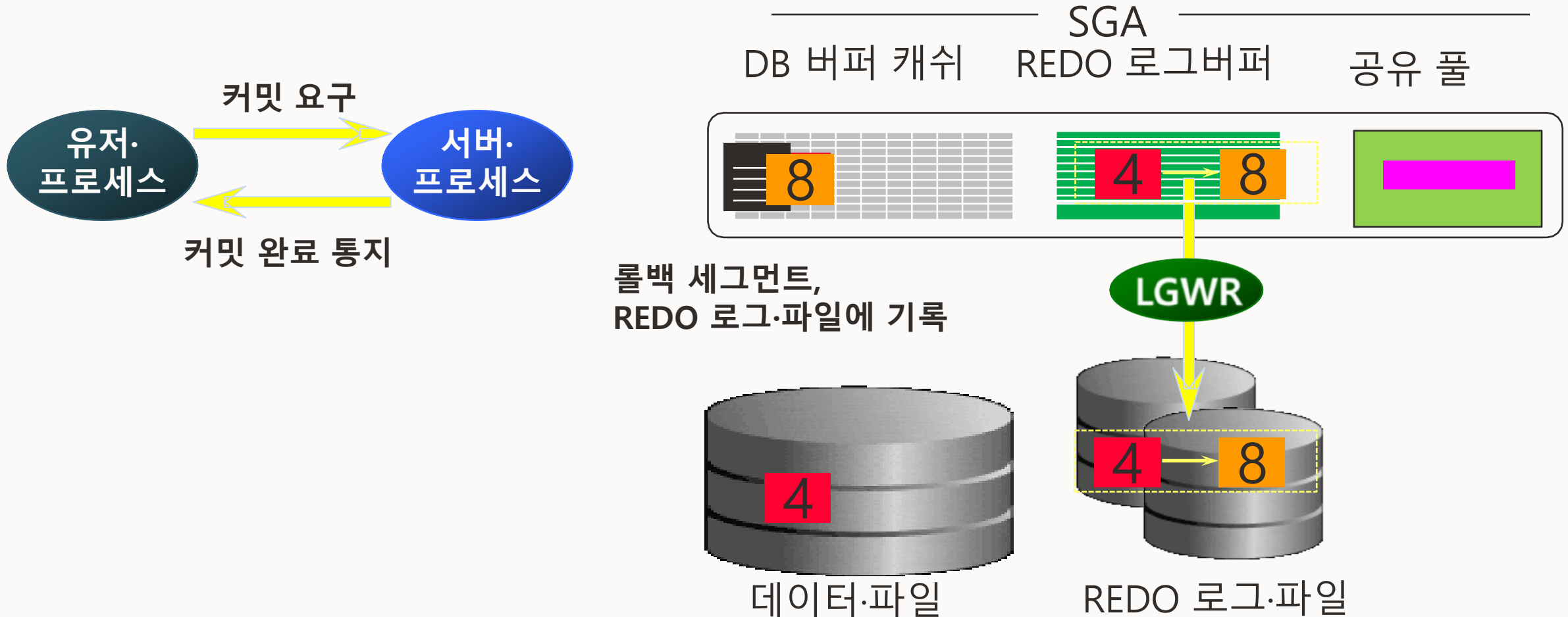
데이터 변경의 지속성(Durability) 보장 방안

데이터 검색/갱신시의 동작 (Disk Read 포함)



