Indexing

K.Karthikeyan

Overview

- Three main techniques
  - Conventional indexes
    - Think of a page table, …
  - B and B+ trees
    - Perform better when records are constantly added or deleted
  - Hashing

Conventional indexes

Indexes

- A **database index** is a data structure that improves the speed of data retrieval operations on a database table at the cost of additional writes and storage space to maintain the index data structure.

Wikipedia

Types of indexes

- An index can be
  - ***Sparse***
    - One entry per data block
    - Identifies the first record of the block
    - Requires data to be sorted
  - ***Dense***
    - One entry per record
    - Data do not have to be sorted

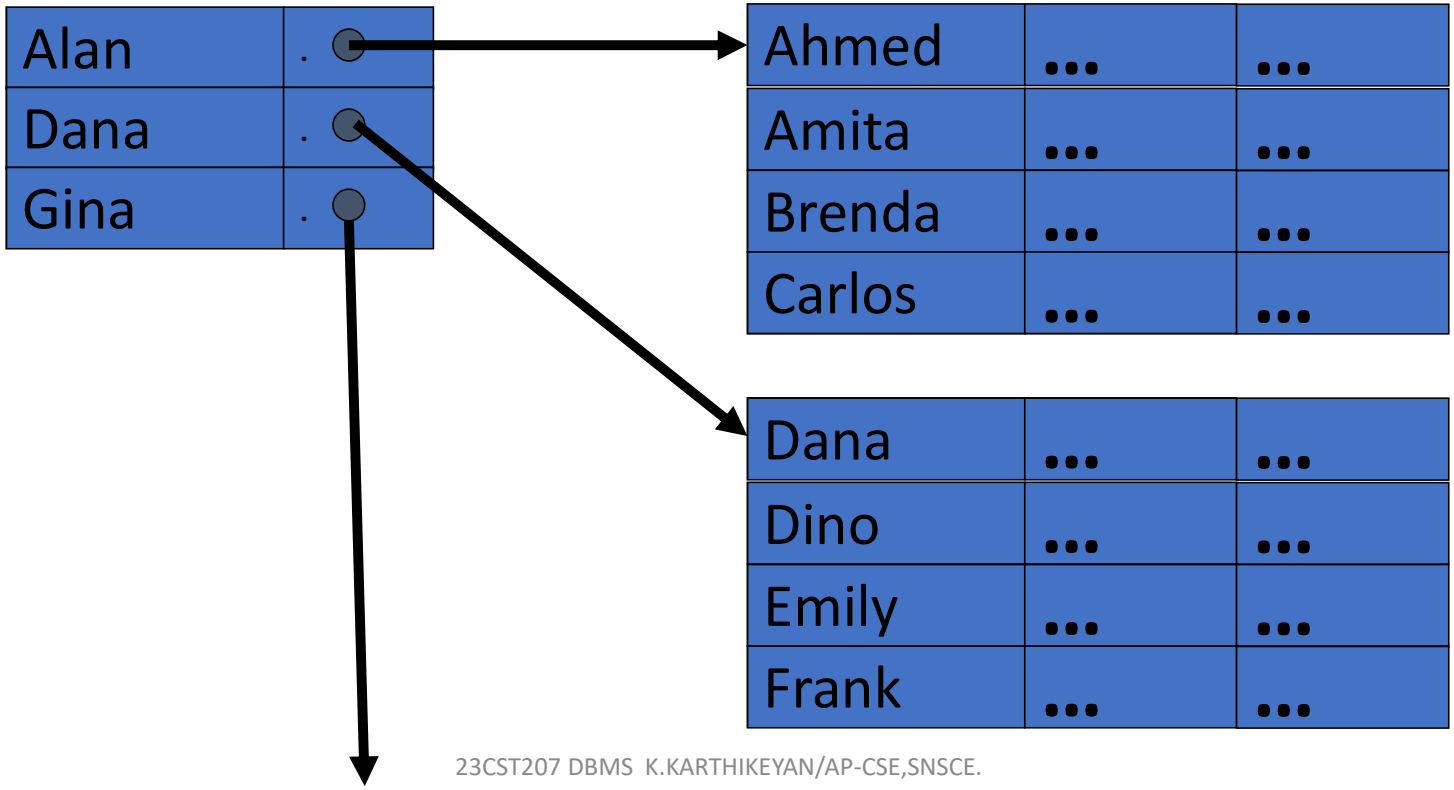
Respective advantages

- ***Sparse***
  - Occupy much less space
  - Can keep more of it in main memory
    - *Faster access*
- ***Dense***
  - Can tell if a given record exists without accessing the file
  - Do not require data to be sorted

