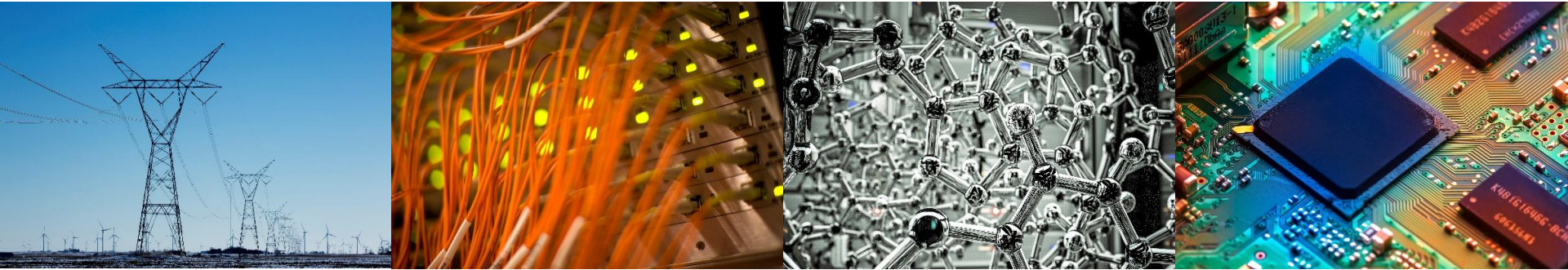


ECE 411 MP4 Presentation

Group PIPT



I ILLINOIS

Electrical & Computer Engineering

GRAINGER COLLEGE OF ENGINEERING

Overview

- Advanced Features
 - Parameterized Four-Way Cache
 - Strided Prefetch
 - Branch Prediction
- Quantitative Evaluation
 - Timing
 - Area
- Improvements

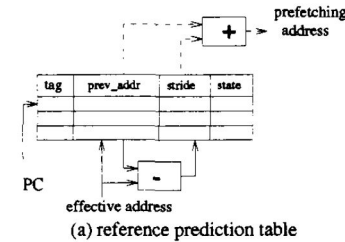
Parameterized Four-Way Cache

- Modified MP3 cache to allow for 1-cycle hits
 - Previously, it only checks for hits on a memory read or write
 - Now, constantly checks for hits
- Modified the cache from two-way to four-way
 - Changed the pseudo-LRU from 1-bit to 3-bits
- Parameterized the cache to allow for changing the number of sets

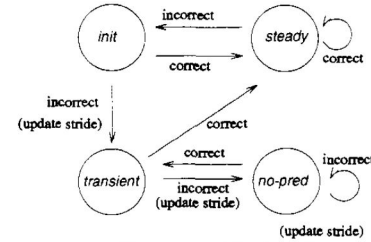
Set Index Bits	Comp1	Comp2	Comp3
4 bits	18.250 seconds	40.200 seconds	20.810 seconds
8 bits	18.020 seconds	40.670 seconds	20.090 seconds

Strided Prefetch

- Designed strided prefetcher for data addresses
- Module sits in parallel with D-Cache in mem_access stage
- Used Reference Prediction Table (RPT) struct to hold:
 - Tag: Instead of the standard instruction address as tag, used 8-bit index of the RPT table array
 - Address: Data Address
 - Stride: $\text{curr_data_addr} - \text{prev_data_addr}$
 - State: {Initial, Transient, Steady, No Predict}
- Prefetch initiated whenever memory is free, a load instruction in mem_access, and in steady state, i.e. two consecutive equal strides



(a) reference prediction table



(b) state transition by PC

Source: Tien-Fu Chen and Jean-Loup Baer, "Effective hardware-based data prefetching for high-performance processors," in IEEE Transactions on Computers, vol. 44, no. 5, pp. 609-623, May 1995, doi: 10.1109/12.381947.

Branch prediction

- Local, global and tournament predictors
- All use 2 saturating bits (SNT, WNT, WT, ST), initially set to weakly not taken
- Tournament predictor saturating bits (SG, WG, WL, SL) initially set to weakly local
- 90% overall accuracy
- Parameterized set number (↑ sets, ↑ accuracy and area)
 - Using 3 bits for indexing: 25% accuracy
 - Using 8 bits for indexing: 90% accuracy

Quantitative Evaluation

	CP1	CP2	CP3	Comp1	Comp2	Comp3
Baseline	0.540 seconds	0.530 seconds	20.500 seconds			
Final Design	0.460 seconds	0.440 seconds	19.840 seconds	18.250 seconds	40.200 seconds	20.810 seconds
Speed-up	1.17391304	1.20454545	1.03326613			

Quantitative Evaluation

	Combinational Area	Non-combinational Area
Baseline	32852	28593
Final Design	317690	313899

- Main increase of area:
 - Increased number of sets/ways in cache
 - RPT table
 - History tables in branch predictor

Improvements

- Setting up RVFI monitor prior to debugging
- Implement sequential instruction prefetching
- Implement pipeline cache to reduce the performance hit of 1-cycle hit cache

Q&A