Overlay and P2P Networks

Introduction

PhD Samu Varjonen

16.1.2017
Contents

• Course Overview
• Lectures
• Assignments/Exercises
Course Overview

• Overlay networks and peer-to-peer technologies have become key components for building large scale distributed systems.

• This course will introduce overlay networks and peer-to-peer systems, discuss their general properties, and applications. The course will cover the following topics:
  – Currently deployed peer-to-peer systems and how they work
  – Distributed Hash Tables as a base for structured peer-to-peer systems
  – Peer-to-peer storage systems
  – Performance issues, legal aspects, and privacy issues
  – Overlay content distribution algorithms
General Info

Advanced course, 5 credits

Requirements: basics of networking

Assignments/exercises done as group work (1-2 persons), idea is to keep the same group structure (but do tell about freeriders!)
Lectures

- Lectures
  - Monday 14-16 in B222 16.1.-27.2.
  - Thursday 12-14 in B222 19.1-2.3.

- Assignments/exercises
  - Wednesday 12-14 D122 18.1-1.3.

- Course based on book
Exam material: additional articles


Lectures

19.1. Unstructured networks I
23.1. Unstructured networks II
26.1. Bittorrent, modelling and evaluation
30.1. Freenet and intro to power-law networks.
2.2. Consistent hashing. Distributed Hash Tables (DHTs)
6.2. DHTs continued.
16.2. Applications I
20.2. Applications II
23.2. Advanced topics
27.2. Conclusions and summary
Grading

Course grading will be based on the final exam and the assignments/exercises.

Course exam 8.3.2017 16:00 B123

Separate exam on 28.4.2017 16:00 in B123

You can use the exercise score with the first separate exam
Assignments/Exercises

- Assignments are done in two-person groups (or alone), groups can change between assignments.
- Assignments give bonus points for the exam.
  - Max exam points 18, max bonus points 4.
- Wednesday 12-14 B222 last session 1.3.
  - 18.1. Reception on questions I
  - 24.1. Answers due to questions I
  - 25.1. Feedback on questions I
  - 1.2. Reception on questions II
  - 7.2. Answers due to questions II
  - 8.2. Feedback on questions II
  - 15.2. Reception on questions III
  - 21.2. Answers due to questions III
  - 22.2 Feedback on questions III
  - 1.3. Generic questions and answers, to prepare for exam.
Consistent hashing alleviates network problems and eventual consistency can be achieved through replication and synchronization.

Examples: Dynamo, Cassandra

Replication, Gossip, etc.

Selective flooding

Consistent hash (O(1) DHT)

Good for arbitrary data and search functions, can aggregate routing info, structure improves scalability.

Examples: Gnutella and Freenet

Example: BitTorrent

Limited flooding / depth first / Bloom filters

Tracker

DHT

Examples: Chord, CAN, Kademlia

Storage: PAST

Rendezvous: Scribe (for multicast), i3

Good for name/value data, note flat address space, one node is responsible, churn is a concern

Consistent hashing alleviates network problems and eventual consistency can be achieved through replication and synchronization.

Examples: Dynamo, Cassandra

Replication, Gossip, etc.

Selective flooding

Consistent hash (O(1) DHT)

Search

Storage

Rendezvous

Search

Storage

Rendezvous

Search

Storage

Rendezvous

Cluster

Wide-area (unstructured)

Wide-area (structured)

Internet (TCP/IP)
<table>
<thead>
<tr>
<th>Main theme</th>
<th>Prerequisites</th>
<th>Approaches learning goals</th>
<th>Meets learning goals</th>
<th>Deepens learning goals</th>
</tr>
</thead>
</table>
| Overlay and peer-to-peer networks: definitions and systems | Basics of data communications and distributed systems (Introduction to Data Communications, Distributed Systems) | Knowledge of how to define the concepts of overlay and peer-to-peer networks, and state their central features  
Ability to describe at least one system in detail | Ability of being able to compare different overlay and p2p networks in a qualitative manner  
Ability to assess the suitability of different systems to different use cases | Ability to give one's own definition of the central concepts and discuss the key design and deployment issues |
| Distributed hash tables                        | Basics of data communications and distributed systems (Introduction to Data Communications, Distributed Systems)  
Big-O-notation and basics of algorithmic complexity | Knowledge of the concepts of structured and unstructured networks and the ability to classify solutions into these two categories  
Knowledge of the basics of distributed hash tables  
Ability to describe at least one distributed hash table algorithm in detail | Ability of being able to compare different distributed hash table algorithms  
Ability of designing distributed hash table-based applications  
Knowledge of key performance issues of distributed hash table systems and the ability to analyze these systems | The knowledge of choosing a suitable distributed hash table design for a problem  
Familiarity with the state of the art |
| Reliability and performance modelling          | Basics of probability theory  
Basics of reliability in distributed systems | Ability to model and assess the reliability of overlay and peer-to-peer networks by using probability theory  
Knowledge of the most important factors pertaining to reliability | Ability of analytically analyzing the reliability and performance of overlay and peer-to-peer networks  
Understanding of the design issues that are pertinent for reliable systems | Familiarity with the state of the art |
| Content distribution                           | Introduction to Data Communications | Knowledge of the basic content distribution solutions  
Ability to describe at least one overlay and p2p network based content distribution solution | Knowledge of different content distribution system and the ability to compare them in detail  
Knowledge of several content distribution techniques | Familiarity with the state of the art |
| Security                                       | Basics of computer security | Knowledge of the basic security issues with overlay and p2p networks  
Knowledge of the sybil attack concept | Ability to discuss how security problems and limitations can be solved  
Knowledge of how to prevent sybil attacks | Knowledge of how to prevent sybil attacks  
Familiarity with the state of the art |
Contact information

Lecturer PhD Samu Varjonen (contact info on homepage)

Assignments: Juhani Toivonen (@cs.helsinki.fi)

Course homepage can be found: www.cs.helsinki.fi/courses
Questions?
Overlay and P2P Networks

Introduction and unstructured networks

PhD Samu Varjonen

16.1.2017