UNIVERSITÄT DES SAARLANDES
Fachbereich 14– Informatik
Prof. Dr.–Ing. G. Weikum
Dipl.–Inform. A. Kraiß

Datenbanksysteme WS 1998/99

Praktische Übung 3

Einführung in Administration, Monitoring und Tuning von Datenbanken

zu bearbeiten bis spätestens 15.1.1999

(Abnahme durch Ihren Übungsgruppenleiter)

In dieser Übung sollen Sie den konzeptionellen und physischen Aufbau einer synthetisch generierten VIDEOTHEKS-Datenbank durch Abfrage des Datenkataloges 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 VIDEOTHEKS-Datenbank zuzugreifen:

1. Ermitteln Sie das konzeptionelle Schema der DB. Welche Integritätsbedingungen existieren in der DB ? Geben Sie die SQL-Queries zur Ermittlung der gesuchten Werte an.
2. Ermitteln Sie das physische Schema der DB. Geben Sie die SQL-Queries zur Ermittlung der gesuchten Werte an.
   a) Wieviele Blöcke belegen die einzelnen Relationen, und wieviele 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. wieviele Blöcke belegen die Indizes, wie hoch sind die B-Bäume, und wieviele Blöcke belegen die Blattknoten? Wieviele unterschiedliche Schlüssel sind im Index enthalten, wieviele Blattknoten und wieviele Datenseiten in der Relation werden im Durchschnitt für einen Schlüssel belegt?
   c) Welche Cluster existieren, und was können Sie über ihren Aufbau aussagen?
3. Ermitteln Sie das Mengengerüst der DB. Wieviele Tupel sind in den Tabellen jeweils enthalten? Wieviele unterschiedliche Werte existieren pro Attribut? Welches sind die kleinsten und größten Werte der numerischen Attribute? Was können Sie über die Verteilungen 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) einen Join zwischen 2 Relationen mit Selektionsbedingungen auf beiden Operanden.
   b) einen Join zwischen 3 Relationen mit Selektionsbedingungen auf allen Operanden.
   c) eine Selektion auf 1 Relation mit einer konjunktiven Verknüpfung von Attributwertebedingungen auf mindestens 3 Attributen
   d) eine Selektion auf 1 Relation mit einer konjunktiven Verknüpfung von Attributwertebedingungen auf mindestens 3 Attributen
   e) eine Gruppierungs–Query mit einer Aggregationsfunktion auf dem Join zwischen 2 Relationen
   f) eine Gruppierungs–Query mit einer Aggregationsfunktion auf dem Join zwischen 3 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 geschätzten und den gemessenen Kosten zu erklären.

Wichtige Hinweise:

- Der ORACLE-Benutzer, der die Datenbankobjekte angelegt hat (*Owner*), heißt *GLOTZKIST*.
- Über die HTML-Seite */home/stud/kraiss/oradoc/Welcome.html* finden Sie die für diese Aufgabe 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 Skriptes */home/stud/kraiss/P3/sp.sql* (sp für show plan) ermitteln, indem Sie in der Kommandozeile *@sp '<q>';* eingeben. Nach dem Aufruf des Skriptes werden Sie aufgefordert, den Identifikator Ihrer Übungsgruppe (also z.B. *49* für *student49*) 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 (Chapter 20) des Teiles *Oracle7 Server Tuning*.
- 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$STATNAME* ist in der ASCII-Datei */home/stud/kraiss/oradoc/statnames.asc* aufgelistet. Beachten Sie, daß der View *V$MYSTAT* die über eine ganze sqlplus-Sitzung akkumulierten Werte beinhaltet, während in 5. die Ausführungskosten einer einzelnen Query zu 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.