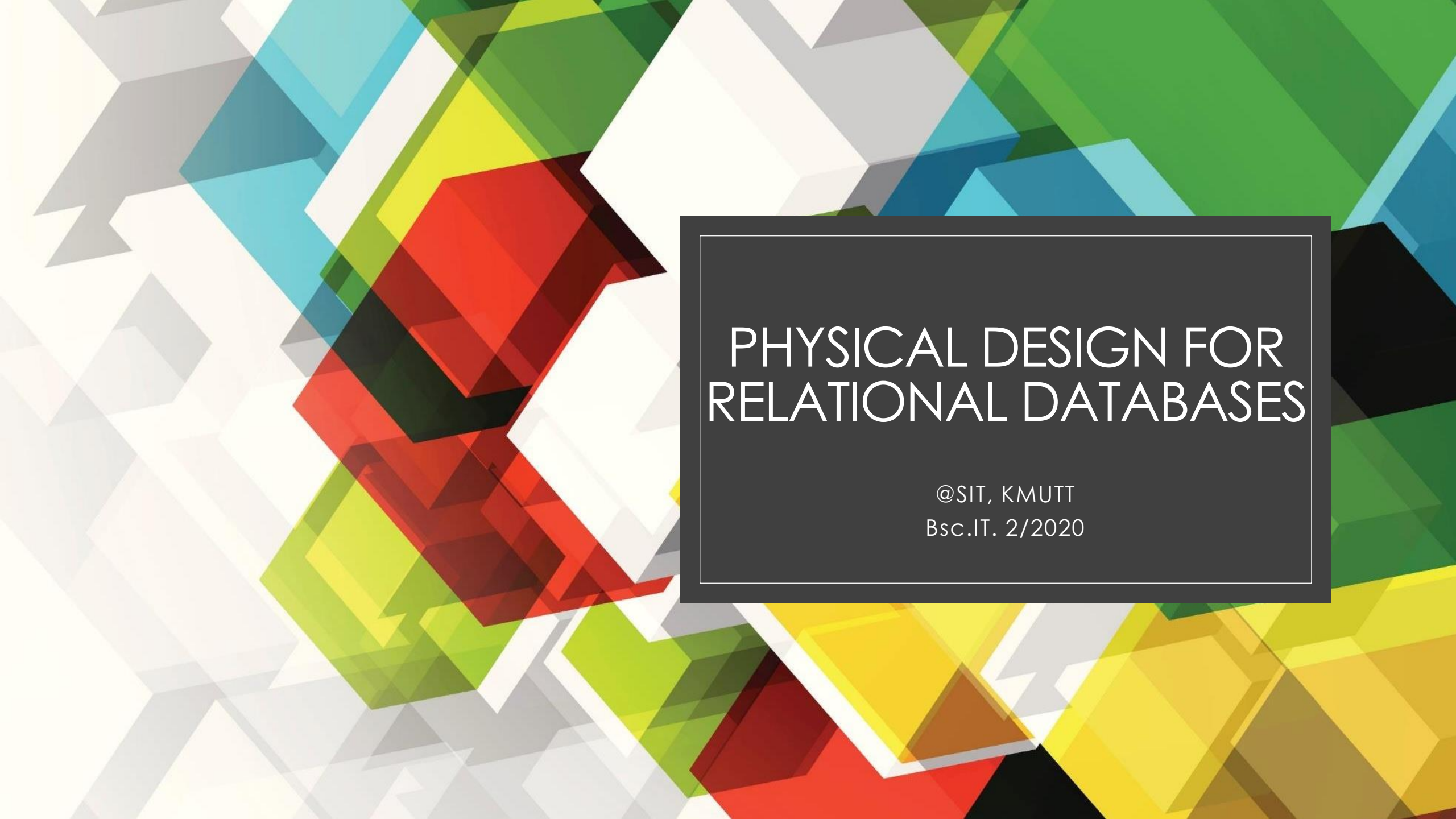
