13 Recovery

- Recovery: Wiederherstellung eines konsistenten Datenzustands nach Fehlersituationen

Drei Arten von Fehlersituationen:

- Transaktionsfehler (lokaler Fehler)
  - Eine Transaktion wird nicht ordnungsgemäß beendet
  - Auswirkung: Daten sind in einem inkonsistenten Zustand
  - Beispiele:
    - Laufzeitfehler im Anwenderprogramm (Division durch 0, Variablenüberlauf, Überschreitung der Indexgrenzen einer Tabelle, etc.)
    - Endlosschleife im Anwenderprogramm
    - Abbruch des Anwenderprogramms durch den Systembediener
    - Time-out (Überschreitung einer CPU-Zeitgrenze) des Anwenderprogramms
    - Kein verfügbarer Plattenplatz bei einem Schreibbefehl im Anwenderprogramm
    - Deadlock (siehe 'Synchronisation paralleler Prozesse')
    - Befehl zum Zurücknehmen der Transaktion im Anwenderprogramm selbst (ROLLBACK WORK)
  - Maßnahme: Transaction Recovery, Transaction Rollback, Dynamic Transaction Backout
    Alle Änderungen, die die Transaktion bis zum Abbruch gemacht hat, müssen im laufenden Systembetrieb zurückgenommen werden.

- Systemfehler (globaler Fehler, Soft-Crash)
  - Der Transaktionsmonitor wird nicht ordnungsgemäß beendet (damit können in einem Multiuserbetrieb mehrere Transaktionen nicht ordnungsgemäß beendet werden)
  - Auswirkung: Daten sind in einem inkonsistenten Zustand
  - Beispiele:
    - Stromausfall
    - Abbruch des Transaktionsmonitors durch den Systembediener
    - Fehler im Betriebssystem (z.B. Endlosschleife)
    - Verlust von Hauptspeicherinhalten
  - Maßnahme: Crash Recovery
    Alle Änderungen der Transaktionen, die zum Zeitpunkt des Systemfehlers aktiv (in flight) waren, müssen beim Systemwiederstart (Restart, Warmstart) zurückgenommen werden.

- Mediumfehler (Speicherfehler, Hard-Crash)
  - Daten sind physikalisch zerstört oder nicht mehr lesbar
  - Auswirkung: Daten sind nicht mehr verfügbar
  - Beispiele:
    - Head-Crash auf Magnetplatte
    - Fehler im Disk-Controller
    - Irrtümliches Löschen von Daten
    - Irrtümliches Formatieren einer Platte
    - Fehler in der Dateiverwaltung des Betriebssystems
    - Katastrophenfälle (Brand, Anschläge, Erdbeben, etc.)
  - Maßnahme: Archive Recovery, Media Recovery, Disaster Recovery
    Ein Sicherungsstand (Backup Copy) der Daten wird eingespielt (reloaded, restored) und die Änderungen aller Transaktionen, die seit dem Zeitpunkt der Sicherung beendet wurden, müssen vom System nachvollzogen werden

- Transaction Recovery und Crash Recovery werden auch unter dem Begriff Backward Recovery zusammengefasst

- Archive Recovery wird auch Forward Recovery genannt

- Techniken für Backward Recovery

  Zwei mögliche Strategien:
  - Logging (Journaling) in Undo-Logs: Vor jeder Veränderung von Daten wird eine Kopie dieser Daten (Before Image) in eine Log-Datei (Undo-Log) eingetragen (WAL = Write Ahead Log - Prinzip). Damit ist für jede Operation (aufnehmen, ändern, löschen) der alte Wert bekannt. Um eine Transaktion rückgängig zu machen, wird das Undo-Log vom Ende beginnend rückwärts gelesen und jeweils der alte Wert wiederhergestellt.
  - Verzögertes Update (Deferred Update): Die Änderungen der Daten werden erst geschrieben, wenn die Transaktion beendet ist. Bis dahin wird auf Kopien der eigentlichen Daten gearbeitet. Rollback heißt dann einfach, diese Kopien zu löschen. In diesem Fall enthalten die Daten selbst die Rollback-Information.

