

Name _____ SAMPLE _____

Last 4 digits of Student #: _____

GRADE: _____/100

CSE 6331, Cloud Computing
Quiz 2, Summer 2012 © DL, UTA, 2012

Please read this:

This is a closed book, closed notes quiz. Use a dark ink (or pencil) and print answers on the test paper following the questions.

Please answer the questions in a **few keywords**, complete sentences are not necessary, **be brief! Write** your answers **legibly**. Unreadable answers will be counted wrong.

Please - no talking! You may write on back. Please turn in quiz promptly, when asked.

1. How many Daemon processes run in Hadoop? Please give a 1 line description of each.
2. In Hadoop, what are combiners?
3. We would like to compete in the "Large Sort" Hadoop contest. We will measure the total time to create and sort ten billion long integers (10^{10} long ints). Assuming it takes 500 secs to create the unsorted list, and we run on 100 nodes (computers), how much **total time** will it take to produce the sorted list? Show/explain your reasoning - is this 100 times faster than one node (x100), or slower, or faster? Why?
4. For Hadoop: Which API/methods do you use to do (any two of) the following:
 - a. Adding input file(s) to the map reduce program
 - b. Letting your program know the key and value for the mapper
 - c. Specify the mapper and the reducer for the program
 - d. Create a new map reduce job