Final Review Database

Physical data independence – not tied to specific physical hardware

Logical Data independence – is data from different systems, make consistent through views.

\*\*\*Normalization\*\*\*

1st Normal Form – all the intersections of rows/columns must be atomic

2nd Normal Form – 1st normal form and there are no partial dependencies on the key.

3rd – is the 1st and 2nd and no multiple dependencies on the key. Every non-key attribute is dependent on the key and nothing but the key.

BCNF – is 3rd normal form and no interior transitive dependencies.

Functional dependencies. A -> B, C

Directional and help us identify the keys.

The difference between primary key and foreign key; Foreign key is a value in one table that must match the primary key in another table or be null. Foreign key is always on the many side.

Lossless joins – decomposition, join back new tables and must have the original data. Trying to not lose information. If you make more information out of nothing then it is meaningless.

Insert Anomaly – when it cannot be null and you don’t know the value you have to make up data to insert the row.

Delete Anomaly – Only one item in one field, if you delete one you remove all other fields. IE, 1 person in 1 zipcode in the same table, if you delete person it removes the zipcode.

Update Anomaly – If you update one value in a column, if its in separate tables, you only have to change once, however the anomaly is when you have them all in one table and you have to change it every time where it is located.

Weak entity – is a table that is entirely dependent on other tables.

Entity subtype- “is a” relationship between two tables. IE, People -> Faculty, Students.

Grant/Revoke privileges.

Database Administration – Control over security policy, look at performance, makes it run faster, query optimization. Concerned about space and time, durability. Responsible for backup and restore.

Nested Loop(n^2), Hashing(takes a value puts it into one of many buckets), Merge(both tables must be sorted)-> O(n)

Clustered Index – physical order of the rows stored on the system.

Logical Index – pointer system in tree architecture.

What is a transaction? Begin -> Commit/Rollback

Group of commands the DB does and does it all or nothing. “EON”

ACID – Atomic (EON), Consistency (Locking,), Independent/Isolated (even though the DB does multiple simultaneous transactions it appears as if they are doing them independently or isolated), Durability (rebuild the server, backup from write-ahead logs; continue what it was doing or rollback ”EON”).

Write-ahead logs – record all of the before images, as the transaction completes, the after images are recorded. When DB goes down it looks at the logs and determine to continue forward or rollback. \*\*CheckPoints\*\*

Commit Consistent Checkpoint- As of this point everything is good and you can delete everything before this point. (Transactions must stop).

Partial Checkpoint – everything is good but record a snapshot of current transactions so you can delete everything before it.

Locks:

Terms – Short/Long

Type – Read(shared) or Write(exclusive)

Granularity – How much we are locking?

* DB
* Table
* Page – not really used anymore
* Row
* Fields

1. Read uncommitted/Dirty Read -> No locks
2. Read committed -> LX/ SS
3. Repeatable Read -> LX/LS
4. Serializable -> LX/LS/ Key-Range Locks/Predicate Locks-> prevents phantoms because it locks the index

\*\*\*Systems Catalog\*\*\*

OLTP – Online Transaction Processing

OLAP/BI/BA – Online Analytical Processing/Business Intelligence/ Business Analytics

ETL – Extraction Transformation Loading

How do we detect deadlock?

Everything is waiting on something else to be finished, so nothing gets done.

Graphs with arrows and edges, “wait for” graph, if that graph has a cycle there is deadlock. Abort a transaction and roll it back.

Analytical DB – take normalization -> star schema

Star schema – dimensional modeling, CFT (Central Fact Table) – all tables combined

Store, Product, Qty, Customer, Ord. No

Transactions BY **customer, products, orders, etc** -> all are dimensions.