

Database: Collection of interrelated data
A very large, integrated collection of data.

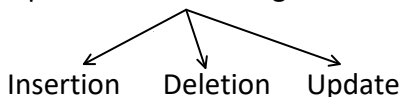
Database Applications Examples:

- A university database might contain information about the following:
 - i. Entities such as students, faculty, courses, and classrooms.
 - ii. Relationships between entities, such as students' enrollment in courses, faculty teaching courses, and the use of rooms for courses.
 - Banking: Customer information, accounts, loans, banking , credit card transactions.
 - Telecommunication: records of calls, texts, and data usage, generating monthly bills, maintaining balances on prepaid calling cards, etc.
-

What is DBMS?

A Database Management System (DBMS) is a software package designed to manage and access databases.

Operations for Management



File systems + OS have the following **limitations**, which are overcome by DBMS.

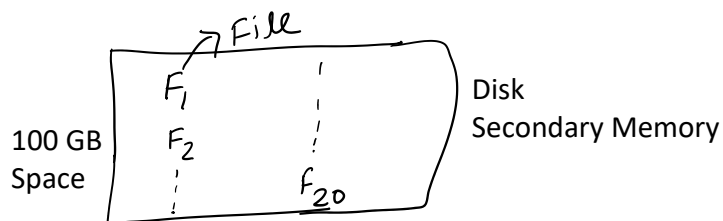
FILE SYSTEMS VERSUS A DBMS

1. Data independence

- a. Application programs should be as independent as possible from details of data representation and storage.
- b. The DBMS can provide an abstract view of the data to insulate application code from such details.

University database
contains

- Students information
- Faculty information
- Admin information



Retrieve all students who got more than 7 CGPA

To get this, we need the following physical details of the file:

- Location of student file
- File Name
- Permission (Accessibility)
- Extension (.txt, .doc, etc)

Application development became too complex as we required many physical details, so File Systems + OS can't handle it.

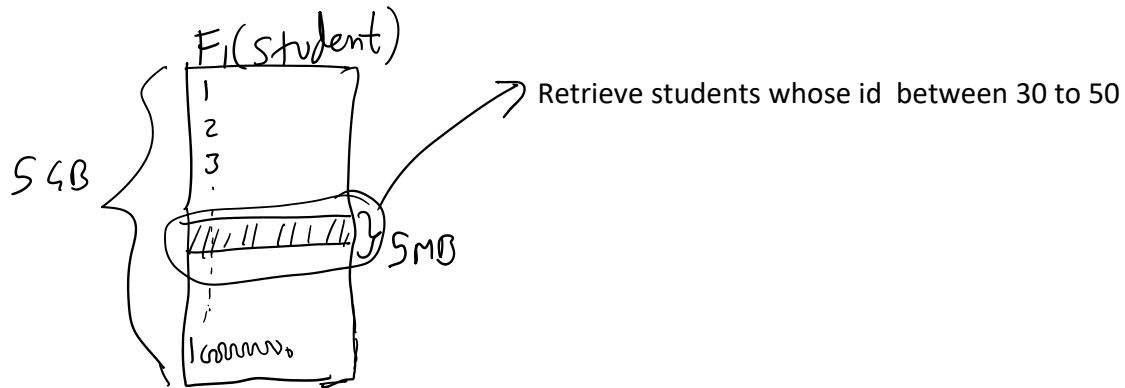
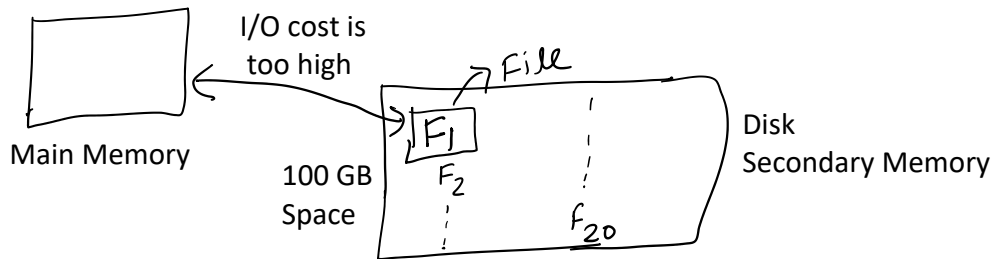
Application programmers should be able to access the database without much knowledge about physical details.