  - In der Praxis werden im wesentlichen Verfahren mit Log-Dateien verwendet

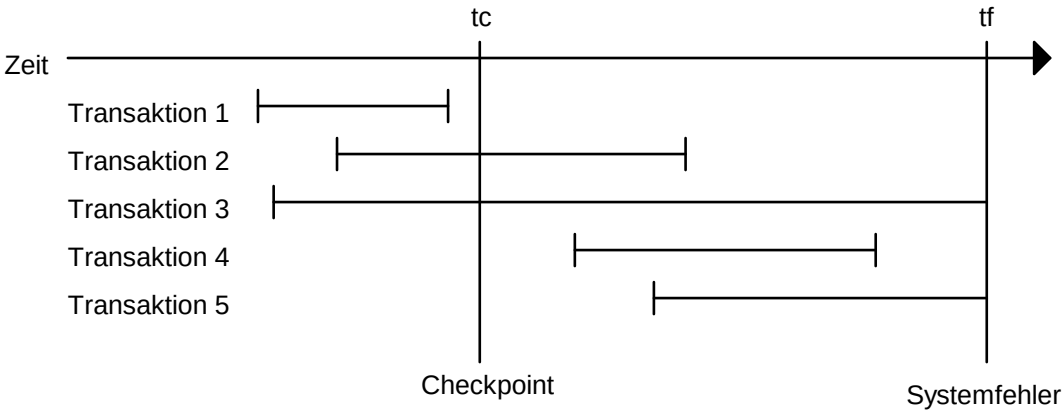
  - Informationen in einer Undo-Logdatei:
    Mindestens:
    - Beginn Transaktion
      Identifikation der Transaktion
    - Kennzeichen der Operation (aufnehmen, ändern, löschen)
      Identifikation der Transaktion
      Identifikation der Daten (z.B. Tabellenname und Zeilennummer)
      Before-Image der veränderten Daten
      Bemerkung: Für die Operation Aufnehmen brauchen nicht alle Daten, sondern nur deren Identifikationen protokolliert werden.
    - Ende Transaktion
      Identifikation der Transaktion

    Zusätzlich:
    - Benutzeridentifikation
    - Terminal- / Workstationbezeichnung
    - Programmname
    - Datum und Uhrzeit

  - Undo-Information braucht nur bis zum erfolreichen Ende der entsprechenden Transaktion aufgehoben zu werden. Meistens wird das Undo-Log aber fortlaufend geführt und an bestimmten Zeitpunkten als Ganzes gelöscht.

  - Bei Transaction Recovery muss das Undo-Log bis zum Beginnkennzeichen der entsprechenden Transaktion nach vorne gelesen werden.

  - Bei Crash-Recovery muss das Undo-Log bis zum Anfang durchgelesen werden, es könnten immer noch Before-Images einer nicht beendeten Transaktion gefunden werden.
    Um den Abschnitt zu begrenzen, der durchsucht werden muss, werden in gewissen Zeitabständen Sicherungspunkte (Checkpoints) auf das Log-File geschrieben. Dabei werden die Identifikationen aller Transaktionen, die zu diesem Zeitpunkt aktiv sind, abgespeichert.



    - Logik: Lese das Log-File rückwärts bis zum jüngsten Checkpoint; jede Transaktion, für die keine Endemarke gefunden wurde, wird mittels Before-Image rückgängig gemacht. Vom Checkpoint weiter rückwärts gehend werden die Transaktionen rückgängig gemacht, die im Checkpoint angeführt sind und für die keine Endemarke gefunden wurde.
    - An den Sicherungspunkten werden üblicherweise auch alle Puffen auf stabile Speicher (Hintergrundspeicher) geschrieben.

- Techniken für Forward Recovery

  - Logging (Journaling) in Redo-Logs: Eine Kopie der veränderten Daten (After Image) wird laufend in eine Log-Datei (Redo-Log) eingetragen. Das Redo-Log muss vom Anfang beginnend vorwärts gelesen und die Änderungen aller beendeten Transaktionen nachvollzogen werden.

  - Informationen in einer Redo-Logdatei:
    Analog Undo-Logdatei, jedoch mit After-Image für alle veränderten Daten.
    Bemerkung: Für die Operation Löschen brauchen nicht alle Daten, sondern nur deren Identifikationen protokolliert werden.

