

# **Architektur von Datenbanksystemen Praktikumsbericht 2**

Benedikt Kappes  
Tuan-Si Tran

31. Oktober 2011

# 1 Einleitung

Im letzten Praktikum haben wir uns auf einer sehr tiefen Ebene mit der DB2-Datenbank beschäftigt, indem wir die Byte-Größe der Indexe und der Daten betrachtet haben. In diesem Praktikum werden die Daten aus dem vorherigen Praktikum als Testdaten verwendet, um verschiedene Anwenderprofile zu simulieren und deren Auswirkung auf die Performance zu beurteilen.

Dafür sollen folgende vier Szenarios durchgespielt werden:

**Szenario 1.1** Viermaliges sequenzielles lesen aller Bestellungsdatensätze in 300er-Portionen mit der Klasse `Statement`

**Szenario 1.2** Viermaliges sequenzielles lesen aller Bestellungsdatensätze in 300er-Portionen mit der Klasse `PreparedStatement`

**Szenario 2** Lesen von 3.000.000 Bestellungsdatensätze in 300er-Portionen in zufälliger Reihenfolge.

**Szenario 3** Lesen von 3.000.000 Bestellungsdatensätze, wobei immer die selben 300 Datensätze gelesen werden.

**Szenario 4** Wie Szenario 1, nur diesmal in 10er-Portionen

Szenario 1.1 und 1.2 dienen dabei als Referenzgröße. Es sollen bei allen Szenarios sichergestellt werden, dass alle bzw. die richtigen Datensätze gelesen werden können.

## 2 Vorbereitung

Wie stellen wir sicher, dass alle Datensätze gelesen werden?

## 3 Durchführung

Als erstes wurde die automatische Leistungsoptimierung der DB2-Datenbank abgeschaltet, um reproduzierbare Ergebnisse zu erhalten. Danach wurde Szenario 1.1 implementiert. Dabei wurde nur mit der Klasse `Statement` gearbeitet. Erst in Szenario 1.2 wurden `PreparedStatement`s verwendet. Das Lesen

von 300er Portionen wurde anhand einer Zählvariable realisiert, die in jedem Durchlauf um 300 erhöht und in der **WHERE**-Klausel angegeben wird. Dabei machen wir uns zu Nutze, dass wir beim Eintragen der Bestellungsdaten für die ID eine fortlaufende Sequenz verwenden.

Szenario 2?

Für das dritte Szenario wurden immer nur die ersten 300 Bestelldatensätze gelesen. Für das vierte Szenario wurden die Datensätze in 10er-Tupel gelesen. Dabei wurden die verschachtelten Schleifen angepasst.

## 4 Ergebnisse und Auswertung

	SSD	mech
Szenario 1.1	54245	49389
Szenario 1.2	27466	20619
Szenario 2	38294	24468
Szenario 3	7156	10869
Szenario 4	126695	121315

Tabelle 4.1: Messwerte für Szenarien

## 5 Zusammenfassung

In diesem Praktikum haben wir gesehen, dass eine leere initiale Datenbank bereits 120 MB einnimmt und mit einer Ausführung eines Systemsbefehls bereits auf 160 MB anwuchs. Das Eintragen der Testdaten ergab, dass diese Daten mehr Speicherplatz einnehmen, als es rein rechnerisch sein sollte, was darauf zurückzuführen war, dass die Speicherung der Indexe einen unerwartet großen Platz einnahm, der größer ist als die Testdaten selbst. Schließlich hat sich noch herausgestellt, dass das Eintragen der Daten in eine Tabelle ohne vorher angelegtem Fremdschlüsselindex wesentlich schneller läuft, da dort nicht abgeglichen werden muss, ob der Key auch existiert.

## **6 Anhang**

# Literaturverzeichnis

[IBM-2010] IBM: *IBM DB2 9.7: Database Administration Concepts and Configuration Reference*. Handbuch, 09.2010, Seite 310f.

[IBM-2011] IBM: *DB2 Solution Information Center* URL: <http://publib.boulder.ibm.com/infocenter/db2luw/v9r7/index.jsp?topic=/com.ibm.db2.luw.sql.ref.doc/doc/r0000927.html>, zuletzt besucht am 20.10.2011