

10. Übungsblatt

Ausgabe: 5. Januar 2015 · Besprechung: 12. Januar und 14. Januar 2015

1 Parallelisierung von Anfragen

Welche Arten von Parallelität gibt es, und wie kann man sie während der Anfragebearbeitung zur Geschwindigkeitssteigerung ausnutzen?

Gegeben seien drei Tabellen Kunde, Produkt und Verkäufe:

- Kunde(kid INT PRIMARY KEY, Name VARCHAR(40), Vorname VARCHAR(40), Alter INT, Geschlecht CHAR);
- Verkäufe(vid INT PRIMARY KEY, kid INT FOREIGN KEY, pid INT FOREIGN KEY, Anzahl INT, Verkaufspreis FLOAT);
- Produkt(pid INT PRIMARY KEY, Produktbezeichnung VARCHAR(60), Produktkategorie VARCHAR(40), Produktpreis FLOAT);

Diskutieren Sie, wie folgende Anfragen parallel ausgeführt werden können. Für welche Techniken, die Sie bereits in der Vorlesung kennen gelernt haben, würden Sie eine Geschwindigkeitssteigerung erwarten und warum?

1. `SELECT * FROM Kunde WHERE Kunde.Alter > 30 AND Kunde.Geschlecht='M';`
2. `SELECT SUM(Verkaufspreis) FROM Verkäufe;`
3. `SELECT Name, Vorname, Wohnort FROM Kunde NATURAL JOIN Verkäufe NATURAL JOIN Produkt FROM Kunde.Alter<40 AND Produkt.Produktkategorie='Fernseher';`

2 Parallele Selektionen

Skizzieren sie einen parallelen Algorithmus für einen einfachen Selektionsoperator. Nehmen Sie an, dass die Eingabetabellen spaltenorientiert (als Arrays) gespeichert sind, und dass die Selektion nur ein einfaches Prädikat auswerten kann (z.B. `COLUMN<5`). Wie ändert sich der Algorithmus, wenn der Selektionsoperator als Ergebnis eine Bitmap oder eine Liste von RowIDs zurückgeben muss?