[DB 2-1] 9th Mar. 2023

Consider it very seriously

Database

(Introduction to Database)

Spring, 2023

Jaewook Byun

Ph.D., Assistant Professor, Department of Software, Sejong University

jwbyun@sejong.ac.kr

https://sites.google.com/view/jack-dfpl/home

https://www.youtube.com/channel/UC988e-Y8nto0LXVae0aqaOQ

Schedule

Week	Title		
1	Course Overview and Environment Setting	Environment Setting and Revisiting Java	-
2	**Introduction to Database**		CH1
3	Relational Model	Relational Algebra	CH2
4	Relational Algebra		CH2
5	Database Language: SQL (DDL, DML, DQL, DCL)		CH3-5
6	Database Language: SQL (DDL, DML, DQL, DCL)		CH3-5
7	Database Language: SQL (DDL, DML, DQL, DCL)		CH3-5
8	Midterm Examination		
9	Physical Database Design: Indexing		CH12
10	Physical Database Design: Indexing		CH12
11	Conceptual Database Design – E-R Data Model		CH6
12	Logical Database Design 1 – Schema Mapping		CH6
13	Logical Database Design 2 – Normalization		CH7
14	Query Processing and Optimization, or View (TBD)		CH13-14, CH3-5
15	Final Examination		

Subject to change

Calendar

- 수업일수는 요일별 15주 이상이며 수업 결손이 발생하지 않도록 진행

요 일 별 수업일수	월	화	수	목	금
	16일	15일	16일	16일	15일

나. 수업주차 : 한 주차는 목요일부터 수요일까지 임

수업주차	기간	수업주차	기간
1주차	03.02.(목) ~ 03.08.(수)	9주차	04.27.(목) ~ 05.03.(수)
2주차	03.09.(목) ~ 03.15.(수)	10주차	05.04.(목) ~ 05.10.(수)
3주차	03.16.(목) ~ 03.22.(수)	11주차	05.11.(목) ~ 05.17.(수)
4주차	03.23.(목) ~ 03.29.(수)	12주차	05.18.(목) ~ 05.24.(수)
5주차	03.30.(목) ~ 04.05.(수)	13주차	05.25.(목) ~ 05.31.(수)
6주차	04.06.(목) ~ 04.12.(수)	14주차	06.01.(목) ~ 06.07.(수)
7주차	04.13.(목) ~ 04.19.(수)	15주차	06.08.(목) ~ 06.14.(수)
8주차 (중간고사)	04.20.(목) ~ 04.26.(수)	16주차 (기말고사)	06.15.(목) ~ 06.21.(수)

Calendar

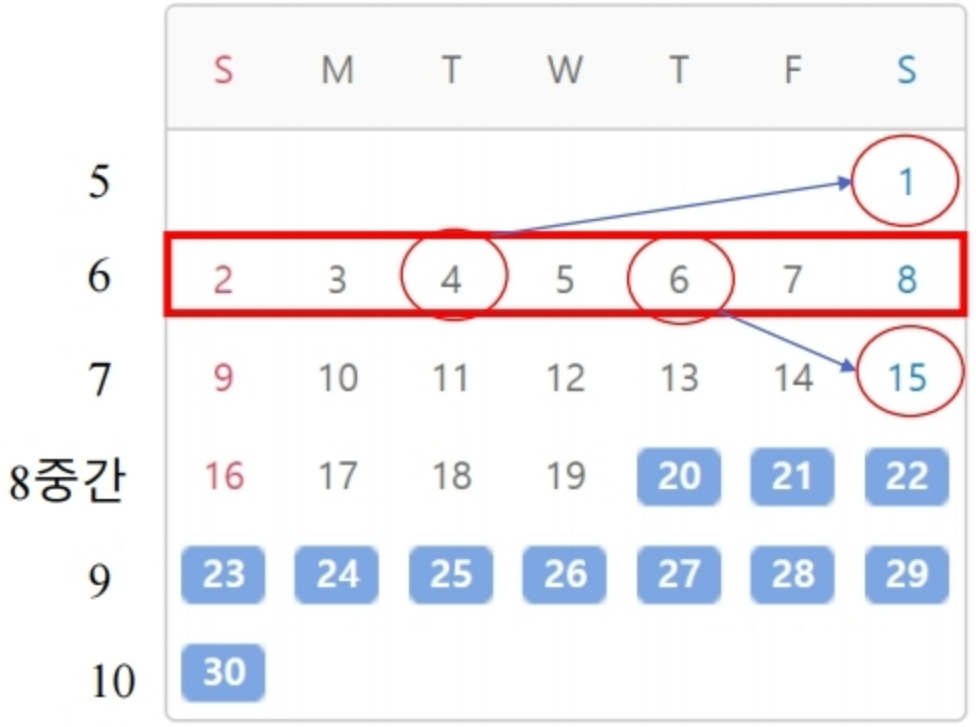
2023년 3월. March

	S	M	T	W	T	F	S
1				1	2	3	4
2	5	6	7	8	9	10	11
3	12	13	14	15	16	17	18
4	19	20	21	22	23	24	25
5	26	27	28	29	30	31	

- 2(목) 1학기 개강
- 3(금) - 8(수) 수강신청 과목 확인 및 변경
- 6(월) - 15(수) 교직신청
- 24(금) - 28(화) 수강신청과목 철회

