DBMS Lec 11 (INDEXING) :

This video explains database indexing and its role in optimizing the performance of data searches. It covers different types of indexes—primary, secondary, and clustering—and shows how they help make data retrieval faster and more efficient.

Introduction to Performance Optimization

Performance optimization in databases involves using techniques that make data access faster and more efficient. Indexing is a key technique where additional data structures are created to help locate data quickly.

Real-World Example of Data Access

Accessing data directly from large files without indexes is slow because the system must search through all records. Indexing helps by providing shortcuts to the data, reducing search time.

How Indexing Works

Indexing creates a separate structure (like a table of contents) that points to the location of data in the main file, allowing for much quicker searches by narrowing down the search area.

Primary Indexing (Clustering) Explained

A primary index is built on a sorted data file using a unique key (like a roll number) not always PK. It allows fast access to data blocks by pointing directly to the first record of each block.
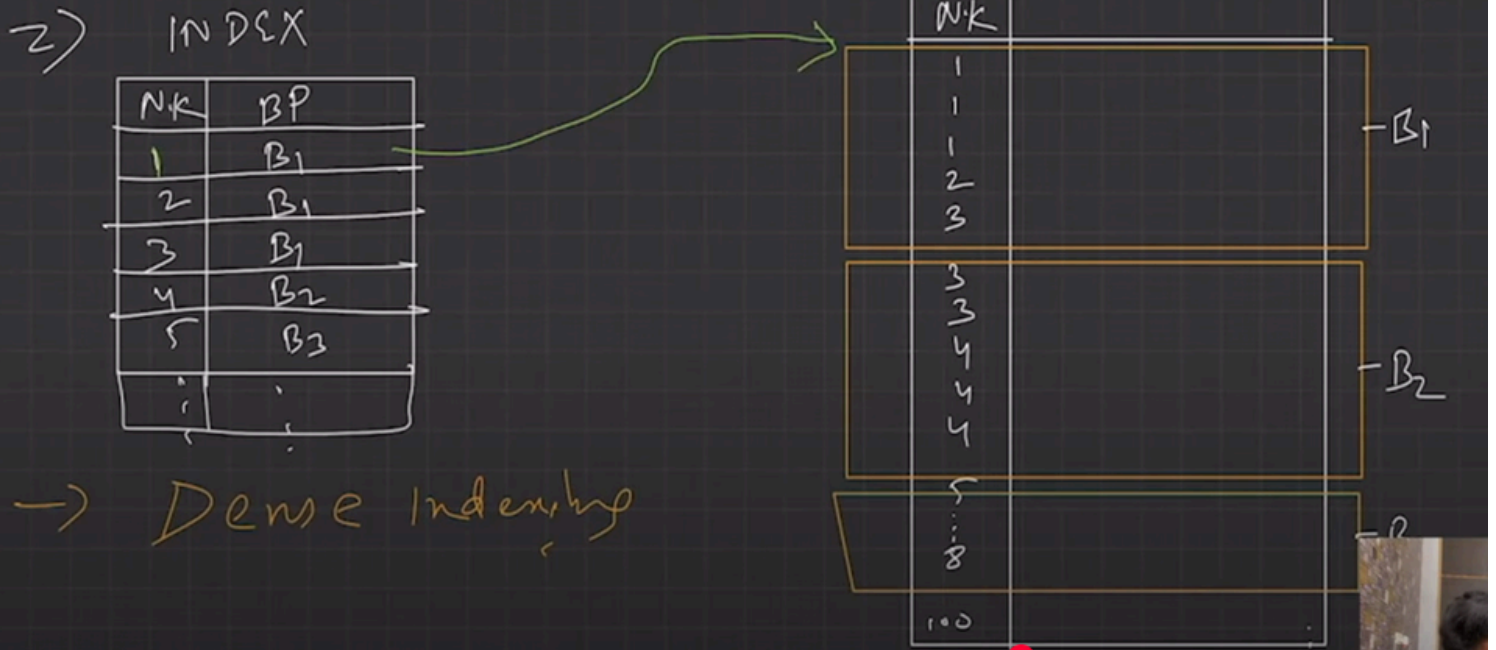
Sparse Indexing and Its Use

Sparse indexing means not every record has an index entry; instead, only some records (like the first in each block) are indexed, saving space while still improving search speed.

Dense Indexing and Its Use

Dense indexing means every record has an index entry, only some records (like the first in each block) are indexed, saving space while still improving search speed. In below image we can see that data in disk under blocks is sorted but repeated.

- We would have to store entry of every unique data from disk into indexing table.

Clustering Indexing

Clustering indexes are used when data is grouped by a non-unique attribute (like department). The index points to the first record of each group, making group-based searches efficient.

Multi-Level Indexing

Multi-level indexing adds another layer of indexes on top of the original index, making searches even faster for very large datasets by further reducing the number of records to scan.

Secondary Indexing

Secondary indexes are used on attributes that are not unique or not sorted, allowing searches on columns other than the primary key. They are especially useful for queries on unsorted data.

Indexing Summary and Practical Benefits

Indexing, whether primary, secondary, or clustering, greatly improves data retrieval speed, especially in large databases. However, they require extra storage and maintenance.