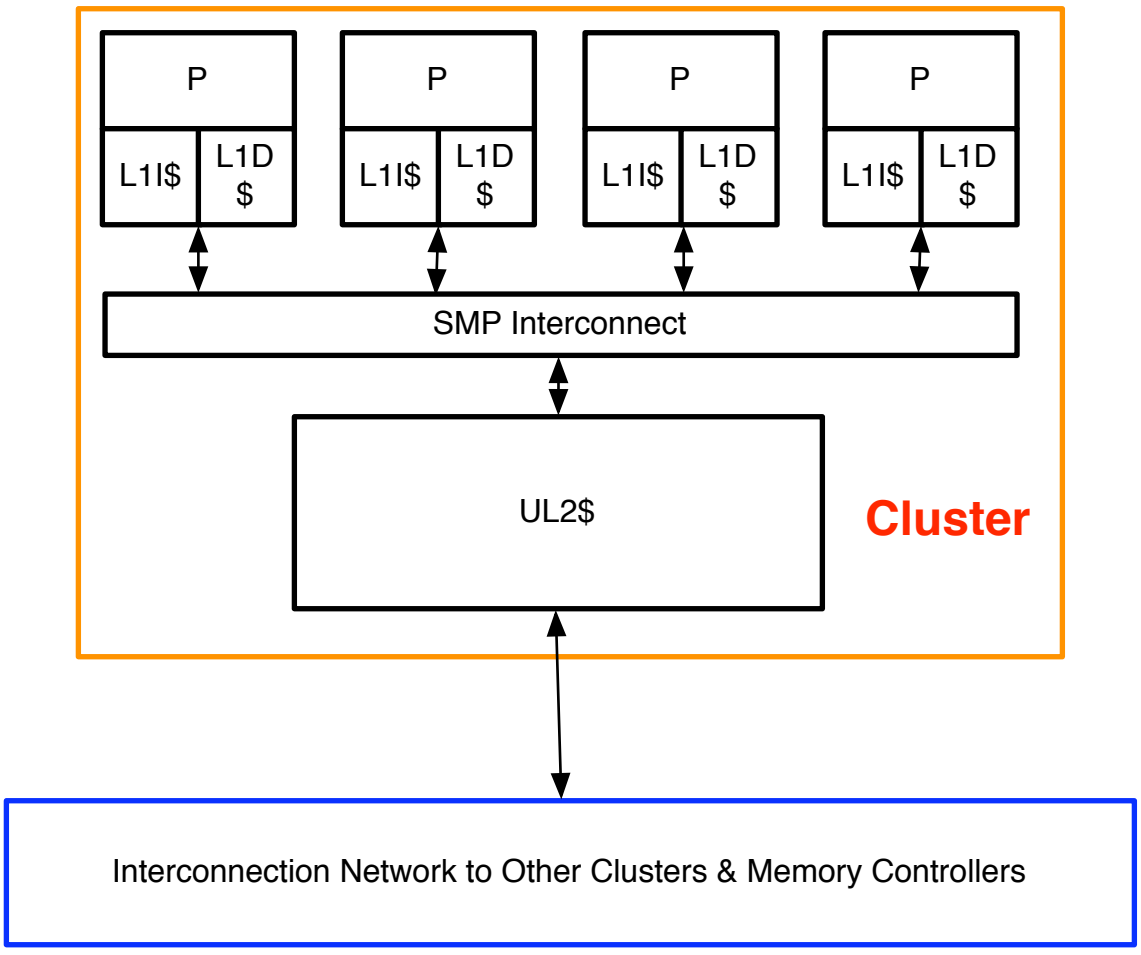
CS 6810 Homework #5

Due: 9:10 a.m. Nov. 12 (no late submissions will be graded)

General instructions: Same as usual. This homework asks you to design a memory system for a multi-core or multiprocessor CC-NUMA machine. Your assignment will be submitted in paper form in class or if you can't attend class by email to Dogan in the form of a .pdf file. NOTE – it's a very good idea to think about this assignment in a group if you are able since there are many possibilities that need to be considered.

