

Advanced Database Systems

Spring 2024

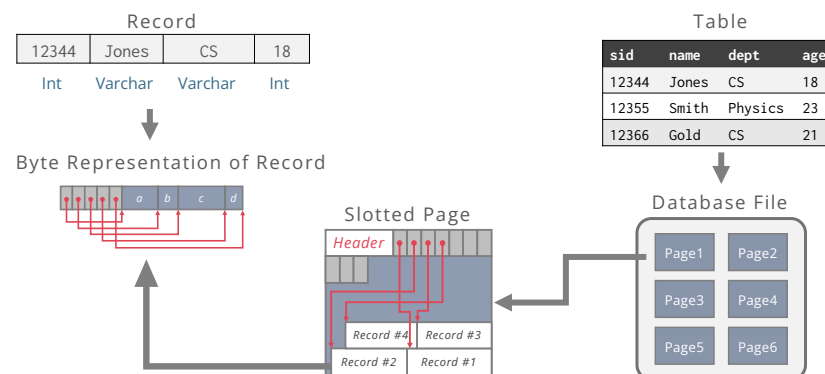
Lecture 12: Files, Pages, Records

R&G: Chapters 9.5-9.7

1

OVERVIEW: REPRESENTATIONS

2

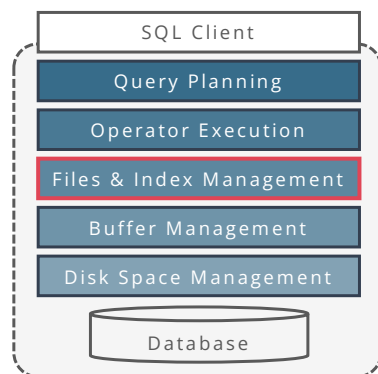


2

OUTLINE

3

- Storage Media
- Disk Space Management
- Buffer Management
- File Layout
- Page Layout
- Record Layout



3

FILES OF PAGES OF RECORDS

4

Tables stored as **logical (database) files**

A file consists of one or more **pages**

Each page contains one or more **records**

Pages are managed

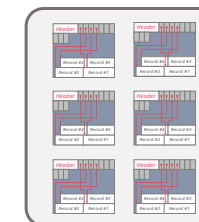
On disk by the **disk space manager**

Pages read/written to physical disk/files

In memory by the **buffer manager**

Higher levels of DBMS only operate in memory

Page management is oblivious to their actual content



Database File

4

FILES OF PAGES OF RECORDS

5

Database file: A collection of pages, each containing a collection of records

Could span multiple OS files and even machines

API for higher levels of the DBMS:

Create/delete a file

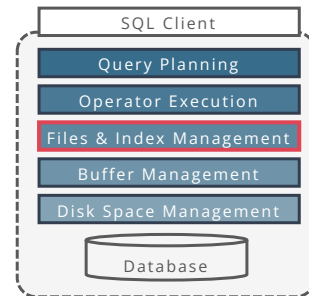
Insert/delete/modify a record

Fetch a particular record by **record ID**

Record ID = (page ID, location on page)

Scan all records

possibly with a predicate on the desirable records



5

DB FILE ORGANISATION

6

Method of arranging a file of records

At this point in the hierarchy, we do not care what is page format

Different types exist, each ideal for some situations & not so good in others:

