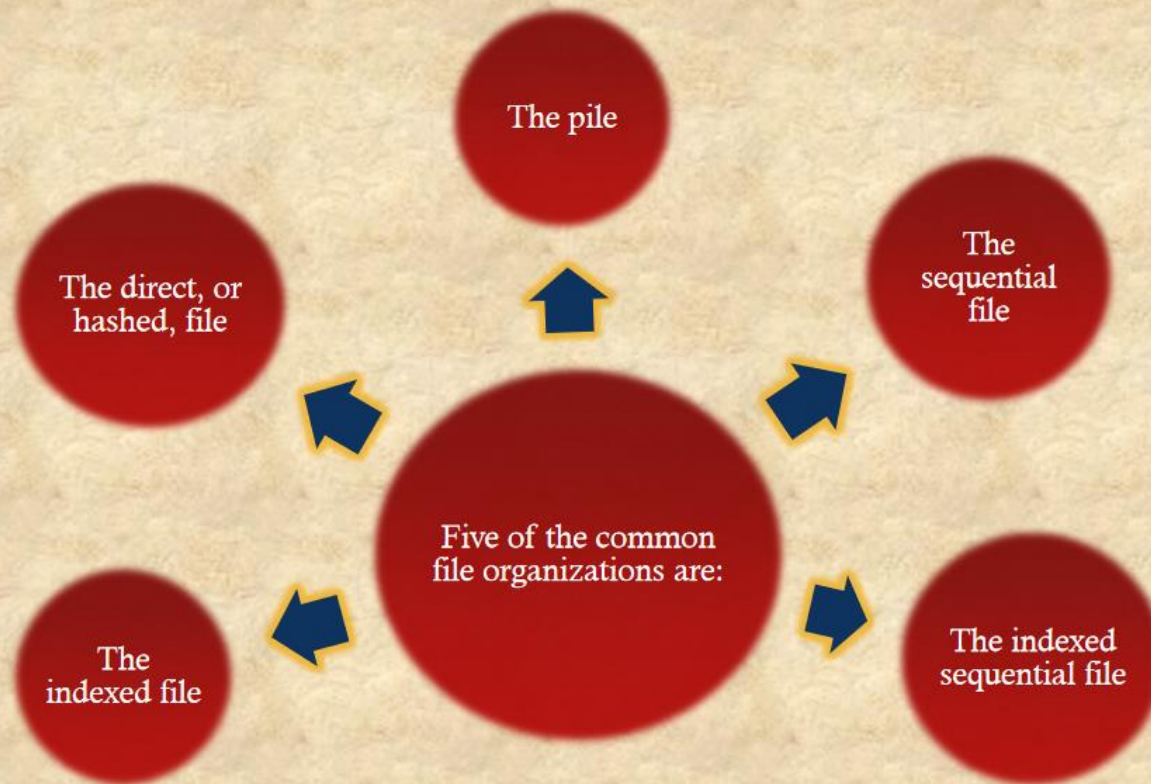


File Organization and Access

- *File organization* is the logical structuring of the records as determined by the way in which they are accessed
- In choosing a file organization, several criteria are important:
 - short access time
 - ease of update
 - economy of storage
 - simple maintenance
 - reliability
- Priority of criteria depends on the application that will use the file

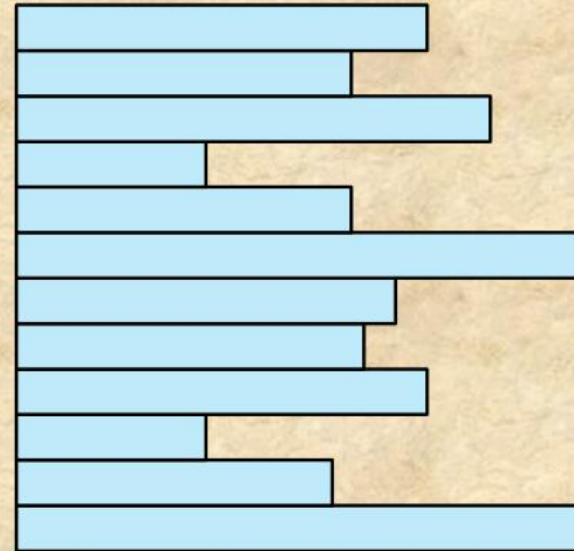


File Organization Types



The Pile

- Least complicated form of file organization
- Data are collected in the order they arrive
- Each record consists of one burst of data
- Purpose is simply to accumulate the mass of data and save it
- Record access is by exhaustive search



