Indexing and Hashing

Content

- Indexing and Hashing
- Ordered Indices

Indexing

- Indexing is used to optimize the performance of a database by minimizing the number of disk accesses required when a query is processed.

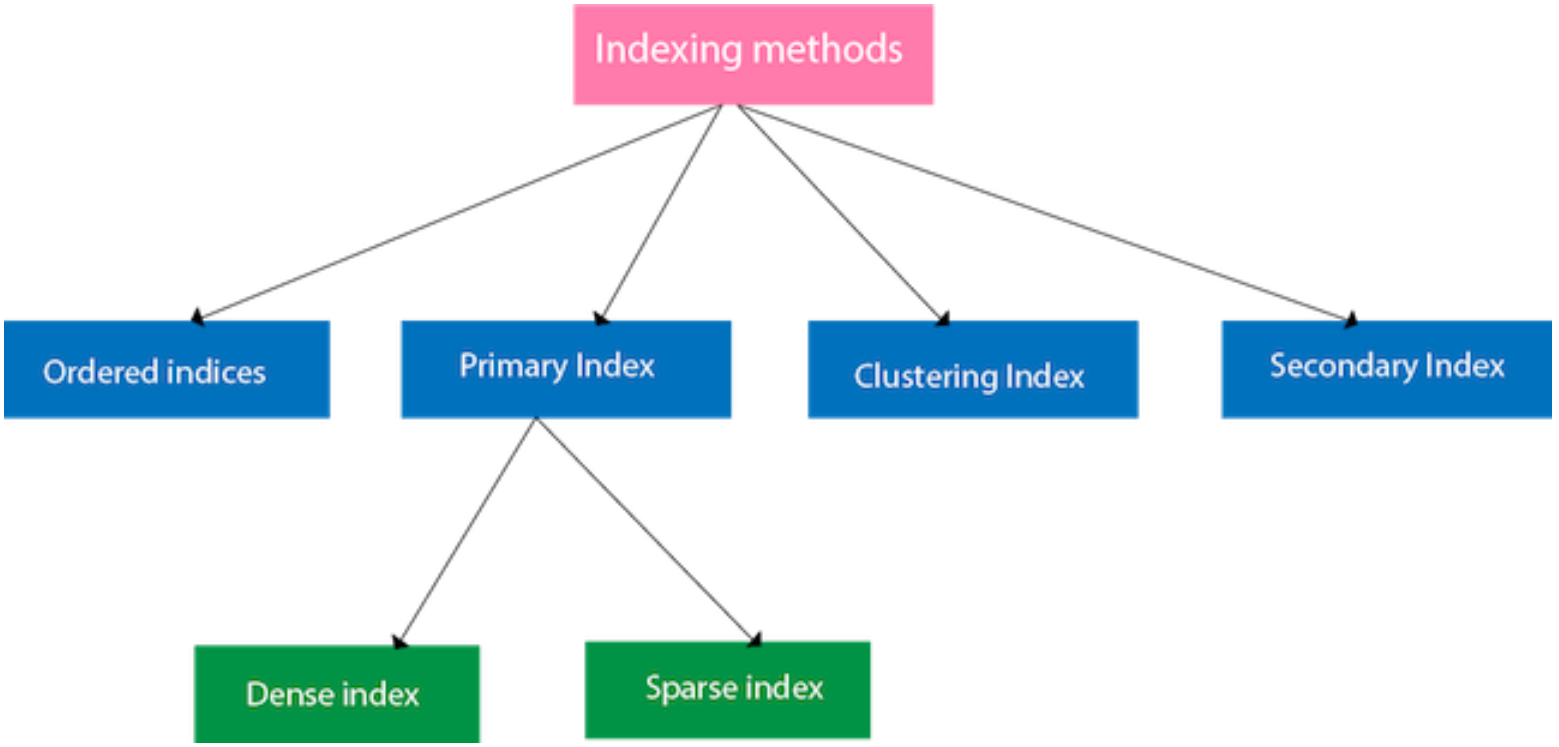
Index structure:

- Indexes can be created using some database columns.

Search key	Data Reference

Fig: Structure of Index

Indexing Methods



Indexing Methods

Primary Index

- Primary index is defined on an ordered data file. The data file is ordered on a **key field**. The key field is generally the primary key of the relation.
- The primary index can be classified into two types: Dense index and Sparse index.

Indexing Methods

Secondary Index

- Secondary index may be generated from a field which is a candidate key and has a unique value in every record, or a non-key with duplicate values.

Clustering Index

- Clustering index is defined on an ordered data file. The data file is ordered on a non-key field.

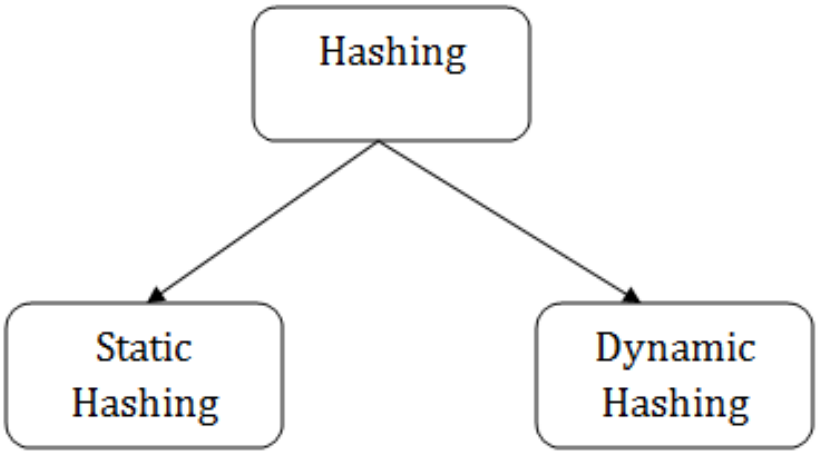
Ordered indices

- The indices are usually sorted to make searching faster. The indices which are sorted are known as ordered indices.
- **Example**: An employee table with thousands of record and each of which is 10 bytes long. If their IDs start with 1, 2, 3....and so on and we have to search student with ID-543.
- Search the disk block from starting till it reaches 543. The DBMS will read the record after reading 543*10=5430 bytes.
- DBMS will read the record after reading 542*2= 1084 bytes which are very less compared to the previous case.

Hashing

- Hashing is an effective technique to calculate the direct location of a data record on the disk without using index structure.
- Hashing uses hash functions with search keys as parameters to generate the address of a data record.

Types of Hashing:



THANK YOU