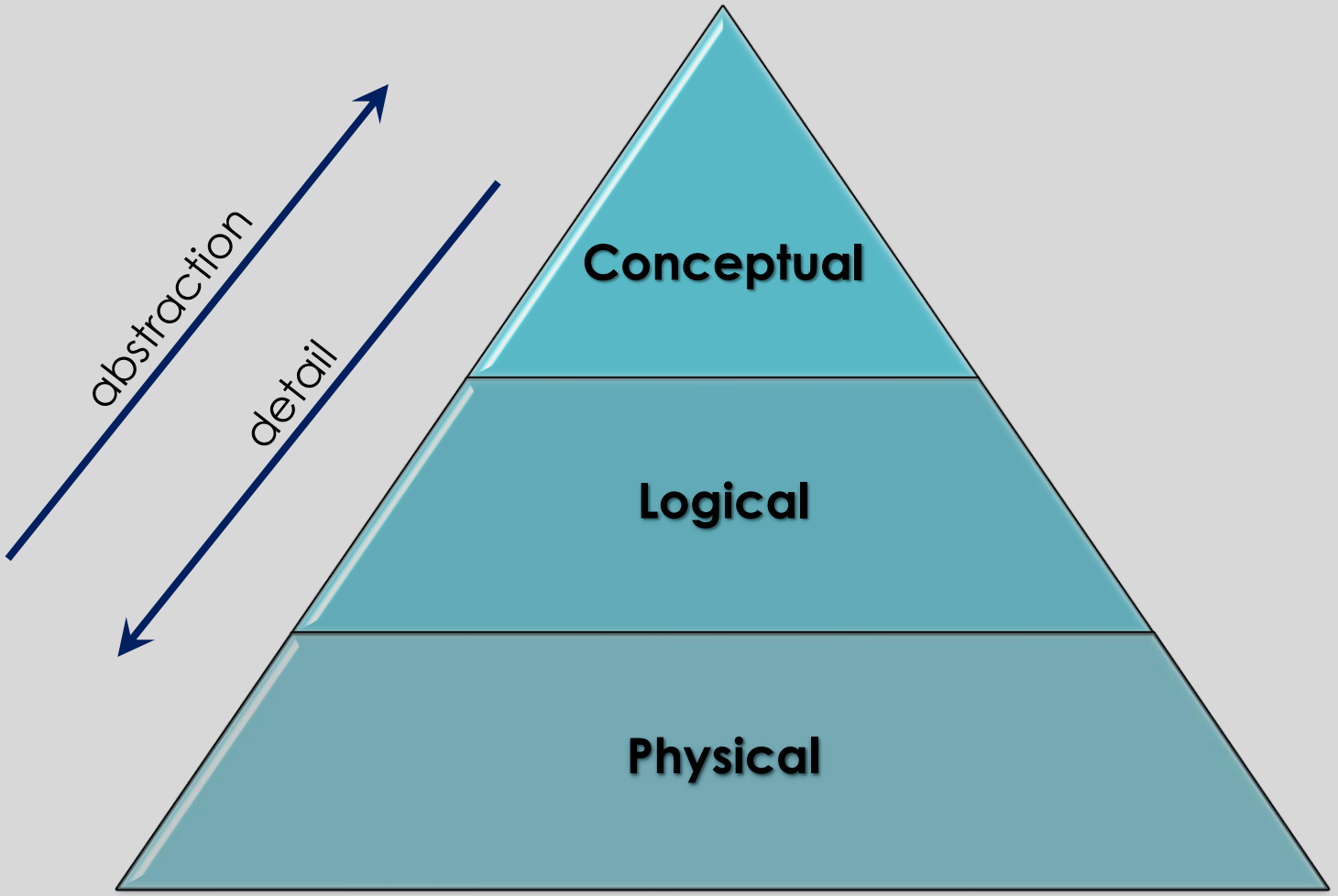


