Assignment 6

Qinglin Li, 5110309074

Problem 1

RAID level 1.Because rebuilding in this case just involves copying data from the failed disk's mirror. In the other levels, rebuilding involves reading the extra contents of other disks.

Problem 2

By allocating related records to blocks, we can often retrieve most of the requested records by a query with one disk access. So this strategy reduces the number of disk accesses for a given operation, and significantly improves performance.

Problem 3

1. Advantages of storing a relation as a file include using the file system provided by the OS, thus simplifying the DBMS, but incurs the disadvantage of restricting the ability of the DBMS to increase performance by using more sophisticated storage structures.
2. By using one file for the entire database, these complex structures can be implemented through the DBMS, but this increases the size and complexity of the DBMS.

Problem 4

It is preferable to use a dense index instead of a sparse index when the file is not sorted on the indexed field (such as when the index is a secondary index) or when the index file is small compared to the size of memory

Problem 5

Hashing distributes search key values uniformly and randomly across the set of buckets available. Because key values do not occupy consecutive buckets, searching for all values within a range may well require reading every bucket — far more inefficient than a B+-tree or even a sorted file without index.

Problem 6

The existence bitmap for a relation can be calculated by taking the union (logical-or) of all the bitmaps on that attribute, including the bitmap for value null.