



JOHNS HOPKINS

WHITING SCHOOL
of ENGINEERING

Introduction to Databases and SQL

Tamás Budavári

Dept of Applied Mathematics & Statistics

Dept of Computer Science

Dept of Physics & Astronomy



Databases

- ❑ Store your bytes
 - ▣ And return them
- ❑ Efficient filtering
 - ▣ Thousands of man-years
 - ▣ Optimally pick from many strategies



Databases

- ❑ SQL basics
 - ▣ filtering, aggregation, joins
- ❑ SQL programming
 - ▣ variables, functions, procedures
 - ▣ data management, transactions
- ❑ Tools



SQL: Structured Query Language

- ❑ Standard declarative language
- ❑ Filter the data
- ❑ Powerful analysis tool
- ❑ Possible to extend





Transactions

ACID Properties

The Elevator Problem

- People on multiple levels
 - ▣ Press the button...



Mutual Exclusion

- ❑ Multiple processes or threads
 - ▣ Access shared resources in critical sections
 - E.g., call the elevator when it's time to go
- ❑ Locking
 - ▣ Elevators, etc...

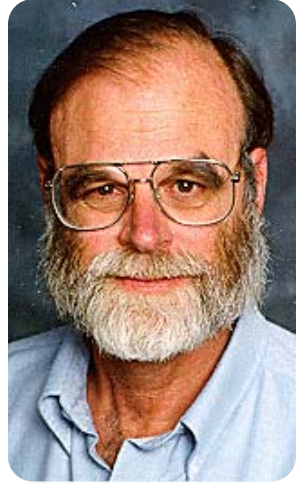
Dining Philosophers

- Five silent philosophers sit at the table
 - ▣ Alternate between eating and thinking
 - ▣ Need both forks left & right to eat
 - Must be picked up one by one!
 - ▣ Infinite food in front of them
- How can they all think & eat forever?



Transactions

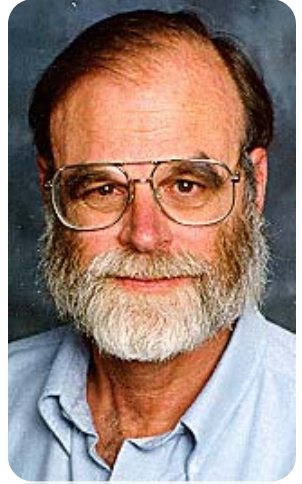
- ACID Properties
 - ▣ Atomicity
 - ▣ Consistency
 - ▣ Isolation
 - ▣ Durability



Defined by Jim Gray

Atomicity

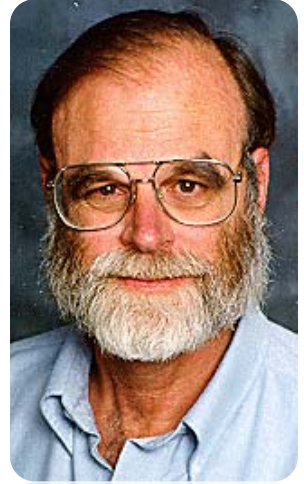
- All parts of a transaction succeed
- Or rollback to previous state



Defined by Jim Gray

Consistency

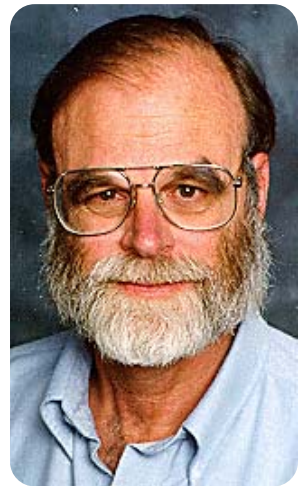
- Data always meets validation rules
 - ▣ Any type of constraints



Defined by Jim Gray

Isolation

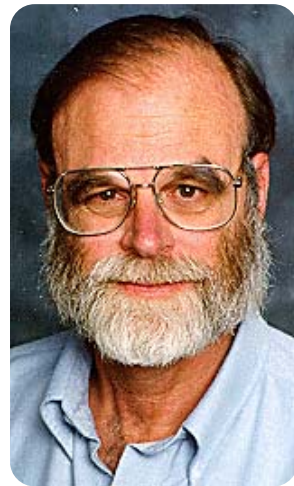
- No interference across transactions
 - ▣ Even if concurrent



Defined by Jim Gray

Durability

- Committed transaction will remain so
 - Even in the event of power failure, errors
 - Caching in harddrives, etc.



Defined by Jim Gray

Database Systems

- ❑ MySQL – open source
- ❑ PostgreSQL and Greenplum
- ❑ Microsoft SQL Server (Express Edition)
 - Same as the one under our LabDB
- ❑ IBM DB2 (Express-C)
- ❑ Oracle Database (XE-Express Edition)

Server – Client

- ❑ Multiple clients
- ❑ Different client apps
 - ▣ Graphical UI
 - ▣ Command line
 - ▣ Your custom analysis

Database Systems

- SQLite – minimalist db
 - ▣ To start with the smallest
 - ▣ Command line executable
 - Single library for coding
 - ▣ Also a pure C# implementation

Database Systems

- MonetDB

- ▣ Column store

- SciDB

- ▣ Array database under development

Research Them!

- Which one?
- Why?
- Things to consider
 - ▣ How much data? Scale to my problem?
 - ▣ Extensibility for scientific analysis?
 - ▣ Hardware requirements? What OS?

Programmatic Interface

- Send SQL commands
- Read out the results
- Standard ODBC
 - ▣ Open Database Connectivity
 - ▣ C interface
- JDBC for Java



SQL by Examples

SQL by Examples

- Interactive session with exercises
 - ▣ See handouts...





JOHNS HOPKINS

WHITING SCHOOL
of ENGINEERING

