Computer Architecture
EECE7352
Prof. Kaeli

Homework 5

PART A. (20 points)
In this problem you will review cache coherency and memory consistency protocols.
   a. Discuss the tradeoffs between using a snoopy cache coherency protocol and a directory-based cache coherency protocol.
   b. Discuss the role of the Exclusive state introduced in the MESI protocol.
   c. Discuss the role of the Ownership state introduced in the MOESI protocol.
   d. Complete problem 5.15 from the Hennessy and Patterson textbook.

PART B. (20 points)
In this problem, find the organization for the TLB present on two different microprocessors, one developed by Intel and the other developed by ARM. Make sure to cite your sources. Provide the details of the organization the whole TLB hierarchy, and the format of a single entry in an L1 TLB (either I or D).

PART C. (40 points)
In this problem you will utilize parallel threads to explore the use of multiple cores. Use the pi.c parallel program that utilizes pthreads to compute the value of PI using integration. The program computes the following equation:

$$PI = 4 * arctan(1)$$

Where the arctan(x) = integral from 0 to x of ($1/(1+x^2)$)

You can adjust the number of intervals chosen for the integration and the number threads used of on the command line. Note that, the more intervals used, the better the approximation.

You will use pthreads to run the pi.c program in parallel. You will need to compile your program with the –lpthread switch and the –lm switch with gcc. Run this on the 64-bit COE Linux systems. Make sure to discuss the CPU you are running on and how many cores/threads are available on the system (look in /proc/cpuinfo).

   a. Select a value for the number of intervals to get reasonable timing results. Plot the speedup you get by increasing the number of threads. Discuss your results. Include results for 1, 2, 4, 8, 16, 32 and 64 threads. Make sure to use a large enough number of intervals to obtain meaningful runtimes and accurate values for pi.
   b. Discuss the trends you are seeing in the graphs in part a.

PART D. (20 points)
Since the turn of the century, power has become a first-rate architectural design constraint in the design of microprocessors. Trevor Mudge describes these challenges in his 2001 paper that is provided on Canvas.
   a. Read Mudge's paper. Then for one of the approaches that he describes, find a current microprocessor available on the market that has adopted that technique. Try to provide as much detail as possible about the implementation.

b. Find performance and power rating information for one current embedded CPU and one current high-performance multi-core CPU. Discuss if it would make sense to combine many embedded CPUs to replace one high-performance CPU, while reducing the power budget. Back up your conclusions with quantitative arguments. Make sure to cite your sources.

Extra credit 1: (30 points to your quiz grade)

Graphics Processing Units are highly parallel single-instruction multi-threaded architectures that achieve impressive speedups for a range of applications. Using the NVIDIA A100 architecture, describe the memory hierarchy present on this device, and compare how it differs from the memory hierarchy of the AMD Ryzen CPU.