Recovery

Database Systems
Connolly/Begg/Strachan
Chapter 17

Fundamental of Database Systems
Elmasri/Navathe
Chapter 19/21
## Definitions

- **Disc Block**
  - The computer reads or writes one block of data when it accesses the disc.

- **Buffers**
  - Area of memory used to store blocks accessed on the disc.

- **Variable**
  - Location in memory used to store data values.

- **Transaction State**
  - Current condition of a transaction, e.g. committing.
Why do DBMSs need recovery?

Rollback means that transactions have different states.

Immediate Update

What would be a poor recovery system?

Deferred Update

There are two main approaches to updating the database.

Transaction Recovery

Can the system log be made more efficient?

System Log

Where are these operations recorded?

Example of checkpoints

What is an example of checkpoints?

Example of a recovery system

To implement recovery, the DBMS must record information about the transaction.

States of a Transaction

What is an example of recovery?

Poor Recovery Procedures

What is an example of recovery system?

Operations in a Transaction

Rollback means that transactions have different states.

Immediate Update

Deferred Update

Transaction Recovery

System Log

Where are these operations recorded?
A DBMS must never allow a transaction to partially complete, that is, to leave some operations completed and others uncompleted. This would destroy the atomicity of the transaction.

Connolly et al defines database recovery as “the process of restoring the database to a correct state in the event of a failure”.

To perform recovery the DBMS must:

1. Recognise that the system is in an inoperable or inconsistent state,
2. Determine the most useful state,
3. Transform the database into the useful state.

The objective of the recovery procedure is to:

Minimise the loss of data,

Return the database as close as possible to the state it was in before the failure.

Ref: Connolly, 17.3; Elmasri, sec 19.1.4
Why is recovery necessary?

<table>
<thead>
<tr>
<th>DBMS</th>
<th>Read disc block into buffer.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Changes contents of buffer.</td>
</tr>
<tr>
<td></td>
<td>Writes the buffer to the disc.</td>
</tr>
</tbody>
</table>

Problem: What if the DBMS fails before the buffer is written to disc?

The DBMS accesses the database by reading and writing disc blocks into and out of buffers held in main memory.

To make a change to the database, the DBMS must:

1. Find the disc block that is to be changed.
2. Read the disc block into a buffer in main memory.
3. Copy the data item from the buffer to a variable
4. Change the contents of the variable.
5. Copy the contents of the variable to the buffer.
6. Write the buffer to a disc block.

The process of writing the contents of a buffer to a block on the disc does not happen when the buffer changes.

The DBMS will wait until the database buffers are full or the transaction that changed the data item is committed before it writes the buffer to disc.

While the database changes are in the buffer and have not been written to disc they are not permanent.

If the DBMS fails then changes made to the buffers but not written to disc may be lost.

Ref: Connolly, 17.3; Elmasri, sec 21.1
Poor Recovery Procedures

- If a transaction fails either:
  - Re-execute the transaction, or
  - Abandon the transaction.
- Both alternatives result in an inconsistent database state.
- Solution
  - Record information about how the transaction has changed the database.

When a transaction fails the DBMS must have a strategy for recovering the database. Two poor strategies are:

1. Re-execute the transaction from the beginning, or
2. Abandon the transaction and leave the database in its current state.

Both strategies fail. For example, assume a transaction transfers £10 from account A (balance £50) to account B (balance £60) but fails after the money is deducted from A but before it is added to B. Using the poor recovery procedures will result in either:

1. Account A with balance £30 (£10 deducted twice) and account B with balance £70 (£10 added once), or
2. Account A with balance £40 (£10 deducted once) and account B with balance £60 (no money added).

Both approaches produce an incorrect state. The correct balances should be account A with £40 and account B with £70.

This problem occurs because the transaction fails after the money is deducted from account A but before it is added to account B.

The solution is to record the fact that money was deducted from A and, if the transaction fails, add the deducted amount back to A.
Overview

Why do DBMS need recovery?

States of a Transaction

Example of a recovery system

Immediate Update

Deferred Update

Checkpoints

Example of checkpoints

Transaction Recovery

System Log

What is an example of recovery?

What would be a poor recovery system?

There are two main approaches to updating the database.

How is the system log used to implement recovery?

Where are these operations recorded?

Rollback means that transactions have different states.

