The University of Michigan
Department of EECS

EECS 573 – Microarchitecture

Prof. Todd Austin

Midterm Exam
November 28th, 2018
10:30am-11:50am
Open Book, Computer and Notes (no Internet or communication!)

Name: ____

Problem 1 – What this course is all about	_______________	/18
Problem 2 – Reliable system design	_______________	/10
Problem 3 – Secure and safe systems	_______________	/10
Problem 4 – Application-specific processors	_______________	/10
Problem 5 – Papers and presentations	_______________	/12
TOTAL	_______________	/60

Attempt to solve **all** the problems in this exam. Use your time wisely, and pay attention to the point distribution. Think before you plunge, if you are spending too much time on one part, move on to another one. All problems are divided in multiple parts; each part is independent from the others.

If you need more space to work out some the problems, use the backside of the exam sheets using a statement like "Go to back of page 6". When in doubt, state any assumption you make. Show all your work, you will get partial credit for partial answers. Good luck!

HONOR PLEDGE:
"I have neither given not received aid on this exam, nor have I concealed any violations of the Honor Code."
Signature:
____

1. What this course is all about – 15 points

[Hint: Please don't write an essay. You can answer with just a few sentences per question.]

1A. (3 points) The "bathtub curve" represents the probability of failure for transistors over the lifetime of a design. (i) Why is probability of failure high immediately after a transistor is manufactured? (ii) What do manufacturers do to prevent most early transistor failures from happening in the field? (iii) As transistors scale to smaller sizes, are they more reliable or less reliable? Explain your answer.

1B. (3 points) Simulation-based reliability studies can be much more computationally expensive than traditional simulation-based performance or power studies. a) List two reasons why this is the case. b) Some researchers have advocated that reliability studies be performed with faster FPGA-based designs. Give one downside of an FPGA-based study compared to a simulation-based reliability study.

1C. (3 points) Many are predicting (including Prof. Austin) that Moore's Law dimensional scaling will soon end -- please answer three questions about the end of Moore's Law. a) What two trends are forcing the end of Moore's Law? b) Why would we have to have to change the way we design systems if Moore's Law ended? c) What style of architectural design has the potential to overcome the end of Moore's Law and why is this so?

1D. (3 points) Please answer the following two questions about secure system design. a) What is a buffer overflow attack? b) Describe one way they can be stopped by the compiler. c) Describe one way they can be stopped with hardware support.

1E. (3 points) Control-flow attacks execute arbitrary code on the victim machine by hijacking the control flow of an applications. a) How does Control-Data Isolation (CDI) protect against control-flow attacks? b) What have the authors done to optimize the performance of their solution?

1F. (3 points) Large-scale datacenter services impose several stringent and competing requirements on the design of accelerators. a) Briefly describe two such competing requirements for Microsoft Catapult and Google TPU. b) Explain two reasons why Catapult used custom multicore processor for the Free Form Expressions (FFE) phase of the Bing's page raking case study. c) Describe why replacing TPU's DDR3 Weight Memory with GDDR5 improve its performance but increasing the clock rate has little effect.

2. Reliable System Design – 10 points

The following questions are related to reliable system design techniques from the papers we read this semester.

2A. (5 points) Prodromou *et al.*'s NoCAlert detects runtime faults within the interconnect in real-time. a) Briefly describe the invariances and the mechanism used by NoCAlert to detect a violation. b) What is one limitation of an invariance checking based fault detection mechanism?

2B. (5 points) Zhang *et al.*'s MELLOW WRITES work to extend the lifetime of non-volatile memory. a) Why does non-volatile memory have a limited lifetime? b) What is a MELLOW WRITE and how does it reduce wear-out?

3. Secure and Safe Design – 10 points

The following questions are related to secure and safe design techniques from the papers we read this semester.

3A. (5 points) Kim *et al.*'s Rowhammer attack causes a disturbance error on data in physically adjacent addresses by repeatedly accessing the same addresses in DRAM. a) Why is it possible to create disturbance errors? b) How does their proposed solution PARA mitigate this problem? c) Compare it with Aweke *et al.*'s solution ANVIL (discussed in the class lecture).

3B. (5 points) Evtyushkin *et al.*'s JUMP OVER ASLR attack uses side channels to de-randomize kernel ASLR. a) Describe the side channel they utilize and how it reveals information about the address of kernel code. b) How could the CPU designer prevent this attack? And indicate with a circle the impact of this protection? Performance: *higher* or *same* or *lower*. Power: *higher* or *same* or *lower*. Area: *higher* or *same* or *lower*.

4. Application-Specific Processor Design - 10 points

The following questions are related to application-specific design techniques from the papers we read this semester.

4A. (5 points) Aga *et al.*'s Compute Caches enables in-place computation for caches, which requires operand locality. a) Briefly describe the methods used by the authors to improve operand locality. b) What other techniques did the authors propose for operations that lack operand locality?

4B. 5 points) The CONVOLUTION ENGINE works to speed up the processing of convolution-like data-flow, such as those found in computational photography, image processing, and video processing applications. a) What is a 2D register file, and why does it work well for the algorithms running on the Convolution Engine? b) In the results, the authors compare designs based on "Ops/mm^2". What is the meaning of this metric, and why is it particularly relevant for applications that run on the Convolution Engine?

5. Papers and Presentations – 12 points

5A. (6 points) Possessing a well-developed skill for presentation will serve you in any walk of life. Please give three pitfalls to avoid when presenting technical research.

5B. (6 points) One of the goals of EECS 573 is to expose you to the research **process**. Please list three things you learned about conducting research? (Specifically, list here what you learned about *conducting* research, not presenting it or writing it up.)