Calendar

2023년 4월. April



	S	M	T	W	T	F	S
5							1
6	2	3	4	5	6	7	8
7	9	10	11	12	13	14	15
8중간	16	17	18	19	20	21	22
9	23	24	25	26	27	28	29
10	30						

- 20(목) - 26(수) 1학기 중간고사
- 27(목) - 5.1(월) 1학기 중간고사 성적 입력

IEEE ICDE 2023, Anaheim, California, US

Calendar

2023년 5월. May

	S	M	T	W	T	F	S
10		1	2	3	4	5	6
11	7	8	9	10	11	12	13
12	14	15	16	17	18	19	20
13	21	22	23	24	25	26	27
14	28	29	30	31			

- 2(화) - 7(일) 1학기 중간고사 성적 열람 및 정정
- 4(목) - 30(화) 복수·부전공, 연계융합전공 신청
- 5(금) 창립 83주년 기념휴일 (창립일 : 1940. 5. 20)
- 29(월) - 31(수) 하계 계절학기 수강신청

Calendar

2023년 6월. June

	S	M	T	W	T	F	S
14					1	2	3
	4	5	6	7	8	9	10
15기말	11	12	13	14	15	16	17
	18	19	20	21	22	23	24
	25	26	27	28	29	30	

- 9(금) - 26(월) 1학기 강의평가
- 15(목) - 21(수) 1학기 기말고사 및 수업결손 보충
- 22(목) - 26(월) 1학기 기말고사 성적 입력
- 22(목) 하계방학 시작 및 계절학기 개강
- 27(화) - 7.3(월) 1학기 기말고사 성적 열람 및 정정

Calendar

2023년 7월. July

S	M	T	W	T	F	S
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

- 4(화) - 5(수)
- 24(월) - 30(일)

1학기 기말고사 성적마감

2학기 복학, 휴학 신청

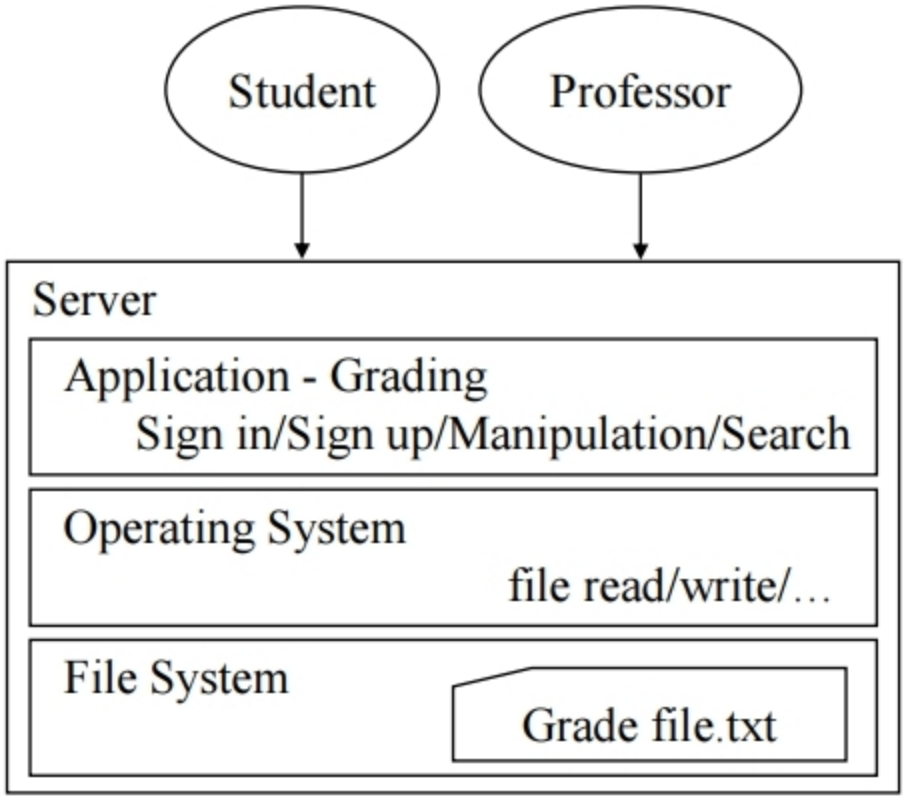
Introduction to Database

- Database
  - An organized collection of data

  - Applications
    - Banking System
    - Airlines
    - Education sector
    - Telecom
    - Sales
    - Industry
    - Online Shopping
    - …

