

Lab 4 Report

Multi-threaded Machine

Name : Raghavendra Vinayak Belapure

T-Square Account Name : rbelapure3

1.0 Benchmarks for Weighted Speedup

First, we will formulate three benchmarks by running trace 1 and 2 together, trace 1 and 3 together, and finally, trace 2 and 3 together. Weighted speedup will be calculated by using the relation –

$$\text{Weighted Speedup} = \sum_{i=0}^n \left\{ \frac{IPC_withSMT_i}{IPC_withoutSMT_i} \right\} / n$$

The default values used to run the simulator are as follows –

MSHR Size : 4	Cache Size : 512KB	Type : 4 way set associative
D-cache latency : 5 cycles	DRAM Row Buffer Hit latency : 100	DRAM Row Buffer Miss latency : 200
DRAM_BANK_NUM : 4	DRAM_PAGE_SIZE : 2KB	Block size = 64
TLB size = 4	Branch predictor : G-share	History Length : 12

1.1 Benchmark 1: Trace 1 and 2

All the traces were run for maximum of 10 million instructions only. The results are as follows:

Trace 1 alone	Trace 2 alone
Total instruction: 7861921 Total cycles: 24092611 Total IPC: 0.326321 Total D-cache miss: 7192 Total D-cache hit: 2865329 Total data hazard: 1536278 Total control hazard : 1320811 Total DRAM ROW BUFFER Hit: 2089 Total DRAM ROW BUFFER Miss: 940 Total Store-load forwarding: 28 Total Branch Predictor Mispredictions: 50706 Total Branch Predictor OK predictions: 1270105 Total DTLB Hit: 2563845 Total DTLB Miss: 154338 THREAD instruction: 7861921 Thread id: 0 THREAD IPC: 0.326321 THREAD D-cache miss: 7192 Thread id: 0 THREAD D-cache hit: 2865329 Thread id: 0 THREAD data hazard: 1536278 Thread id: 0 THREAD control hazard : 1320811 Thread id: 0 THREAD Store-load forwarding: 28 Thread id: 0 THREAD Branch Predictor Mispredictions: 50706 Thread id: 0 THREAD Branch Predictor OK predictions: 1270105 Thread id: 0 THREAD DTLB Hit: 2563845 Thread id: 0 THREAD DTLB Miss: 154338 Thread id: 0	Total instruction: 3354758 Total cycles: 9950141 Total IPC: 0.337157 Total D-cache miss: 52464 Total D-cache hit: 974961 Total data hazard: 634663 Total control hazard : 656859 Total DRAM ROW BUFFER Hit: 6107 Total DRAM ROW BUFFER Miss: 1110 Total Store-load forwarding: 25 Total Branch Predictor Mispredictions: 36617 Total Branch Predictor OK predictions: 620242 Total DTLB Hit: 768061 Total DTLB Miss: 129682 THREAD instruction: 3354758 Thread id: 0 THREAD IPC: 0.337157 THREAD D-cache miss: 52464 Thread id: 0 THREAD D-cache hit: 974961 Thread id: 0 THREAD data hazard: 634663 Thread id: 0 THREAD control hazard : 656859 Thread id: 0 THREAD Store-load forwarding: 25 Thread id: 0 THREAD Branch Predictor Mispredictions: 36617 Thread id: 0 THREAD Branch Predictor OK predictions: 620242 Thread id: 0 THREAD DTLB Hit: 768061 Thread id: 0 THREAD DTLB Miss: 129682 Thread id: 0

Trace 1 and 2 with SMT	Speedup calculation
Total instruction: 10000000 Total cycles: 30056336 Total IPC: 0.332709 Total D-cache miss: 38067 Total D-cache hit: 3605488 Total data hazard: 945244 Total control hazard : 1772578 Total DRAM ROW BUFFER Hit: 7296 Total DRAM ROW BUFFER Miss: 3030 Total Store-load forwarding: 34 Total Branch Predictor Mispredictions: 89107 Total Branch Predictor OK predictions: 1683471 Total DTLB Hit: 2730118 Total DTLB Miss: 456687 THREAD instruction: 6645242 Thread id: 0 THREAD IPC: 0.221093 THREAD D-cache miss: 6475 Thread id: 0 THREAD D-cache hit: 2532093 Thread id: 0 THREAD data hazard: 745750 Thread id: 0 THREAD control hazard : 1115719 Thread id: 0 THREAD Store-load forwarding: 18 Thread id: 0 THREAD Branch Predictor Mispredictions: 48352 Thread id: 0 THREAD Branch Predictor OK predictions: 1067367 Thread id: 0 THREAD DTLB Hit: 2039554 Thread id: 0 THREAD DTLB Miss: 249507 Thread id: 0 THREAD instruction: 3354758 Thread id: 1 THREAD IPC: 0.111616 THREAD D-cache miss: 31592 Thread id: 1 THREAD D-cache hit: 1073395 Thread id: 1 THREAD data hazard: 199494 Thread id: 1 THREAD control hazard : 656859 Thread id: 1 THREAD Store-load forwarding: 16 Thread id: 1 THREAD Branch Predictor Mispredictions: 40755 Thread id: 1 THREAD Branch Predictor OK predictions: 616104 Thread id: 1 THREAD DTLB Hit: 690564 Thread id: 1 THREAD DTLB Miss: 207180 Thread id: 1	Thus, weighted speedup calculated is $WS_{12} = \left(\frac{0.221093}{0.326321} + \frac{0.111616}{0.337157} \right) / 2$ Therefore, $WS_{12} = 0.504291383$