PHYSICAL DESIGN FOR RELATIONAL DATABASES

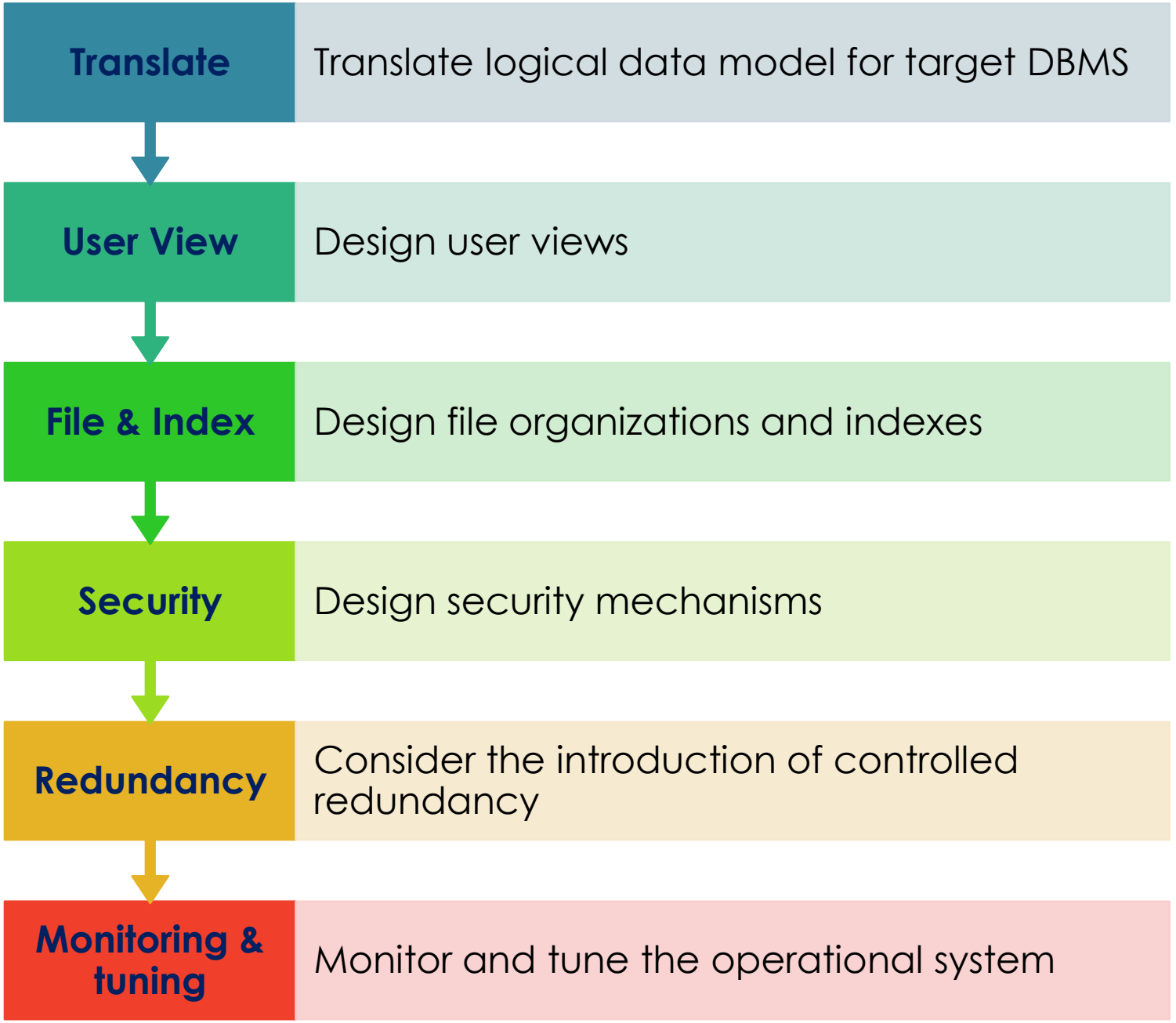
@SIT, KMUTT

Bsc.IT. 2/2020

Database Design



Steps of the **physical** database design



How to design file organizations and indexes

Analyze transactions

Choose file organizations

Choose indexes

Estimate disk space requirements

Choose file organizations

- The physical arrangement of data in a file into records and pages on secondary storage to **store and access data in an efficient way**.
- Types of file organizations
  - **Heap**
  - **Sequential**
  - **Hash**
  - Indexed Sequential Access Method (ISAM)
  - Clusters

