IMPERIAL

Profiling Assisted Memory Dependency Prediction Using Aliasing and Distance Information

Johan Jino[1], **Luke Panayi**[1, 2], **Paul Kelly**[1]

Background: The Memory Ordering Challenge

Out-of-Order (OoO) processors must schedule memory operations before their target addresses are known. Stalling for addresses is safe but slow.

A LOAD can't bypass a STORE it depends on, so **processors must predict dependencies** to maintain both speed and correctness.

The Challenge: False Dependencies

Memory Dependency Predictors (MDPs) suffer from **Index Collisions** when unrelated LOAD instructions map to the same entry in a prediction table.

These collisions create **False Dependencies**, causing unnecessary pipeline stalls and slowing down the processor.

Our Approach: Smarter Prediction Bypassing

We leverage **Predict No Dependency (PND) labels**, which allow certain LOADs to bypass the MDP and execute aggressively, avoiding potential false dependencies.

Key Finding: Distance is all you need

We show that using **Store Distance**, the number of store instructions between a dependent LOAD-STORE pair, is a far more effective for labeling PND LOADs.

Performance Gains from profile assisted PND labelling

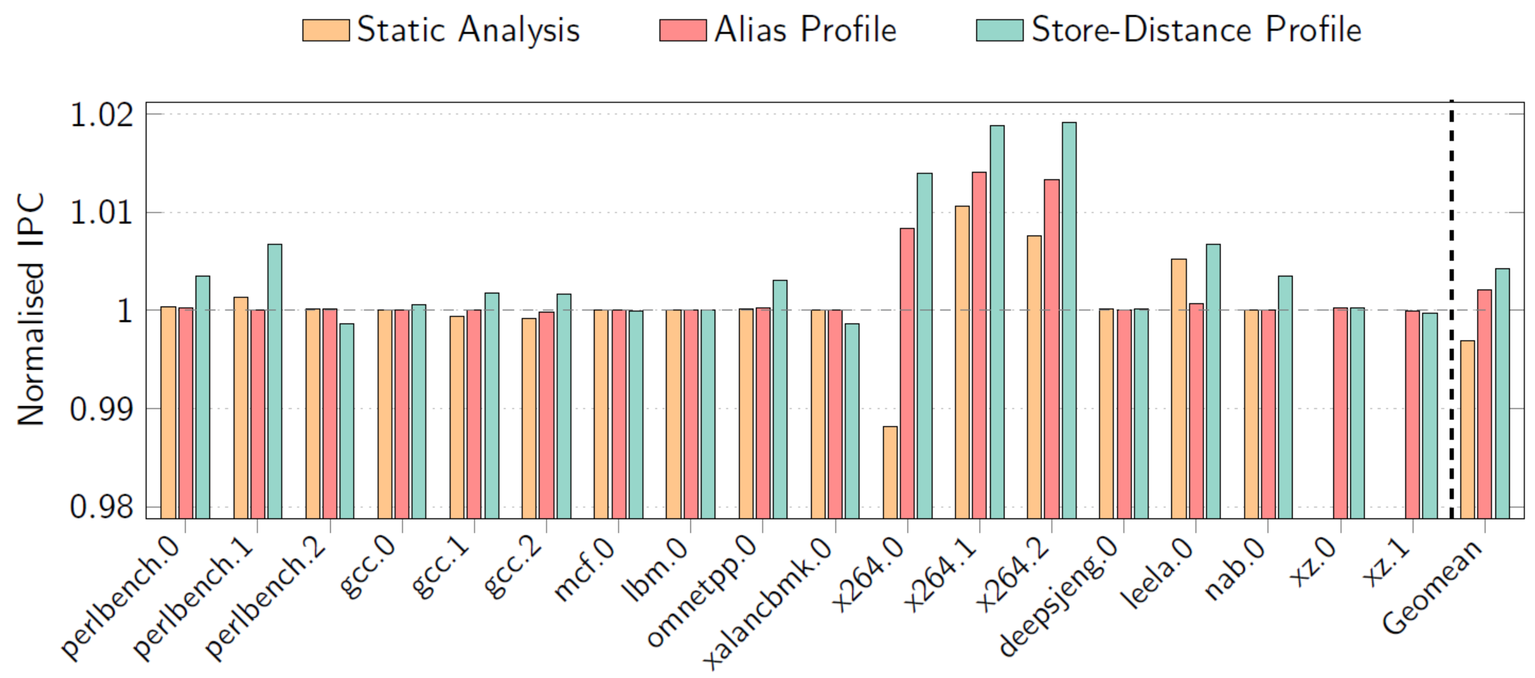


Figure 3: IPC of different labelling strategies normalised to base unlabelled binary (on A14). Higher is better. Numerical suffix of each benchmark indicate different input from reference input set. Bars above 1 mean speed-up compared to native CPU.

