CSCI 460—Operating Systems

Lecture 9

File Management

Textbook: Operating Systems
by William Stallings

1. Basic Concepts

- File Manager controls all the files in the system.

- Several factors determine the efficiency of a file manager

    - 1. How are the files organized: sequential, direct, indexed sequential.

    - 2. How are the files stored: contiguously, non-contiguously, indexed.

    - 3. How are each file's records structured: fixed-length, variable length.

- When you are creating, deleting, modifying and controlling access to a file, a file manager is functioning.

- Responsibilities of a file manager

    - 1. keep track of where each file is stored.
    - 2. Allocate each file with an access right and record its use.
    - 3. Deallocate a file when it is deleted.
    - 4. Maintain available storage space for future use.

- Some basic definitions
    - 1. A **field**
    - 2. A **record**
    - 3. A **file**

Student-No	Name	GPA	address
-02970017	J. Paxton	3.3	- - -.
- · · -			
_00234568	B. Zhu	3.2	- - -.

A field: a cell

A record: a line

A file: a group of records.

2. File Organization

- **Record Format:** Fixed-length records and variable-length records.

  - 1. **Fixed-length records**: easy to access, ideal for data files. If the size is too small some 'left-over' characters are truncated. If the size is too large then spaces are wasted.

  - 2. **Variable-length records**: do not leave empty space and do not truncate characters.

- **Physical File Organization**

  - 1. Physical file organization is decided by the way records are arranged.

  - 2. Physical file organization is also decided by what kind of medium are used to store the files.

    tape vs. disk

disk

- Physical File Organization is composed of

  - 1. Sequential record organization
  - 2. Direct record organization
  - 3. Indexed sequential record organization

index: think of index at the end of a textbook.

- To decide which way you use to store the data, you need to look at the following parameters

  - 1. Volatility of the data: the frequency with which additions and deletions are made.
  - 2. Activity of the file: the percentage of records processed during a given run.
  - 3. Size of the file.
  - 4. Response time: the amount of time user is willing to wait.

- Sequential record organization

  - 1. Records are stored and retrieved sequentially
  - 2. An optimization version of it is to select a key and sort all the records by the key, of course, we have to pay for the overhead.

- Direct record organization

  - 1. Records are identified by relative addresses — their addresses relative to the beginning of the file.
  - 2. We can again identify a **key** — which uniquely identifies a record. Then we can use a hashing technique, say $h(i) = \lfloor m \times [(i \times A) mod 1] \rfloor, A = \frac{\sqrt{5}-1}{2}$.
  - 3. Although *collision* is not avoidable, hashing is widely used in practice.

- Indexed sequential record organization

  - 1. It is created and maintained through an Indexed Sequential Access Method. *This is very much like the contents in a textbook, if you want to search for a specific topic you look up the index and then go to the physical location indicated by that entry.*
  - 2. Again, overhead is the drawback.

B-tree

height is usually small.