

Implementierung von Datenbanksystemen

Klausur WS 2011/12 Braindump

22. Februar 2012

Fehler oder fehlende Fragen bitte melden: he29heri@stud.informatik.uni-erlangen.de

110% Klausur, allerdings recht wenig Zeit.

1 Wissensfragen

Was ist Datenabstraktion? Siehe Vorlesungsfolien 1 (Einführung).

Beschreiben Sie das Schichtenmodell. Wie ist eine Schicht beschrieben? Siehe Vorlesungsfolien 1 (Einführung).

Welche Methoden gibt es, SQL-Anfragen aus einer Anwendung ausführen? Vorübersetzer, Call-Level-Interface. Siehe Vorlesungsfolien 7 (Programmzugriff).

2 Speicherungsstrukturen

Beschreiben Sie die Abstraktion von Blöcken zu Sätzen und den wesentlichen Unterschied. Blöcke: hardwarespezifisch, Sätze: anwendungsspezifisch; Blöcke: feste Länge, Sätze: variable Länge. Siehe Vorlesungsfolien 3 (Sätze).

TIDs: Sätze einfügen, Sätze vergrößern, Fragmentierung, neue TIDs angeben. TIDs ändern sich *nicht*. Siehe Übung 2 (Sätze).

Was ist C-Store? Nennen Sie zwei Komprimierungsverfahren die eingesetzt werden. Spaltenweises Speichern der Daten. Siehe Vorlesungsfolien 9 (Speicherung), Übung 8 (Speicherung).

Wie kann min, max effizient in C-Store implementiert werden? Siehe Vorlesungsfolien 9 (Speicherung), Übung 8 (Speicherung).

3 Indizes

Lineares Hashing am Beispiel mit Überlaufbuckets durchführen. Siehe Übung 4 (Schlüssel II).

In B-Baum einfügen und löschen. Siehe Übung 5 (Schlüssel III).

Wie findet man einen Eintrag im B-Baum? „Binäre Suche“. Siehe Vorlesungsfolien 5 (Schlüssel II).

Wie kann ein B*-Baum Bereichsanfragen effizient ausführen? Durch Verkettung der sortieren Blattknoten. Siehe Vorlesungsfolien 5 (Schlüssel II).

4 Systempuffer

Wie funktioniert ein Systempuffer und wozu braucht man ihn? Siehe Vorlesungsfolien 6 (Puffer).

Nennen Sie zwei Ersetzungsstrategien. z. B. LRU, LFU, FIFO, CLOCK. Siehe Vorlesungsfolien 6 (Puffer).

CLOCK-Verfahren anwenden. Siehe Übung 6 (Puffer).

5 Transaktionen

Nennen Sie vier Eigenschaften von Transaktionen und beschreiben Sie sie kurz. ACID. Siehe Vorlesungsfolien 8 (Transaktionen).

Geben Sie ein Beispiel zu Lost-Update an. $r_1(A), r_2(A), w_1(A), w_2(A), c_1(A), c_2(A)$. Siehe Vorlesungsfolien 8 (Transaktionen).

Zeichnen Sie einen Abhängigkeitsgraphen zu dem gegebenen Schedule und begründen Sie ob er serialisierbar sind. Nicht serialisierbar da Zyklus im Abhängigkeitsgraphen. Siehe Übung 11 (Synchronisation).

S und X Sperrern am Beispiel des vorherigen Schedules. Hier ohne Deadlock. Siehe Übung 11 (Synchronisation).

6 Anfrageoptimierung

SQL-Statement in unoptimierten Anfragebaum umwandeln. Siehe Übung 9 (Anfrageverarbeitung).

```
SELECT
  A.id, B.id, C.id
FROM
  A, B, C
WHERE
  A.b_id = B.id AND B.c_id = C.id
  AND
  A.id > 100;
```

```
PROJ(SEL(CROSS(A, B, C),
  (A.b_id = B.id AND B.c_id = C.id AND A.id > 100)),
  (A.id, B.id, C.id))
```

(Als Baum zu zeichnen.)

Nennen Sie eine Möglichkeit diesen Baum zu optimieren. CROSS in JOIN umwandeln ((A.b_id = B.id AND B.c_id = C.id)), SEL vor JOIN ziehen (A.id > 100). Siehe Übung 9 (Anfrageverarbeitung).

7 Überschrift unbekannt TODO

Beschreiben Sie logische und physikalische Konsistenz und die Unterschiede. Siehe Vorlesungsfolien 8 (Transaktionen).

Definieren Sie Primär- und Sekundärorganisation und wie häufig diese vorkommen dürfen. Primärorganisation einmal, Sekundärorganisation beliebig oft. Siehe Vorlesungsfolien 5 (Schlüssel II).

8 Weiterführende Fragen

Was ist ein Systemkatalog und nennen Sie zwei Beispiele, was darin gespeichert wird?
Informationen über Felder und ihre Reihenfolge. Siehe Vorlesungsfolien 9 (Speicherung).

Erstellen Sie eine Tabelle im Systemkatalog mit Hilfe der geprüft werden kann ob ein Bitmap-Index existiert, genutzt werden kann und wo er in der Datenbank auffindbar ist? Ungeprüft: Tabelle (wo der Index genutzt wird), Feldname (Index über dieses Feld), Blocknummer (erste Blocknummer wo Index gespeichert wird), Anzahl (Anzahl der Blöcke).