

Name _____

Last name _____

Student number

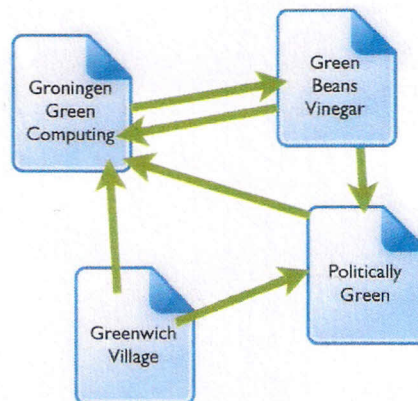
--	--	--	--	--	--	--	--

Web and Cloud Computing academic year 2013-14

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 document graph in Figure where each document contains $2/3$ words and has links to other documents.
 - (a) What is the average degree of such graph?
 - (b) What is the diameter of the graph?
 - (c) Provide a vector representation of each document in the space of words.
 - (d) In such space, what are the closest documents in terms of cosine similarity?
 - (e) Consider the PageRank Algorithm. Provide the H and G matrices assuming $\alpha = 0.5$. Compute the PageRank of each page up to at least step 3.
 - (f) What should be the result of a query "Green" to a system implementing PageRank ranking and a vector representation of documents?
2. What is the Command pattern? How is it used in the context of web development? What are its benefits/drawbacks?
3. What is index-free adjacency property? Why is it so important for graph databases?
4. What are R, W, and N parameters for data replication? What can you tell about the database consistency level looking at these parameters? What are the values of R, W, and N for a traditional RDBMS master-slave replication?
5. How would you implement a distributed locking algorithm using Zookeeper?
6. Let say, I have to build a recommendation system, and I am interested to answer the following question: "What is liked by the people that like the same things as me that I dont already like?". How would you model the data? What kind of query do you need to write to answer the question for a particular user? *Hint: use graph databases.*
7. How does Apache Giraph (an open source implementation of Google Pregel) deal with faults? What will happen if one of the workers crashes? What will happen if master crashes?
8. What is an actor model in distributed computing? How does it deal with faults?