582615 - Overlay and P2P Networks, Spring 2017

Exercise package 2 (20 points)

The exercises are intended to be done working in pairs. This package contains four exercises and an optional turbo challenge. During the course there will be three sets of exercises. The course book and the lectures contain some answers, but searching for outside sources too is strongly encouraged.

Schedule

There are two types of exercise sessions: Clarification sessions, where you can ask questions about the exercises or other matters about the course; and Answer sessions, where some answers to the returned exercises are presented and discussed.

- Clarification session: Wednesday 1.2. at 12:15
- **Exercise deadline**: Tuesday 7.2. at 12:00
- Answer session: Wednesday 8.2. at 12:15

Submission

Return your answers by email to juhani.toivonen@cs.helsinki.fi as an attached PDF or TXT document. Use "Overlay Exercise 2" as the subject line. The document should include:

- The title "Overlay exercise package 2"
- The name and student number of the writer/writers
- The answers to the exercises

Assignments

Assignment 1 - Bloom filters (5 points)

A Bloom filter (named after its inventor Burton Howard Bloom) is a deliberately non-error-free data-structure. Concentrate on the standard Bloom filter and describe:

- What are Bloom filters and what are they used for?
- What is the difference between a Bloom filter and a regular hash table?
- Describe the steps of inserting an item into a Bloom filter, and querying the presence of that item.
- Can elements be removed from a standard Bloom filter? Explain why.
- Bloom filters sometimes return false results. Doesn't that make them useless? What do we know about the results?

Assignment 2 - Freenet (5 points)

The Freenet Project has created a model and software for an overlay network with the goal of enhancing online privacy and preventing censorship.

- What is Freenet? What services does it provide to its users? (not what content can be found)
- How does Freenet find files that are requested?
- How does it guarantee privacy?

The basic operation of Freenet is described in the journal article "Protecting Free Expression Online with Freenet" by Ian Clarke, Scott G. Miller, Theodore W. Hong, Oskar Sandberg and Brandon Wiley. It was published in IEEE Internet Computing journal, volume 6, issue 1, year 2002. https://freenetproject.org/papers/freenet-ieee.pdf

Assignment 3 - Distributed Hash Tables (DHT) (5 points)

Distributed Hash Tables, like hash tables, are systems for storing (key, value) pairs, but in a setting that can span multiple hosts. Describe:

- What are typical operations that a DHT provides?
- What challenges does spreading a hash table over multiple nodes impose?
- Choose a DHT (e.g., Kademlia, CAN, Tapestry or Chord) and describe the steps of inserting a (key, value) pair into that DHT. What can you say in general about retrieving that value from the DHT?

Assignment 4 - Consistent hashing (5 points)

Consistent hashing is a technique used for load balancing and minimizing the effort of redistributing keys when changing the amount of nodes in a DHT. Describe:

- How does consistent hashing work?
- What happens when you add/remove a node in a system that uses consistent hashing? Why is it important to use a well balanced hash function?
- How can replication be done with consistent hashing?

Turbo challenge (optional, up to 4 points)

The turbo challenge allows you to recover lost points from other assignments, but will not increase the maximum points available. You can get full points from the exercise set without the turbo challenge. The turbo challenges may step outside the themes discussed on the course.

Tor is another overlay system for increasing anonymity and circumventing censorship on the internet. Unlike Freenet, Tor can be used to anonymise browsing of the regular World Wide Web.

- How does Tor anonymise the browsing?
- How does Onion routing work? i.e. What happens while a message travels across the Tor network.
- In addition to regular WWW, Tor can be used to access so called hidden services. How are they different from browsing regular WWW sites?
- Tor provides anonymity. Does it also mean security? How can anonymity get compromised on Tor?