Similarly, we calculate the weighted speedup for other benchmarks as well with 10M instructions.

1.2 Benchmark 2: Trace 1 and 3

Trace 1 alone	Trace 3 alone
Total instruction: 7861921 Total cycles: 24092611 Total IPC: 0.326321 Total D-cache miss: 7192 Total D-cache hit: 2865329 Total data hazard: 1536278 Total control hazard : 1320811 Total DRAM ROW BUFFER Hit: 2089 Total DRAM ROW BUFFER Miss: 940 Total Store-load forwarding: 28 Total Branch Predictor Mispredictions: 50706 Total Branch Predictor OK predictions: 1270105 Total DTLB Hit: 2563845 Total DTLB Miss: 154338 THREAD instruction: 7861921 Thread id: 0 THREAD IPC: 0.326321 THREAD D-cache miss: 7192 Thread id: 0 THREAD D-cache hit: 2865329 Thread id: 0 THREAD data hazard: 1536278 Thread id: 0 THREAD control hazard : 1320811 Thread id: 0	Total instruction: 4951261 Total cycles: 14333556 Total IPC: 0.345431 Total D-cache miss: 52546 Total D-cache hit: 1430577 Total data hazard: 672352 Total control hazard : 924487 Total DRAM ROW BUFFER Hit: 6119 Total DRAM ROW BUFFER Miss: 1106 Total Store-load forwarding: 26 Total Branch Predictor Mispredictions: 51021 Total Branch Predictor OK predictions: 873466 Total DTLB Hit: 1132355 Total DTLB Miss: 175336 THREAD instruction: 4951261 Thread id: 0 THREAD IPC: 0.345431 THREAD D-cache miss: 52546 Thread id: 0 THREAD D-cache hit: 1430577 Thread id: 0 THREAD data hazard: 672352 Thread id: 0 THREAD control hazard : 924487 Thread id: 0

THREAD Store-load forwarding: 28 Thread id: 0 THREAD Branch Predictor Mispredictions: 50706 Thread id: 0 THREAD Branch Predictor OK predictions: 1270105 Thread id: 0 THREAD DTLB Hit: 2563845 Thread id: 0 THREAD DTLB Miss: 154338 Thread id: 0	THREAD Store-load forwarding: 26 Thread id: 0 THREAD Branch Predictor Mispredictions: 51021 Thread id: 0 THREAD Branch Predictor OK predictions: 873466 Thread id: 0 THREAD DTLB Hit: 1132355 Thread id: 0 THREAD DTLB Miss: 175336 Thread id: 0
---	--

Trace 1 and 3 with SMT	Speedup calculation
Total instruction: 10000000 Total cycles: 29858268 Total IPC: 0.334916 Total D-cache miss: 38725 Total D-cache hit: 3600416 Total data hazard: 433732 Total control hazard : 1757504 Total DRAM ROW BUFFER Hit: 7362 Total DRAM ROW BUFFER Miss: 2967 Total Store-load forwarding: 34 Total Branch Predictor Mispredictions: 98810 Total Branch Predictor OK predictions: 1658694 Total DTLB Hit: 2457041 Total DTLB Miss: 590871 THREAD instruction: 5048739 Thread id: 0 THREAD IPC: 0.16909 THREAD D-cache miss: 6830 Thread id: 0 THREAD D-cache hit: 2011838 Thread id: 0 THREAD data hazard: 238373 Thread id: 0 THREAD control hazard : 833017 Thread id: 0 THREAD Store-load forwarding: 19 Thread id: 0 THREAD Branch Predictor Mispredictions: 40793 Thread id: 0 THREAD Branch Predictor OK predictions: 792224 Thread id: 0 THREAD DTLB Hit: 1462032 Thread id: 0 THREAD DTLB Miss: 278188 Thread id: 0 THREAD instruction: 4951261 Thread id: 1 THREAD IPC: 0.165825 THREAD D-cache miss: 31895 Thread id: 1 THREAD D-cache hit: 1588578 Thread id: 1 THREAD data hazard: 195359 Thread id: 1 THREAD control hazard : 924487 Thread id: 1 THREAD Store-load forwarding: 15 Thread id: 1 THREAD Branch Predictor Mispredictions: 58017 Thread id: 1 THREAD Branch Predictor OK predictions: 866470 Thread id: 1 THREAD DTLB Hit: 995009 Thread id: 1 THREAD DTLB Miss: 312683 Thread id: 1	Thus, weighted speedup calculated is $WS_{13} = \left(\frac{0.16909}{0.326321} + \frac{0.165825}{0.345431} \right) / 2$ Therefore, $WS_{13} = 0.499111607$