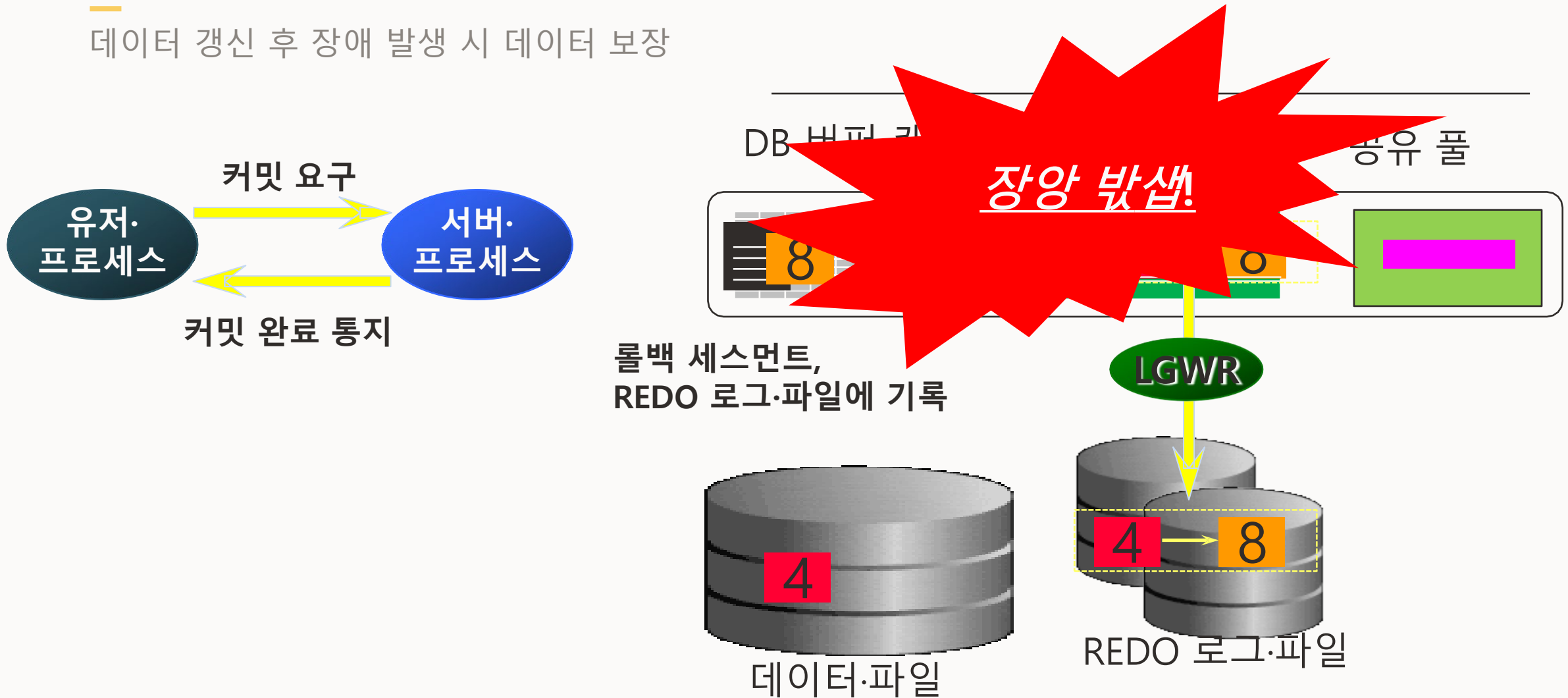
데이터 변경의 지속성(Durability) 보장 방안

데이터 갱신시의 동작 (로그 기록)