Variable-length records
Variable set of fields
Chronological order

(a) Pile File

The Sequential File

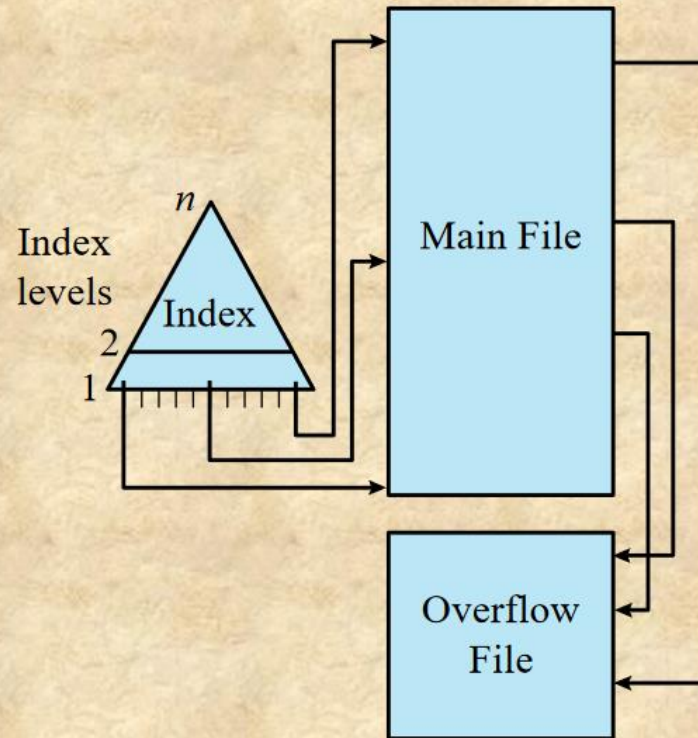
- Most common form of file structure
- A fixed format is used for records
- Key field uniquely identifies the record
- Typically used in batch applications
- Only organization that is easily stored on tape as well as disk

Fixed-length records
Fixed set of fields in fixed order
Sequential order based on key field

(b) Sequential File

Indexed Sequential File

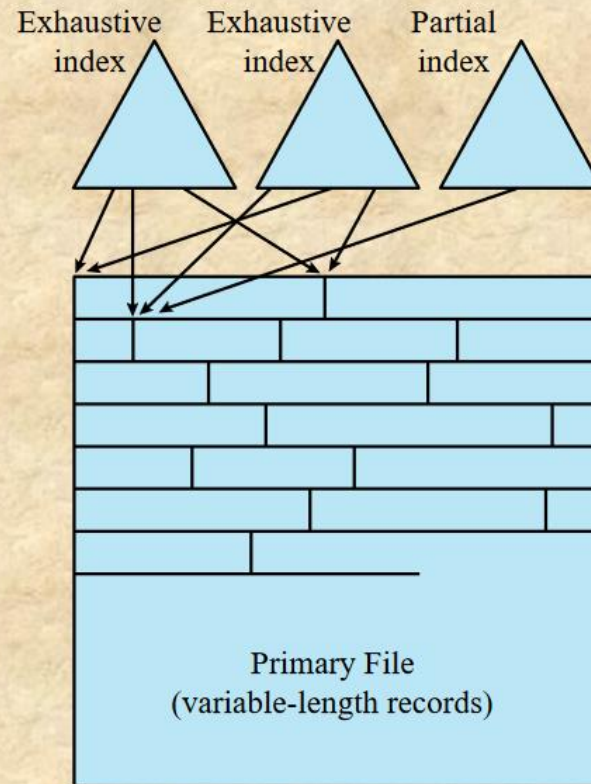
- Adds an index to the file to support random access
- Adds an overflow file
- Greatly reduces the time required to access a single record
- Multiple levels of indexing can be used to provide greater efficiency in access



(c) Indexed Sequential File

Indexed File

- Records are accessed only through their indexes
- Variable-length records can be employed
- Exhaustive index contains one entry for every record in the main file
- Partial index contains entries to records where the field of interest exists
- Used mostly in applications where timeliness of information is critical
- Examples would be airline reservation systems and inventory control systems



(d) Indexed File

Direct or Hashed File

- Access directly any block of a known address
- Makes use of hashing on the key value
- Often used where:
 - very rapid access is required
 - fixed-length records are used
 - records are always accessed one at a time

Examples are:

- directories
- pricing tables
- schedules
- name lists