Heap Files

Records placed arbitrarily across pages

Sorted Files

Pages and records are in sorted order

Index Files

B+ trees, hash-based files

May contain records or point to records in other files

6

HEAP FILE

7

Most important type of files in a database

Collection of records in **no particular order**

Not to be confused with "heap" data-structure

As file shrinks/grows, pages allocated/deallocated

To support record level operations, we must

Keep track of the pages in a file

Keep track of free space on pages

Keep track of the records on a page

7

HEAP FILE: LINKED LIST

8

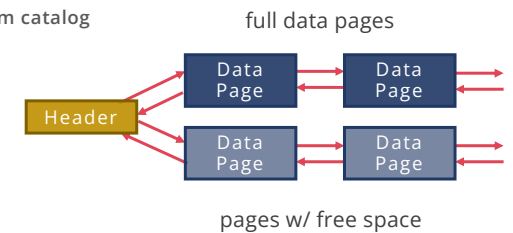
Doubly linked lists of pages

Header page allocated when the file is created

Header page ID stored in the **system catalog**

Initially both page lists are empty

Each page keeps track of the free space in itself



Easy to implement, but

Most pages end up in the free space list

Finding a page with sufficient empty space may search many pages

8

HEAP FILE: PAGE DIRECTORY

9

Directory = set of special pages storing metadata about data pages

Each directory entry

Identifies a data page & records #free bytes on it

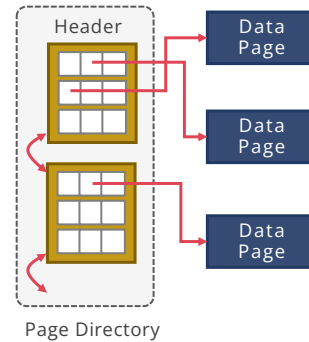
Free space search more efficient

Far fewer pages read to find a page to fit a record

One header page reveals free space of many pages

Header pages accessed often \Rightarrow likely in cache

Small memory overhead to host the directory



9

OUTLINE

10

Storage Media

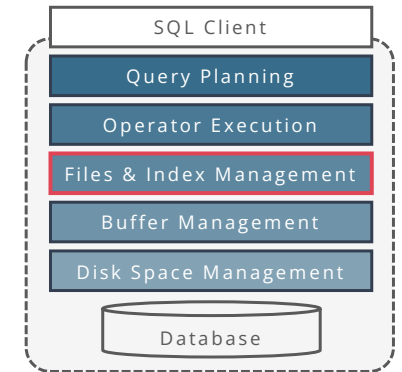
Disk Space Management

Buffer Management

File Layout

Page Layout

Record Layout



10

PAGE LAYOUT

11

How to organize the data stored inside the page

Header stores metadata about the page

page ID, #records, free space, next/prev pointers, etc.

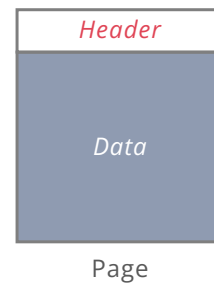
Find record by record ID, insert, delete records

record ID (rid) = (page ID, offset in page)

Things to consider:

Record length? Fixed or variable

Page packing? Packed or unpacked



11

FIXED-LENGTH RECORDS

12

Records are made up of multiple **fields**

Fields = values for columns in a table

We have fixed-length records when field lengths are consistent

The first field always has N bytes, the second field always has M bytes, etc.

\Rightarrow Record lengths are fixed

Every record is always the same number of bytes

Notice that the implication might not be true the other way

We can store fixed-length records in two ways: packed and unpacked

12

FIXED-LENGTH RECORDS, PACKED

13

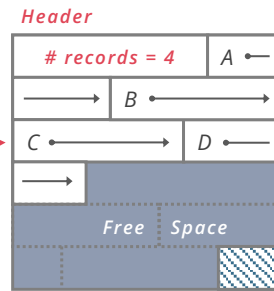
No gaps between records

Record ID = (page ID, position in page)

Easy to compute the offset of a record in the page:

$$\text{sizeof(Header)} + (\text{position} - 1) * \text{sizeof(Record)}$$

Record ID:
(Page 2, 3)



Page

Unused space

Note:

Data is always stored in linear order

For presentation, we "wrap around" the linear order into a rectangle

13

FIXED-LENGTH RECORDS, PACKED

14

No gaps between records

Record ID = (page ID, position in page)

Easy to compute the offset of a record in the page:

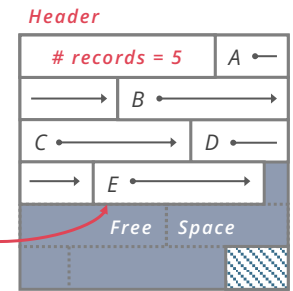
$$\text{sizeof(Header)} + (\text{position} - 1) * \text{sizeof(Record)}$$

Insert record

Just append a new record to the end

$$\text{sizeof(Header)} + \# \text{ records} * \text{sizeof(Record)}$$

Record ID:
(Page 2, 5)



Page

Unused space

14

FIXED-LENGTH RECORDS, PACKED

15

No gaps between records

Record ID = (page ID, position in page)

Easy to compute the offset of a record in the page:

$$\text{sizeof(Header)} + (\text{position} - 1) * \text{sizeof(Record)}$$

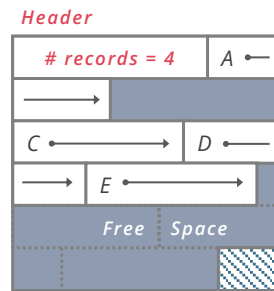
Delete record

Move the last record to the emptied slot

But this changes the last record's ID!