데이터 변경의 지속성(Durability) 보장 방안

데이터 갱신 후 장애 발생 시 데이터 보장

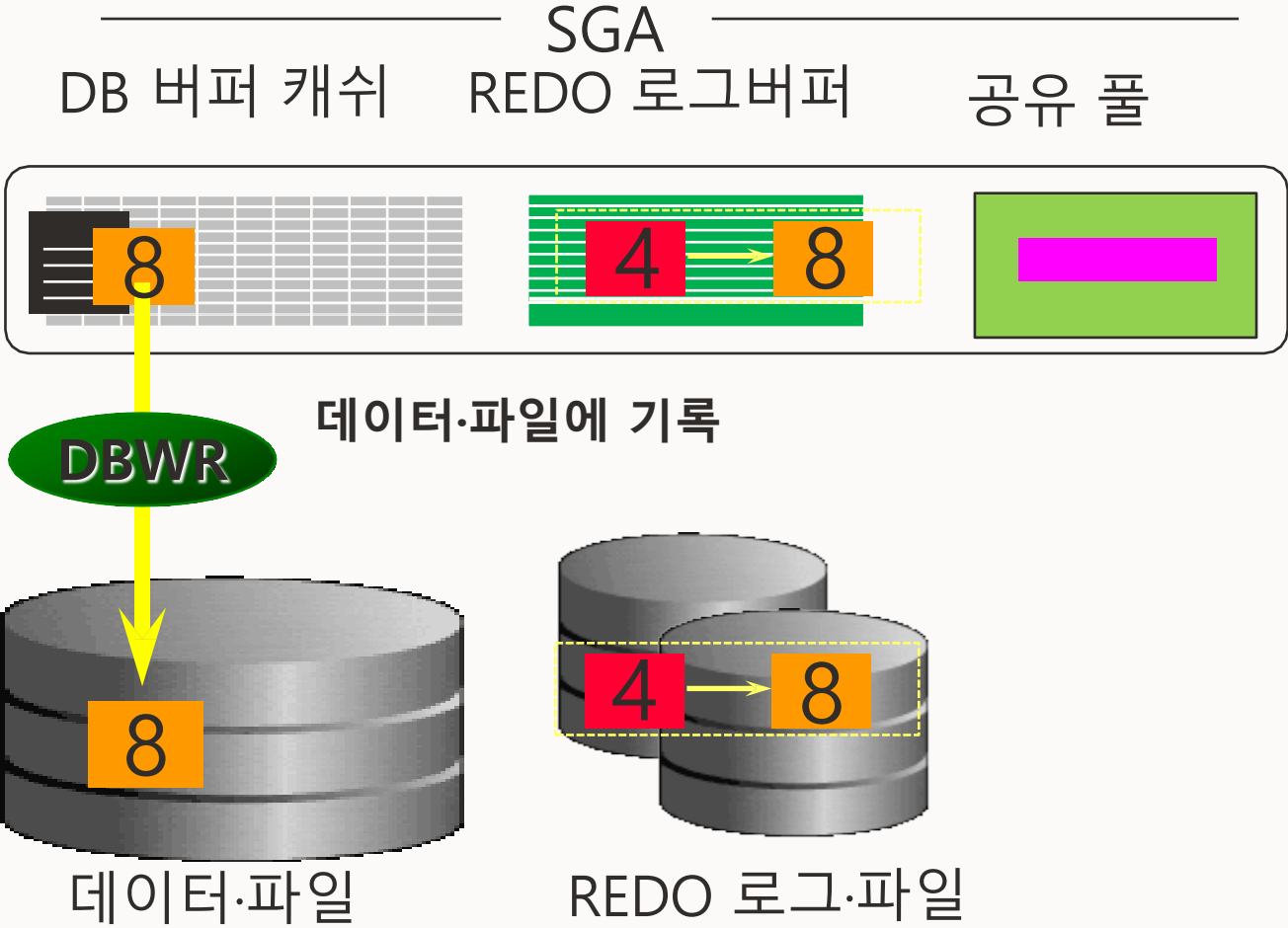


데이터 변경의 지속성(Durability) 보장 방안

데이터 갱신시의 동작 (데이터 기록)

