Select * from employee e where e.manager=10;
Physical database design

Elmasri/Navathe ch 17 and 18
Padron-McCarthy/Risch ch 21 and 22

DATABASE DESIGN I - 1DL300 Spring 2013
Sobhan Badiozamany
Department of Information Technology, Uppsala University

Select *
from employee e
where e.manager=10;

The physical database is a collection of stored records that have been organized in files on the hard disk.

Query parser and optimizer

Execution plan: An storage/retrieval program

Data files

Index Files

Hard Disk System
Select *
from employee e
where e.manager=10;

The physical database is a collection of stored records that have been organized in files on the hard disk.
Hard Disk System

Buffer: portion of main memory available to store copies of disk blocks.

The DBMS engine calls the buffer manager when it needs a block from disk.
- The DBMS is given the address of the block in the main memory, if it is already present in the buffer.
- If the block is not in the buffer:
  - the buffer manager allocates space in the buffer for the block.
  - If buffer is full, it has to flush out some other block.
- The block that is thrown out is written back to disk only if it was modified since the most recent time that it was written (or fetched) from the disk.
- Once space is allocated in the buffer, the buffer manager reads the block from the disk to the buffer, and returns the address of the block in main memory to the requester.

Block transfer is slow
The goal:
- Minimize the number of blocks transferred.
- Minimize the arms movements.

The subsystem responsible for allocating buffer space in main memory.

- Database system tries to minimize the number of block transfers between the disk and main memory.
- We can reduce the number of disk accesses by keeping as many blocks as possible in main memory.
Disks

- Preferred secondary storage device for high storage capacity and low cost.
- Data stored as magnetized areas on magnetic disk surfaces.

- Disks are divided into concentric circular tracks on each disk surface.
  A track is divided into smaller blocks or sectors:
  - The block size $B$ is fixed for each system.
    - Typical block sizes range from $B=512$ bytes to $B=4096$ bytes.
  - Whole blocks are transferred between disk and main memory for processing.

A read-write head moves to the track that contains the block to be transferred.
- Disk rotation moves the block under the read-write head for reading or writing.
Reading or writing a disk block is time consuming because of:
  - the seek time: the time it takes to move the arm.
  - rotational delay: the time it takes for the disk to rotate such that the head reaches the desired sector.
Disks

- Preferred secondary storage device for high storage capacity and low cost.
- Data stored as magnetized areas on magnetic disk surfaces.

Figure 17.2
Different sector organizations on disk. (a) Sectors subtending a fixed angle. (b) Sectors maintaining a uniform recording density.
Disks are divided into concentric circular tracks on each disk surface.

A track is divided into smaller blocks or sectors:

- The block size $B$ is fixed for each system.
  - Typical block sizes range from $B=512$ bytes to $B=4096$ bytes.
- Whole blocks are transferred between disk and main memory for processing.
A read-write head moves to the track that contains the block to be transferred.

- Disk rotation moves the block under the read-write head for reading or writing.
A read-write head moves to the track that contains the block to be transferred.

- Disk rotation moves the block under the read-write head for reading or writing.

Reading or writing a disk block is time consuming because of:

- the seek time: the time it takes to move the arm.
- rotational delay: the time it takes for the disk to rotate such that the head reaches the desired sector.
**Hard Disk System**

**Buffer:**
portion of main memory available to store copies of disk blocks.

**Buffer manager**
The subsystem responsible for allocating buffer space in main memory.

- The DBMS engine calls the buffer manager when it needs a block from disk.
  - The DBMS is given the address of the block in the main memory, if it is already present in the buffer.
  - If the block is not in the buffer:
    - the buffer manager allocates space in the buffer for the block.
    - if buffer is full, it has to through out some other block.
  - The block thus obtained is written back to disk only if it was modified since the most recent time that it was written back to disk.
  - Once space is allocated in the buffer, the buffer manager reads the block from the disk to the buffer, and passes the address of the block in main memory to the requester.

- Database system tries to minimize the number of block transfers between the disk and main memory.
- We can reduce the number of disk accesses by keeping as many blocks as possible in main memory.

**Block transfer is slow**
The goal:
- Minimize the number of blocks transferred.
- Minimize the arms movements.
The DBMS engine calls the buffer manager when it needs a block from disk.