Introduction to Database

- Implementing an information service with file processing



Introduction to Database

- Implementing an information service with file processing
  - P1: Grade printing service

grade.txt

```
jack|010-1234-5678|jwbyun@gmail.com|database|6|98
jack|010-1234-5678|jwbyun@gmail.com|database programming|7|96
jack|010-1234-5678|jwbyun@gmail.com|data analysis|7|92
jack|010-1234-5678|jwbyun@gmail.com|advanced database|8|99
james|010-4321-8765|jms@gmail.com|database|6|89
james|010-4321-8765|jms@gmail.com|database programming|7|84
james|010-4321-8765|jms@gmail.com|data analysis|7|81
james|010-4321-8765|jms@gmail.com|advanced database|8|82
```

Printing Service

```
<terminated> P1 (1) [Java Application] C:₩Program Files₩java₩jdk-14.0.2₩bin₩javaw.exe (2021. 1. 17. 오후 11:4
jack 010-1234-5678 jwbyun@gmail.com database 6 98
jack 010-1234-5678 jwbyun@gmail.com database programming 7 96
jack 010-1234-5678 jwbyun@gmail.com data analysis 7 92
jack 010-1234-5678 jwbyun@gmail.com advanced database 8 99
james 010-4321-8765 jms@gmail.com database 6 89
james 010-4321-8765 jms@gmail.com database programming 7 84
james 010-4321-8765 jms@gmail.com data analysis 7 81
james 010-4321-8765 jms@gmail.com advanced database 8 82
```

Introduction to Database

- Implementing an information service with file processing
  - P2: Grade average service

grade.txt

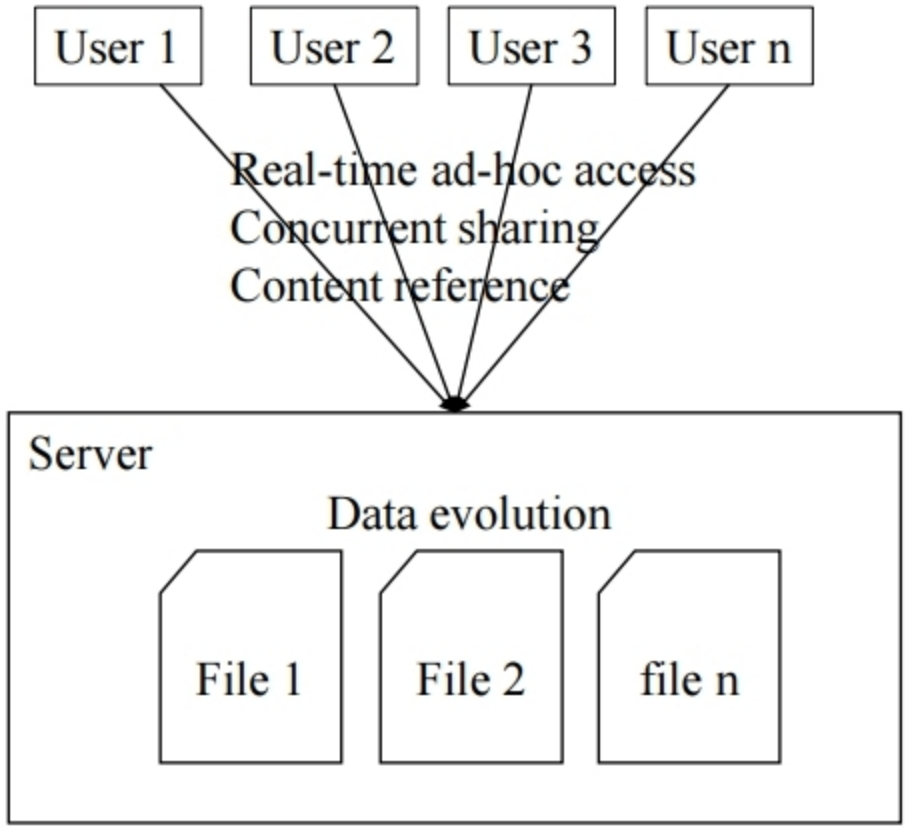
```
jack|010-1234-5678|jwbyun@gmail.com|database|6|98
jack|010-1234-5678|jwbyun@gmail.com|database programming|7|96
jack|010-1234-5678|jwbyun@gmail.com|data analysis|7|92
jack|010-1234-5678|jwbyun@gmail.com|advanced database|8|99
james|010-4321-8765|jms@gmail.com|database|6|89
james|010-4321-8765|jms@gmail.com|database programming|7|84
james|010-4321-8765|jms@gmail.com|data analysis|7|81
james|010-4321-8765|jms@gmail.com|advanced database|8|82
```

Grade average service

```
<terminated> P2 (1) [Java Application] C:₩Pr
{james=84.0, jack=96.25}
```

