

Datenbanksysteme II
SS 2015

Übungsblatt 7: Relationale Anfragebearbeitung, I/O-Kosten

Besprechung: 29.06.-01.07.2015

Aufgabe 7-1 Implementierung der Join-Operatoren: CPU-Kosten

Es soll nun der Equi-Join der im Folgenden abgebildeten Relationen R und S berechnet werden.

R	S
1	1
2	2
3	3
6	4
7	5
9	6
10	8
11	10
12	13
13	14

- Führen Sie den Join mittels des Nested-Block-Loop-Join durch. Wieviele Paare von Tupeln müssen dabei auf Erfüllung des Joinprädikates untersucht werden?
- Führen Sie den Join mittels des Sort-Merge-Join durch. Wieviele Paare von Tupeln müssen dabei auf Erfüllung des Joinprädikates untersucht werden?
- Führen Sie den Join mittels des einfachen Hash-Join mit Hashfunktion $h(x) = x \bmod 5$ durch. Wieviele Paare von Tupeln müssen dabei auf Erfüllung des Joinprädikates untersucht werden?
- Führen Sie den Join mittels des Hash-Partitioned-Join (GRACE) durch. Dabei wird als Hashfunktion für die Blockpartitionierung die Funktion $h_B(x) = x \bmod 3$ verwendet. Auf den einzelnen Blöcken soll der Join mittels eines einfachen Hash-Joins mit der Hashfunktion $h(x) = x \bmod 2$ durchgeführt werden. Wieviele Paare von Tupeln müssen dabei auf Erfüllung des Joinprädikates untersucht werden?

Aufgabe 7-2 *I/O-Kosten: NBL-Join*

Gegeben seien zwei Relationen R und S , die jeweils eine Größe von 10.000 Blöcken besitzen. Im folgenden soll der Join $R \bowtie S$ mittels eines Nested-Block-Loop-Joins berechnet werden. Dabei wird als Cachelstrategie Variante 3 (siehe Skript) verwendet.

- (a) Berechnen Sie die benötigte Anzahl an Plattenzugriffen bei einer Cachegröße von 1.000 Blöcken.
- (b) Berechnen Sie die benötigte Cachegröße in Blöcken, um das Joinergebnis mit höchstens 100.000 Plattenzugriffen zu berechnen.
- (c) Berechnen Sie die benötigte Cachegröße in Blöcken, um das Joinergebnis mit höchstens 20.000 Plattenzugriffen zu berechnen.

Aufgabe 7-3 *Äquivalenzregeln*

Beweisen oder widerlegen Sie folgende Äquivalenzen:

- (a) $\sigma_{p_n \wedge p_{n-1} \wedge \dots \wedge p_1}(R) = \sigma_{p_n}(\sigma_{p_{n-1}}(\dots(\sigma_{p_1}(R))\dots))$
- (b) $\sigma_p(R_1 \times R_2) = \sigma_p(R_1) \times R_2$, falls $\text{attr}(p) \subseteq \text{attr}(R_1)$
- (c) $\pi_l(R_1 \cap R_2) = \pi_l(R_1) \cap \pi_l(R_2)$
- (d) $\pi_l(R_1 \cup R_2) = \pi_l(R_1) \cup \pi_l(R_2)$
- (e) $\pi_l(R_1 - R_2) = \pi_l(R_1) - \pi_l(R_2)$