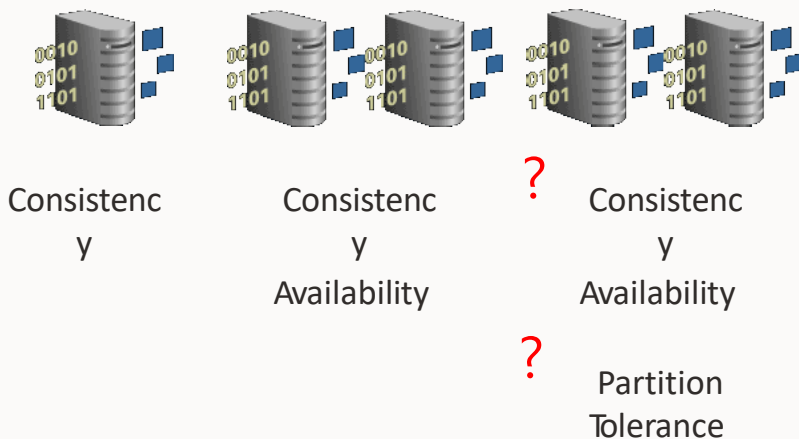
유저.
프로세스

서버.
프로세스



CAP이론

‘적절한 응답시간내” 세가지 속성을 만족시키는 분산시스템은 불가능



Consistency

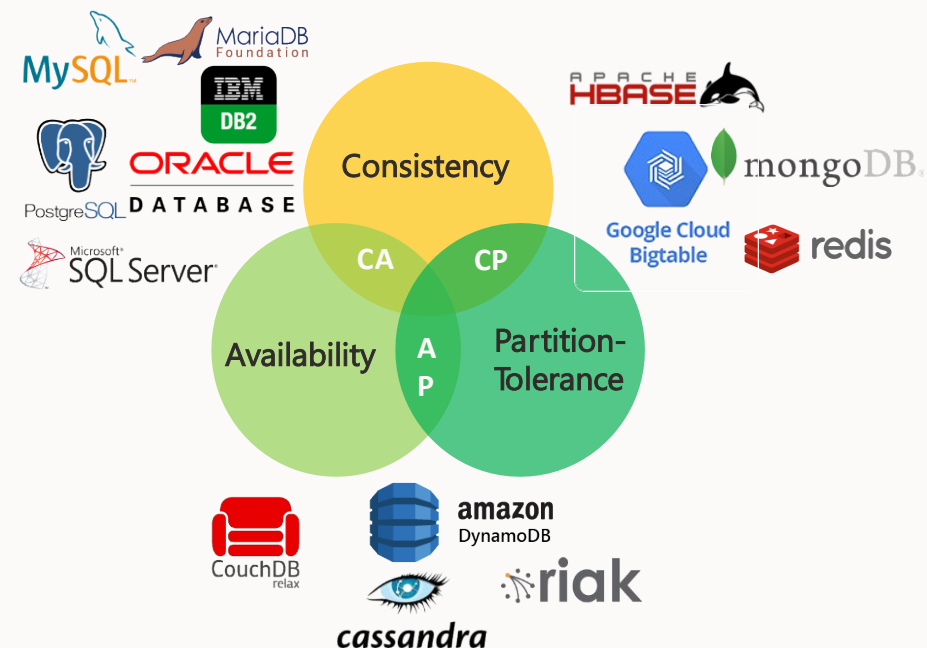
- 모든 client가 언제나 같은 상태의 데이터를 바라봐야 한다

Availability

- 모든 client가 노드 장애시 언제나 데이터를 접근할 수 있어야 한다(Replica필요)

Partition Tolerance

- Data가 물리적으로 다른 Network 공간에 분산되어 있으며, Network 단절후에도 시스템은 동작해야 한다.



Strict Consistency (CA)

VS

Eventually Consistency (AP)

조회 시점마다 다른 Data가 조회될수 있으며, 이는 일정시간이 지나야 일관성이 유지됨

관계형 데이터베이스 모델링

데이터 정규화의 필요성

- 데이터의 중복, 변경 시 모든 데이터를 찾아서 변경

사원 정보			키패
주원ID	주원명	직책명	
02530	유재석	코미디언	
02531	정준하	코미디언	
02532	정형돈	코미디언	
02540	이병현	탤런트	
02541	소지섭	탤런트	
02542	이민우	탤런트	
02550	김주하	아나운서	

정규화 (Normalization) 화

	Objectives	Rule Definition
제 1 정규형	Non-Attribute Repetition	Entity 내의 모든 Attribute는 반복성 (Repetition)의 가능성을 배제한다.
제 2 정규형	Primary Key 종속성	Entity 내의 Primary Key를 제외한 모든 Attribute는 Primary Key 전체에 대해서 종속성(Dependency)를 가져야 한다.
제 3 정규형	Non-Primary Key 종속성	Entity 내의 Primary Key를 제외한 모든 Attribute들 사이에는 어떠한 종속성 (Dependency)을 가져서는 안된다.

관계형 데이터베이스 모델링

테이블

- 키(key)와 값(value)들의 관계를 정규화 시키어 저장

사원 (TABLE)

사원의 특성을 나타내는 속성들(Attribute)

증원ID	증원명	부서ID
02530	유재석	100
02531	정준하	100
02532	정형돈	100
02540	이병현	110
02541	소지섭	110
02542	이민우	110
02550	김주하	120

컬럼(COLUMN)

부서(TABLE)

부서의 특성을 나타내는 속성들(Attribute)

부서ID	부서명
100	코미디언
110	탤런트
120	아나운서

관계(RELATION)

로우(ROW)

관계형 데이터베이스 모델링

인덱스

- 조회속도를 향상시키기 위한 데이터베이스 검색 기술

고객(TABLE)

약 2천만 건

이름	생년월일	주소
장동건	1970/06/18	서울...
원빈	1977/10/05	부산..
이병헌	1972/03/01	인천..
원빈	1956/01/05	강원..
....		

이름(INDEX)

이름	ROWID
...	
장동건	000E9B...
원빈	000A95...
원빈	000A95...
원빈	000E62...
이병헌	000A96...
....	