Introduction to Database

- Implementing an information service with file processing
  - Challenges from
    - Data redundancy
    - Data inconsistency
    - Data integrity
    - Inefficient data access
    - Atomicity
    - Concurrent access
    - Security
    - Data dependency
    - …



Introduction to Database

- Implementing an information service with file processing
  - Challenges from
    - Data redundancy

redundancy

Jack	010-1234-5678	jwbyun@gmail.com	Database	6	98
Jack	**010-1234-5678**	**jwbyun@gmail.com**	Database programming	7	96
Jack	**010-1234-5678**	**jwbyun@gmail.com**	Database analysis	7	92
Jack	**010-1234-5678**	**jwbyun@gmail.com**	Advanced database	8	99
James	010-4321-8765	jms@gmail.com	Database	6	89
James	**010-4321-8765**	**jms@gmail.com**	Database programming	7	84
James	**010-4321-8765**	**jms@gmail.com**	Database analysis	7	81
James	**010-4321-8765**	**jms@gmail.com**	Advanced database	8	82

Introduction to Database

- Implementing an information service with file processing
  - Challenges from
    - Data redundancy

No redundant data
- Split data into multiple files
- Merge them for generating original data

Jack	**010-1234-5678**	jwbyun@gmail.com
James	010-4321-8765	jms@gmail.com

010-1234-5678	Database	6	98
010-1234-5678	Database programming	7	96
010-1234-5678	Database analysis	7	92
010-1234-5678	Advanced database	8	99
010-4321-8765	Database	6	89
010-4321-8765	Database programming	7	84
010-4321-8765	Database analysis	7	81
010-4321-8765	Advanced database	8	82

Normalization

Introduction to Database

- Implementing an information service with file processing
  - Challenges from
    - Data inconsistency

What if the phone number of Jack changes?
- Data inconsistency

Jack	010-9876-5678	jwbyun@gmail.com
James	010-4321-8765	jms@gmail.com

010-9876-5678	Database	6	98
010-9876-5678	Database programming	7	96
010-9876-5678	Database analysis	7	92
010-9876-5678	Advanced database	8	99
010-4321-8765	Database	6	89
010-4321-8765	Database programming	7	84
010-4321-8765	Database analysis	7	81
010-4321-8765	Advanced database	8	82

Introduction to Database

- Implementing an information service with file processing
  - Challenges from
    - Data inconsistency

Applications have to manage data **consistently**

Update both

Jack	010-9876-5678	jwbyun@gmail.com
James	010-4321-8765	jms@gmail.com

010-9876-5678	Database	6	98
010-9876-5678	Database programming	7	96
010-9876-5678	Database analysis	7	92
010-9876-5678	Advanced database	8	99
010-4321-8765	Database	6	89
010-4321-8765	Database programming	7	84
010-4321-8765	Database analysis	7	81
010-4321-8765	Advanced database	8	82

Introduction to Database

- Implementing an information service with file processing
  - Challenges from
    - Data integrity

P3
Database should be
5th semester

Jack	010-9876-5678	jwbyun@gmail.com
James	010-4321-8765	jms@gmail.com

010-9876-5678	Database	**6**	98
010-9876-5678	Database programming	7	96
010-9876-5678	Database analysis	7	92
010-9876-5678	Advanced database	8	99
010-4321-8765	Database	6	89
010-4321-8765	Database programming	7	84
010-4321-8765	Database analysis	7	81
010-4321-8765	Advanced database	8	82

Introduction to Database

- Implementing an information service with file processing
  - Challenges from
    - Inefficient data access

P2: Grade average service is only built for providing average grade per person

P4

What if printing students, who have over 90 GPA, is required?
What if printing the number of students who takes database is required?
…

P5

Jack	010-1234-5678	jwbyun@gmail.com	Database	6	98
Jack	010-1234-5678	jwbyun@gmail.com	Database programming	7	96
Jack	010-1234-5678	jwbyun@gmail.com	Database analysis	7	92
Jack	010-1234-5678	jwbyun@gmail.com	Advanced database	8	99
James	010-4321-8765	jms@gmail.com	Database	6	89
James	010-4321-8765	jms@gmail.com	Database programming	7	84
James	010-4321-8765	jms@gmail.com	Database analysis	7	81
James	010-4321-8765	jms@gmail.com	Advanced database	8	82