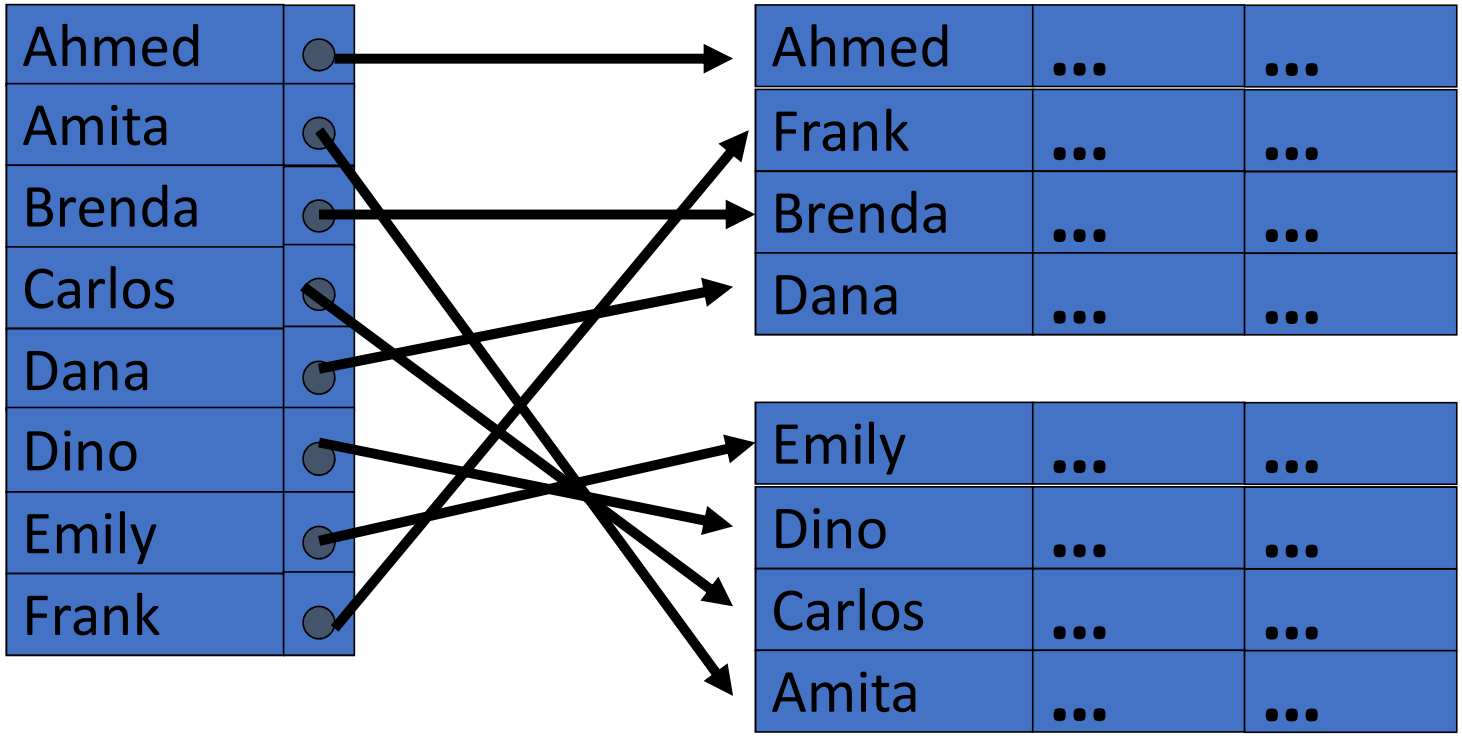
Indexes based on primary keys

- Each key value corresponds to a specific record
- Two cases to consider:
  - Table is sorted on its primary key
    - Can use a sparse index
  - Table is either non-sorted or sorted on another field