To implement recovery the DBMS must record information about the transaction.

Why do DBMS need recovery?

Why?

Poor Recovery Procedures

Transactions have different states.
A transaction goes through a number of states as it is processed. At any point in time the transaction is in one of the states.

The transaction starts in the active state and remains in the active state while all its statements are being executed.

If one of the statements fails, for example, an insert cannot be executed, then the transaction enters the ‘failed’ state.

If the transaction executes all the statements successfully then it enters the ‘partially committed’ state. During this stage all the buffers are written to disc.

    If the database writes are successful then the transaction enters the ‘committed’ state.
    If the database writes are not successful then the transaction enters the ‘failed’ state.

A transaction in the failed state has not executed properly and must be rolled back, that is, all its changes to the database must be undone.

    Once the changes have been undone, the transaction enters the ‘abort’ state.

    An aborted transaction can be abandoned or re-run.
In this example, six transactions have been executing when the system fails for some reason.

The DBMS may take one of two actions regarding each transaction:

1. The transaction may be *undone* (rolled back).
2. The transaction may be *redone* (rolled forward).

Without any further information about the transactions, the DBMS must take the following actions:

- **T<sub>1</sub>** has not committed when the failure occurred, so it must be *undone*.
- **T<sub>2</sub>** has committed but the DBMS will not know if the disc has been updated or if the transaction’s updates were still in the buffer and lost during the failure. **T<sub>2</sub>** must be *redone*.
- **T<sub>3</sub>** has been committed but, as with **T<sub>2</sub>**, it must be *redone*.
- **T<sub>4</sub>** has been committed but, as with **T<sub>2</sub>**, it must be *redone*.
- **T<sub>5</sub>** has been committed but, as with **T<sub>2</sub>**, it must be *redone*.
- **T<sub>6</sub>** has not been committed when the failure occurred, so it must be *undone*.

*Ref:* Connolly, sec. 17.3.
Overview

Why do DBMS need recovery?

Rollback means that transactions have different states.

What would be a poor recovery system?

There are two main approaches to updating the database.

Can the system log be made more efficient?

Where are these operations recorded?

To implement recovery the DBMS must record information about the transaction.

States of a Transaction

Example of a recovery system

Operations in a Transaction

System Log

Example of checkpoints

Checkpoints

Immediate Update

Deferred Update

Why?

Why do DBMS need recovery?

What is an example of checkpoints?

What is an example of recovery?

Transaction Recovery

What would be a poor recovery system?
To be able to support transaction processing and database recovery the DBMS keeps track of certain operations:

1. **When a transaction begins.**
   If the transaction fails then the database must be returned to the state it was in before the transaction began.

2. **Read and write operations.**
   These changes must be undone if the transaction is rolled back.

3. **The end of the transaction.**
   The transaction can end in two ways:
   a. When the transaction succeeds, called COMMIT, or
   b. When the transaction fails, called ROLLBACK.

When a transaction is committed all the changes it has made to the database become permanent and are seen by other transactions.

When a transaction is rolled back all the changes it has made to the database are undone and the database is returned to the state it was in before the transaction began. Effectively, the transaction never occurred.
The system log is a list of all the operations which have been performed on the database. It is updated by the DBMS each time the database changes. The log can be used to undo a set of operations for a particular transaction. For example, if a transaction deletes a tuple from a relation and then fails, the system log can be used to insert the tuple again.

Each transaction has a unique identifier (T) which is recorded in the log alongside each operation. The operations recorded in the log include:

1. [start_transaction, T] marks the start of transaction T.
2. [write_item, T, X, old, new] records a write of item X by transaction T which changed value ‘old’ to ‘new’.
3. [read_item, T, X] records a read of item X by transaction T.
4. [commit T] records the end of database accesses by transaction T.

All transactions which have been started ([start_transaction] in the log) but have not been committed (no [commit] in the log) must be undone if there is a system failure.

Ref: Elmasri, 19.2.2; Connolly sec. 17.3.3.
The system log is used to undo the uncommitted changes in the database and redo committed changes that have been lost.

For example,

1. \([\text{start\_transaction} \ T1]\)
2. \([\text{read\_item}, \ T1, \ x]\) - reads \(x = 10\)
3. \([\text{write\_item}, \ T1, \ x, \ 10, \ 20]\) - writes \(x = 20\)

After (3) if transaction \(T1\) fails the \([\text{write\_item}]\) operation must be undone by replacing \(x=20\) with \(x=10\) (the old value).