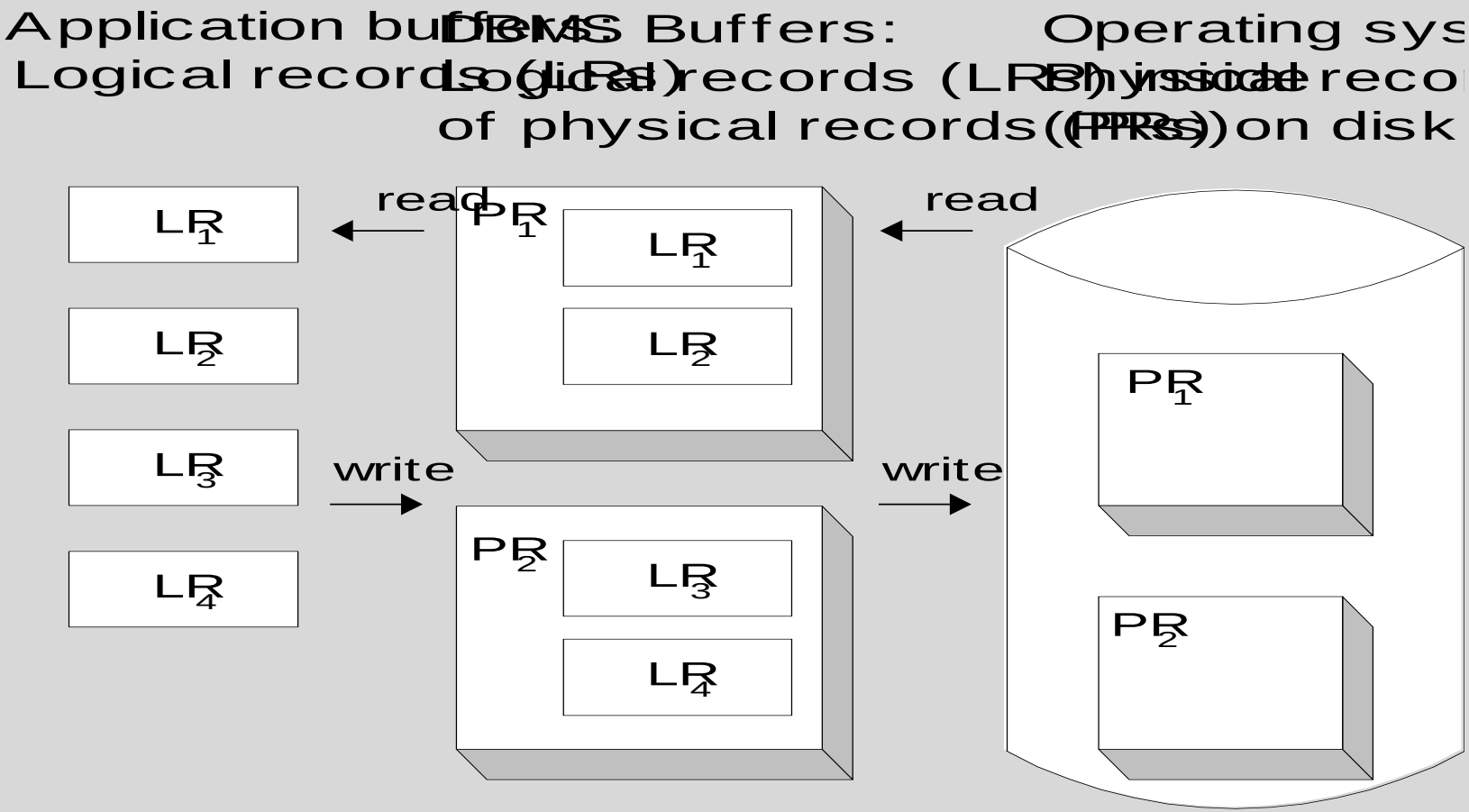
Logical Record

- It is a record designed in the phase of logical database design.
- It contains fields which have values of a particular type, which may be fixed length or variable length

Physical Record

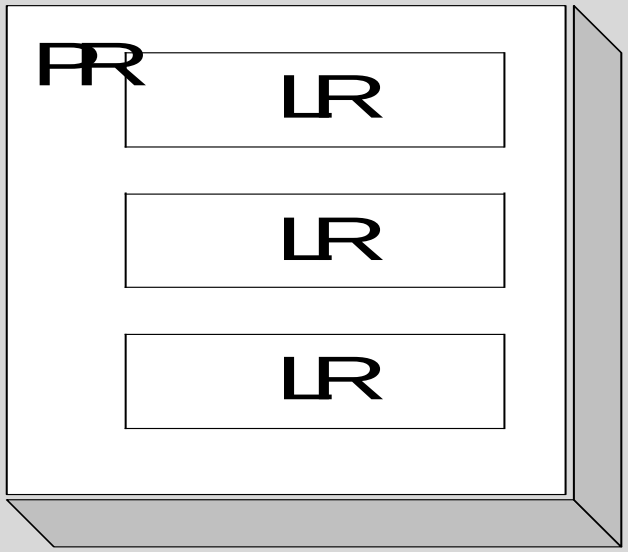
- It is known as block or page
- It is a unit of transfer between disk and memory
- Normally, one block stores a number of records.
  - Blocking factor refers to the number of records per block.
- There may be empty space in a block if an integral number of records do not fit in one block.