Introduction to Database

- Implementing an information service with file processing
  - Challenges from
    - Transaction Management and Concurrency Control

a.txt
50

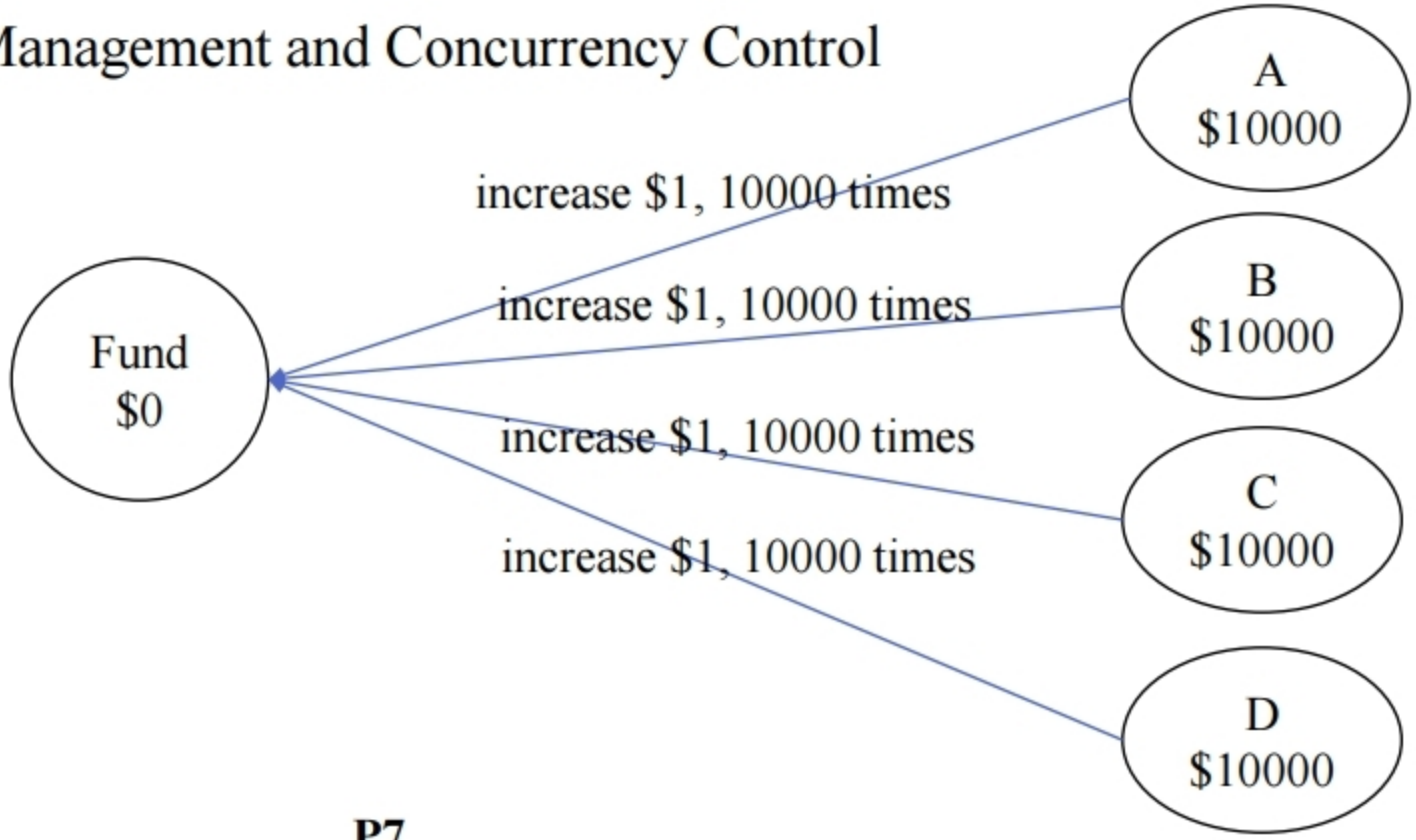
b.txt
250

1. a = a – 50
2. b = b + 50

P6
What if the system terminates before step 2?
Total balance should be same but different
Send $50 is not atomic operation

Introduction to Database

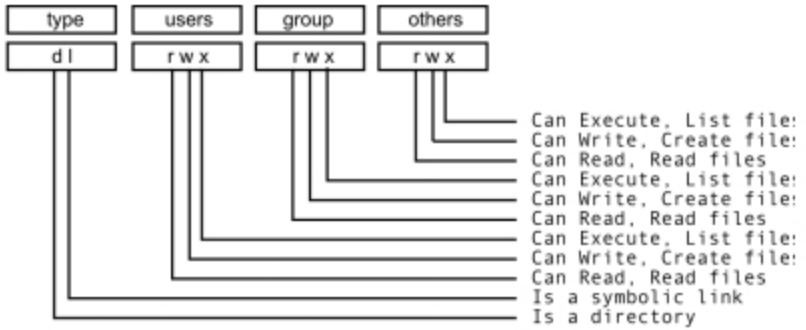
- Implementing an information service with file processing
  - Challenges from
    - Transaction Management and Concurrency Control



P7
Non atomic operation and concurrent access
Might yield synchronization problem
Send $1 is not atomic operation

