Final Project Presentations — Part 1

Rodolfo Azevedo

Institute of Computing, University of Campinas (UNICAMP), Brazil

rodolfo.azevedo@unicamp.br

http://www.ic.unicamp.br/~rodolfo/mo801

Goal of this class

Module 6, Class 2: teams present Project 3 results.

At the end of this class, you should be able to:

- Present your accelerator design choices, implementation, and measurement results in a structured 15-minute talk.
- Demonstrate KWS inference running on the Tang Nano 9K (live or recorded).
- Ask focused technical questions to evaluate peer projects during the Q&A segment.
- Articulate the relationship between your design decisions and the measured speedup.

Presentation format

- **Duration**: 15 minutes presentation + 5 minutes Q&A per pair.
- **Demo**: live on the Tang Nano 9K (preferred) or recorded video with narration.
- **Slides**: optional but encouraged for the measurement and analysis sections.

Order is randomized and announced at the start of class. All pairs must be present for all presentations — peer Q&A is part of the grade.

Required content

Every presentation must cover these four sections:

1. Design choices (~3 min)

- Interface: why memory-mapped? What alternatives did you consider?
- Datapath width: how many MAC units? Why that number?
- Memory: local BRAM buffer? How many entries? Why?

2. Measurement results (~5 min)

- Baseline cycles per layer (from M04A03 profiling).
- Accelerated cycles per layer.
- Speedup per layer and system-wide.
- Resource utilization: LUTs, FFs, BRAMs, DSPs (nextpnr report).

3. Roofline analysis (~4 min)

- Arithmetic intensity of your design (MACs per bus byte).
- Where does your design sit on the roofline?
- Is the speedup compute-bound or memory-bound? What would need to change to improve it?

Grading criteria

Criterion	Weight	Notes
Correctness	40%	Bit-exact match with C kernel on all test vectors
Speedup achieved	20%	Measured, not estimated; honest about overhead
Roofline analysis	20%	Correct arithmetic intensity calculation; correct bound identification
Demo	10%	Working on hardware or convincing simulation evidence
Q&A	10%	Both partners must be able to answer questions about any part

A working **1-wide scalar** accelerator with correct results and honest analysis scores well. A broken 4-wide accelerator with wrong output scores zero on correctness regardless of claimed speedup.

Q&A guidelines

The audience (instructor + peers) will ask questions from this pool:

- "Why did you choose memory-mapped over a custom instruction?"
- "What is the arithmetic intensity of your design, and how did you calculate it?"
- "If you doubled the number of MAC units, would the system speedup double? Why or why not?"
- "Show me the part of your SystemVerilog that implements the accumulator. Walk me through it."
- "Your speedup is 3.5×. What is the maximum achievable system speedup given Amdahl's Law?"
- "What was the hardest bug you encountered, and how did you find it?"

Both partners must be able to answer any of these. If one partner cannot answer, it signals that only one partner did the work.

Today's presentation schedule

(Filled in at the start of class from the randomized list)

Slot	Pair	Project title
1		
2		
3		
4		
5		

Remaining pairs present in the next class.