Eidgenössische
Technische Hochschule
Zürich

Ecole polytechnique fédérale de Zurich
Politecnico federale di Zurigo
Swiss Federal Institute of Technology Zurich

Institut für Informationssysteme, Dr. M. Tresch

U. Röhm

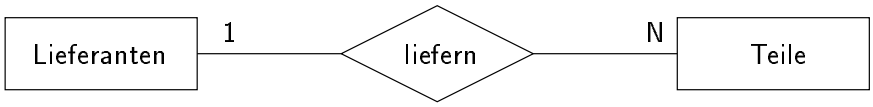
Verteilte Objektverwaltung: Client, Server & Middleware
Aufgabenblatt 1

Aufgabe 1: Data-, Query- und Hybrid-Shipping

Einführung

In der Vorlesung wurden die drei Möglichkeiten der Aufgabenteilung („Task-Split“) in Client/Server Systemen vorgestellt: Data-, Function- und Hybrid-Shipping. Wir wollen in dieser Übung einen Vergleich der drei Methoden anhand eines konkreten (relationalen) Beispiels erarbeiten. Welche Freiheitsgrade gibt es, unter welchen Voraussetzungen eignet sich eine der drei Shipping-Methoden am besten? Doch zuerst zu unserem Beispielszenario:

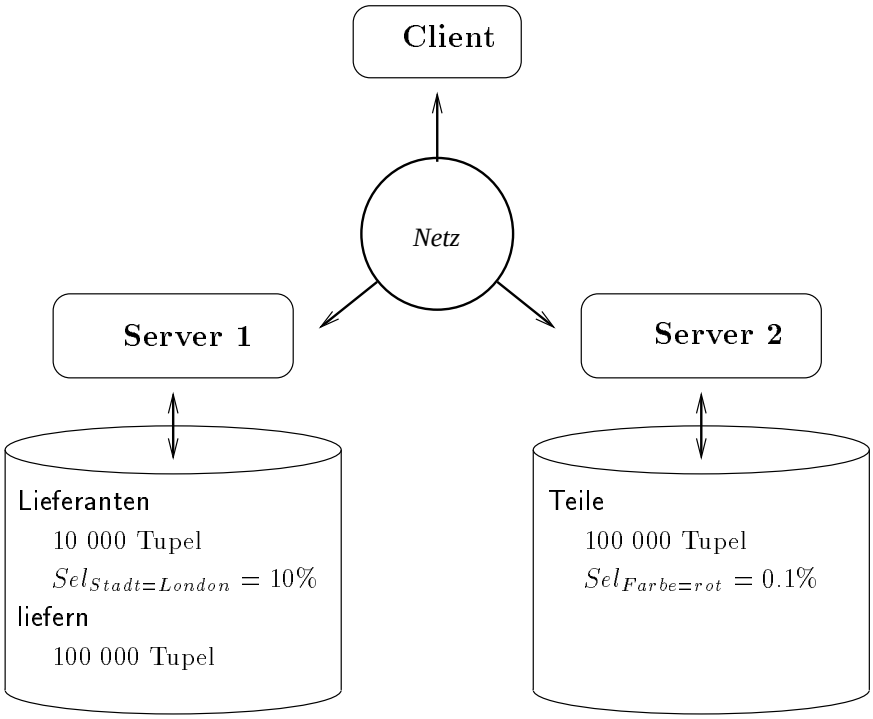
ER-Diagramm



Schema-Definition

Lieferanten (LiefNr integer, Name char(30), Stadt char(30))
Teile (TeilNr integer, Bezeichnung char(30), Farbe char(10))
liefern (LiefNr integer, TeilNr integer)

Systemverteilung



Kostenmodell

Die Daten werden zwischen Festplatte und Hauptspeicher, bzw. zwischen den Rechnern seitenweise ausgetauscht („Page-Server"). Eine Seite ist 4 KByte = 4096 Bytes groß. Sobald ein benötigtes Tupel nicht verfügbar ist, wird die Seite, auf dem es sich befindet, von der Festplatte oder über das Netz eingelagert. Netz- und Diskzugriffe setzen sich aus der Zugriffszeit und der größenabhängigen Transferzeit zusammen. Alle Operationen (Disk- oder Netzzugriff und die Anfrageberechnung) benötigen CPU-Zeit.

CPU: Rechenleistung: 50 MIPS. Eine Anfrageoperation (Scan, Join, Select, Project) benötige 10 Instruktionen pro Tupel.

Disk: Transferrate: 2 MB/sec. Blockgröße: 4096 Byte. Zugriffszeit: 5ms
Je Blockzugriff 5000 CPU-Instruktionen.

Netz: Ethernet mit 1 MB/sec Transferrate und 0,2ms Zugriffszeit pro Nachricht.
Je Nachrichtenversand/-empfang 40000 CPU-Instruktionen.

RAM: Client: 3 MByte Datenbankpuffer.
Datengrößen von *integer* 4 Byte, *date* 8 Byte und *boolean* 2 Byte.

Fragen

1. Am Client wird folgende Anfrage gestellt:
   „Welche Lieferanten aus London liefern rote Teile?"
   Berechnen Sie die Kosten (Datenvolumen, CPU-Belastung, Ausführzeiten) der Auswertung obiger Anfrage im Falle von Data-, Query- und Hybrid-Shipping. Wir nehmen an, daß alle Datenbankpuffer zu Anfang leer sind.

2. Wenn dieselbe Anfrage danach erneut gestellt wird, wie sehen dann die Auswertungskosten der drei Methoden aus?

3. Vergleichen Sie jeweils die Lastverteilung bezüglich CPU-, Disk- und Speichergebrauch auf die einzelnen Rechner und das Netz. Welche Auswirkungen haben die Annahmen bezüglich Rechen-, Disk- und Netzleistung auf die Resultate der verschiedenen Ausführungsmethoden?

Geben Sie zu jeder Teilaufgabe bitte die Auswertungspläne mit an. Bei der Join-Auswertung können Sie von einem Nested-Loop-Algorithmus ausgehen. Speziell bei Hybrid-Shipping sollten Sie sich auf „vernünftige" Auswertungspläne beschränken.

Viel Erfolg! — May the calculator with you!

Ausgabe und Vorbesprechung: 29.10.1996
Abgabe und Besprechung: 05.11.1996