- The DBMS is given the address of the block in the main memory, if it is already present in the buffer.
- If the block is not in the buffer:
  - the buffer manager allocates space in the buffer for the block.
  - if buffer is full, it has to throw out some other block.
- The block that is thrown out is written back to disk only if it was modified since the most recent time that it was written to/fetched from the disk.
  - Once space is allocated in the buffer, the buffer manager reads in the block from the disk to the buffer, and passes the address of the block in main memory to the requester.
memory.

- Database system tries to minimize the number of block transfers between the disk and main memory.
- We can reduce the number of disk accesses by keeping as many blocks as possible in main memory.
Hard Disk System

Buffer: portion of main memory available to store copies of disk blocks.

The DBMS engine calls the buffer manager when it needs a block from disk.
- The DBMS is given the address of the block in the main memory, if it is already present in the buffer.
- If the block is not in the buffer:
  - the buffer manager allocates space in the buffer for the block.
  - If buffer is full, it has to free out some other block.
- The block that is thrown out is written back to disk only if it was modified since the most recent time that it was written (flushed) from the disk.
- Once space is allocated in the buffer, the buffer manager reads the block from the disk to the buffer, and passes the address of the block in main memory to the requester.

Block transfer is slow
The goal:
- Minimize the number of blocks transferred.
- Minimize the arms movements.

The subsystem responsible for allocating buffer space in main memory.

- Database system tries to minimize the number of block transfers between the disk and main memory.
- We can reduce the number of disk accesses by keeping as many blocks as possible in main memory.
The physical database is a collection of stored records that have been organized in files on the hard disk.
The physical database is a collection of stored records that have been organized in files on the hard disk.

Records are used for physical storage of:

- Tuples, where each attribute in the tuple is stored as a field.
- Objects in object-oriented databases.

Records contain **fields** which have values of a particular type.
Data files

Similarly, on the level:
File header

- Contains information that is needed for record access:
- Block addresses
- record format etc.
Files with fixed length record.

For example, consider a file containing records of this type:
All fields have fixed length.

<table>
<thead>
<tr>
<th>Name</th>
<th>personnummer</th>
<th>Salary</th>
<th>Job_code</th>
<th>Department</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anders, Andersson</td>
<td>861115-5874</td>
<td>30,000</td>
<td>F-34174</td>
<td>Fincance</td>
</tr>
</tbody>
</table>

Advantages:

- Easy to locate record r, given its record number n and size of each record s:
All fields have fixed length.

<table>
<thead>
<tr>
<th>Name</th>
<th>personnummer</th>
<th>Salary</th>
<th>Job_code</th>
<th>Department</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anders, Andersson</td>
<td>861115-5874</td>
<td>30,000</td>
<td>F-34174</td>
<td>Fincance</td>
</tr>
</tbody>
</table>

**Advantages:**

- Easy to locate record $r$, given its record number $n$ and size of each record $s$:
  - $\text{position}(r) = n \times s$
- Easy to locate field values

**Disadvantage:**

Space is wasted
**mix of fixed and variable length records**

<table>
<thead>
<tr>
<th>Name</th>
<th>personnummer</th>
<th>Salary</th>
<th>Job_code</th>
<th>Department</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smith, John</td>
<td>861115-5874</td>
<td>20,000</td>
<td></td>
<td>Information Technology</td>
</tr>
</tbody>
</table>

Field separator:  

**Advantages:**

- space is wasted, but only where NULL
<table>
<thead>
<tr>
<th>Name</th>
<th>personnummer</th>
<th>Salary</th>
<th>Job_code</th>
<th>Department</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smith, John</td>
<td>861115-5874</td>
<td>20,000</td>
<td></td>
<td>Information Technology</td>
</tr>
</tbody>
</table>

Field separator:  

Advantages:

- space is wasted, but only where NULL values occur.

Disadvantage:

- locating records and fields needs more processing:
  - For example, location of the salary field depends on the length of name in the record.
files with Variable-field records

Name=Smith, John  Personnummer=861115-5874  Department=Information Technology

Field separator: □  Record terminator: ⊥

Advantages:

• Less space is wasted, specially if there
Field separator: | Record terminator:  

Advantages:

• Less space is wasted, specially if there might be many NULL values.

