VITMAB04 – Databases

Lecturer: Dr Levente Erős | **Teaching Assistant:** Dr János Varga

18 November 2021 – 5th tutorial: Physical Storage

1. We store a relation of 1,000 records using a sparse index. Size of a record is 850 bytes, block capacity (ignoring header) is 4,000 bytes. Key is 50 bytes; a pointer is 18 bytes.
    a. How many records fit in a block?
    b. How much disk space does the index occupy and what is the size of the entire relation?
    c. On which level, in which blocks and between which records do we need ordering?
    d. How long does it take (at most) to read the contents of a record if we suppose that the index is already in the RAM?
    e. How long does it take (at most) to read the contents of a record if we suppose that the index does not fit into the RAM?
2. A hash table with 7 buckets uses the following hash function: $h(k) = k \bmod B$. The storage engine receives the following records to store: 56, 91, 27, 19, 36, 52, 79.
    a. Assuming that one block can contain one record, what is the average record access time?
    b. Picking arbitrary hash functions what is the theoretically possible best and worst random record access time? (With same record count.)
3. A relation of 15,000 records is stored using a bucket hash. The size of one record is 120 bytes, the size of one block is 4,000 bytes, a key is 25 bytes, a pointer is 8 bytes. The hash table has 10 buckets. (Assume that the hash function distributes the values evenly.)
    a. What is the average bucket size?
    b. How much disk space does this structure occupy? (True size, Used disk space?)
    c. What is the average record access time given block access time of 5ms?
    d. What should be the minimum number of buckets if we wanted to achieve an average record access time of 5×[block access time]?

4. A relation can be stored in two ways: either a dense index combined with a single-level sparse index or hash algorithm. Interval lookups will sometimes be performed on the relation. Which storage structure should we choose?

   Calculate lower bound of the number of necessary blocks given that:

   - Relation has 3,000,000 records
   - Size of a record is 300 bytes
   - Size of a block is 4,000 bytes
   - A key is 45 bytes
   - A pointer is 5 bytes.

5. A relation of 10,000,000 records will be stored using a B*-tree. One record is 850 bytes, a block is 4,000 bytes. Key is 50 bytes and a pointer is 18 bytes. What is the minimum number of blocks needed? What is the average record access time if we can only store one block in the RAM at a time? (Block access time is 5ms.)
6. Fill the cells with the following words: „block” / „one”:

	How many records does one entry represent?	How many records does one pointer point to?
Dense index		
Sparse index		