테이블의 row와 하나씩 대응되는 별도의 객체로 독립적인 저장공간을 보유

자료를 빠르게 접근할 수 있는 수단을 제공

구조적 특성상 압축을 권하지 않음

인덱스를 생성할수록 데이터 입력/수정/삭제 시 성능이 저하됨.



제1정규형 (1NF)

원자성



1. 테이블 반정규화

- 테이블 병합/분할/추가

데이터 관리 시스템 **OLTP vs OLAP**

1. **OLTP System(Online Transaction Processing)**

- 실시간으로 데이터를 갱신하거나 조회 작업 처리
- 대부분의 기간계 업무처리 시스템이 해당
- 주간 단기 배치, 야간 장기 배치 작업과 병행
- 응답시간과 가용성이 가장 중요

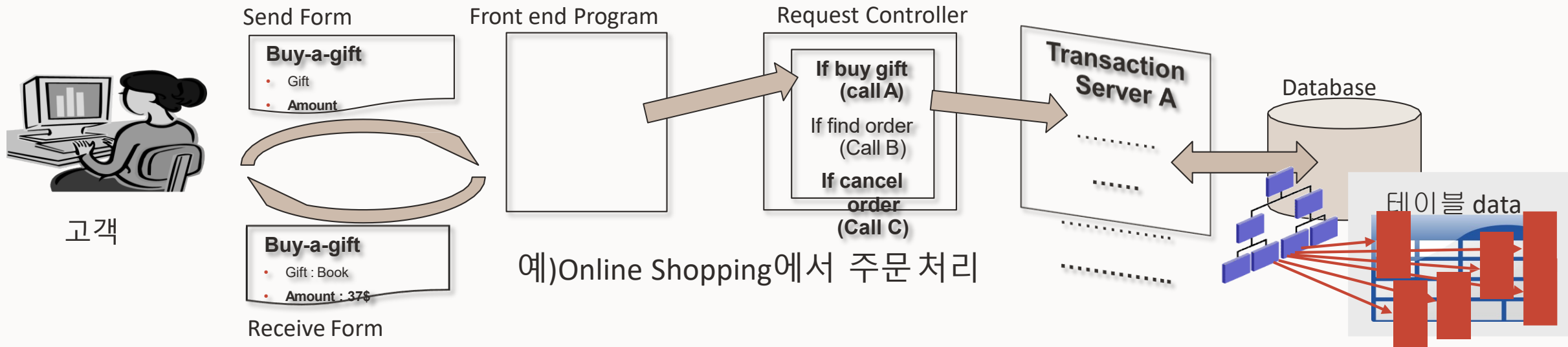
2. **OLAP System(Online Analytical Processing)**

- 최종 사용자가 다차원 정보에 직접 접근하여 대화식으로 정보를 분석하고 의사 결정에 활용하기 위한 시스템
- DSS, EIS, DW, DM
- 대규모 데이터 조회 성능

데이터 관리 시스템

OLTP 업무 유형 : 단건 처리 업무

- 다수의 사용자에게 의해 사용되며 DBMS내부에서 대량의 Transaction을 동시에 처리(동시성, Lock 메커니즘 중요)
 - 대부분 조회 업무 소수의 Data의 변경작업
- ➔ “인덱스”를 통한 성능 향상