Disadvantage:

• More processing is needed for:
  • locating records/fields.
  • Identifying NULL values
    • need to see the whole record to identify a NULL value
There might be left over space in the file blocks

Record 1

Record 2

**Blocking factor**

The highest number of records that can be contained in a block is called the block factor (here Bf) for the file of records.

\[ Bf = \text{Block size} \div \text{Record size} \]

where R is the record size and B the block size.

In this example, assume a block size B is 512 bytes, and record size R is 200 bytes.

\[ Bf = 512 \div 200 \approx 2.56 \]

Rounded off downwards gives \( Bf = 2 \), i.e., we can store 2 records per block.

There is space, but only two records can fit, so part of the block is wasted.
Blocking factor

The highest number of records that can be contained in a block is called the block factor (here: bfr) for the file of records.

\[ bfr = \text{floor}[B/R] \]

where \( R \) is the record size and \( B \) the block size.

In this example, assume a block size \( B \) is 512 bytes, and record size \( R \) is 200 bytes.

- \( B / R = 512/200 = 2.56 \)
- Rounded off downwards gives \( bfr = 2 \), i.e. we can store 2 records per block.
similarly, on file level:

A file with \( r \) no. of records therefore require:

\[ b = \text{ceiling}[r/bfr] \]

blocks.

in our example, \( bfr=2 \), so a file with \( r=605 \) records needs \( 605/2=302.5 \) \( \sim \) 303 blocks.

Q: How much space is wasted in total?
File organization

• The database is stored as a collection of files. Each file is a sequence of records.
• Records can have constant (simplest) or variable length
• A file can store records of the same type (simplest) or of different type.
• Specific files can be used to store specific relations (simplest) or the same file can store different relations (maybe even the whole database).
• The database is stored as a collection of files. Each file is a sequence of records.
• Records can have constant (simplest) or variable length.
• A file can store records of the same type (simplest) or of different type.
• Specific files can be used to store specific relations (simplest) or the same file can store different relations (maybe even the whole database).

Heap files
- New records are added at the end of the file. Such an organization is called a heap file.
- Records are added in the order they are inserted into the file.
- Search for a specific record is expensive (linear in the worst case).
- Delays of a record can be expensive (searches must involve all the records ahead of it).
- Records in a physically ordered sequence may not be the same as in a file. Relational DBMSs require specially organized data files.
- Modification of a record of variable length can be hard.
- Records are added by a certain order requires that the file must be read which is expensive.

Ordered files
- The records in the file are ordered according to the value of a certain field in the figure according to "term.
- Records are physically stored in ascending order of the value of the field.
- Improved efficiency in range searches.
- Improved efficiency in sequential searches.
- Improved efficiency in all searches requiring comparisons of fields.

External Hashing
- The blocks are divided into the designated (hashed) numbers of partitions, ... (hashes).
- Typically, a function corresponds to one or a few numbers of each block.
- Each block is assigned a unique number in function 'hashing function'.
- The record is placed in the bucket in which the record is stored.
A record can be placed anywhere in the file where there is space
Heap files

- New records are added to the end of the file. Such an organization is called a heap file.
  - Suitable when we don’t know how data shall be used.
- Insert of a new record is very efficient.
- Search after a specific record is expensive (linear to the size).
- Delete of a record can be expensive (search - read into - delete - write back).
  - Instead of physically removing a record one can mark the record as deleted. Both methods require a periodically reorganization of the file.
- Modification of a record of variable length can be hard.
- Retrieval according to a certain order requires that the file must be sorted which is expensive.
store records in sequential order, based on the value of the search key of each record.

A hash function is computed on some attribute of each record; the result specifies a block of the file the record is stored in the file.
Ordered files

The records in the file are ordered according to the value of a certain field, in the figure according to name.

- Ordered retrieval very fast (no sorting needed).
- Next record in the order is found on the same block (except for the last record in the block).
- Search is fast (binary search - log2b).
- Insert and delete are expensive since the file must be kept sorted.
- Suitable for applications that require sequential processing of the entire file.
- Need to reorganize the file from time to time to restore sequential order.

