

### Tutorial 5

**Question 1** Consider the following configuration:

- A disk with block size (B) = 512 bytes
- A block pointer (P) = 6 bytes

An EMPLOYEE file has the following fields:

NAME (30 bytes), SSN (9 bytes), DEPARTMENTCODE (9 bytes), ADDRESS (40 bytes), PHONE (9 bytes), BIRTHDATE (8 bytes), SEX (1 byte), JOBCODE (4 bytes), SALARY (4 bytes)

The EMPLOYEE file has  $N = 30,000$  EMPLOYEE records of fixed-length format and there is an additional byte used as a deletion marker in each record.

- A. Calculate the record size (R) in bytes.
- B. Calculate the blocking factor (bfr) and the number of file blocks (b) assuming an unspanned organization.
- C. Suppose the file is ordered by the key field SSN and we want to construct a primary index on SSN.
  - a. Calculate the index blocking factor (bfri).
  - b. Calculate the number of index entries and the number of index blocks.
- D. Suppose the file is not ordered by the key field SSN and we want to construct a secondary index on SSN.
  - a. Calculate the index blocking factor (bfri).
  - b. Calculate the number of index entries and the number of index blocks.
- E. Suppose the file is ordered by the non-key field DEPARTMENTCODE and we want to construct a clustering index on DEPARTMENTCODE that uses block anchors (every new value of DEPARTMENTCODE starts at the beginning of a new block). Assume there are 1000 distinct values of DEPARTMENTCODE, and the EMPLOYEE records are evenly distributed among these values.
  - a. Calculate the index blocking factor (bfri).
  - b. Calculate the number of index entries and the number of index blocks.
- F. Now we want to extend this single-level clustering index to a multi-level index.
  - a. Calculate the number of levels needed if we make it a multi-level index.
  - b. Calculate the total number of blocks required by the multi-level index.
  - c. Calculate the number of block accesses needed to search for and retrieve all records in the file having a specific DEPARTMENTCODE value using the multi-level index. Assume that multiple blocks in a cluster are contiguous. That is, data records with the same DEPARTMENTCODE are stored in adjacent blocks.

## INFS2200/7903 – Relational Database Systems

School of Information Technology and Electrical Engineering (ITEE), UQ

**Question 2** Consider an unsorted file which has  $N = 10,000$  Movie records. It is stored on a disk with the following configuration:

- Block size ( $B$ ) = 1000 bytes
- Block pointer size ( $P$ ) = 6 bytes

- A.** Suppose that a secondary index is constructed on the primary key MovieID (integer, 4 bytes). Calculate the number of block accesses needed for an equality search on the primary key using the secondary index.
- B.** Suppose that a B+ tree index is constructed on the primary key MovieID (integer, 4 bytes). Calculate the maximum number of block accesses needed for an equality search on the primary key using the B+ tree index.