    - Must use a dense index

Sparse Index



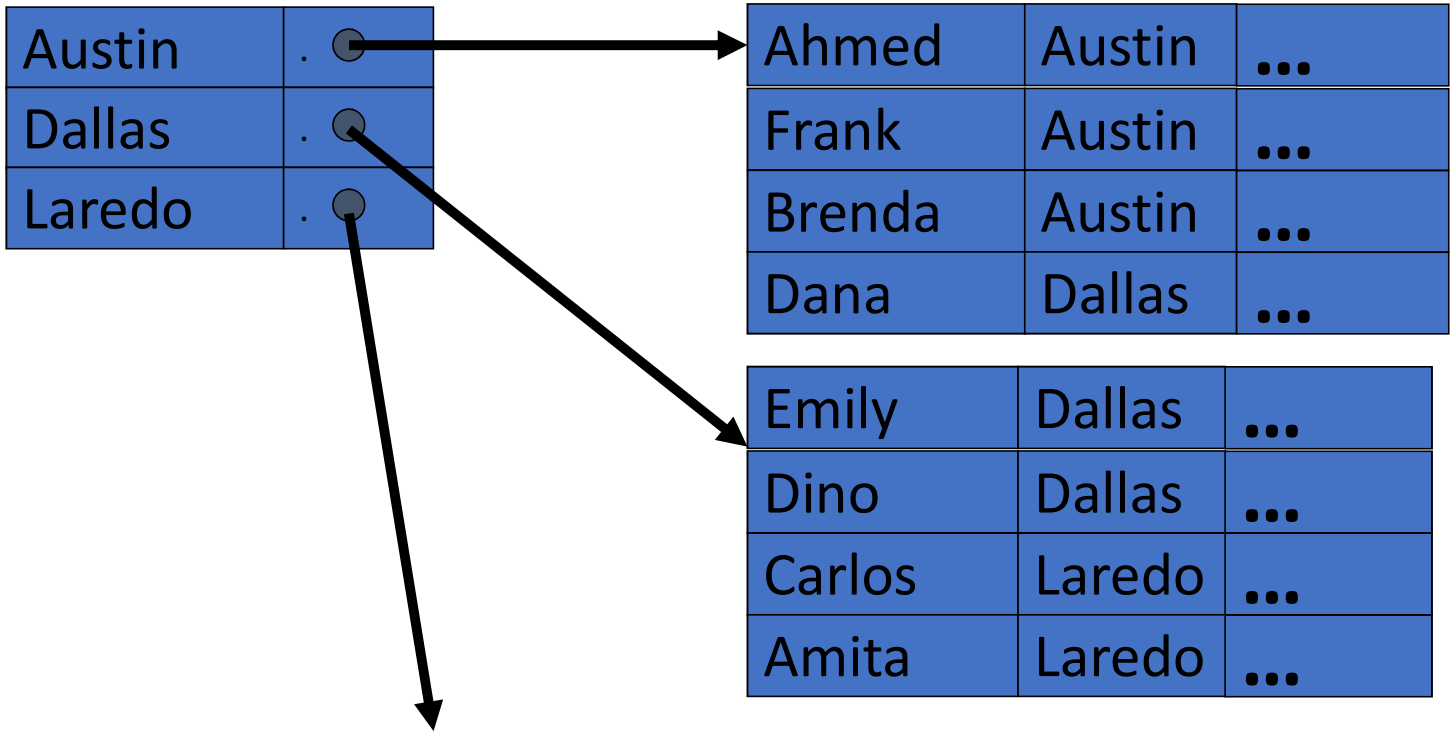
Dense Index


Ahmed
Amita
Brenda
Carlos
Dana
Dino
Emily
Frank
Ahmed ... ...
Frank ... ...
Brenda ... ...
Dana ... ...
Emily ... ...
Dino ... ...
Carlos ... ...
Amita ... ...

