

Aufgabenblatt 4 Benchmarking

- Abgabetermin: **Dienstag, 19.1.2010 (23:59 Uhr)**
- Zur Prüfungszulassung muss ein Aufgabenblatt mit mind. 25% der Punkte bewertet werden und alle weiteren Aufgabenblätter mit mindestens 50% der Punkte.
- Die Aufgaben sollen in Zweiergruppen bearbeitet werden.
- Abgabe:
 - per E-Mail an dbs2-200910@hpi.uni-potsdam.de mit Subject
Abgabe DBS II: Aufgabenblatt <n> Namen
 - ausschließlich pdf-Dateien
 - eine Datei mit folgendem Dateinamen: `blatt<aufgabenblattNr><Nachnamen>.pdf`
Bitte **keine Leerzeichen, Unterstriche, Umlaute, Sonderzeichen**, ... im Dateinamen!
 - **jedes Blatt beschriftet mit Namen**
 - Wir korrigieren die Abgaben aufgabenweise. Das beschriebene Verfahren vereinfacht uns die Arbeit erheblich!
- **Die DB2-Dokumentation findest du unter:**
<http://publib.boulder.ibm.com/infocenter/db2luw/v9r7/index.jsp>

Aufgabe 1: TPC-H

Optimiere deine Datenbank in Bezug auf den Workload des TPC-H.

Im Rahmen des Benchmarks werden zunächst neue Tupel in die Tabellen `lineitem` und `orders` eingefügt, anschließend 22 Anfragen abgesetzt und zuletzt Daten aus den genannten Tabellen gelöscht. Diese drei Schritte können mit Hilfe des Skriptes

```
/mnt/tpchData-1G/run_benchmark.sh
```

ausgeführt werden. Zum Zurücksetzen der Datenbank können die beiden Skripte

```
/mnt/tpchData-1G/updates/revert_insert.sh  
/mnt/tpchData-1G/updates/revert_delete.sh
```

genutzt werden. Um letzteres zu ermöglichen, muss noch *vor der Ausführung der Updates* mit Hilfe von `create_backup_schema.sql` und `create_backup.sh` ein Backup der betroffenen Daten erstellt werden.

Die TPC-H Spezifikation liegt unter `lehrveranstaltungen\DBSII_naumann\tpc-h`.

Hinweis: Das Tool `db2advis` kann bei der Bearbeitung der Aufgabe helfen!

Folgende Grundregeln müssen eingehalten werden:

- Die Tablespace – insbesondere der Index-Tablespace – dürfen nicht vergrößert werden!
- Die Datenbankparameter `AUTO_TBL_MAINT` und `SELF_TUNING_MEM` müssen auf OFF gesetzt sein!
- Es dürfen keine Materialisierten Sichten erstellt werden.

Dokumentiere und begründe deine Maßnahmen.

- Warum wurde ein bestimmter Index angelegt?
- Warum wurde ein bestimmter Index *nicht* angelegt?

Wie schnell wird der Workload auf deiner Datenbank bearbeitet? Berechne das TPC-H Maß `TPC-H Power@Size`.

Das Aufgabenblatt gilt als bestanden, wenn die (ergriffenen oder verworfenen) Maßnahmen nachvollziehbar dokumentiert und begründet sind. Es sollten mindestens 10 Maßnahmen in Erwägung gezogen werden!

Das beste Team gewinnt das Buch zur Vorlesung (eins für jeden ;-):

Hector Garcia-Molina, Jeffrey D. Ullman, Jennifer: Database Systems - The Complete Book, Pearson Education International.