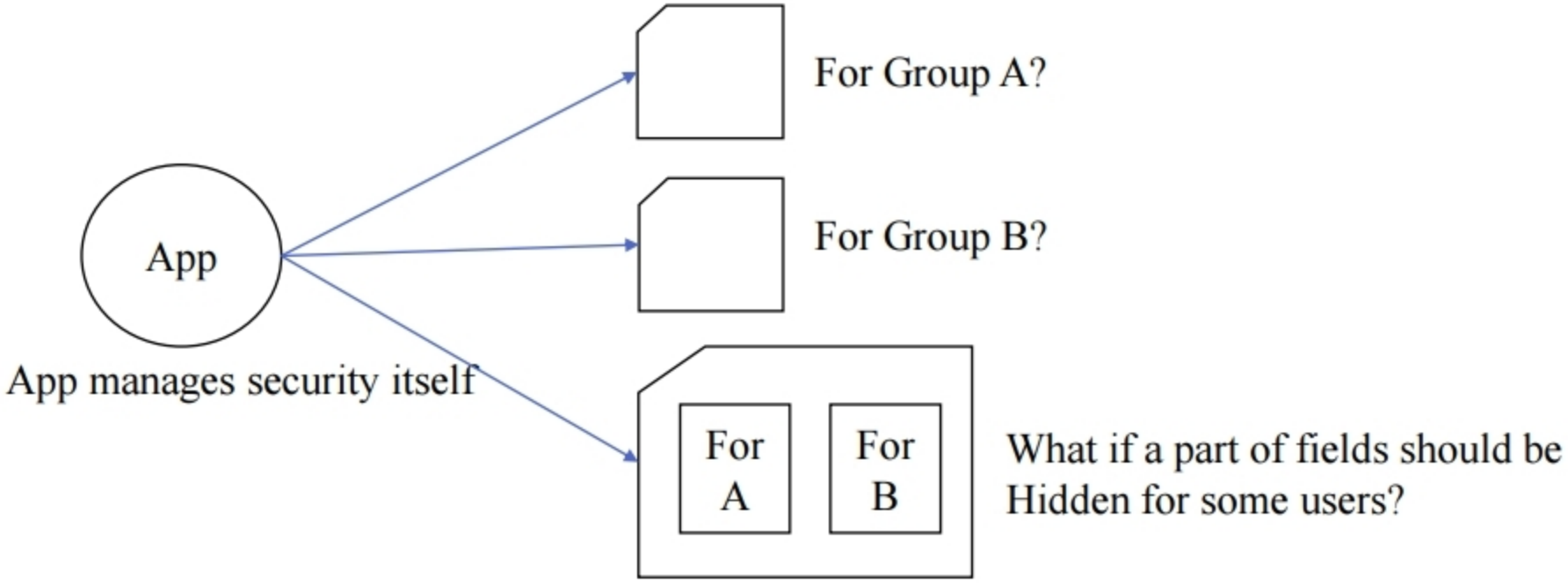
Introduction to Database

- Implementing an information service with file processing
  - Challenges from
    - Security