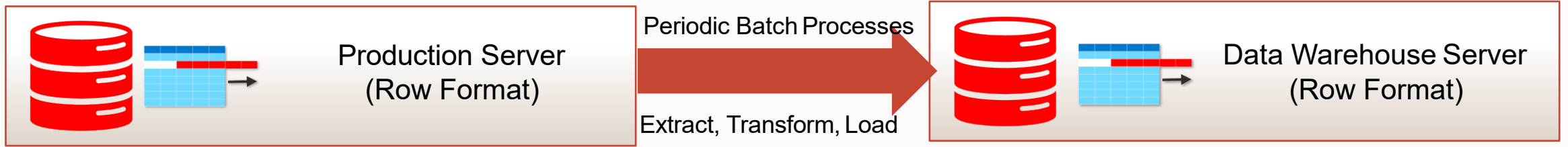


데이터 관리 시스템

Batch/OLAP 업무 유형: Batch처리 업무/주어진 시간 이내의 데이터 처리가 중요

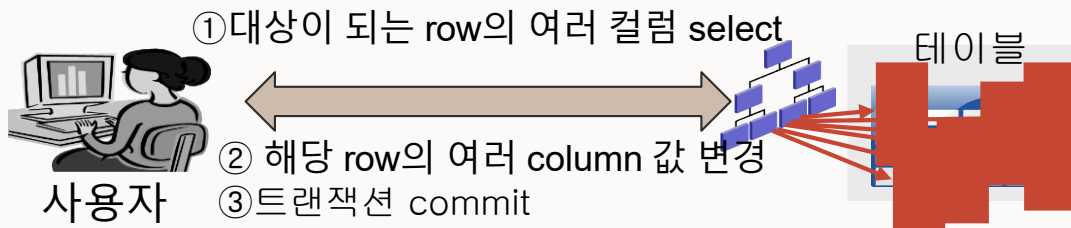
- 특정 프로그램이 주기적으로 동작하여 대규모 Delete, Update작업을 수행, 경우에 따라 복잡한 Logic
- 대규모 조회 및 변경 작업
 - ➔ “병렬처리”를 통한 성능 향상

데이터 관리 시스템의 분류 OLTP vs OLAP using Row format



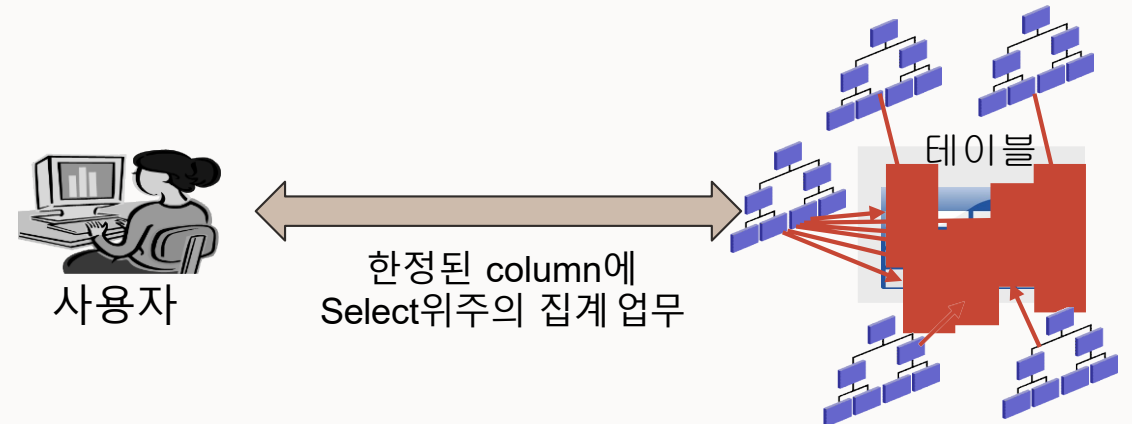
OLTP System(Online Transaction Processing) 위주

- 응답 시간을 빠르게 하기 위해 인덱스 생성

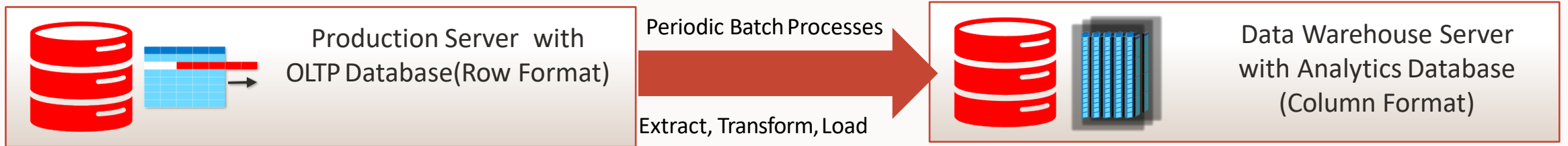


OLAP/DW System(Online Analytical Processing)

- 성능향상을 위해 병렬처리 및 인덱스를 많이 생성



데이터 관리 시스템 OLTP vs OLAP using column format



Source Data

CO_ID	Ticker	Stock Price	No of Trades	Exchange	Date
1	GM	30.77	1,000	NYSE	09/11/2020
2	GM	30.77	10,000	NYSE	09/11/2020
3	GM	30.79	12,500	NYSE	09/11/2020
4	ORCL	93.24	9,000	NQDS	09/11/2020

