Exercise 5

Peer-to-Peer Networks, Spring 2008

Guidelines

- Group work, groups of up to 3 people allowed
- Due date: 18.4. at 10:00 (start of exercises)
- Exercise is worth 12 points
- What to return: Bring answers on paper to the exercises on 18.4.
- Don't forget to write all the names and student numbers on the answer!

Question 1

Consider a DHT-based P2P storage system where different objects are stored. Suppose that the system has I peers which are uniformly distributed in the hash space. Suppose that a file is stored in the node which is responsible for the file's ID and that for each file, we create k copies of the file (in the nodes closest to the actual responsible node). Furthermore, suppose that each peer is up at a given time with probability p. Peer up probabilities are assumed to be independent of each other.

Peer A stores file F at time t_0 on the k closest peers. Later, at time t_1, peer B wants to retrieve file F. Derive the expression for the probability that at least one of the k peers is up at time t_1.

Question 2

Is the probability of Question 1 equal to the probability of peer B getting file F? If not, why not?

Question 3

Consider a system where the files are stored on the peers who want to store them, and the responsible nodes store only pointers to these nodes (e.g., like in Tapestry). These pointers are stored on the k closest nodes. All the other assumptions from Question 1 are still valid.

Peer A publishes a file F at time t_0, i.e., informs the responsible node that A has F. Later, at time t_1, peer B wants to retrieve F. Assuming that between t_0 and t_1, peers have only gone down or crashed (i.e., case 1 from lecture slides), what is the probability of B getting file F?

Question 4

Compare the probabilities from Question 1 and Question 3 (assume case 1 from the lecture slides). What conclusions can you make to the differences between a storage system that stores objects and a system which stores pointers to objects?

Question 5

Let $I = 1000$, $k = 5$, and $p = 0.8$. What are the probabilities of Question 1 and Question 3?

Question 6

Consider the system in Question 1. Suppose that between times t_0 and t_1, 20 new peers join the system. Furthermore, suppose that peer B can query only the k closest peers and that the new peers do *not* receive the data they should when they come up (i.e., only the original k nodes have the file). What is the probability of peer B getting file F? Use the numbers in Question 5. (Note that any of the k peers may be down at time t_1.)