Recovery

From System Crash

Types of Failure

• Catastrophic -- Physical Damage to data
  – Restore thru last backup from archived data
  – Restore to present using Log
• Non-catastrophic
  – Restore to present using Log
Logs, Buffers and Disks

Updates written to log and buffers

Commit written to log

Blocks read from disk

Forcewriting Log

- Why is log written first?
- Why delay in writing buffers?
- Promise can be kept after write to log
- Many changes preserved with one log write
- Other transactions can read buffers instead of disk
- No hurry to write to data disk
Requirements on the Log

- Cannot tell from log whether data got to disk.
- After crash, state of the disk is unknown.
- After crash, all transactions are dead
- Must undo changes from uncompleted transactions
  - If uncommitted writes allowed to get to disk
- Must redo changes from committed transactions
- No cascading rollback
Note on Reads in Log

- Read information used only when cascading rollback possible
- Not used for system crash recovery
- Used only for normal processing.

Conceptual Phases of Recovery

- ROLLBACK -- go backward thru log
  - Note COMMITS
  - UNDO writes of uncommitted transactions
- ROLL FORWARD -- now go forward thru log
  - REDO writes of committed transactions
Recovery Example

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30</td>
<td>15</td>
<td>40</td>
<td>20</td>
</tr>
</tbody>
</table>

- **Start, T3**
- **Read, T3, C**
- **Write, T3, B, 15, 12**
- **Start, T2**
- **Read, T2, B**
- **Write, T2, B, 12, 18**
- **Start, T1**
- **Read, T1, A**
- **Read, T1, D**
- **Write, T1, D, 20, 25**

**Commit, T1**

**Read, T2, D**

**Write, T2, D, 25, 26**

**Read, T3, A**

**Write, T3, A, 30, 35**

**End of Log**

---

**Note on Buffers and Disk**

- **UNDOING** and **REDOING** operate on buffers read in from disk.
- Several **REDOs** to same data item will all be written at once.
- System may crash during recovery.
- Still don’t know if data got to disk so whole recovery process must proceed again.
Longer and longer Recoveries

- Say backups daily at 4 AM.
- Log in morning is small and recovery quick
- Log in evening is large and recovery slow
- Same data REDONE or UNDONE many times
- Solve with Checkpoints