

## File Organization

### Types of Access

**Serial:** Records are accessed in the order in which they were stored.

**Sequential:** Records are accessed in descending or ascending key sequence.

**Storage Medium:** Magnetic Tape. **Information Processing Mode:** Batch.

**Random/Direct:** Records are accessed without any regard to the order in which they were stored.

**Storage Medium:** Magnetic disk, floppy, hard disk, optical disk, DVD.

**Information Processing Mode:** Realtime, online

### Terms

**File/Database:** A logical collection of tables.

**Table:** A logical collection of records.

**Record:** A logical collection of fields.

**Field:** The smallest section of a record which can be manipulated separately.

**Key:** A unique field used to identify a record. For example, Student/Employee/ Account/ Registration Numbers.

**Source Data Automation:** The use of automated methods of data entry.

**Index:** A file used to access a large data file quickly. It contains a key field from the file the index is for and the address of the record.

**Master File:** The permanent source of data for a particular computer application area (e.g. account master file in a bank) which is kept up-to-date and stores the main information.

**Transaction File:** A temporary file which is used to update the master file.

**Change File or Transaction Log:** This file keeps a record of the changes made to the transaction file.

**Transactions:** Events that occur as part of doing business. These include sales, purchases, deposits, withdrawals and payments.

### **Sequential File Organization**

Records of the file are organized by physical sequence only. To access a data record all records before the one in question must be passed first, that is serial access. However, if the records are stored in key sequence, then sequential access is possible.

#### **Features**

- Records are stored and processed in a predetermine order which is based on their physical location.
- Magnetic tape, magnetic disk or punch cards may be used for this file organization.
- Files are not available for processing until the next update. This file organization method is often used by batch processing systems.

## Updating a Record

1. Both the master file and the transaction file must be sorted on the same key sequence.
2. All old master file records with a key less than the record to be changed are read and written to the new master file.
3. The record to be updated is read and edited. The updated record is written to the new master file.
4. All the remaining old master file records are written to the new master file.

## Inserting a Record

1. All records before the record to be inserted are written to the new master file.
2. The new record is written to the new master file.
3. The remaining old master file records are written to the new master file.

## Deleting a Record

1. Compare each record with the key of the record to be deleted.
2. All old master file records with a key less than the record to be deleted are written to the new master file.
3. When the record to be deleted is encountered it is not written to the new master file.
4. The remaining old master file records are written to the new master file.

## Processing Errors

- Trying to delete or update a record that does not exist.
- Trying to insert a record that is already in the file
- No data in the transaction file to update the master file.

## **Direct File Organization**

Records are organized on the basis of the key of the record. The key is used to determine where the record is stored. The key is converted to a relative address by using a special algorithm. This process is called Hashing. With a direct processing system transactions are processed as they occur without having to be pre-sorted into batches. Such processing may be used with Airline/Hotel reservations, Accounting, and Law enforcement systems.

### **Features**

- Data is processed continuously and not in batches.
- When updating occurs, the original content of the updated record is lost .This method of updating is called destructive or overlay updating.
- Records are not stored or access in any given sequence.
- Magnetic disk is the storage medium used.
- Records can be updated without accessing other records and it is not necessary to create a new file.
- Used with files which require frequent updates.
- Used with processes which require fast access to records.

## **Index Sequential File Organization**

This allows for both direct and sequential processing. Data is written to the disk as a sequential file but some sections of the disk are reserved to record the address of data records. These addresses may be used as an index to obtain data randomly or the index may be ignored and the data access sequentially.

### **Features**

- Used where periodic and random updates occur. For example, a savings and loan company.
- Instead of having an index entry for each record, there is an entry for each block of records, starting from the lowest to the highest.

### **Searching for a Record**

1. Compare the search key with the index keys to find the highest index key that comes before the key you are searching for.
2. Perform a linear search from the highest index key located at step 1 until the search key is found or until the record pointer to the next index entry is reached