  - Redo-Information muss bis zur erfolgreichen Erstellung der nächsten Backup-Copy aufgehoben werden.

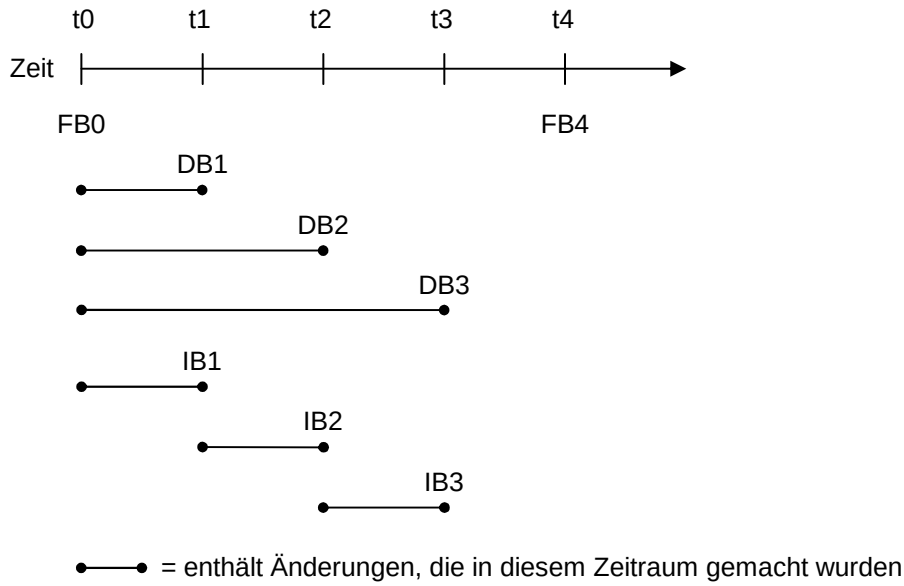
  - Die Redo-Logs können daher, wenn Backup-Copies aus Gründen des Datenumfanges oder der Verfügbarkeit nur in großen Zeitabständen gemacht werden, ebenfalls sehr umfangreich werden. Mit Dienstprogrammen können ein oder mehrere Redo-Logs komprimiert werden; es werden dabei nur die endgültigen (jüngsten) Änderungsstände abgespeichert (Change Accumulation), damit schnellerer Wiederanlauf.

  - Alternative zur Archive Recovery: Gespiegelte Platten (Mirror Disks), Gespiegelte Datenbanken (Mirror Databases, Shadow Databases), RAID-Technologie
    Die Daten werden doppelt gehalten und laufend parallel geändert.
    Vorteil: schneller Wiederanlauf, weniger Backup-Copies notwendig, hohe Verfügbarkeit ('24 Stunden an 7 Tagen')
    Nachteil: große Hardwarekapazität, räumlich ausgelagerte Backups und Logs müssen für Katastrophenfälle trotzdem zusätzlich verwendet werden

  - Redo-Logs und Spiegelplatten müssen natürlich auf physisch anderen Disk-Drives als die Basis-Daten gehalten werden.

- Backup-Techniken

  - Full Backup (Gesamt-Backup, Voll-Backup) (FB)
  - Partial Backup (Teil-Backup)
    - Differential Backup (Differenzielles-Backup) (DB): Alle Änderungen gegenüber dem letzten Full Backup
    - Incremental Backup (Zuwachs-Backup) (IB): Alle Änderungen gegenüber dem letzten Partial Backup



  - Vorteile Differential Backup:
    - sicherer (wenn Sicherungsmedien nicht mehr lesbar)
    - kürzere Restorezeiten
  - Vorteile Incremental Backup:
    - Weniger Speicherplatzverbrauch
    - kürzere Backupzeiten

  - Paralleles Backup (Striped Backup)
    Gleichzeitig mit mehreren Sicherungslaufwerken auf mehrere Sicherungsmedien