Application loads logical records from the physical records

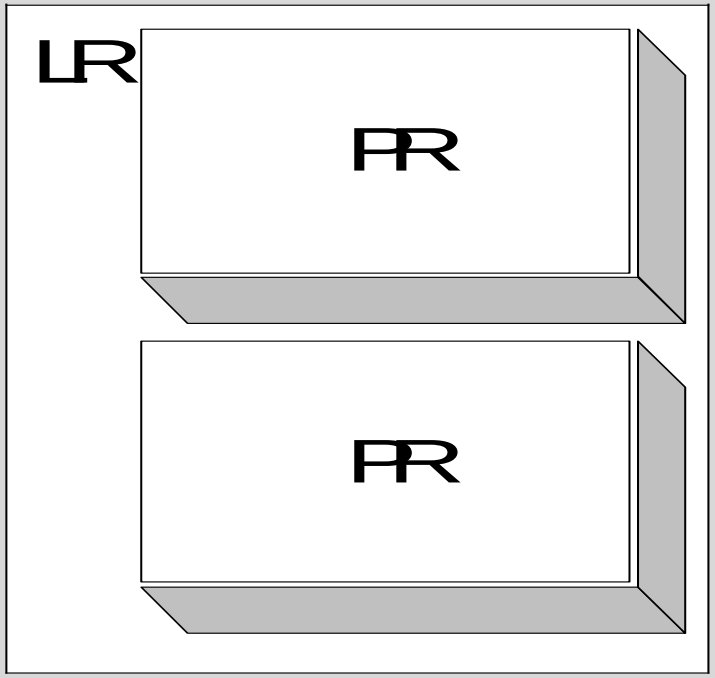


Logical and physical record

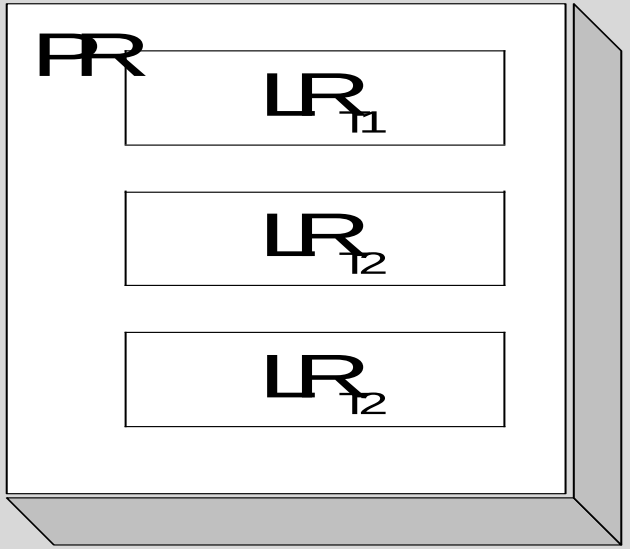
(a) Multiple LRs per PR



(b) LR split across PRs



(c) PR containing LRs from different tables



How to design file organizations and indexes

Analyze transactions

Choose file organizations

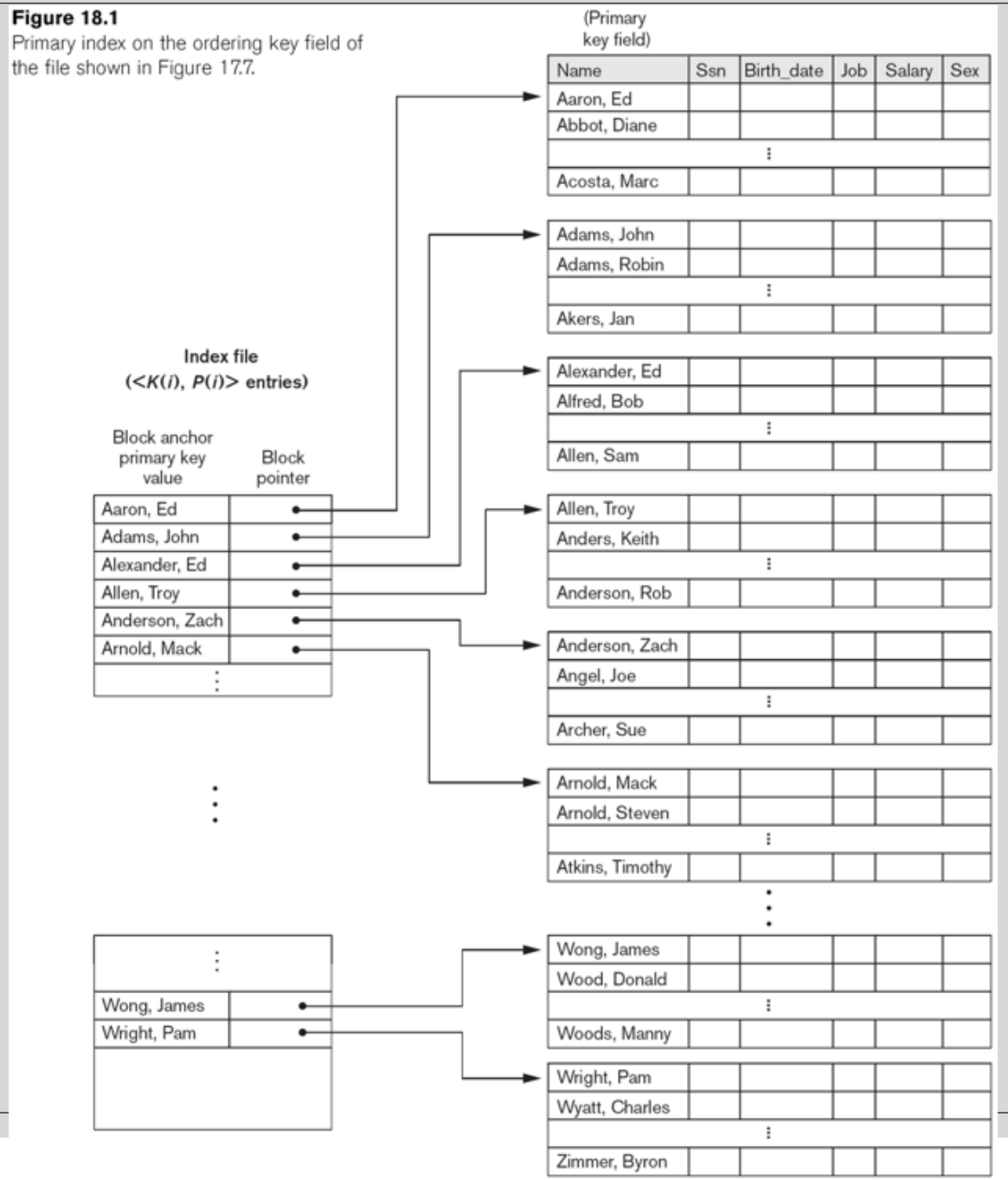
Choose indexes

Estimate disk space requirements

What is an Index?

- To locate records in a file more quickly
- Is an auxiliary structure file (**index/key file**) associated with a file (**data file**)
- An index structure (index file) consists of:
  - A **key** value of the selected field.
  - The **address** of logical record.
- Types of Index
  - Primary Index is ordered by an ordering key field to guarantee to have a unique value
  - Secondary Index is ordered by a non-ordering field, and may not contain unique value
  - Clustering Index is ordered by on a non-key field
- A file can have at least one primary index or one clustering index, zero or several secondary indexes

