Sheet 1, SS 2015
Verteilte Systeme/ Distributed Systems
Published 23.4.2015

Katinka Wolter
Simon Schmitt
Inst. für Informatik
Freie Universität Berlin

Distributed Systems Lecture - Discussion 30.4.2015

Assignment 1. Distributed Hash Tables

Please explain what distributed hash tables are and how they work. Explain two of the methods for keyspace partitioning that are used in systems such as

1. Self-Chord
2. Kademlia (Mainline DHT)
3. Koorde
4. Pastry
5. Tapestry
6. Apache Cassandra
7. Riak

Assignment 2. Properties of Distributed System Architectures

Please discuss the following problems in the context of the architectures and properties we have discussed.

1. A service is implemented by several servers. Explain why resources might be transferred between them. Would it be satisfactory for clients to multicast all requests to the group of servers as a way of achieving mobility transparency for clients?
2. The host computers used in peer-to-peer systems are often simply desktop computers in users offices or homes. What are the implications of this for the availability and security of any shared data objects that they hold and to what extent can any weaknesses be overcome through the use of replication?
3. Consider a hypothetical car hire company and sketch out a three-tier solution to the provision of their underlying distributed car hire service. Use this to illustrate the benefits and drawbacks of a three-tier solution considering issues such as performance, scalability, dealing with failure and also maintaining the software over time. Discuss the difference between tiers and layers.
4. Consider a simple server that carries out client requests without accessing other servers. Explain why it is generally not possible to set a limit on the time taken by such a server to respond to a client request. What would need to be done to make the server able to execute requests within a bounded time? Is this a practical option?

5. For each of the factors that contribute to the time taken to transmit a message between two processes over a communication channel, state what measures would be needed to set a bound on its contribution to the total time. Why are these measures not provided in current general-purpose distributed systems?

Distributed Systems Seminar - Discussion 07.05.2015

Assignment 1. Looking at Frameworks

Have a look at these three frameworks which focus on distributed and/or parallel applications:

- akka - `http://akka.io/`
- OpenMPI (Open Message Passing Interface) - `http://www.open-mpi.org/`
- OpenMP (Open Multi-Processing) - `http://openmp.org/`

1. Describe these frameworks (i.e. attributes, applications, ...).
2. Develop a small chat-room like application using

   akka **or** OpenMPI

   with multiple participants on the same and on remote machines.
3. Whenever a sender sends a string message (consecutive letters of arbitrary length) into the chat room, the receivers shall count all occurances of the consecutive letters "DSS" (abbr. of *Distributed Systems Seminar*) in that message, and send the result back.
4. Try to minimize the latency of the result in your network by using whatever methods provided by those frameworks. If you chose OpenMPI to implement your solution, you should use OpenMP for parallelization.

You can use any appropriate programming language for your task. Please be prepared to present your solution and design choices.