Background: Your job is to design a cache coherent memory system for a future multi-core processor. Some decisions have already been cast in concrete. The new processor will contain 64 cores. Each core has an 8KB L1 data cache and a 4KB L1 instruction cache. The cores are arranged into 16 clusters, where each cluster contains 4 cores. Each cluster has a 128 KB unified L2 cache. All cores in a cluster communicate over a switched interconnect fabric (viz. it's not a bus). The general idea is shown in the following diagram.



The chip will also have 4 memory controllers, each controlling 2 DRAM channels and 64 GB of main memory. The total main memory size is therefore 256 GB. Each memory controller also holds a directory SRAM to maintain the global state of any cache coherent lines for which it is "home". Note that in order to keep this assignment from becoming ridiculous, there are a number of decisions that are made to simplify the design exercise.

1) Block size for both L1 and L2 caches is 64 bytes and all caches are direct mapped.
2) The L1 caches are physically addressed and physically tagged.
3) The caches are all write-back & write-allocate.
4) The caches are inclusive – meaning for a given cluster the L2 cache has a copy of any line held in any of the L1 caches.

The L1 and L2 caches are coherent and use the MESI write-invalidate protocol. The L1 caches are kept coherent using a snooping protocol via the SMP interconnect. The L2 caches are kept coherent in DSM fashion using the home directory located at the appropriate memory controller. On an L1 read miss the line will be supplied by the L2. There are some difference from the MESI protocol that was discussed in class. Namely:

1) On a write-hit to a shared line (MESI state S), the cache controller will issue an upgrade request in either the DSM or SMP case.
2) If a shared & modified line (MESI state M) is victimized, the cache controller will issue a write-back request in either the DSM or SMP case.
3) If a clean shared line (MESI state S) is victimized the cache controller will issue an unshared request in the DSM case.
4) The other thing to keep in mind is that this system needs to support both unshared and shared memory which is allocated on a per-page basis so you will need to keep track of transactions to private lines as well.
5) If a shared modified (MESI state M) line exists on core A in a cluster and a write miss happens on core B in the same cluster then core A will invalidate the line and write-back the line to the SMP interconnect and it will be copied into both the L2 cache and the L1 cache of core B. Namely there is no need to do a global invalidate.
6) Similarly L2 misses to shared lines by cluster A will be satisfied as a cache to cache transfer when possible in the DSM case. [**hint**: think about this one, it is a bit trickier than you might initially think.]

Problem 1: [10 points]

What is the physical address size that will be needed for this system? [2 points]

What is the worst-case size for the directory? [2 points]

Briefly describe a method that would allow you to have a directory that is smaller than the worst-case size. [6 points]

Problem 2: [30 points] L1 cache controller

1) Enumerate the set of SMP bus transactions that your L1 cache controller will send and receive. [5 points]
2) Enumerate the kinds of requests that the L1 cache controller will receive from the processor – note that this is NOT read-hit, read-miss, etc. – the cache controller is responsible for determining whether or not a request hits or misses. [5 points]
3) Describe what the L1 cache tags will look like both status and address tags. [5 points]
4) To simplify things you may have a state in your state machine that says "Find Victim". Draw the L1 cache controller state machine. [15 points]

Problem 3: [10 points] Switched Interconnect Messages

Enumerate the kinds of messages that will be carried over the switched interconnect, and provide formats for those messages.

Problem 4: [30 points] L2 caches

1) Enumerate the transactions that the L2 cache controller will send and receive over the switched interconnect. [5 points.]
2) Describe what the L2 cache tags look like – both status and address. [5 points]
3) Draw the L2 cache controller state machine. REMEMBER – that there is a delay over the switched interconnect between when messages are sent and received and things may arrive out of order. You will need to account for this possibility.

Problem 5: [20 points] Directory

1) Describe the structure of your directory entries including the tags. [5 points]
2) Draw the state machine for your directory controller.