Not always desirable

Updating record ID pointers could be expensive if they're in other files



Page

Unused space

15

FIXED-LENGTH RECORDS, PACKED

16

No gaps between records

Record ID = (page ID, position in page)

Easy to compute the offset of a record in the page:

$$\text{sizeof(Header)} + (\text{position} - 1) * \text{sizeof(Record)}$$

Delete record

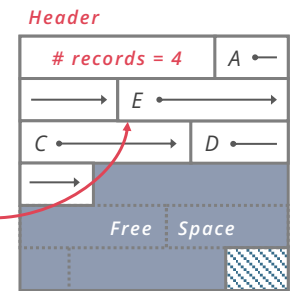
Move the last record to the emptied slot

But this changes the last record's ID!

Not always desirable

Updating record ID pointers could be expensive if they're in other files

Record ID:
(Page 2, 2)



Page

Unused space

16

FIXED-LENGTH RECORDS, UNPACKED (BITMAP)

17

Allow gaps between records

Record ID = (page ID, position in page)

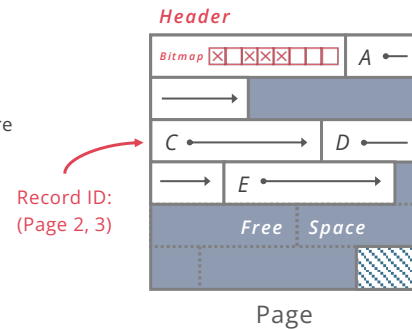
Use a **bitmap** to keep track of where the gaps are

Insert record

Find first empty slot by scanning the bitmap

Delete record

Clear bit in the bitmap



FIXED-LENGTH RECORDS, UNPACKED (FREE LIST)

18

Alternative to using bitmap

Link all free slots into a **free list**

Each link points to the beginning of a free slot, last is null

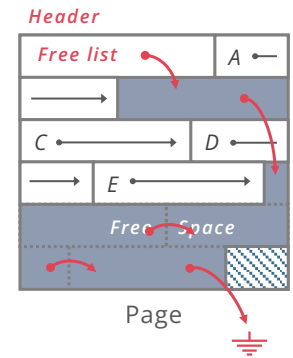
Insert record

Insert into slot pointed by head of free list

Set next free slot as new head

Delete record

Set slot of deleted record as new head



17

18

FIXED-LENGTH RECORDS, UNPACKED (FREE LIST)

19

Alternative to using bitmap

Link all free slots into a **free list**

Each link points to the beginning of a free slot, last is null

Insert record

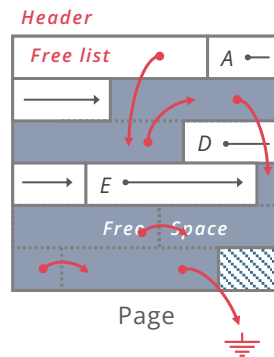
Insert into slot pointed by head of free list

Set next free slot as new head

Delete record

Set slot of deleted record as new head

Example: after deleting record C



19

VARIABLE-LENGTH RECORDS

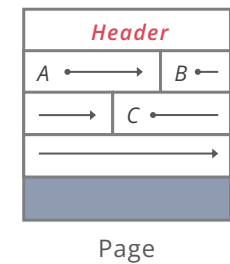
20

Variable-length records may not have field lengths consistent

E.g.: The third field may take 0 to 4 bytes

How do we know where each record begins?

What happens when we add and delete records?



