UNIVERSITÄT DES SAARLANDES
Fachbereich 14 – Informatik
Prof. Dr.-Ing. G. Weikum
Dipl.-Inform. R. Schenkel

Datenbanksysteme WS 1999/2000

Praktische Übung 3

Einführung in Administration, Monitoring und Tuning von Datenbanken

Zu bearbeiten bis spätestens 17. Januar 2000

(Diese Übung wird *nicht* bewertet.)

In dieser Übung sollen Sie den konzeptionellen und physischen Aufbau einer synthetisch generierten VIDEOTHEK-Datenbank durch Abfrage des Datenbankkatalogs ermitteln. Melden Sie sich dazu wie gewohnt via SQLPlus unter der Kennung Ihrer Gruppe bei Oracle an.

Beantworten Sie die folgenden Fragen, ohne auf die VIDEOTHEK-Datenbank zuzugreifen:

1. Ermitteln Sie das konzeptionelle Schema der Datenbank. Welche Integritätsbedingungen existieren in der Datenbank? Geben Sie die SQL-Queries zur Ermittlung der gesuchten Werte an.
2. Ermitteln Sie das physische Schema der Datenbank. Geben Sie SQL-Queries zur Ermittlung der gesuchten Werte an.
   a) Wie viele Blöcke belegen die einzelnen Relationen, und wie viele freie Blöcke sind bereits reserviert? Wie ist der durchschnittliche Füllgrad der belegten Blöcke in den einzelnen Relationen? Wie groß ist die durchschnittliche Länge der Tupel in den Relationen? Welches ist das Attribut, das in seiner Relation am meisten Platz belegt?
   b) Welche Indizes existieren, und welche Attribute sind in welcher Reihenfolge indiziert? Wie sind die Indizes aufgebaut, d.h. wie viele Blöcke belegen die Indizes, wie hoch sind die B-Bäume, und wie viele Knoten belegen die Blattknoten? Wie viele unterschiedliche Schlüssel sind im Index enthalten, wie viele Blattknoten und wie viele Datenseiten in der Relation werden im Durchschnitt für einen Schlüssel belegt?
3. Ermitteln Sie das Mengengerüst der Datenbank. Wie viele Tupel sind in den Tabellen jeweils enthalten? Wie viele unterschiedliche Werte existieren pro Attribut? Welches sind die kleinsten und größten Werte der numerischen Attribute? Was können Sie über die Verteilung von Attributwerten aussagen? Geben Sie die SQL-Queries zur Ermittlung der gesuchten Werte an.
4. Ermitteln Sie für jeweils eine – von Ihnen wählbare – Query aus jeder der folgenden Kategorien den Ausführungsplan und dessen Kosten:
   a) Ein Join zwischen zwei Relationen mit Selektionsbedingungen auf beiden Operanden.
   b) Ein Join zwischen drei Relationen mit Selektionsbedingungen auf allen Operanden.
   c) Eine Selektion auf einer Relation mit einer konjunktiven Verknüpfung von Attributwertebedingungen auf mindestens drei Attributen.
   d) Eine Selektion auf einer Relation mit einer disjunktiven Verknüpfung von Attributwertebedingungen auf mindestens drei Attributen.

e) Eine Gruppierungs-Query mit einer Aggregationsfunktion auf dem Join zwischen zwei Relationen.

f) Eine Gruppierungs-Query mit einer Aggregationsfunktion auf dem Join zwischen drei Relationen.

g) Eine Gruppierungs-Query mit HAVING-Klausel.

5. Messen Sie die Ausführungskosten (Anzahl I/O's usw.) Ihrer unter 4. studierten Queries. Überlegen Sie sich dazu zunächst einen Meßaufbau, um die Ausführungskosten einer einzelnen Query zu bestimmen (siehe Hinweise). Warum können Sie hier nicht den View *V$FILESTAT* verwenden?
6. Versuchen Sie, etwaige Unterschiede zwischen den gemessenen und den geschätzten Kosten zu erklären.

Wichtige Hinweise:

- Der Oracle-Benutzer, der die Datenbankobjekte angelegt hat, heißt *GLOTZKIST*.
- Über die HTML-Seite

  http://www-dbs.cs.uni-sb.de/~schenkel/oradoc/

  finden Sie die für diese Aufgaben interessanten Teile der Oracle-Dokumentation.
- Allgemeine Informationen über den physischen Aufbau von Tabellen, Clustern, Indizes usw. finden Sie in dem Teil *Oracle7 Server Concepts Manual*.
- Oracle stellt Datenkatalog-Views (*Data Dictionary Views*) zur Verfügung, um Informationen über den Aufbau von Datenbanken abfragen zu können. Die Views sind in dem Teil *Oracle7 Server Reference Manual* beschrieben.
- Den Ausführungsplan einer Query <*q*> können Sie in SQLPlus durch Aufruf des Skripts *sp.sql* (sp für show plan) ermitteln, indem Sie in der Kommandozeile *@sp '<q>';* eingeben. Nach dem Aufruf des Skripts werden Sie aufgefordert, den Identifikator Ihrer Übungsgruppe (also z.B. *42* für *student42*) einzugeben. Anschließend bekommen Sie den Ausführungsplan der Query <*q*> ausgegeben. Das Skript ruft den Oracle-Befehl *EXPLAIN PLAN* auf. Hinweise zur Interpretation der Ausgabe finden Sie unter *The EXPLAIN PLAN Command* in Kapitel 20 des Teils *Oracle7 Server Tuning*. Sie finden das Skript *sp.sql* unter der folgenden URL:

  http://www-dbs.cs.uni-sb.de/dbs99/uebungen/sp.sql
- Zur Messung von Ausführungskosten stellt Oracle sogenannte dynamische Performance-Tabellen (*Dynamic Performance Tables*) zur Verfügung, die permanent mit gemessenen Werten aktualisiert werden und auf die Sie über entsprechende Views zugreifen können. Für Ihre Zwecke sind insbesondere die Views *V$MYSTAT* und *V$STATNAME* interessant, deren Aufbau im Teil *Dynamic Performance Tables* des *Oracle7 Server Reference Manual* beschrieben ist. Der Inhalt des Views *V$STATNAME* ist unter der folgenden URL zu finden:

  http://www-dbs.cs.uni-sb.de/~schenkel/oradoc/statnames.asc
- Beachten Sie, daß der View *V$MYSTAT* die über eine ganze SQLPlus-Sitzung akkumilierten Werte beinhaltet, während in 5. die Ausführungskosten einer einzelnen Query yu bestimmen sind. Um die Ausführungskosten einer einzelnen Query zu bestimmen, muß also die Differenz der entsprechenden Werte in *V$MYSTAT vor* und *nach* der Ausführung einer Query gebildet werden.