During rollback the system log is read in reverse order. That is, the most recent operations are undone first.
Overview

Why?

Why do DBMS need recovery?

Rollback means that transactions have different states.

States of a Transaction

What is an example of recovery?

Example of a recovery system

Immediate Update

What would be a poor recovery system?

Deferred Update

There are two main approaches to updating the database.

Operations in a Transaction

To implement recovery the DBMS must record information about the transaction.

Example of recovery system

System Log

Where are these operations recorded?

Transaction Recovery

How is the system log used to implement recovery?

Checkpoints

Can the system log be made more efficient?

Example of checkpoints

What is an example of checkpoints?

The first major method of implementing database recovery is *deferred updates*. Changes to the database only occur when the transaction has committed, that is, successfully completed.

During the processing of a deferred update transaction changes to the database are recorded in the system log.

- If the transaction completes successfully the DBMS uses the log to apply the updates to the database.
- If the transaction fails the DBMS removes entries belonging to the transaction from the system log. The database remains unchanged.

After a failure the DBMS may have to re-apply updates which were being applied to the database when the failure occurred. That is, some transactions will have been committed but the DBMS failed when it was trying to apply the updates to the database.
The second major method of database recovery is *immediate updates*. Changes are made to the database when the transaction is executing.

Because a transaction can fail it is important to be able to undo changes made during the execution of the transaction. All database updates are recorded in the system log.

The log must be updated on disc before the database changes. Otherwise if a failure occurs between writing to the database and writing to the log, the DBMS will have no record of the changes.

If the transaction completes successfully then no changes need to be made to the database.

If the transaction fails then the information in the system log must be used to undo changes to the database.

All writes must be undone using the [write_item, T, X, old, new] entry in the log.

The DBMS applies each of the ‘write_item’ entries in reverse by replacing the new value with the old value.
Overview

Why do DBMS need recovery?

Rollback means that transactions have different states.

What is an example of recovery?

Example of a recovery system

To implement recovery
the DBMS must record
information about the
transaction.

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery procedures

Operations in a Transaction

System Log

How is the system log
used to implement
rollback?

Where are these operations recorded?

Transaction Recovery

Checkpoints

Can the system log be made more
efficient?

What is an example of checkpoints?

What do DBMS need recovery?

What would be a poor recovery system?

Example of recovery system

Operations in a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
updating the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery procedures

States of a Transaction

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.

Example of recovery system

Immediate Update

Deferred Update

What would be a poor recovery system?

There are two main approaches to
update the database.
When a transaction finishes, all the changes it has made to the database may not be written to disc immediately.

The DBMS may only update the database at fixed periods, for example, when the database buffers are filled.

This is okay if the system log has been written to disc as it can be used to recover any lost data.

However, the system log may become very large. When the system must be recovered the whole of the system log must be checked for operations that must be undone or redone. This will be a slow process.

As database changes that have been written to disc will never be undone/redone, the DBMS writes a [checkpoint] marker to the system log when it writes a set of updates to the database.

All transactions which have committed before a checkpoint will be written to disc and are permanent.

If a failure occurs then the DBMS only needs to undo operations up to the last checkpoint. All transactions which have been committed before the checkpoint will have been written to disc. Hence, to rollback the database the DBMS can stop the process when it reaches a checkpoint.

Ref: Connolly, sec. 17.3.3.
In this example, six transactions have been executing when the system fails for some reason. Unlike the previous example, this example includes checkpoints to indicate when the database buffers have been written to disc.

The DBMS must take the following actions:

- **T₁** has not committed when the failure occurred, so it must be *undone*.
- **T₂** has committed before the checkpoint and must have been written to disc. **T₂** has been committed successfully.
- **T₃** has also been committed successfully.
- **T₄** has been committed after the checkpoint and so the DBMS cannot be certain that **T₄**’s updates were written to disc. **T₄** must be *redone*.
- **T₅** has been committed but, as with **T₄**, it must be *redone*.
- **T₆** has not been committed when the failure occurred, so it must be *undone*. 

![Example with Checkpoints Diagram](image-url)