

VerteilteWebInf Hausaufgabe 2

Gruppe 6

October 21, 2014

Aufgabe 1

- a) Socket-basierte Kommunikation bringt einige Nachteile für verteilte Systeme mit sich. Zunächst ist ein Socket eine Zwei-Wege-Verbindung zwischen einem Client und einem Server. Parallelität bei verteilten Anwendungen muss also explizit vom Programmierer berücksichtigt werden, da Sockets per se keine Abstraktion / Transparenz für parallele Anwendungen liefern. Bei der Verwendung von Datagram-Sockets welche Daten per UDP übertragen, muss (manuell) sichergestellt werden, dass die UDP Nachteile wie z.B. Packetloss ausgeglichen werden. Bei Sockets die eine TCP Verbindung aufbauen, muss der Verbindungsauf- und Abbau explizit programmiert werden. Bei verteilten Systemen muss die Verbindung zu jedem der beteiligten Server sichergestellt und überwacht werden.
- b) `c++ -o server server.cpp -lzmq`, Server starten: `./server`
`c++ -o client client.cpp -lzmq`, Client starten: `./client 5678`
- c) ZeroMQ und ähnliche Systeme haben den Vorteil, dass sie einem Entwickler viel Zeit und Arbeit abnehmen können. Bei verteilten Systemen kann es z.B. vorkommen, dass verschiedenen Systemen auch eine unterschiedliche Byte-Reihenfolge (Little vs. Big Endian) zugrunde liegt. Außerdem sind verteilte Systeme oft dynamisch aufgebaut, d.h. abhängig von den benötigten Ressourcen, werden mehr oder weniger Rechner bereitgestellt um die Last zu verarbeiten. ZeroMQ liefert bereits entsprechende Abstraktionen und Features wie Message Ques, um die Komplexität der genannten Dinge zu verstecken. Außerdem ist ZeroMQ ein etabliertes und funktionierendes System, mit dem Entwickler sich auf die Entwicklung ihrer eigentlichen Anwendung fokussieren können und nicht um den Overhead der Socket-Programmierung. Nachteil eines solchen Systems ist, dass man an den Hersteller gebunden ist und einen gewissen "Vendor Lock-In" hat.

Aufgabe 2

- a) r Tupel t von A nach B übertragen:

$$t_{A \rightarrow B} = \frac{r \cdot k_r}{b} + r \cdot l = r \left(\frac{k_r}{b} + l \right)$$

Menge von Tupeln (j Stück) von B nach A übertragen:

$$t_{B \rightarrow A} = \frac{j \cdot k_s}{b} + r \cdot l$$

\Rightarrow

$$t_{ges} = t_{A \rightarrow B} + t_{B \rightarrow A} = r \left(\frac{k_r}{b} + l \right) + \frac{j \cdot k_s}{b} + r \cdot l$$

- b)
- bei verteilten Systemen, bei denen nicht sämtliche Daten/die gesamte Relation übertragen wird (z.B. aus Datenschutzgründen)
 - Join kann dadurch parallelisiert werden, allerdings Performanceverlust durch höhere Latenz
 - Verbindung zweier autonome Systeme

Aufgabe 3

- a) Man könnte dem Joinoperator noch die Joinbedingung übergeben. Dadurch sind nicht nur natürliche Joins möglich (sondern allgemeine Joins).
- b)
- Vorteil: geteilte Arbeit
 - Nachteil: langsamer aufgrund langsamer Kommunikation