Homework 2
ITCS-6010/8010: Cloud Computing for Data Analysis
Due: Friday, March 2, 2012

Homeworks are due at the beginning of class on March 2, and are to be done individually. Homeworks will be graded on the basis of clarity and legibility. See course syllabus for late submission policy.

1. The *seek time* is the time needed to position the disk head in a new position. The *transfer time per byte* is the rate of transfer from disk to memory when the head is in the right position. Assume the seek time is 5 milliseconds and the transfer time per byte is 0.02 microseconds. What is the time to transfer 64 MB from disk to memory if it is stored as one chunk, versus if it is stored in 100 noncontiguous chunks (requiring the disk head to move up to 100 times)?

2. If in a non-fault-tolerant system with 1000 nodes, each node has an 99.9% uptime, what is the percentage of time that all nodes in the system are up?

   What is the percentage of time that 5 or fewer nodes are down?

3. Consider a web graph with three nodes 1, 2, 3. The links are as follows: $1 \rightarrow 2$, $3 \rightarrow 2$, $2 \rightarrow 1$, $2 \rightarrow 3$. Write down the transition probability matrices for the surfer's walk with teleporting, for the following three values of the teleport probability:

   (a) $\alpha = 0$,

   (b) $\alpha = 0.25$,

   (c) $\alpha = 1$.

4. You must submit a reading report on the following paper, which describes the original Google search engine.

   The Anatomy of a Large-Scale Hypertextual Web Search Engine, Sergey Brin and Lawrence Page. Proceedings of the Seventh World Wide Web Conference (WWW7), pp. 107-117, 1998.

   The paper is available at `http://infolab.stanford.edu/~backrub/google.html` or off the course web page at `http://www.cs.uncc.edu/~sakella/courses/cloud/papers/`.

   For this homework, your reading report may be up to **two** pages in length, and must be written on a computer.

   (a) Your reading report should provide a concise summary of the paper (e.g., what was the problem addressed, what did they do and what was their approach, what were their contributions).

(b) It should also provide a brief technical critique of the paper (e.g., the importance of the problem, what are the merits of the approach, what could have been done better or described more clearly, what might the next steps be).