
data
distributed
database
DBMS
parallel
shared
architecture
system
systems
network
processing
DBMSs
Distributed
DDBMS
processors
access
centralized
performance
memory
multiple
sites
site
processor
single
disk
computer
applications
application
nothing
technology
Parallel
stored
may
fragments
require
physically
local
provide
one
number
logically
decentralized
typically
major
users
computers
Database
control
scalability
approach
improve
logical
distinction
development
definition
nodes
Shared
databases
Figure
bottleneck
interconnection
transparent
management
organizational
administered
consists
sometimes
computing
split
order
scalable
architectures
islands
resources
large
disks
remote
defined
controlled
based
directly
Date
com
make
replicated
also
discuss
many
often
reliability
eliminates
across
fact
communication
distribution
referred
operational
global
way
used
least
capable
must
System
user
dis
per
makes
Management
within
machines
like
partitioned
associated
structure
collection
incompatible
provides
mode

Database System Concept (CSE 3103)

Lecture 08-Day 03

Nazmus Sakib, Assistant Professor, Dept. of CSE, AUST

Serializability

- **Basic Assumption** – Each transaction preserves database consistency.
- Thus, serial execution of a set of transactions preserves database consistency.
- A (possibly concurrent) schedule is serializable if it is equivalent to a serial schedule. Different forms of schedule equivalence give rise to the notions of:
  1. **conflict serializability**
  2. **view serializability**

Simplified view of transactions

- We ignore operations other than **read** and **write** instructions
- We assume that transactions may perform arbitrary computations on data in local buffers in between reads and writes.
- Our simplified schedules consist of only **read** and **write** instructions.

Conflicting Instructions

- Let I_i and I_j be two Instructions of transactions T_i and T_j respectively. Instructions I_i and I_j **conflict** if and only if there exists some item Q accessed by both I_i and I_j, and at least one of these instructions wrote Q.
  1. I_i = **read**(Q), I_j = **read**(Q). I_i and I_j don't conflict.
  2. I_i = **read**(Q), I_j = **write**(Q). They conflict.
  3. I_i = **write**(Q), I_j = **read**(Q). They conflict
  4. I_i = **write**(Q), I_j = **write**(Q). They conflict
- Intuitively, a conflict between I_i and I_j forces a (logical) temporal order between them.
  - If I_i and I_j are consecutive in a schedule and they do not conflict, their results would remain the same even if they had been interchanged in the schedule.

Conflict Serializability

- If a schedule *S* can be transformed into a schedule *S´* by a series of swaps of non-conflicting instructions, we say that *S* and *S´* are **conflict equivalent**.
- We say that a schedule *S* is **conflict serializable** if it is conflict equivalent to a serial schedule

Conflict Serializability (Cont.)

- Schedule 3 can be transformed into Schedule 6 -- a serial schedule where T_2 follows T_1, by a series of swaps of non-conflicting instructions. Therefore, Schedule 3 is conflict serializable.

T_1	T_2
read (A)	
write (A)	
	read (A)
	write (A)
read (B)	
write (B)	
	read (B)
	write (B)

Schedule 3

T_1	T_2
read (A)	
write (A)	
read (B)	
write (B)	
	read (A)
	write (A)
	read (B)
	write (B)

Schedule 6

Conflict Serializability (Cont.)

- Example of a schedule that is not conflict serializable:

T_3	T_4
read (Q)	
	write (Q)
write (Q)	

- We are unable to swap instructions in the above schedule to obtain either the serial schedule < T_3, T_4 >, or the serial schedule < T_4, T_3 >.