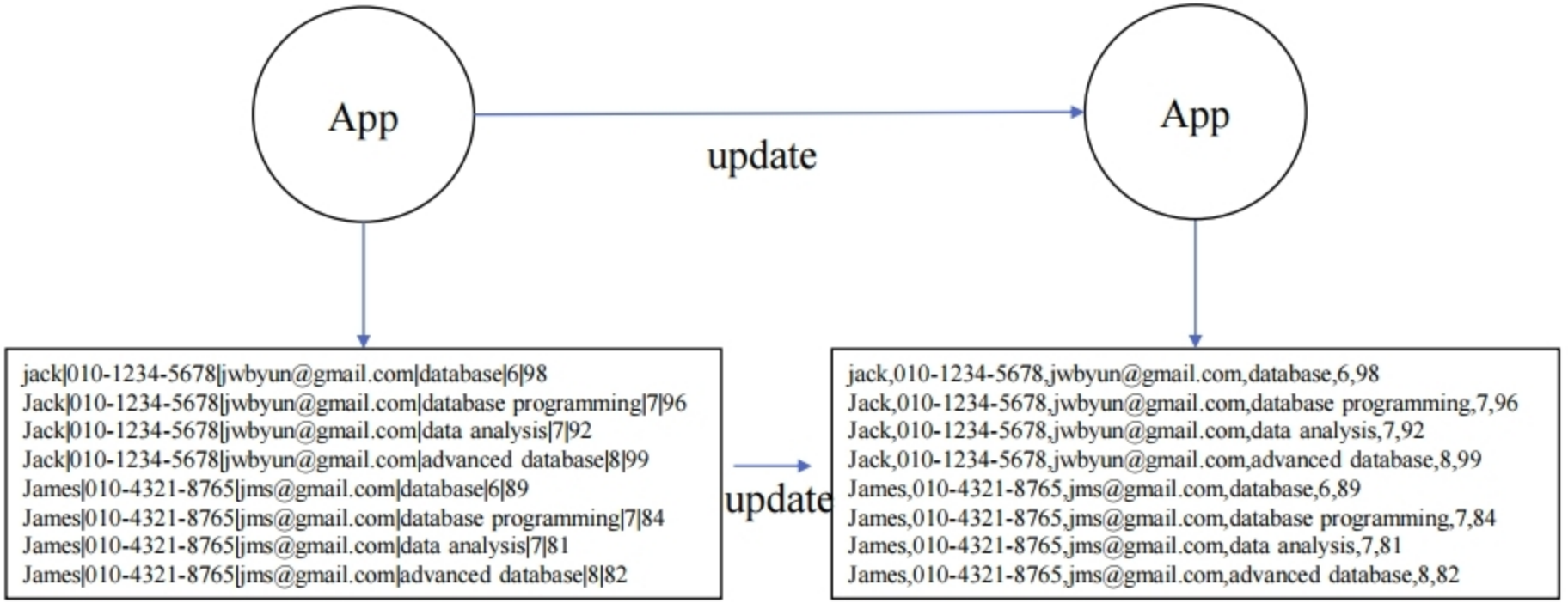
The overview of UNIX permissions field

https://www.samba.org/samba/docs/old/Samba3-HOWTO/AccessControls.html



Introduction to Database

- Implementing an information service with file processing
  - Challenges from
    - Data dependency
      - between applications and data
      - changes of internal structure affect the application



Introduction to Database

- Database Management System (DBMS)
  - A software system enabling to define, create, maintain and control access to the database

  - Resolving
    - Data redundancy: logical database design, …
    - Data inconsistency: concurrency control, integrity constraints, …
    - Data integrity: integrity constraints, …
    - Inefficient data access: SQL, …
    - Atomicity: transaction, …
    - Concurrent access: concurrency control, …
    - Security: privilege, …
    - Data dependency: dbms, view, …
    - …

  - Consists of
    - 3-level architecture
    - Data language
    - User
    - DBA
    - …

Introduction to Database

- Database Management System (DBMS)
  - 3-level architecture
    - Student Information

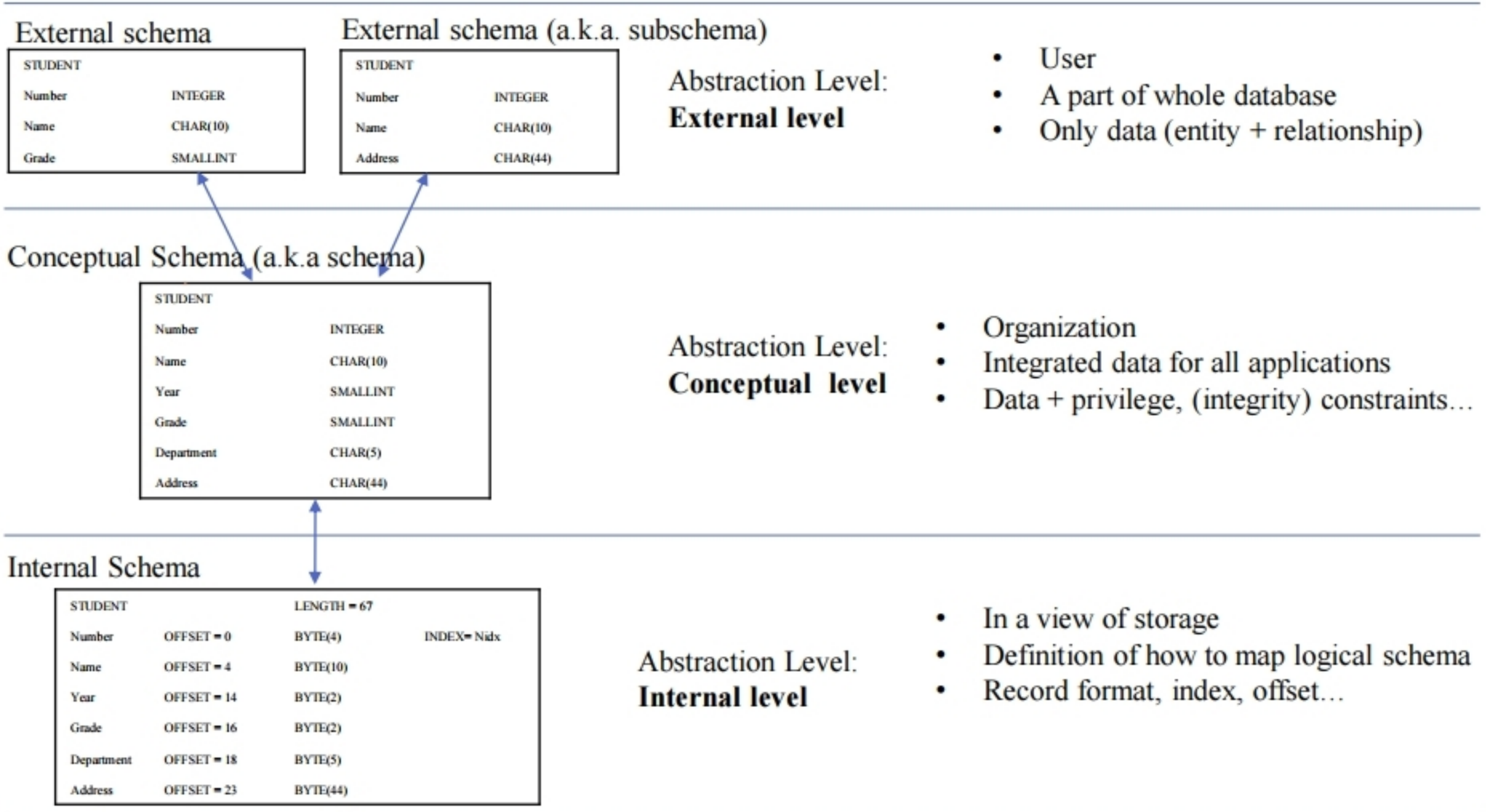
number	name	year	grade	department	address
615458	Jaewook Byun	2	A+	Software	604
20180123	Gildong Hong	3	A0	CS	613

