COMP2411 HW2

JIANG Guanlin (21093962D)

QA:

a):
Hash Functions relative to B-tree indices:
Advantage:
1. Hash Function better and effective to search than B-tree, also stable.
2. Hash Function have better synchronization.

Disadvantage:
1. Hash Function can't do the range search, but B-tree can search in a range.
2. If Hash Function have collisions, the hash is not that good for save space.
3. Hash Function can't have null value.

The index to point to the data location within the database more faster.

b):
The Dence indexes as compared to the sparse index are faster, but use more memory. So, when the file is not ordered on the indexed field it is better to use the Dence index.

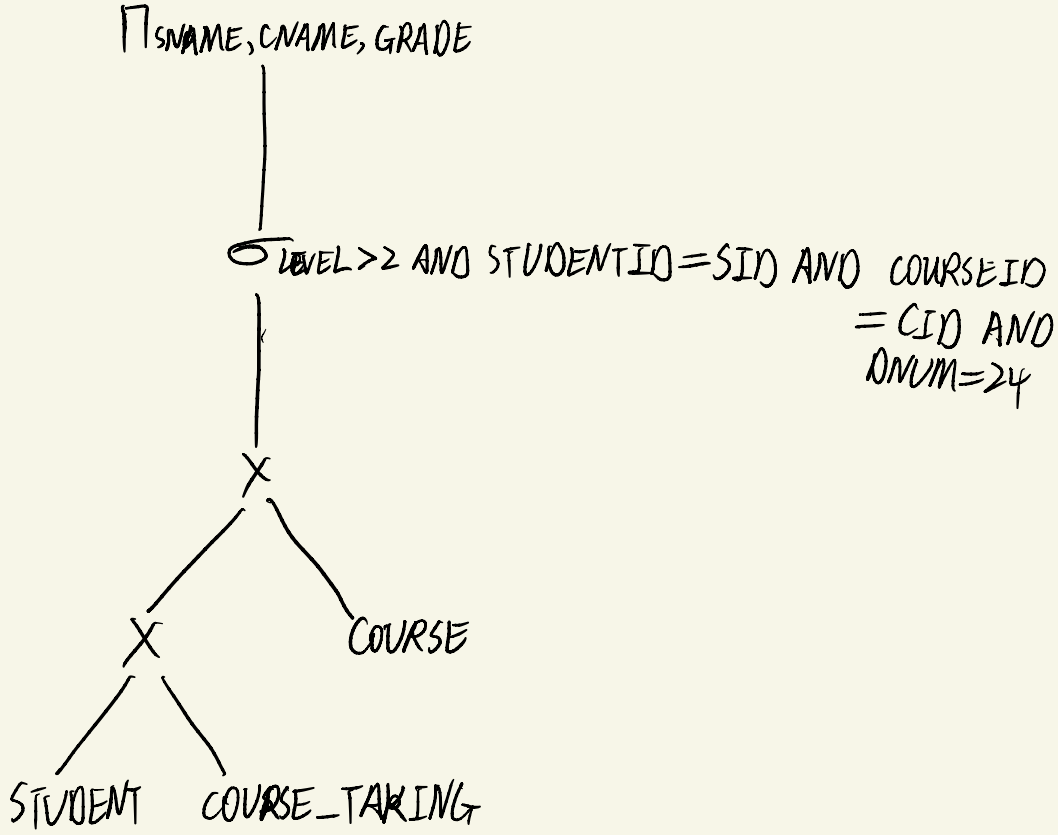
c):
Primary Index: The set of fields that includes the unique primary key and is will not be repeat.

Secondary Index: The set of fields will be not the primary index and may have repeat key.

d):
The property of hash functions have:
1. No same hash key (Unique)
2. Input content can map to hash value

QB:

a):



b):
Necessary Assumptions:

1. Among the JOIN operations, reranking the tree to do the one with the lowest join selectivity first.

2. Move the SELECT and PROJECT operations down to bottom of the query tree.

3. Among the SELECT operations, reranking the tree to do the one with the lowest selectivity factor first.

QC:

a):

Yes, I agree. The schedule is non-recoverable and cannot ensure transaction atomicity. Even we undo the command we do, but the logs that we update can't be undo. Also, undo the command will be let logs and database inconsistent.

b):

First will be undo in the reverse of the order, and after redo all the write operations of the transactions from the logs.

c):

When the system again recovers, the same system logs will be used, also the system will be do the same step in second recover, because first recovers was did before the second crash.