## Datenbanken II und Implementierungskonzepte für Datenbanksysteme

Prüfer: Prof.Dr.Schlageter

Datum: 06.10.11 Dauer: ca.20 min

Leider kann ich mir nicht mehr an allen Fragen erinnern. Einige davon sind.......

## Implementierungskonzepte für Datenbanksysteme:

Warum werden ein Datensatz manchmal getrennt in verschiedene Seiten gespeichert werden? (Mir zu helfen malt er die Datensätze und da sehe ich, dass die einige Spalte getrennt gespeichert werden.)

Anwort is...... Datensätze sehr großer Länge, Beispiel: Name, Adresse in einer Seite und Bilder, Video in anderen Seite zu speichern.

## **Datenbaken II**

Prinzip der Serialisierbarkeit

Ein paralleles System von Transaktionen is dann korrekt synchronisiert, wenn es serialosierbar ist, dh. wenn es mindestens eine (gedachte) serielle Ausführung derselben Transaktionen gibt, die denselben Datenbankzustand und dieselben Ausgabedaten der Transaktionen liefert.

Wie viele Klassen von Synchronisationsverfahren gibt es?

Verifizierende(optimistisch) Verfahren

Präventive(pessimitisch) Verfahren

Über beide Verfahren erklären.

optimistisch: zuerst nicht sperren, erst Commit -Zeit auf Serialisierbarekeit prüfen. pessimitisch: zuerst sperren

Sperrprotokolle

Zwei-Phasen-Sperrprotokoll, Strikte Zweiphasigkeit, Preclaiming

Was passiert bei Lock Unlock Lock Unlock?

Inkonsistente Sicht (Ich musste malen mit einem Beispiel, Banktransaktion)

Wie sieht ein Index aus?

Ich musste malen und erklären was die Spalte hier bedeuten.

Kurse: 01664 Implementierungskonzepte für Datenbanksysteme, 01672 Datenbanken II

Termin: 23.9.2011 Dauer: ca. 15 Min.

Prüfer: Prof. Dr. Schlageter

Note: 1.0

 Was muss getan werden, nachdem Sie die Entities im konzeptionellen Modell festgelegt haben, um diese in der Datenbank zu speichern?
 Primäre und Sekundäre Organisation festlegen.

- Welche Formen der Primären Organisation gibt es?
   Sequentielle, Index Sequentielle, Baum-Strukturen, Hash-Verfahren
- Wozu wird eine Sekundäre Organisation benötigt? Was ermöglicht uns das?
   Um Selects auf Nicht-Schlüssel-Attributen zu beschleunigen.
- Wie sieht so eine sekundäre Organisation aus?
   Es handelt sich um eine invertierte Organisation. Es ist im Grunde eine Liste aus Werten und Identifikatoren der dazu gehörenden Datensätzen.
- Ist ein Index in jedem Fall sinnvoll?
   Nein, dies ist abhängig von der Selektivität. Je höher die Selektivität, desto sinnvoller ist ein Index. Nur dann, wenn der durschnittliche Zugriff weniger als 10% der Tupel der Relation zurückliefert, macht ein Index Sinn.
- Warum ist das so? Könnte nicht sagen, dass bei einer Selektivität von 50% ein Index nicht immer noch besser ist als garkein Index?
   Nein, die Wahrscheinlichkeit ist sehr hoch, dass die selektierten Tupel sowieso über fast alle Seiten verteilt sind, so dass insgesamt kaum weniger Seiten geladen werden müssten, als wenn kein Index verwendet wird.
- Was muss bei einem Systemfehler getan werden?
   Bei einem Systemfehler kann es sein, dass sich die Datenbank nicht in einem konsistenten
   Zustand befindet. Es können Änderungen von Transaktionen in der Datenbank stehen, die nicht abgeschlossen wurden, und es können Änderungen von abgeschlossenen Transaktionen fehlen. Ziel ist es, die Datenbank in den letzten konsistenten Zustand zu versetzen. Um die Änderungen nicht abgeschlossener Transaktionen rückgängig zu machen, werden die Before-Images rückwärts bis zum letzten Checkpoint wieder eingespielt. Um sicher zu gehen, dass alle Änderungen abgeschlossener Transaktionen in der Datenbank stehen, werden die After-Images dieser Transaktionen vom letzten Checkpoint an, vorwärts wieder eingespielt.
- Was passiert, wenn ein Checkpoint erstellt wird?
  Die aktiven Transaktionen werden erfasst, damit im Fall eines Systemfehlers die aktiven
  Transaktionen leicht ermittelt werden können. Ansonsten müsste das gesamte Log
  durchlaufen werden, um diese zu ermitteln. Außerdem werden alle After-Images der
  Transaktionen, die seit dem letzten Checkpoint abgeschlossen worden sind, in die Datenbank
  geschrieben. Zudem werden alle Before-Images der Transaktionen, die seit dem letzen
  Checkpoint abgebrochen worden sind, in die Datenbank geschrieben.

- Warum werden die geänderten Seiten, mit dem Commit nicht einfach alle auf den Externspeicher geschrieben?
   Da es auch andere Transaktionen geben kann, die auf diese Seiten zugreifen. Wenn die Seite im Systempuffer steht, ist der Zugriff deutlich schneller.
- Bei welcher Art von Daten ist das besonders wichtig?
   Bei hotspots. Dies sind Bereiche in der Datenbank auf die ständig von vielen Transkationen zugegriffen wird.
- Wovor schützt das Zwei-Phasen-Sperrprotokoll? Inkonsistenz
- Wie kommt Inkonsistenz zu Stande?
  Hier habe ich zunächst gesagt, dass Inkonsistenz entsteht, wenn Daten aus einer
  Inkonsistenten Sicht weiter verarbeitet werden. Allerdings wollte er das genauer erklärt
  haben, und hat mir einen Hinweis gegeben, dass es ein Beispiel mit zwei Konten gäbe. Das
  habe ich dann erklärt.
- Was ist eine inkonsistente Sicht?
   Hier habe ich das andere Beispiel mit den Konten erläutert.
- Dann hat Herr Schlageter eine Transaktion grafisch skizziert und wollte wissen, wann die Sperre von Objekt a aufgehoben werden darf.

Nach dem Zwei-Phasen-Sperrprotokoll darf eine Transaktion keine Sperre mehr anfordern, nachdem sie eine Sperre freigegeben hat. Die Transaktion darf a daher erst freigeben, nachdem sie b gesperrt hat.

- Was kann bei Hotspots bzgl. der Sperren getan werden?
   Für Hotspots können besondere atomare Sperren verwendet werden, die das Objekt nur für diese eine Operation schützen. Das ist möglich auf Grund der Kommutativität einiger Operatoren. Beispiele für solche Sperren sind der Add-Lock und der Sub-Lock.
- Was ist nun wenn eine Transaktion eine Multiplikation auf solch einem Objekt durchführen möchte?

Die Transaktion muss dann eine ganz normale Sperre verwenden.

Herr Schlageter ist als Prüfer sehr zu empfehlen. Seine ruhige Art macht die Prüfung sehr angenehm. Zu Anfang haben wir ein bisschen Small Talk gehalten, bevor es dann richtig los ging. Die anfängliche Nervösität war so schnell verflogen. Die Fragen waren - bis auf wenige Ausnahmen - sehr verständlich und kamen auch bereits in anderen Protokollen vor. Die Prüfung war bereits nach ca. 15 Minuten vorbei und ich habe direkt meine Note bekommen. Zur Vorbereitung haben die Protokolle zu Kurs 01665 Datenbanksysteme sehr geholfen.