- To make record insertions cheaper:
  - Create a temporary unsorted file a so called overflow file or transaction file (the main file is then called “the master file”).
  - Update the master file periodically in accordance with the transaction file.
  - These measures improve insertion time but search for records becomes more complicated.
- Ordered files are not used that often in databases.
  - Exception: when extra access paths are created, so called primary indexes.
#### Block 1

<table>
<thead>
<tr>
<th>NAME</th>
<th>SSN</th>
<th>BIRTHDATE</th>
<th>JOB</th>
<th>SALARY</th>
<th>SEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aaron, Ed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abbott, Diane</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acosta, Marc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Block 2

<table>
<thead>
<tr>
<th>NAME</th>
<th>SSN</th>
<th>BIRTHDATE</th>
<th>JOB</th>
<th>SALARY</th>
<th>SEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adams, John</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adams, Robin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Akers, Jan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Block 3

<table>
<thead>
<tr>
<th>NAME</th>
<th>SSN</th>
<th>BIRTHDATE</th>
<th>JOB</th>
<th>SALARY</th>
<th>SEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alexander, Ed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alfred, Bob</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allen, Sam</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Block 4

<table>
<thead>
<tr>
<th>NAME</th>
<th>SSN</th>
<th>BIRTHDATE</th>
<th>JOB</th>
<th>SALARY</th>
<th>SEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allen, Troy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anders, Keith</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anderson, Rob</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Block 5

<table>
<thead>
<tr>
<th>NAME</th>
<th>SSN</th>
<th>BIRTHDATE</th>
<th>JOB</th>
<th>SALARY</th>
<th>SEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anderson, Zach</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angeli, Joe</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Archer, Sue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Block 6

<table>
<thead>
<tr>
<th>NAME</th>
<th>SSN</th>
<th>BIRTHDATE</th>
<th>JOB</th>
<th>SALARY</th>
<th>SEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arnold, Mack</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arnold, Steven</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atkins, Timothy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

... (Continues with more blocks)

#### Block n

<table>
<thead>
<tr>
<th>NAME</th>
<th>SSN</th>
<th>BIRTHDATE</th>
<th>JOB</th>
<th>SALARY</th>
<th>SEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wong, James</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wood, Donald</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Woods, Manny</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NAME</th>
<th>SSN</th>
<th>BIRTHDATE</th>
<th>JOB</th>
<th>SALARY</th>
<th>SEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wright, Pam</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wyatt, Charles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zimmer, Byron</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
records in the file are ordered according to the value of a certain field, in the figure according to name.

- Ordered retrieval very fast (no sorting needed).
- Next record in the order is found on the same block (except for the last record in the block)
- Search is fast (binary search - \( \log_2 b \))
- Insert and delete are expensive since the file must be kept sorted.
- Suitable for applications that require sequential processing of the entire file
- Need to reorganize the file from time to time to restore sequential order

• To make record insertions cheaper:
  - Create a temporary unsorted file a so called overflow file or transaction file (the main file is then called “the master file”)
  - Update the master file periodically in accordance with the transaction file.
  - These measures improve insertion time but search for records becomes more complicated.
Binary search

Consider element to be searched for i.e.
item = 60

int F(int item, int low, int high, int a[])
{
    int mid;
    if (low > high) return -1;
    mid = (low + high)/2;
    if (item == a[mid]) return mid;
    if (item < a[mid]) F(item, low, mid-1, a);
    return F(item, mid+1, high, a);
}

Watch between 0:50 to 01:32
the block)
– Search is fast (binary search - $\log_2 b$)
– Insert and delete are expensive since the file must be kept sorted.
– Suitable for applications that require sequential processing of the entire file
– Need to reorganize the file from time to time to restore sequential order

• To make record insertions cheaper:
  – Create a temporary unsorted file a so called overflow file or transaction file
    (the main file is then called “the master file”)
  – Update the master file periodically in accordance with the transaction file.
  – These measures improve insertion time but search for records becomes more complicated.

• Ordered files are not used that often in databases.
  – Exception: when extra access paths are created, so called primary indexes.
A hash function is computed on some attribute of each record; the result specifies in which block of the file the record should be placed.
External Hashing

The file blocks are divided into M equal-sized buckets, numbered bucket0, bucket1, ..., bucketM-1.

- Typically, a bucket corresponds to one (or a fixed number of) disk block.

One of the file fields is designated to be the hash key of the file. The record with hash key value K is stored in bucket i, where i=h(K), and h is the hashing function.

Properties

- Search is very efficient on the hash key.
- Collisions occur when a new record hashes to a bucket that is already full.
  - An overflow file is kept for storing such records.
  - Overflow records that hash to each bucket can be linked together.

