Name:

EE666 Mid-Term #1

Please answer all questions. You have 120 minutes starting from now.
Please write clearly and concisely. Please **do not dump core**!

Programming Model (10 points)

1. Define each of shared address space, message passing, and data parallel programming models. (5 points)

   Name two advantages of the message passing model against the shared address space model. (5 points)

Name:

Artifactual Communication (10 points)

2. What is artifactual (a.k.a. useless) communication? Give four instances of artifactual communication in a shared-memory multiprocessor. (6 points)

   For each of the above four instances, describe an architectural modification that would reduce the artifactual communication. (4 points)

Name:

Synchronization (15 points)

3. What is the base hardware requirement to implement Test&Set on a shared-memory multiprocessor? What is the base hardware requirement to implement MCS locks (queue-based locks in software)? Name three fundamental problems with Test&Set. (10 points)

   Briefly describe Reactive Synchronization. (5 points)

Name:

Cost-Effective Computing (10 points)

4. You are kick-starting a computer architecture research project to design high-performance multiprocessors. You have WWT-II at your disposal which is parallel simulator of multiprocessors. You want to find out if it is a good idea to buy a multiprocessor or to buy a fancy uniprocessor for your simulation studies. Assume the cost of a Sun UltraEnterprise 450 SMP can be modeled according to $cost_{SMP} = base_{SMP} + n * processor + memory$ where $base_{SMP}$ is the base system cost (i.e., cabinet, memory bus backplane, etc.), n is the number of processors, processor is cost per processor, and memory is the cost for memory. Similarly, the cost of Sun Ultra60 uniprocessor workstation is $cost_{Uni} = base_{Uni} + processor + memory$. Assume the SMP and uniprocessor use the same processor and memory technology.
   a) Assume a 300-MHz Ultra2 processor costs $600 a piece and memory costs $5 a Meg. Your simulations require 512Meg of memory and memory costs $5 a Meg. Assume that base price for SMPs is $30000 and for the fancy workstation is $5000. How much should WWT-II speed up with two processors before it is worth buying an SMP? How about four processors? (5 points)

   b) Assume that simulation typically exhibits large primary cache working sets; the software does not run well unless the primary working set fits in the cache. Assume the primary cache working set cache size for WWT-II is 1/100 of the memory size. Assume processors have a single-level cache of 2 Meg. Conjecture as to if WWT-II will meet the speedup requirement for part (a). (5 points)

Name:

Application Scaling Models (10 points)

5. Communication-to-computation ratio is a useful metric when evaluating scalability of applications on multiprocessor platforms. The ratio is defined to be the time an application spends on average communicating (and therefore not performing computation) over the time it spends computing. Explain the impact of problem-size constrained scaling, memory-size constrained scaling, and time constrained scaling on communication-to-computation ratio.

Name:

Snoopy FSM Design (20 points)

6. Sun SPARC processors implement a write-invalidate coherence protocol for a write-back cache that includes all of the MOESI states. As you recall, M stands for modified or dirty, O stands for owner, E stands for exclusive copy, S stands for shared, and I stands for invalid. How would you implement SPARC's protocol using these states? Draw the FSM diagram consisting of the states. For each state, draw a complete set of transitions both from outside (due to processor-initiated events) and between states (due to both processor-initiated and bus-initiated events). For each transition, specify an initiating event/resulting event pair. Specify which of Inhibit and/or Shared bus lines are required for your design. Specify where in your FSM such lines are used and how. Use the following set of events:

   PrRd— Processor reads
   PrWr— Processor writes
   BusRd— Bus read transaction
   BusRdX— Bus read exclusive (a.k.a read-and-invalidate)
   BusInv— Bus invalidate other copies
   BusCache— Bus cache-to-cache transfer
   BusWB— Bus writeback

Name:

Name:

Snoopy Protocol Flavors (10 points)

7. What is a write-invalidate protocol? What is a write-update protocol? Name two advantages of write-invalidate protocols over write-update protocols. (5 points)

What type of sharing pattern in application would benefit from a write-invalidate protocol? What type of sharing pattern would benefit from a write-update protocol? (5 points)

Name:

Snoopy Protocol Optimization (15 points)

8. An application often exhibits different sharing patterns across different phases. You would like to design a protocol (also referred to as competitive update/invalidate) that optimizes communication and switches between a write-invalidate and a write-update protocol depending on the sharing patterns in a given application phase. Describe the modifications required for the cache (controller) hardware to dynamically select and switch between the two base protocols. Comment on the best-case and worst-case performance of your new competitive protocol with respect to the original base protocols.