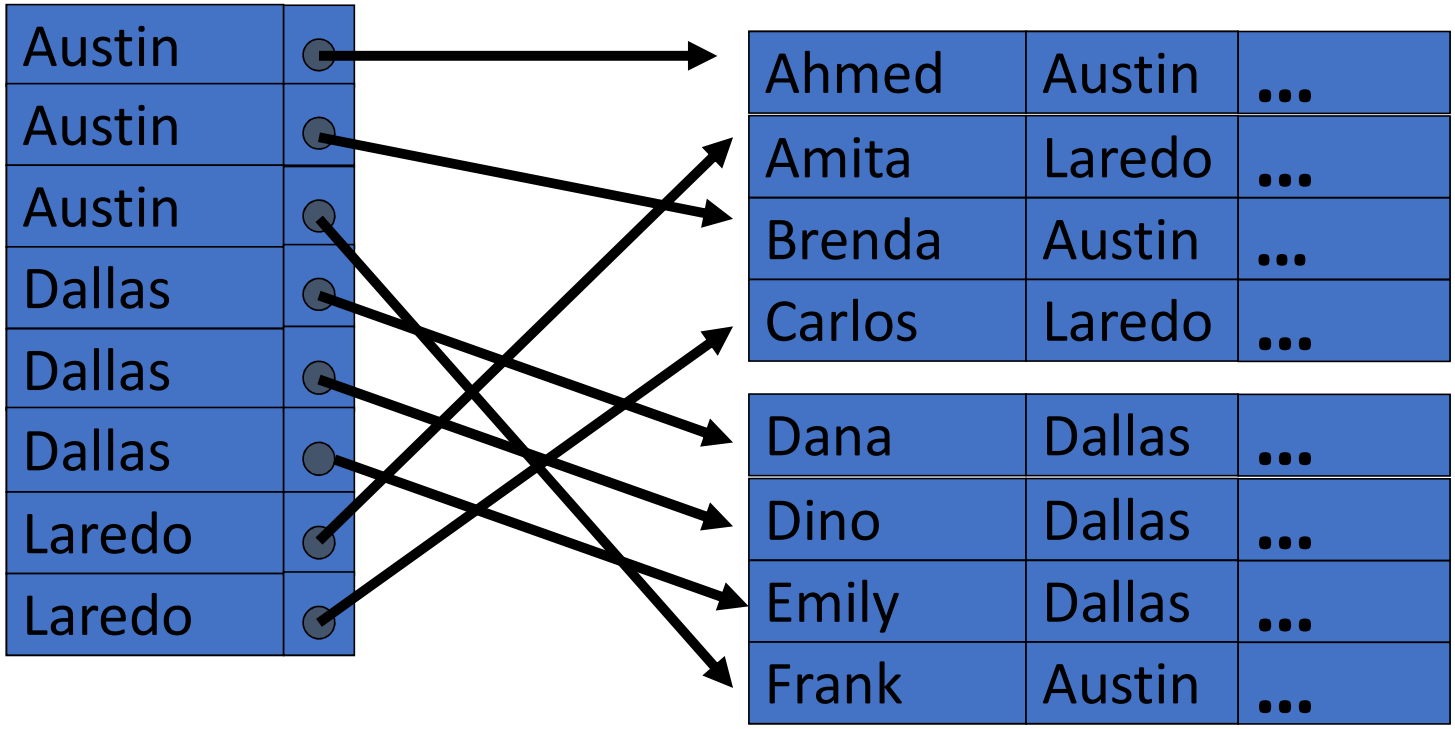
Indexes based on other fields

- Each key value may correspond to more than one record
  - ***clustering index***
- Two cases to consider:
  - Table is sorted on the field
    - Can use a sparse index
  - Table is either non-sorted or sorted on another field
    - Must use a dense index