Reduced Dynamic Predictions through MDP bypass

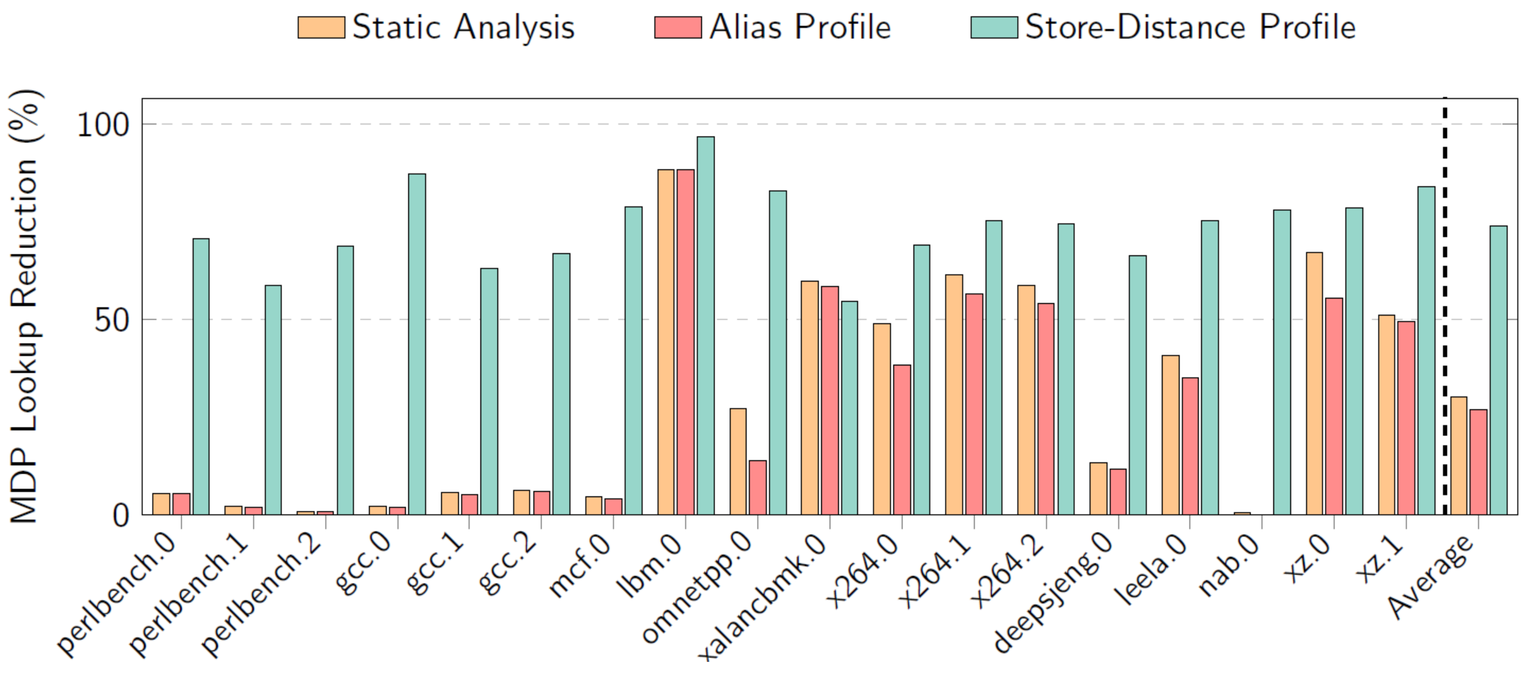


Figure 4: MDP Reduction Percentage using different labelling strategies (on A14). Higher is better. Numerical suffix of each benchmark indicate different input from reference input set. Directly corresponds to reduced dynamic energy consumption.