Primary Index



A Clustering

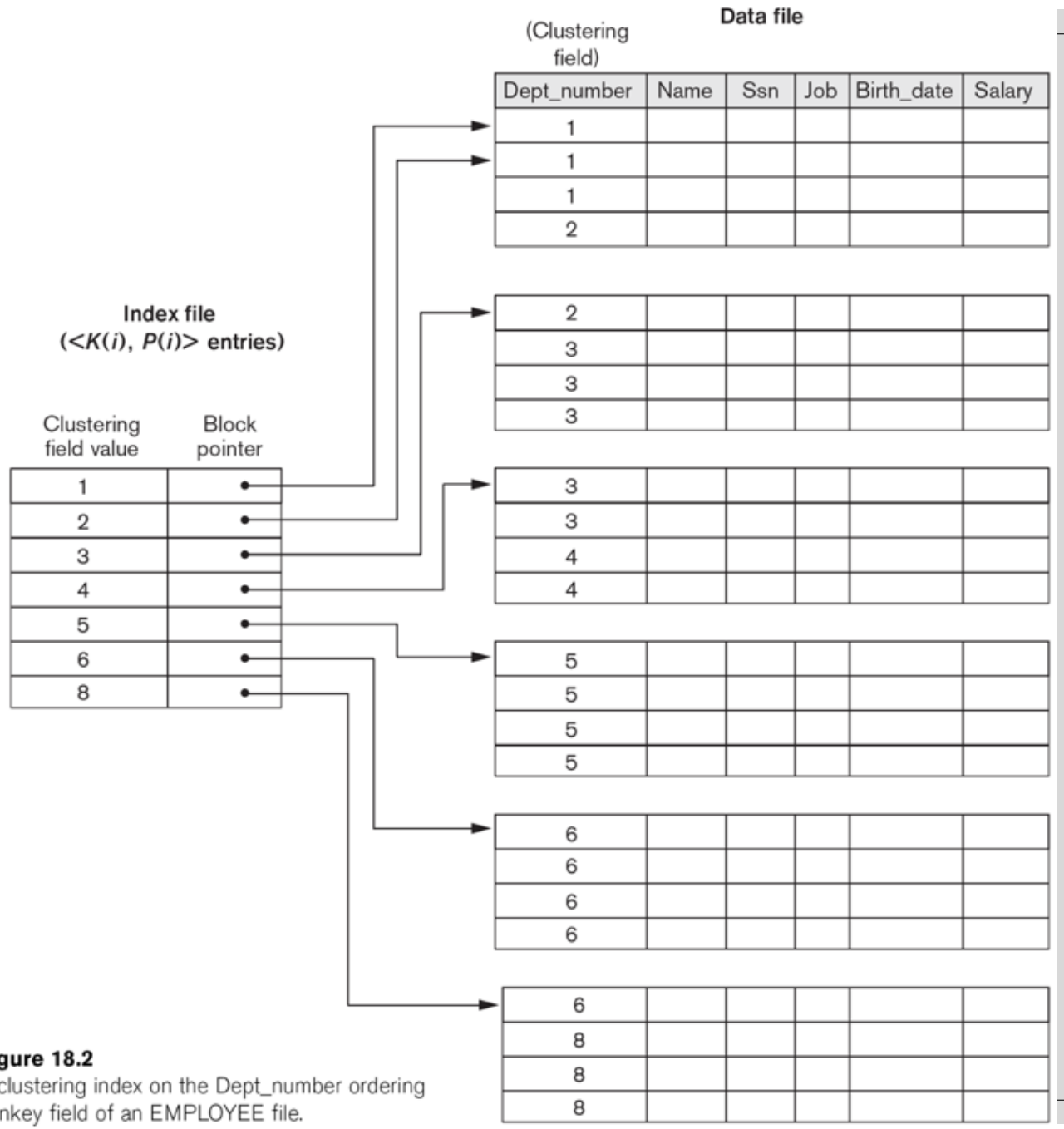
