

CS/CmpE 6761: Parallel Computer Architecture II
Final Exam: June 11, 1998 2:50 to 4:50 (MAX = 90 pts)
OPEN BOOK AND NOTES
CONCISE BULLETS GET MORE CREDIT THAN WORDY SENTENCES
LEGIBILITY IS A REQUIREMENT NOT AN OPTION!

1. (1 point, 1 min) Pick one of the following events as the top one with far reaching international consequence for the future:
 - (a) The relationship between a certain Washington individual and Monica Lewinsky.
 - (b) The economic crisis in Asia.
 - (c) The political crisis in the Indian subcontinent.
 - (d) The political crisis in Kosovo.
 - (e) Global Warming.
 - (f) El Nino.
2. (9 points, 10 min) Explain the *pros* and *cons* of keeping the directory information of a CC-NUMA machine at the memory or the cache.
3. (10 points, 10 min) An SMP node is to be used as a building block for a large-scale cache-coherent machine. If the SMP employs a weak memory model (such as PSO) at the hardware level, is it possible to implement a stronger memory model (say SC) at the hardware level of the large-scale machine? Explain why, or why not.
4. (20 points, 20 min) Give a flow chart for the path of a processor read operation in a virtually indexed physically tagged *flat* COMA organization. Make the following assumptions about the organization of the attraction memory:
 - page size 4Kbytes (this is the granularity at which the occupancy information of the attraction memory is kept);
 - block size 128 bytes (this is the granularity at which the validity of the data is maintained in the attraction memory, and it is also the granularity at which coherence actions are performed);

In your answer, be specific as to the actions taken upon replacements in the attraction memory.

5. (20 points, 20 min) Each thread of a certain parallel computation executes the following code:

```
loop
  compute vector  $V_{m,y,d}$  using all  $V_i$ 
  barrier sync
  check for convergence
end
```

V is a shared data structure, with each thread reading all of it, and writing a portion of it in each iteration of the loop as shown above. Assume that the base machine uses a write-invalidate cache coherence protocol. The machine has mechanisms for both a non-binding prefetch and a poststore.

- (a) Modify the above code to use
 - i. the prefetch mechanism
 - ii. the poststore mechanism
 - (b) Comment on which of the two mechanisms is expected to result in improving the performance of the original loop, and why.
6. (30 points, 30 minutes) Consider an 8 node hypercube. An algorithm requires a computation time of 1 second on a single processor. The problem can be parallelized such that the computation time is equally divided among either 4 or 8 processors. However, in the parallel solution each processor would have to send a round-trip message to each of the other processors once every 1 millisecond of computation time on the average.

What can be the maximum latency on each link of the hypercube such that it is still worthwhile parallelizing the problem on 8 processors? Assume that the messages can be routed without any contention on the network.