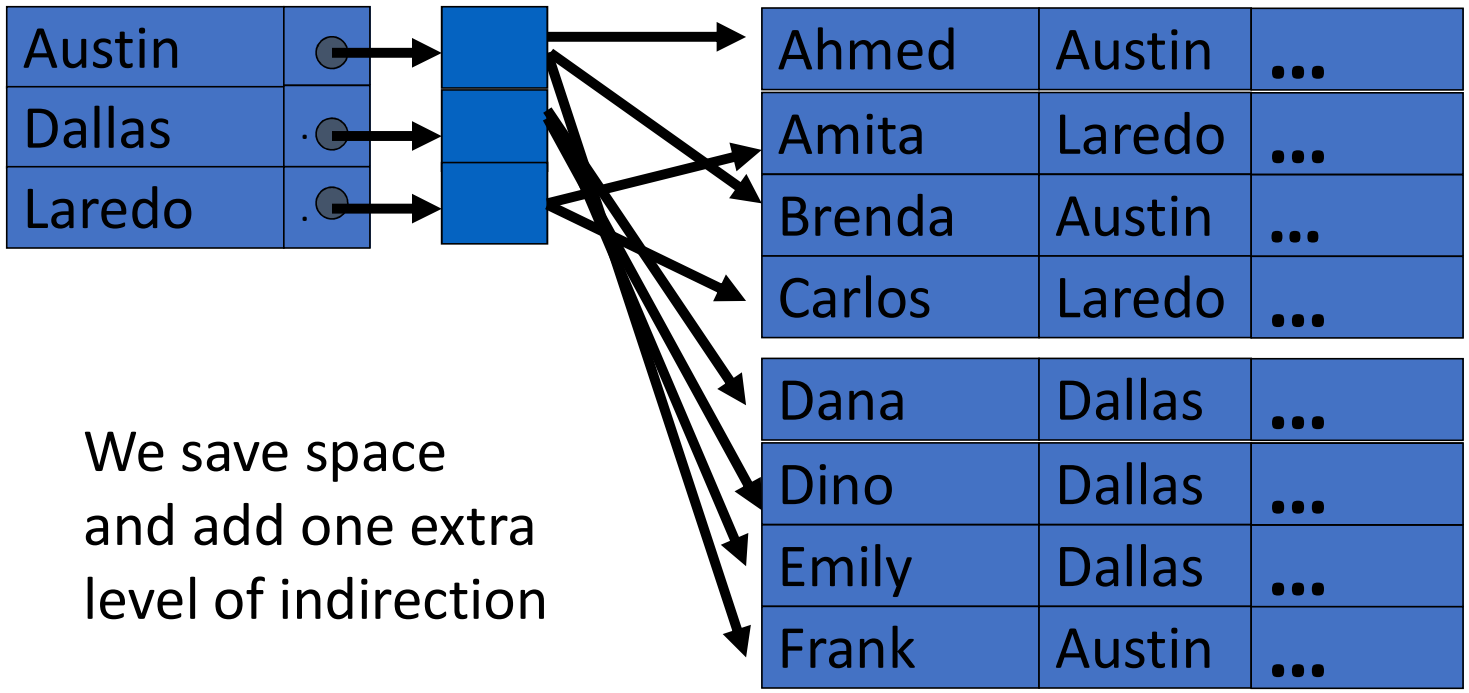
Sparse clustering index



Dense clustering index



Another realization



We save space
and add one extra
level of indirection

Activity

Fill in the blanks

- Types of indexes
- A)_____
- B)_____

Thank you