  - Offline Backup (Cold Backup)
  - Online Backup (Hot Backup, Dynamic Backup)
    Im laufenden Betrieb (Hochverfügbarkeitssysteme, High-Availability-Systems)
    Strategie zur Sicherstellung der Konsistenz des gesicherten Datenbestandes:
    - sequentielles Sichern der Zeilen, jede Zeile wird als gesichert gekennzeichnet
    - wenn eine Transaktion eine Änderung machen will, wird das sequentielle Sichern unterbrochen, der alte Stand gesichert, die Änderung durchgeführt, die Zeile als gesichert gekennzeichnet und das sequentielle Sichern wieder fortgesetzt
    - Es wird damit ein konsistenter Stand, wie er am Zeitpunkt des Sicherungsbeginns vorlag, gesichert

- Undo-Log und Redo-Log können auch in einer gemeinsamen Log-Datei (Audit Trail) gespeichert sein. Da das After-Image gleich dem Before-Image der nächsten Änderung ist, können hier Komprimierungen erreicht werden.

- Bei kombinierten System- und Mediumfehlern ist Forward- und Backward-Recovery (in dieser Reihenfolge) durchzuführen.
  Bemerkung: Bei gewissen Strategien werden die After-Images erst bei Transaktionsende auf die Log-Datei geschrieben, dann kann Backward-Recovery entfallen.

- Die oben beschriebenen Mechanismen müssen noch genauestens abgestimmt werden mit der meist gepufferten Datenein-/-ausgabe zwischen Haupt- und Externspeicher (logisches / physikalisches Schreiben) sowie der Modifikation von Indizes. Außerdem werden die Log-Dateien oft nicht satz-/zeilenweise, sondern block-/seitenweise geführt.

- Übung:

Tabelle P (vor Beginn der Transaktionen)

Zeilen#	PersNr	PName	Gehalt
01	27	Fuchs	2100
03	15	Maus	1500
05	47	Hase	800
06	75	Fisch	1600
09	32	Wolf	1500

Transaktionen

	T17	T18
t1	BEGIN TRANSACTION	
t2	UPDATE P SET PName='Gans', Gehalt=Gehalt+100 WHERE PName='Hase'	
t3		BEGIN TRANSACTION
t4		DELETE FROM P WHERE PersNr=15
t5	INSERT INTO P VALUES (83,'Adler',1900)	
t6		UPDATE P SET Gehalt=Gehalt*2 WHERE PName LIKE 'F%'
t7	COMMIT TRANSACTION	
t8		UPDATE P SET Gehalt=Gehalt+300 WHERE Gehalt<1000
t9		COMMIT TRANSACTION

Undo-Logdatei

	Aktion	Trans#	Tab	Zeilen#	Before Image
1	B	17			
2	U	17	P	05	47 / Hase / 800
3	B	18			
4	D	18	P	03	15 / Maus / 1500
5	I	17	P	10	
6	U	18	P	01	27 / Fuchs / 2100
7	U	18	P	06	75 / Fisch / 1600
8	E	17			
9	U	18	P	05	47 / Gans / 900
10	E	18			

Tabelle P (nach Ende der Transaktionen)

Zeilen#	PersNr	PName	Gehalt
01	27	Fuchs	4200
05	47	Gans	1200
06	75	Fisch	3200
09	32	Wolf	1500
10	83	Adler	1900

a) Was passiert, wenn das UPDATE von t8 zwischen t6 und t7 durchgeführt wird?

b) Welchen Inhalt hat für obiges Beispiel die Redo-Logdatei?

c) Wieviele Zeilen haben die Undo- und Redo-Logdatei im Verhältnis zueinander?

d) Wie sieht die Redo-Logdatei aus, nachdem sie von einem Change-Accumulation Utility bearbeitet wurde?

e) Welchen Inhalt hat ein Checkpoint, der zwischen t1 und t2 / t6 und t7 / t7 und t8 / nach t9 geschrieben wurde?

f) Zwischen t8 und t9 tritt ein Systemfehler auf, die untere Tabelle soll den Stand nach dem Systemfehler darstellen. Skizzieren Sie P nach erfolgter Recovery.

g) Zwischen t6 und t7 tritt ein Systemfehler auf, die untere Tabelle soll den Stand nach dem Systemfehler darstellen. Skizzieren Sie P nach erfolgter Recovery.