20

SLOTTED PAGES

21

Most common layout scheme is called **slotted pages**

Slot directory maps “slots” to the records’ starting position offsets

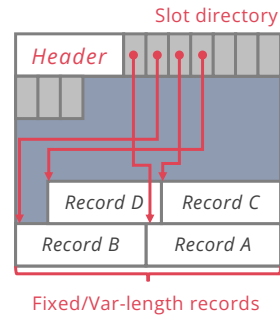
Record ID = (page ID, slot ID)

Header keeps track of:

The number of used slots

The offset of the last slot used

Records stored at the end of page



21

SLOTTED PAGES

22

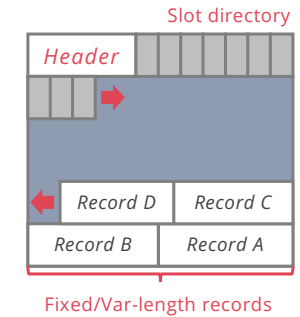
Records can be moved without changing rid

Delete record

Set slot offset to -1, delete slot only if last
Move records to fill up the whole or defragment space periodically

Insert record

Find a slot with offset -1 or create if none
Allocate just the right amount of space
Defragment if not enough free space



22

OUTLINE

23

Storage Media

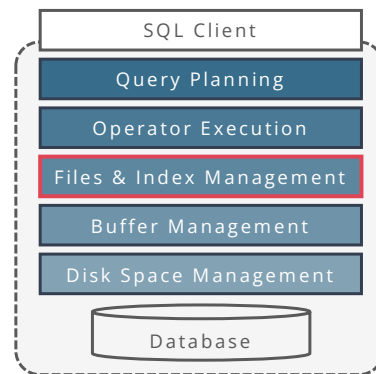
Disk Space Management

Buffer Management

File Layout

Page Layout

Record Layout



23

RECORD LAYOUT

24

Relational model

Each record in table has some fixed type

We do **not** need to store metadata about the schema

Information about field types is store in the **system catalog**
System catalog is just another set of tables

Goals:

Records should be compact in memory & disk format
Fast access to fields (why?)

Easy case: Fixed-length fields

Interesting case: Variable-length fields

24

RECORD LAYOUT: FIXED-LENGTH RECORDS

Each field has a **fixed** length

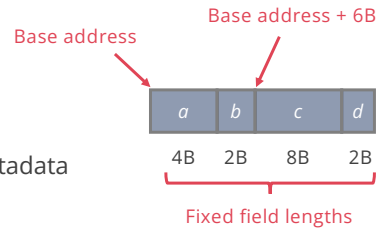
Direct access to record fields

Done via arithmetic (fast)

Each record can have a header storing metadata

E.g., bitmap for **NULL** values

No need to store information about the schema



VARIABLE-LENGTH RECORDS

Some fields have **variable** length

Two ways to store variable length records:

1. Fields delimited by special symbols

Access to fields requires a scan of the record

Special symbols in fields require "escaping"



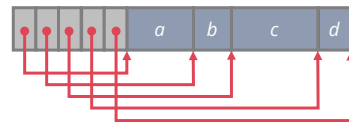
VARIABLE-LENGTH RECORDS

2. Array of field offsets

Direct access to fields & no "escaping"

Useful for fixed-length records too

Clean way of dealing with **NULL** values



Aside: this is actually not sufficient for storing NULL values for all types

Cannot distinguish between empty string ("") and NULL

Need some extra metadata (e.g. bitmap in record header or special char in field), which varies widely between different DBMSs

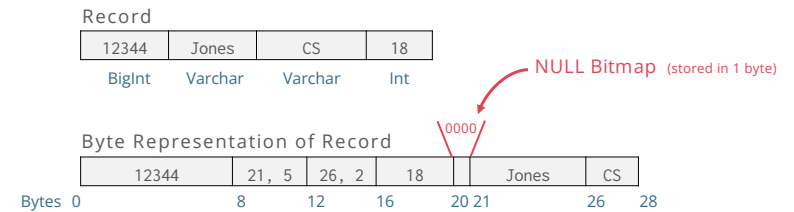
VARIABLE-LENGTH RECORDS

Possible implementation

Fields are stored in order

Variable-length fields represented by fixed size (**offset, length**), with actual data stored after all fixed-length fields

NULL values represented by NULL-value bitmap



SUMMARY

DB file contains pages, and records within pages

Heap files: unordered records organized with directories

Page layout

Fixed-length packed and unpacked

Variable length records in slotted pages

Variable-length record layout

Direct access to i -th field and NULL values