Example:

- Hash key:
  - `emp_num`
  - Number of buckets M=100
- Hash function:
  - `h(emp_num) = emp_num mod 100`

To find the block that contains the record for employee number 102, we should compute

```
h(102) = 102 mod 100 = 2
```

Figure 17.9

Matching bucket numbers to disk block addresses.
The file blocks are divided into $M$ equal-sized buckets, numbered $bucket_0$, $bucket_1$, ..., $bucket_{M-1}$.

- Typically, a bucket corresponds to one (or a fixed number of) disk block.

One of the file fields is designated to be the hash key of the file. The record with hash key value $K$ is stored in bucket $i$, where $i = h(K)$, and $h$ is the hashing function.

**Example:**

Hash key:
- employee number $emp_num$

Number of Buckets $M = 100$

Hash function:
Example:

Hash key:
- employee number emp_num

Number of Buckets M=100

Hash function:
- \( h(emp\_num) = emp\_num \mod 100 \)

To find the block that contains the record for employee number 102 we should compute \( h(102)=102 \mod 100 =2 \)

Figure 17.9
Matching bucket numbers to disk block addresses.
This is designated to be the hash key of the file.

The hash key value $K$ is stored in bucket $i$, where $i$ is the hash function.
Properties

• Search is very efficient on the hash key.
• Collisions occur when a new record hashes to a bucket that is already full.
  • An overflow file is kept for storing such records.
  • Overflow records that hash to each bucket can be linked together.
Indexing: Book example

Indexing in DBMS's

- To find records, one or several blocks transferred to (one or several buffers) to main memory. These blocks can then be searched to find the records that were sought.
- But which block?
- If the address to the block containing the record is unknown one has to search through all block in the file (so called linear search).

If there was no index/table-of-content the student had to "scan" the whole book!
O o!
I need to start studying for the exam
:-(

Where is it in the book that the relational algebra is covered?
Table of contents

Chapter 1: introduction
...
...
Chapter 6: Formal Relational Languages: the Algebra and Calculus.
...
...
Index of key words

...  
...  
relational data model .... 55-80  
relational algebra .......... 143-170  
...  
...
Relational algebra

bla bla bla bla bla
bla bla bla bla bla
bla bla bla bla bla
bla bla bla bla bla
bla bla bla bla bla
143
Indexing: Book example

Indexing in DBMS's
- To find records, one or several blocks transferred to one or several buffers in main memory. These blocks can then be searched to find the records that were sought.
- But which blocks?
- If the address to the block containing the record is unknown one has to search through all block in the file (so called linear search).

If there was no index/table-of-content the student had to "scan" the whole book!
Indexing in DBMS's

- To find records, one or several blocks transferred to (one or several buffers in) main memory. These blocks can then be searched to find the records that were sought.
- But which blocks?
  - If the address to the block containing the record is unknown one has to search through all block in the file (so called linear search).
Select *
from employee e
where e.manager=10;

The physical database is a collection of stored records that have been organized in files on the hard disk.
Index Files

An index (or index file) is an extra file structure that is used to make the retrieval of records faster.

An index file consists of records (called index entries) of the form: 

Search key  pointer

Search key (or index field): Attribute or set of attributes (data fields) used to look up records in a file.

pointer (or address): The address of blocks that contain records with the given search key.

• These entries determine the physical address for records having a certain value in their index field.
• Index files are typically much smaller than the original file.
• The file that should be indexed is called the data file.
Index Files

An index (or index file) is an extra file structure that is used to make the retrieval of records faster. An index file consists of records (called index entries) of the form: [Search key | pointer]

Search key (or index field): Attribute or set of attributes (data fields) used to look up records in a file.

pointer (or address): The address of blocks that contain records with the given search key.

- These entries determine the physical address for records having a certain value in their index field.
- Index files are typically much smaller than the original file.
An index file consists of records (called index entries) of the form:

- **Search key**: Attribute or set of attributes (data fields) used to look up records in a file.
- **Pointer** (or address): The address of blocks that contain records with the given search key.

**•** These entries determine the physical address for records having a certain value in their index field.

**•** Index files are typically much smaller than the original file.

**•** The file that should be indexed is called the **data file**.
Primary index

The first record in each block is called the Anchor Record of the block.

