# Midterm 2 Preparation
## CS 5513 Computer Architecture, Fall 2024

### Wei Wang

Computer Science
University of Texas at San Antonio

# Goals and Topics

► The goal is to help you systematically review the basic knowledge in Computer Architecture.

  – Only basic knowledge is tested.
  – No trick questions.

► Topics for this exam:

  – Pipelining
  – Speculative Execution (branch prediction)
  – Scoreboarding
  – Tomasulo's Algorithm
  – Reorder Buffer
  – ILP Limitation and Multithreading

# Location, Time and Logistics

► Nov 9th 2024, Saturday, in-class
  – Exam starts from 2nd session at 2:30pm
► Close-book, close-notes, close everything
► The exam have Conceptual questions and Problems.
  – The problems will be similar as those in assignments 3 and 4.

# Materials to Review

- ► Slides, all questions are from slides
- ► Assignments 3 and 4.
- ► You can check out the textbooks, but it is not required.
  - – There are some differences in the details between my slides and the textbooks. Please follow my slides in those cases.

# Pipelining

► Know the solutions to all types of hazards
  – Slide 37 has a summary of these solutions.
  – You also need to know whether a solution can properly solve the hazards or not. In particular,
    ► Why only branch prediction is the only practical solution to control hazard? Why other solutions do not work well?
    ► Does data bypassing/forwarding eliminate stalls? And why?

► The definition superscalar CPUs.

► At least one problem about pipelining. The problem is similar to those in assignment 3.
  – In particular, you need to know how the pipeline works when stalling is the only solution to the hazards.

# Branch Prediction

- ► The basic two-bit saturate counter for branch prediction.
    - – You need to memorize the state machine.
- ► Know the implementation of Branch Prediction Buffer and Branch Target Buffer.
    - – What components do these two buffers have?
    - – How does a branch locates its entries in these two buffers?
- ► Correlating branch prediction
    - – Why does correcting branch prediction work for some branches?
    - – Why does correlating branch prediction not work form some branches?
    - – The implementation of correlating branch prediction with Global Branch History Register (GBHR) and two-bit saturate counters.
- ► There will be one problem about branch prediction, similar to the last question of Assignment 3.

# OoO: Scoreboarding, Tomasulo's Algorithm and Reorder Buffer

► Be able to carry out the algrithms for OoO scheduling

► Why can OoO improve performance?

► What are the differences of scoreboards (SB), reservation station (RS) and reorder buffer (ROB)?

- Scoreboard is a buffer for storing the status of the instructions
- Reservation station is a buffer for holding the status and the source operands of the instructions.
- Reorder buffer is a buffer for holding the status and the results of the instructions.
- RS replaces SB, but ROB cannot replace RS.

► Common Data Bus (CDB)

- What is the use of CDB? (sending results and source FU)
- What is difference between CDB and the buses in pipelining? (CDB is a broadcasting bus)

# ILP Limitation and Multithreading

▶ If we have unlimited resources in the CPU, what limits ILP?
- RAW hazards

▶ What are the practical limitation on ILP?
- Renaming register count, ROB size (insn window size), Branch prediction accuracy, memory aliasing accuracy, memory latency, FU count and latency

▶ Why Multithreading (MT) is adopted?
- Most techniques to increase ILP failed
- Many application naturally has TLP in them.
- TLP is cheaper to exploit than ILP at the moment.

▶ Know the characteristics of the three types of MT: fine-grained, coarse-grained and SMT.