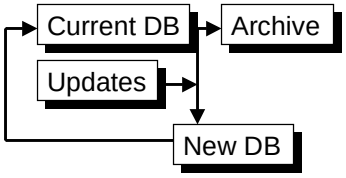
Transactions: Changing A DB Dynamically

CSE 100

Some databases are static, e.g. the taxonomy of flowering plants. Others add new records slowly over time, e.g. employees. The most valuable databases are those that change continually, e.g. credit card accounts. These databases use process *transactions*.

CSE 100 Organizing Business Data

- Companies, universities, government agencies, etc. have many database applications in common
  - Employee records
  - Payroll records
  - Customers/clients/students records
  - Products/services listings
  - …



- Databases, as described thus far, are completely adequate for representing and managing this data
- Database changes can be controlled
  - Changes can be made by "authorized" employees
  - Changes can be made periodically, in batches

CSE 100 Maintaining On-line Databases

- Many databases are only useful if they are on-line, i.e. can be changed interactively
  - Airline reservations
  - Credit card accounts, ATMs and other banking
  - Catalog merchandising
  - eCommerce
  - …
- On-line changes are called transactions
- The “transactions concept” not only makes interactive DBs possible, it is a good metaphor for other database processing as well

CSE 100 Transactions ...

- A transaction is a single operation (reference or change) to a database usually involving only one or a few records
  - Credit card purchase
  - ATM withdrawal of cash
  - Flight reservation
  - …
- Many transactions are taking place at once, typically
- Keeping the DB “correct” is a problem!

VISA processed 110,086,395 transactions on December 14, 1998, a 1-day world’s record

CSE 100 Correct Database?

- Because transactions take place simultaneously, there is a possibility that two computers can be making changes to the same data at the same instant, possibly corrupting it …
- Consider two transactions
  - T1: Deposit $100 into Account #12345
  - T2: Withdraw $100 from Account #12345
- When transactions are over, the balance *should be* unchanged
- But what if the transactions take place “at once”?

CSE 100 Tale Of Two Transactions

Balance: $500

Time

T1: Deposit $100	T2: Withdraw $100
T1.1: Fetch balance for Acct #12345: $500.00	**T2.1: Fetch balance for Acct #12345: $500.00**
	T2.2: Balance >= $100?
Balance: $400	**T2.3: Yes, Bal = Bal - $100**
	T2.4: Set DB balance to $400
	T2.5: End transaction
	Balance: $600
T1.2: Bal = Bal + $100	
T1.3: Set DB balance to $600	
T1.4: End transaction	

- The depositor might be pleased, but the bank won’t be!

CSE 100 Correctness

- The DB System must assure that every change to the database happens as if the transactions took place one-at-a-time
- The one-at-a-time protocol solves the “problem” with T1 and T2:
  - T1 applied first, then T2: $500 --> $600 --> $500
  - T2 applied first, then T1: $500 --> $400 --> $500
- Transaction processing systems make sure that such problems do not arise by “locking” the data (only one computer at a time can unlock the data) and using a special protocol to perform the locking

CSE 100 Transactions Are A Metaphor

- Recall the Students and Classes database design

Classes : Table

Field Name	Data Type
ClassID	AutoNumber
ClassName	Text
DepartmentID	Number
SectionNumber	Number
InstructorID	Number
Term	Text
Units	Text
Year	Number
Location	Text
DaysAndTimes	Text
Notes	

Students : Table

Field Name	
StudentID	Auto
FirstName	Text
LastName	Text
Address	Text
City	Text
StateOrProvince	Text
PostalCode	
PhoneNumber	
Major	
StudentNumber	

Departments : Table

Field Name	Da
DepartmentID	AutoN
DepartmentName	Text
DepartmentNumber	Numbe
DepartmentManager	Text
DepartmentChairperson	Text

Assignments : Table

Field Name	Data Type
AssignmentID	AutoNumber
AssignmentDescription	Text
ClassID	Number
Exam	Yes/No
PercentOfGrade	Number
MaximumPoints	Number

Students And Classes : Table

Field Name	Data Ty
StudentClassID	AutoNumber
ClassID	Number
StudentID	Number
Grade	Text

Taking a class is a transaction

CSE 100 Reliability

- What happens to the database when …
  - The power goes out
  - Someone spills coffee into the disk drives
  - The computer crashes with all the changes to the DB for the last three hours in its (volatile) RAM
  - A new employee accidentally deletes the payroll file before printing the pay checks?
  - A virus cleans off the corporate disks
  - A hacker infiltrates the enterprise and begins transferring funds to a Swiss bank account
  - A disgruntled employee deletes the retirement plans and stock option accounts for senior management
  - ...

CSE 100 Basic Mechanisms

- There are a series of techniques that preserve the integrity of the data
  - Error detection/correction in the hardware
  - Passwords and authentication assist in verifying that the person(s) making changes are legitimate
  - Validation … verify that changes to the DB are "plausible"
  - "Commitment" … keep record in a safe place of all changes to the database, and then when it has been verified, make the actual change effective; deletions never actually result in the data being removed

Backup copies of a DB must be made regularly, and kept off-site

CSE 100 Redundancy

- To protect against computer crashes, disk failures, loss of power, etc. duplicate the hardware, disks, power sources, etc.
- The duplicate systems can compare answers as a means of detecting errors
- RAID systems are arrays of disks that contain “hot spares” and special data encodings to recover from disk failures
- By keeping a snapshot of the database and a record of all of the transactions, it is possible in case of catastrophic disaster to reconstruct the database by applying all of the transactions to the old database