- A primary index has one index record for each block in the data file, and therefore is called a sparse index (or “nondense index”).
- A dense index consists of one record for each record in the data file.

The data file is sorted on the primary key field.

Primary index - pros and cons

- Require much less space than the data file.
  - a) There is much fewer index records than records in the data file.
  - b) Every index record need less space (⇒ fewer memory blocks).
- Problem with insertion and deletion of records.
  - If anchor records are changed the index file must be updated.
<table>
<thead>
<tr>
<th>Search key</th>
<th>Block anchor primary key value</th>
<th>Block pointer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aaron, Ed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adams, John</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alexander, Ed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allen, Troy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anderson, Zach</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arnold, Mack</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

the same type as the ordering field (index field) for the data file. pointer to a block (block pointer)
The first record in each block is called the Anchor Record of the block.

**Figure 18.1**
Primary index on the ordering key field of the file shown in Figure 17.7.

- A primary index has one index record for each block in the data file, and therefore is called a **sparse index** (or “nondense index”)
- A **dense index** consists of one record for each record in the data file.
Primary index - pros and cons

• Require much less space than the data file.
  – a) There is much fewer index records than records in the data file.
  – b) Every index record need less space (⇒ fewer memory blocks).

• Problem with insertion and deletion of records.
  – If anchor records are changed the index file must be updated.
Cluster index

Cluster index is defined for files that are ordered according to a non-key field (the cluster field), i.e. several records in the data file can have the same value for the cluster field.

![Diagram of a cluster index](image)

*Figure 18.2*
A clustering index on the Dept_number ordering nonkey field of an EMPLOYEE file.
Cluster index is defined for files that are ordered according to a non-key field (the cluster field), i.e. several records in the data file can have the same value for the cluster field.

<table>
<thead>
<tr>
<th>Clustering field value</th>
<th>Block pointer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

Index file
(<\textit{K}(i), \textit{P}(i)> entries)

- **Search key**: the same type as the Cluster field for the data file.
- **Pointer**: pointer to a block (block pointer)
<table>
<thead>
<tr>
<th>Clustering field value</th>
<th>Block pointer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

- **Search key**
- **pointer**

The same type as the Cluster field for the data file. Pointer to a block (block pointer).
Secondary index

The data file is not sorted according to the index field.

There are two different cases:
1. The index field has unique values for each record (the figure).
2. Several records in the data file can have the same values for the index field (see Elmasri/Navathe Fig 18.5)
Secondary index

The data file is not sorted according to the index field.

Figure 18.4
A dense secondary index (with block pointers) on a nonordering key field of a file.

There are two different cases:
1. The index field has unique values for each record (the best case).
2. Several records in the data file can have the same values.
<table>
<thead>
<tr>
<th>Search key</th>
<th>pointer</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td></td>
</tr>
</tbody>
</table>

the same type as the indexing field (any field in the data file) pointer to a block (block pointer)
There are two dimensions:
1. The indexing field values for each index file entry.
2. Several related data file entries can have the same indexing field value (as in Fig 18.5)
Secondary index

The data file is not sorted according to the index field.

There are two different cases:
1. The index field has unique values for each record (the figure).
2. Several records in the data file can have the same values for the index field (see Elmasri/Navathe Fig 18.5)
Index evaluation metrics

Indexing techniques evaluated on basis of:

• Access types supported efficiently. E.g.,
  – records with a specified value in an attribute e.g. where salary=2000;
  – or records with an attribute value falling in a specified range of values.
  e.g. where salary between 2000 and 3000;
• Access time
• Insertion time
• Deletion time
• Space overhead
Indexing in SQL

CREATE INDEX index_name ON table_name (column_name);
Creates an index on a table.
Duplicate values are allowed.

CREATE UNIQUE INDEX index_name ON table_name (column_name);
Creates a unique index on a table.
Duplicate values are not allowed.
Physical database design

Elmasri/Navathe ch 17 and 18
Padron-McCarthy/Risch ch 21 and 22

DATABASE DESIGN I - 1DL300 Spring 2013
Sobhan Badiozamany
Department of Information Technology, Uppsala
University

Select *
from employee e
where e.manager=10;

The physical database is a collection of stored records that have been organized in files on the hard disk.

Physical database

Execution plan:
An storage/retrieval program

Data files

Index Files

Hard Disk System