From Source to Binary: The Toolchain

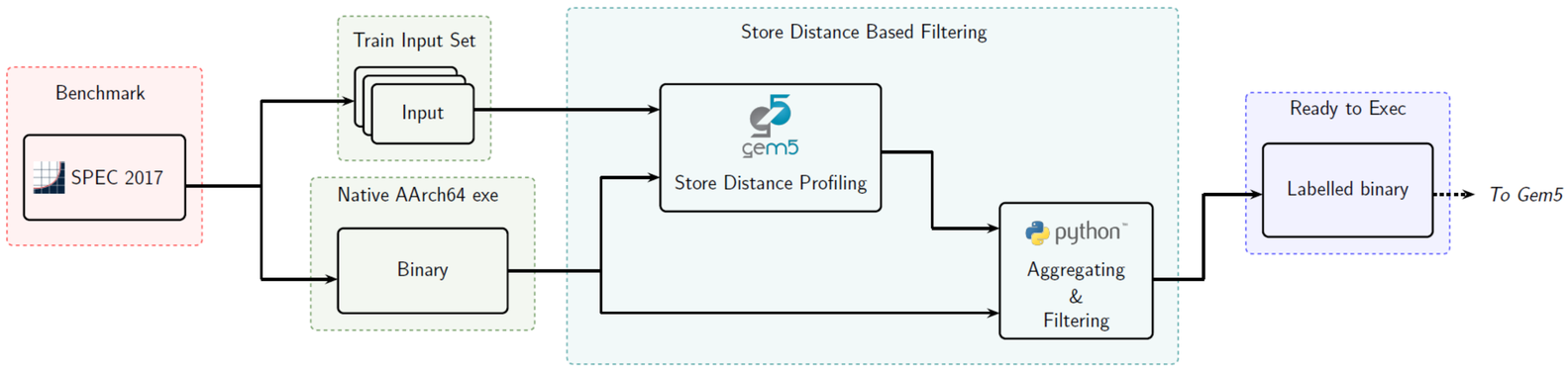


Figure 1: Toolchain for Store Distance based LOAD labelling

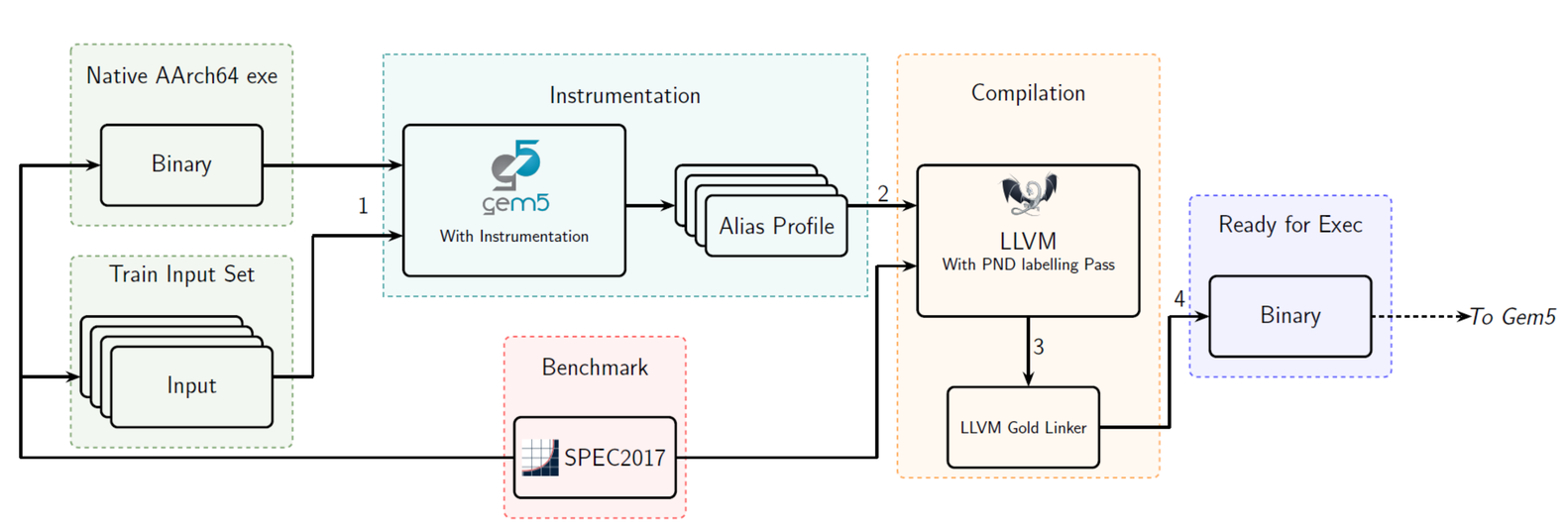


Figure 2: Toolchain for LOAD-STORE aliasing information based LOAD labelling

Comparison Across Different CPU Sizes

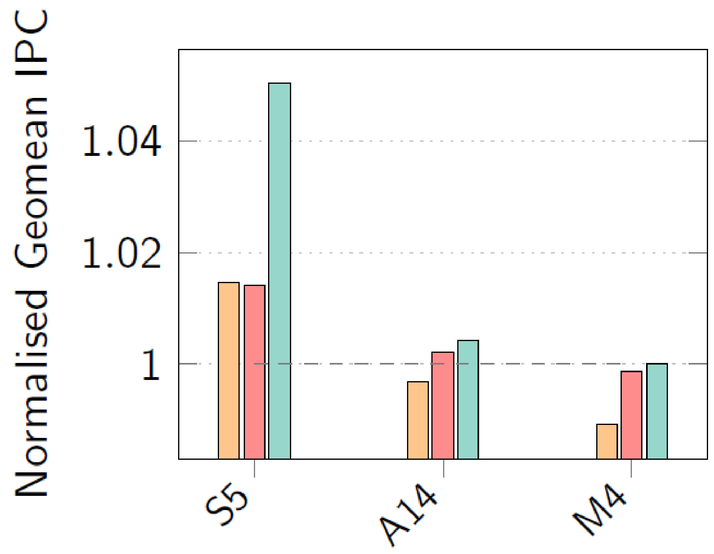


Figure 5: Geometric Mean IPC difference for each CPU model for each labelling strategy. Legend identical to Figure 3/4. CPU sizes are Gem5 models to resemble Apple's wearable, mobile and workstation class CPUs respectively. Note, S5 model does not have an MDP.

Results in Figure 5 show diminishing returns as CPU size is scaled. This is expected since Index Collisions reduce with increased CPU sizes.

However, future work could show how much MDP size can be reduced to still maintain same performance through PND labels.

Affiliations
[1] Imperial College London, UK
[2] Huawei R&D UK Ltd