For example, in the below query, we do not need physical details.

```
Select * from student where CGPA >= 7
```

2. Efficient data access

- a. A DBMS utilizes a variety of sophisticated techniques to store and retrieve data efficiently.
- b. This feature is especially important if the data is stored on external storage devices.



Whenever we require data, we first have to transfer the entire big file into the main memory to search the data.

Instructions and data should be stored in high-speed main memory.

I/O Cost: The number of blocks required to transfer from secondary memory to main memory is too high.

Sophisticated Indexing: index the data within the file.

Due to the index, we have to transfer only a few blocks from SM to MM.

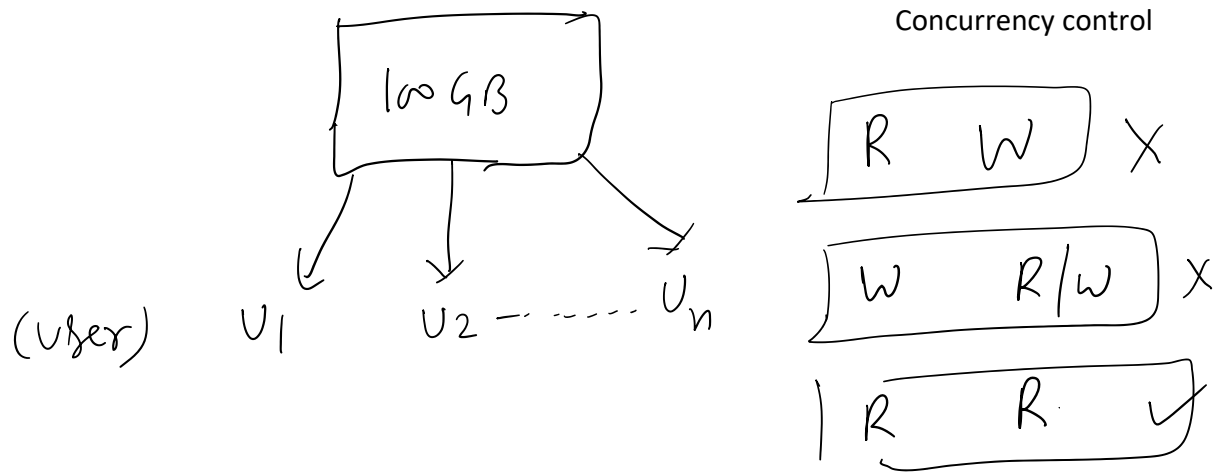
I/O cost reduces drastically.

Indexing techniques are used in DBMS. Therefore, the I/O cost is too low.

The above problem will not arise if the file size is small.

3. Concurrent access and crash recovery

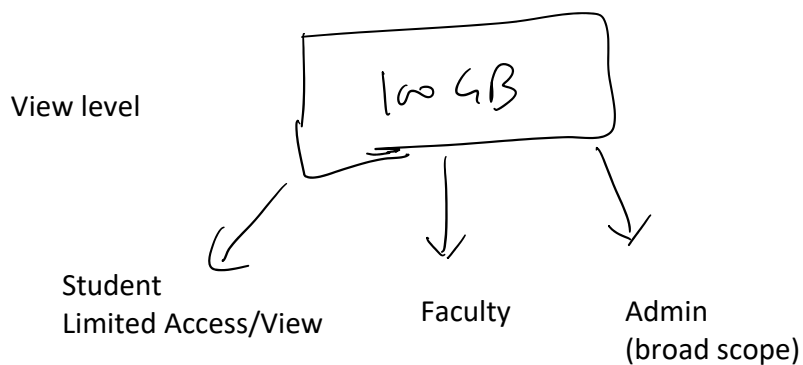
- Must protect data from inconsistency due to multiple concurrent users.
- A DBMS schedules concurrent accesses to the data in such a manner that users can think of the data as being accessed by only one user at a time. Further, the DBMS protects users from the effects of system failures.



DBMS provides concurrency control by using various protocols.

4. Data integrity, security and access control

- If data is always accessed through the DBMS, the DBMS can enforce integrity constraints on the data. For example, before inserting salary information for an employee, the DBMS can check that the department budget is not exceeded.
- Also, the DBMS can enforce access controls that govern what data is visible to different classes of users.



Different levels of security by DBMS

File systems + OS can not handle it as it does not provide different views.