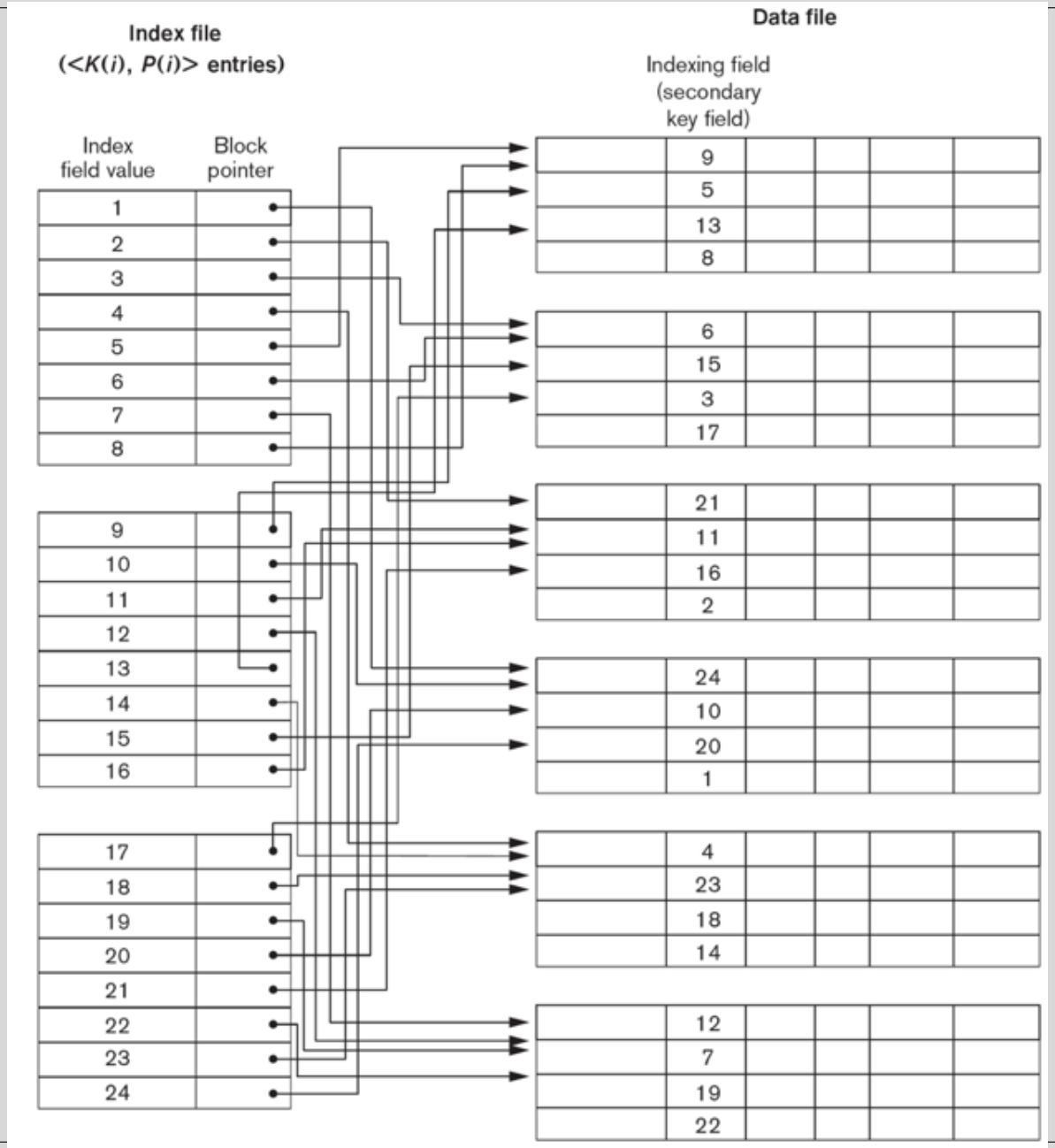
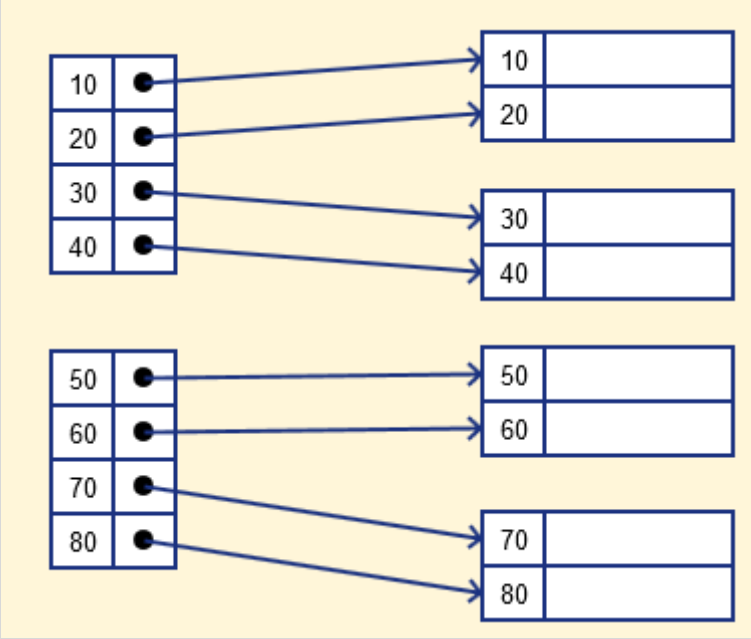


