

Tutorübung zur Vorlesung Grundlagen Rechnernetze und Verteilte Systeme Übungsblatt 12 (8. Juli – 12. Juli 2013)

Hinweis: Die mit * gekennzeichneten Teilaufgaben sind ohne Kenntnis der Ergebnisse vorhergehender Teilaufgaben lösbar.

Aufgabe 1 Domain Name System (DNS)

Eine zentrale Aufgabe des Domain Name Systems (DNS) ist es, menschenlesbare Namen auf IP-Adressen abzubilden, die dann für die Wegewahl auf der Netzwerkschicht verwendet werden können. Bei dem aus der 3. Programmieraufgabe bekannten Namen `farina.net.in.tum.de` handelt es sich um einen sog. *Fully Qualified Domain Name (FQDN)*.

a)* Benennen Sie *alle* Bestandteile dieses FQDNs.

In Abbildung 1 sind ein PC sowie eine Reihe von DNS-Servern dargestellt. Wir nehmen an, PC1 nutze einen DNS-Server von Google unter der IP-Adresse 8.8.8.8 zur Namensauflösung. Ferner nehmen wir an, dass der Google-Server gerade neu gestartet wurde (also insbesondere keine DNS-Einträge gecached hat) und rekursive Namensauflösung anbietet. Der Server `D.ROOT-SERVERS.NET` sei ein DNS-Rootserver während `F.NIC.de` einer der autoritativen Namensserver für `de`-TLDs ist.

b)* Welche Funktion erfüllen `D.ROOT-SERVERS.NET` und `F.NIC.de`?

c)* Für welche Zonen sind die Server `ns.tum.de`, `ns.in.tum.de` und `ns.net.in.tum.de` (vermutlich) autoritativ?

d) Zeichnen Sie in Abbildung 1 alle DNS-Nachrichten (Request / Response) ein, die ausgetauscht werden, sobald PC1 auf `farina.net.in.tum.de` zugreift. Nummerieren Sie die Nachrichten gemäß der Reihenfolge, in der sie zwischen den einzelnen Computer ausgetauscht werden.

e) Erklären Sie den Unterschied zwischen iterativer und rekursiver Namensauflösung.

f)* Wie wird im DNS sichergestellt, dass kein bösartiger DNS-Server Anfragen für andere Domänen beantwortet? (Wir gehen davon aus, dass keine Man-in-the-Middle-Angriffe möglich sind.)

Der FQDN eines DNS-Servers muss nicht notwendiger Weise mit der Domain korrelieren, für die dieser Server autoritativ ist. Beispiele hierfür sind bereits die Rootserver sowie TLD-Server.

g)* Überprüfen Sie mittels des Kommandozeilen-Tools `dig` (Linux / OSX) bzw. `nslookup` (Windows), welche DNS-Server tatsächlich für die Domain `tum.de` sowie deren Subdomains zuständig sind. **Hinweis:** Bedingen Sie Ihre abfragen auf den Resource Record Type NS (Name Servers).

Aufgabe 2 MapReduce

Wir haben von verschiedenen Verkäufern Preislisten und möchten für jeden Artikel den günstigsten Anbieter und Preis bestimmen. Eine Preisliste ist eine Tabelle, in der in jeder Zeile ein Produktname und ein Preis steht. Am Ende hätten wir gerne eine Tabelle, in der für jeden Produktnamen der günstigste Anbieter und der günstigste¹ Preis steht.

Die gesamte Transformation lässt sich wie folgt formulieren:

$$(\text{Anbieter} \times (\text{Produkt} \times \text{Preis}) \text{ list}) \text{ list} \rightarrow (\text{Produkt} \times (\text{Anbieter} \times \text{Preis}) \text{ list}) \text{ list}$$

Schreiben Sie geeignete Map- und Reduce-Funktionen (z. B. in Pseudo-Code oder mathematischer Notation), die das gewünschte Ergebnis im MapReduce-Framework berechnen. Die Eingabe für die Map-Funktion besteht dabei aus dem Anbieternamen und der Preisliste.

Aufgabe 3 Amdahl's Gesetz

Ein Programm soll in einem Cluster mit n Knoten verteilt ausgeführt werden. Das Programm lässt sich in einen Anteil zerlegen, der optimal parallelisierbar ist (p) und einen sequentiellen Anteil ($1 - p$), der nur auf einem Knoten ausgeführt werden kann.

Mit Hilfe von Amdahl's Gesetz kann man den maximalen Speedup s bestimmen, den dieses Programm auf diesem Cluster erreichen kann:

$$s = \frac{1}{(1 - p) + \frac{p}{n}}$$

Um die maximal erreichbare Effizienz mit anderen Programmen und Clustern erreichen zu können, ist der Speedup relativ zur Gesamtzahl an Knoten $s^* = s/n$ interessanter als der absolute Speedup.

a)* Was ist der Unterschied zwischen Speedup s und relativem Speedup s^* ?

b)* Bestimmen Sie mit Hilfe von Amdahl's Gesetz den Anteil des Programms, der parallelisierbar sein muss (p), um einen bestimmten, relativen Speedup (s^*) zu erreichen!

c) Bestimmen Sie den parallelisierbaren Anteil des Programms für verschiedene relative Speedups und Anzahlen an Knoten:

| p | $s^* = 50\%$ | $s^* = 80\%$ |
|-----------|--------------|--------------|
| $n = 2$ | | |
| $n = 10$ | | |
| $n = 100$ | | |

¹Wir ignorieren den Fall, dass mehrere Verkäufer zum günstigsten Preis anbieten.



PC1



google-public-dns-a.google.com
(8.8.8.8)



D.ROOT-SERVERS.NET
(128.8.10.90)



F.NIC.de
(81.91.164.5)



ns.tum.de



ns.in.tum.de



ns.net.in.tum.de

Abbildung 1: Loesungsblatt zu Aufgabe 1.