Homework 1
ITCS-6010/8010: Cloud Computing for Data Analysis
Due: Friday, February 24, 2012

Homeworks are due at the beginning of class on February 24, and are to be done individually. Homeworks will be graded on the basis of clarity and legibility. See course syllabus for late submission policy.

You are to submit a reading report on the following two papers.

1. *MapReduce: Simplified Data Processing on Large Clusters, Jeffrey Dean and Sanjay Ghemawat, OSDI'04: Sixth Symposium on Operating System Design and Implementation, 2004.*

2. *The Hadoop Distributed File System, Konstantin Shvachko, Hairong Kuang, Sanjay Radia, Robert Chansler, 2010 IEEE 26th Symposium on Mass Storage Systems and Technologies (MSST '10), pages 1-10, May 2010.*

The papers are available off the course web page at
`http://www.cs.uncc.edu/~sakella/courses/cloud/papers/`.

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

1. 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).

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