Exam questions for Cluster computing DIKU 2007

The exam for the Cluster Computing class is based on the programming exercises, either the original cluster computing exercises or the new eScience assignments. There are five assignments in total, as shown in figure 1 below.

Category	**Cluster Computing**	**eScience**
1 Shared Memory	Road Map	Tumor radiation
2 PVM	Race Trap	Protein Folding
3 MPI	Freeze Trap	Successive Over Relaxation
4 Remote Memory	Wind Trap	Lattice Gas Automaton
5 Tuple Space	Clone Machine	N-Body simulation

Figure 1. The programming assignments

For the exam you must choose 3 of assignments for your final report. You may only choose one assignment from each category so that you end up with a report that has code based on 3 APIs. The report that describes the 3 assignments must be kept below 15 pages and must include performance figures including speedup or CPU utilization graphs.

The description of each programming assignment and a sequential code for the assignment may be found on www.diku.dk/~vinter/CC. You should ignore assignment descriptions regarding deadlines and number of CPUs to test on, this present document holds the valid information for deadline and test environments.

You must test your code on 1,2,4 and 8 CPUs in each of the assignments. Hardware is available through the Minimum intrusion Grid, MiG, and you need a certificate to access these resources, read more on www.migrid.org.

You are allowed to work in groups of up to 3 persons. In addition to the programming assignments each person must answer, individually, a theoretical assignment, the length of the reply of this question should be limited to one page. The allocations of theoretical assignments are as follows:

The member of the group who has his or her birthday first in a given year, i.e. lowest day, month, answers question a., the second answers question b. and the third answers question c. Individuals who work alone always answer question a. and groups of two naturally answers questions a. and b.. You should mark clearly who has replied to which individual assignment on top of the page.

Theoretical questions.

a. **Latency hiding**
   Overlapping communication and calculation is called latency hiding. Explain the purpose of latency hiding and the advantages one may hope to achieve. Discuss the premises that must be met for latency hiding to work and the limitations to latency hiding as a technique for improving performance.
b. **Task Granularity**
   When we make applications parallel we divide the work into subtasks for each processor to handle, the size of such subtasks are also referred to as task granularity. Discuss the

implications of finer task granularity, i.e. smaller subtasks. What are the consequences for performance and the requirements towards the hardware as parallelism becomes more fine-grained.

c. **Homogeneity vs. heterogeneity**
   In this class we have argued that clusters must be homogeneous. Discuss the problems that arise with the introduction of heterogeneity in a cluster. Given the applications you have written for the practical portion of this class, discuss which codes could be made to work with heterogeneous systems and which require homogeneous clusters.

Deadline for hand in of the exam report is July 2nd at 12.00.