

UNIVERSITY OF EDINBURGH
SCHOOL OF INFORMATICS
INFR11199 - ADVANCED DATABASE SYSTEMS (SPRING 2024)

Tutorial Sheet 2

1. (Files, Pages, Records) Consider the following relation:

```
CREATE TABLE Customer (  
    customer_id INTEGER PRIMARY KEY,    — cannot be NULL!  
    age INTEGER NOT NULL,  
    name VARCHAR(10) NOT NULL,  
    address VARCHAR(20) NOT NULL  
);
```

Assume that `INTEGER`s are 4 bytes long and `VARCHAR(n)` can be up to n bytes long.

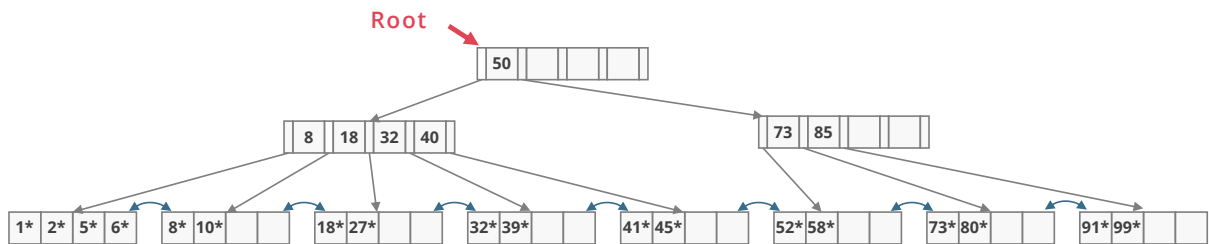
- (a) As the records are variable length, we will need a *record header* in the record. How big is the record header? You may assume pointers are 4 bytes long, and that the record header only contains pointers for variable-length values.
- (b) Including the record header, what is the smallest possible record size (in bytes) in this schema? Note: `NULL` is treated as a special value by SQL, and an empty string `VARCHAR` is different from `NULL`, just like how a 0 `INTEGER` value is also different from `NULL`.
- (c) Including the record header, what is the largest possible record size (in bytes) in this schema?
- (d) Suppose we are storing these records using a slotted page layout with variable length records. The page header contains an integer storing the record count and a pointer to free space, as well as a slot directory storing, for each record, a pointer and length. What is the *maximum* number of records that we can fit on a 8KB page?
- (e) Suppose we stored the maximum number of records on a page, and then deleted one record. Now we want to insert another record. Are we guaranteed to be able to do this? Explain why or why not.
- (f) Now suppose we deleted 3 records. Without reorganizing any of the records on the page, we would like to insert another record. Are we guaranteed to be able to do this? Explain why or why not.

2. (Buffer Management) We are given a buffer pool with 4 pages, which is empty to begin with. Answer the following questions given this access pattern:

A B C D E B A D C A E C

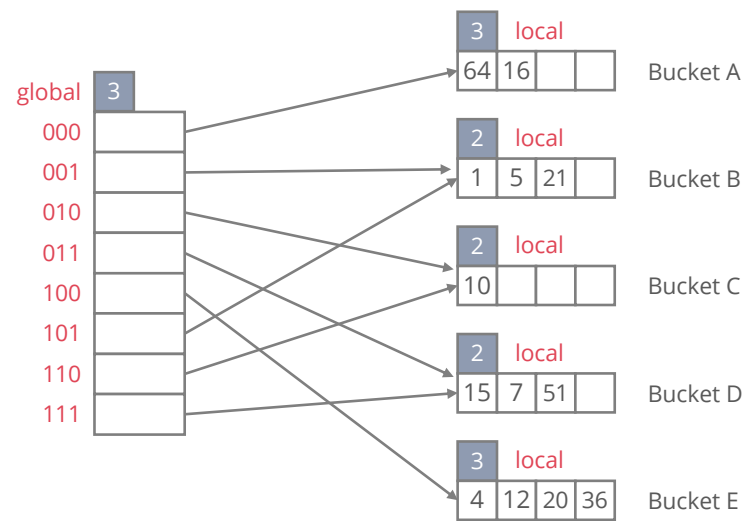
- What is the hit rate for MRU?
- For MRU, which pages are in the buffer pool after this sequence of accesses?
- What is the hit rate for clock? Assume the clock hand initially points at the first frame.
- For clock, which pages are in the buffer pool after this sequence of accesses?
- Which pages have their reference bits set?
- Which page is the hand of the clock pointing to?

3. (B+ Tree) Consider the following B+ tree index of order $d = 2$:



- Show the B+ tree that would result from inserting an index entry with key 9 into this tree.
- Show the B+ tree that would result from inserting an index entry with key 3 *into the original tree*. Assume no redistribution between siblings. How many page reads and page writes does the insertion require? Justify your answer.
- Show the B+ tree that would result from deleting the index entry with key 8 *from the original tree*, assuming that the left sibling is checked for possible redistribution.
- Show the B+ tree that would result from deleting the index entry with key 8 *from the original tree*, assuming that the right sibling is checked for possible redistribution.

4. (Extendible Hashing) Consider the following Extendible Hashing index:



- Draw the index after inserting an entry with hash value 68.
- Draw the index after inserting entries with hash values 17 and 69 into the original index.