Introduction to Database

Schema:
a blueprint of how the data is organized

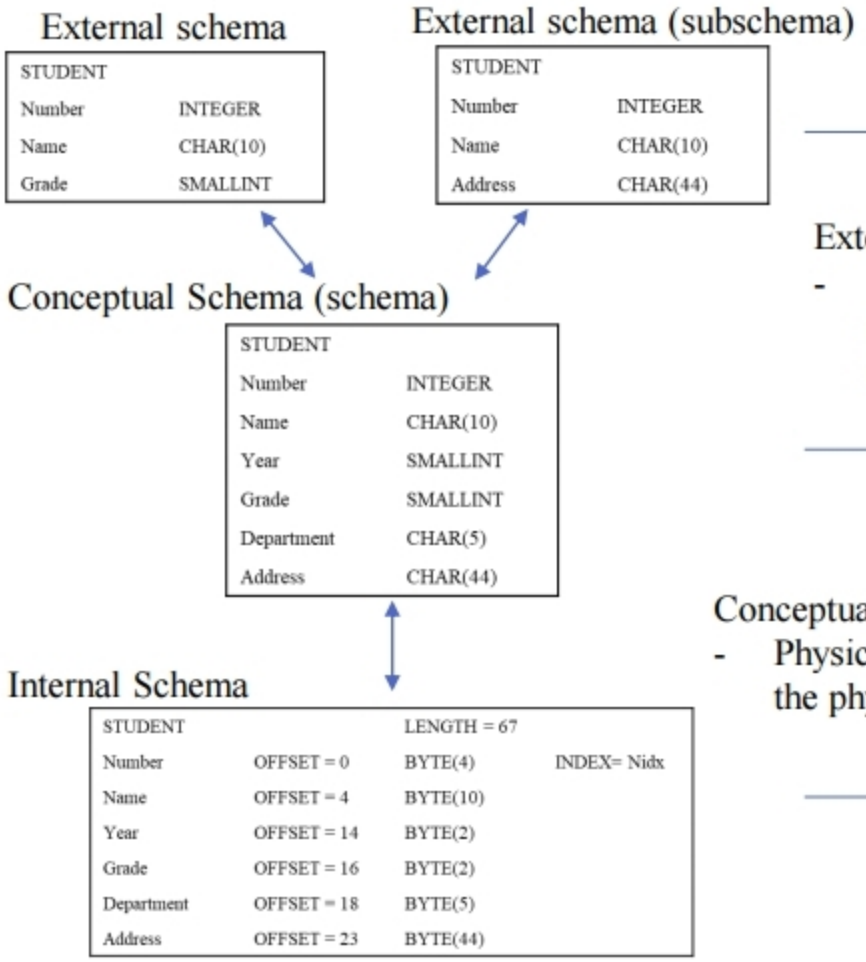
Instance:
a set of the data stored in DB at a specific time

- Database Management System (DBMS)
  - 3-level architecture
    - Student Information



Introduction to Database

- Database Management System (DBMS)
  - Data Independence



External-Conceptual Level Mapping

- Logical Data Independence : the ability to change the conceptual schema without changing applications or external schema (View)

Conceptual-Internal Level Mapping

- Physical Data Independence: the ability to easily change the physical storage structures with an effect on the conceptual schema

Introduction to Database

- Database Management System (DBMS)
  - Data language
    - A way to communicate with DBMS

    - Data Definition Language (DDL)
      - Define conceptual schema and subschema
      - Define internal schema with Data Storage Definition Language

    - Data Manipulation Language (DML)
      - Search, Insert, Delete, Update data
      - Procedural DML
        - What and how
        - Raw level DML: process records one by one
      - Non-procedural DML
        - What (declarative)
        - High level DML: process multiple records at a time
        - Query Language

    - Data Control Language
      - Security
      - Integrity
      - Recovery
      - Concurrency

Summary

- Introduction to Database

- Next
  - Relational Model and Algebra