Name_____ Last name_____ Student number | | | | | | | |

Web and Cloud Computing
academic year 2012-13

Rules

- This is a closed books exam.
- The operation of any electronic device is prohibited (e.g, no calculator, phone or PDA).
- Answer the questions being *precise, complete,* and *formal.*
- Write as *clearly* as possible, both in terms of handwriting and wording.

Questions

1. Consider the following food-chain: Cats eat Goldfish, Goldfish eat Mosquitos, Cats eat Frogs, Frogs eat Mosquitos, Humans eat Frogs. Represent the food chain as a directed graph.
   (a) In such graph what is the average indegree, outdegree, and degree?
   (b) What is the diameter?
   (c) Consider the PageRank algorithm applied to the food-chain graph and provide the G matrix (assume that species deviate from the food chain behavior with probability 10%).
   (d) Which species is most important in the food chain (based on PageRank)?
   (e) Add a new eating behavior to the food-chain so that Humans become more important according to PageRank.
2. Compare Model-View-Controller and View-First approaches to web development. Give and example of a view and a controller, considering a typical web application (ex., online chat). Is it possible to combine MVC pattern with AJAX communication model? If yes, provide an example of a view and a controller in that case.
3. Provide a description of document-oriented databases. What are the difference between them and traditional (relational) databases? What are pro et contra of document-oriented databases?
4. What is the difference between CP and AP (in terms of the CAP theorem)? Provided with a replication strategy for some arbitrary database, how to find out if the database with the given replication strategy is CP or AP? What should you pay your attention to?
5. How would you implement a leader election algorithm using Zookeeper?
6. Cassandra database is considered to be very efficient for applications with a large amount of writes. What makes Cassandra to be efficient (support your answer with an overview of Cassandra architectural/data model design decisions)?
7. How the locality of machines can be taken into account for the MapReduce task to increase its performance?
8. What is the synchronization point (the barrier) in the MapReduce task? Why does the barrier exist?