Figure 18.2
A clustering index on the Dept_number ordering nonkey field of an EMPLOYEE file.

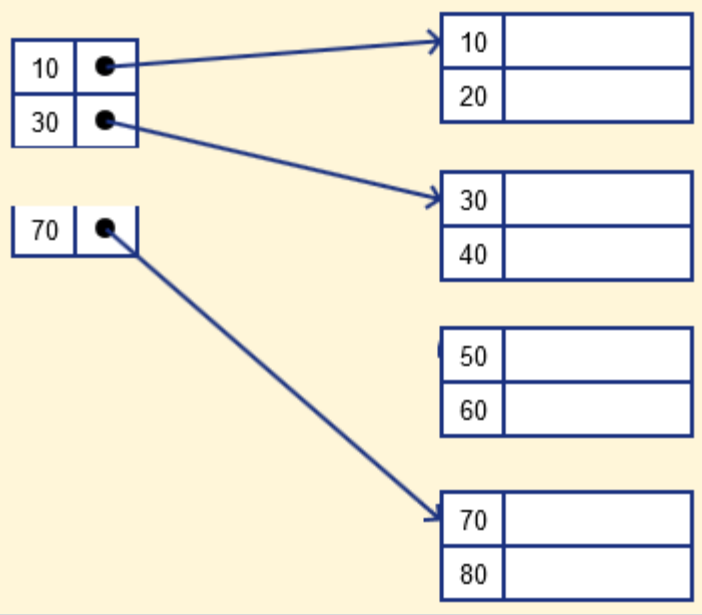
Secondary Index



Dense & Sparse Index



Dense Index



Sparse Index

B-tree index v.s. Bitmap index

B-Tree	Bitmap
· Suitable for high-cardinality columns	· Suitable for low-cardinality columns
· Updates on keys relatively inexpensive	· Updates to key columns very expensive
· Inefficient for queries using OR predicates	· Efficient for queries using OR predicates
· Useful for OLTP	· Useful for data warehousing