Row Store

- 여러 개의 컬럼이 포함된 질의에 우수함
- OLTP 트랜잭션 처리에 효과적

GM	30.77	1,000	NYSE	1/17/2007
GM	30.77	10,000	NYSE	1/17/2007
GM	30.79	12,500	NYSE	1/17/2007
AAPL	93.24	9,000	NQDS	1/17/2007

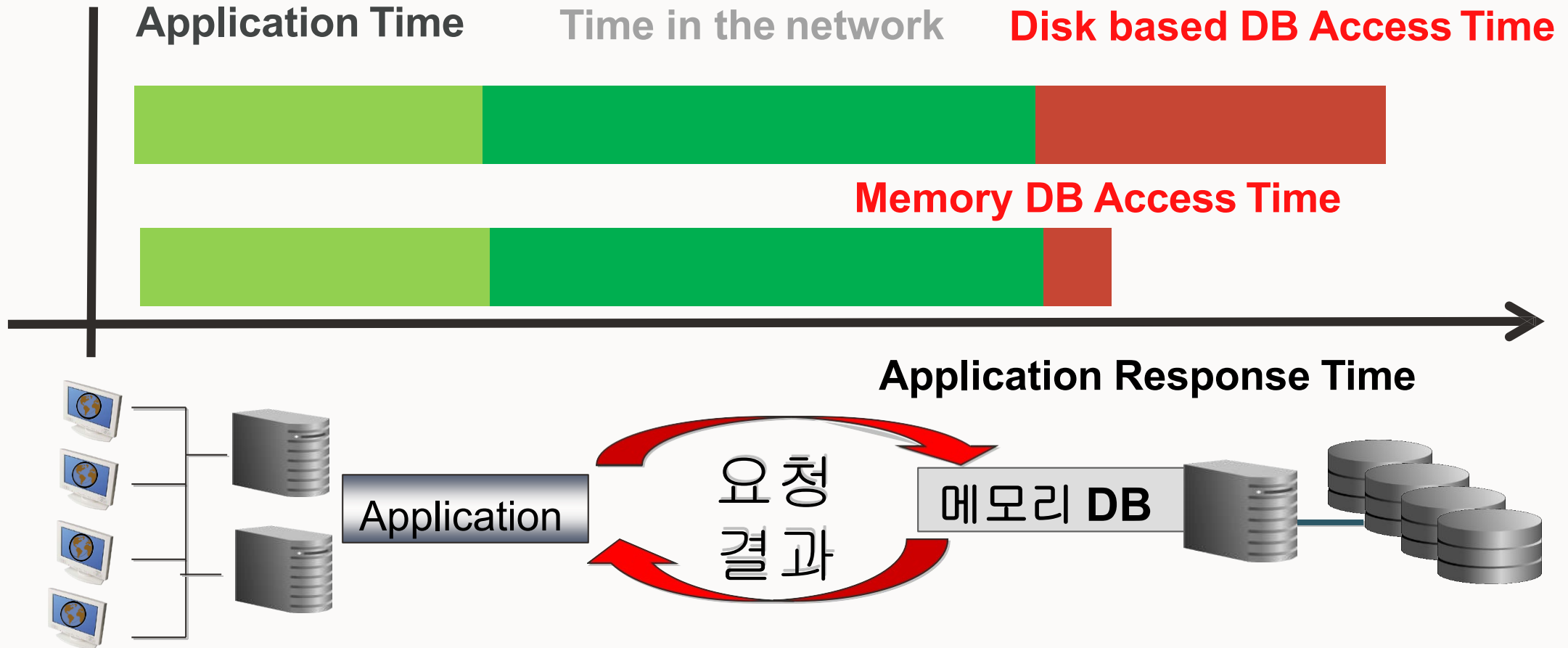
Column Store

- 필요한 컬럼만 접근, 한정적인 컬럼만을 질의하는 경우 우수
- 여러 개의 컬럼이 포함된 질의 시 성능저하 발생
- 압축률이 우수

GM	30.77	1,000	NYSE	1/17/2007
GM	30.77	10,000	NYSE	1/17/2007
GM	30.79	12,500	NYSE	1/17/2007
AAPL	93.24	9,000	NQDS	1/17/2007

인메모리 데이터베이스

I/O성능을 향상 시키기 위한 방안



DBMS의 확장(+NoSQL)