1.3 Benchmark 3: Trace 2 and 3

Trace 2 alone	Trace 3 alone
Total instruction: 3354758 Total cycles: 9950141 Total IPC: 0.337157 Total D-cache miss: 52464 Total D-cache hit: 974961 Total data hazard: 634663 Total control hazard : 656859 Total DRAM ROW BUFFER Hit: 6107 Total DRAM ROW BUFFER Miss: 1110 Total Store-load forwarding: 25 Total Branch Predictor Mispredictions: 36617 Total Branch Predictor OK predictions: 620242 Total DTLB Hit: 768061 Total DTLB Miss: 129682	Total instruction: 4951261 Total cycles: 14333556 Total IPC: 0.345431 Total D-cache miss: 52546 Total D-cache hit: 1430577 Total data hazard: 672352 Total control hazard : 924487 Total DRAM ROW BUFFER Hit: 6119 Total DRAM ROW BUFFER Miss: 1106 Total Store-load forwarding: 26 Total Branch Predictor Mispredictions: 51021 Total Branch Predictor OK predictions: 873466 Total DTLB Hit: 1132355 Total DTLB Miss: 175336

THREAD instruction: 3354758 Thread id: 0 THREAD IPC: 0.337157 THREAD D-cache miss: 52464 Thread id: 0 THREAD D-cache hit: 974961 Thread id: 0 THREAD data hazard: 634663 Thread id: 0 THREAD control hazard : 656859 Thread id: 0 THREAD Store-load forwarding: 25 Thread id: 0 THREAD Branch Predictor Mispredictions: 36617 Thread id: 0 THREAD Branch Predictor OK predictions: 620242 Thread id: 0 THREAD DTLB Hit: 768061 Thread id: 0 THREAD DTLB Miss: 129682 Thread id: 0	THREAD instruction: 4951261 Thread id: 0 THREAD IPC: 0.345431 THREAD D-cache miss: 52546 Thread id: 0 THREAD D-cache hit: 1430577 Thread id: 0 THREAD data hazard: 672352 Thread id: 0 THREAD control hazard : 924487 Thread id: 0 THREAD Store-load forwarding: 26 Thread id: 0 THREAD Branch Predictor Mispredictions: 51021 Thread id: 0 THREAD Branch Predictor OK predictions: 873466 Thread id: 0 THREAD DTLB Hit: 1132355 Thread id: 0 THREAD DTLB Miss: 175336 Thread id: 0
---	---

Trace 2 and 3 with SMT	Speedup calculation
Total instruction: 8306019 Total cycles: 24026310 Total IPC: 0.345705 Total D-cache miss: 132589 Total D-cache hit: 2629110 Total data hazard: 159334 Total control hazard : 1581346 Total DRAM ROW BUFFER Hit: 2473 Total DRAM ROW BUFFER Miss: 13158 Total Store-load forwarding: 35 Total Branch Predictor Mispredictions: 86884 Total Branch Predictor OK predictions: 1494462 Total DTLB Hit: 1649972 Total DTLB Miss: 555476 THREAD instruction: 3354758 Thread id: 0 THREAD IPC: 0.139629 THREAD D-cache miss: 70425 Thread id: 0 THREAD D-cache hit: 1076156 Thread id: 0 THREAD data hazard: 68843 Thread id: 0 THREAD control hazard : 656859 Thread id: 0 THREAD Store-load forwarding: 15 Thread id: 0 THREAD Branch Predictor Mispredictions: 37340 Thread id: 0 THREAD Branch Predictor OK predictions: 619519 Thread id: 0 THREAD DTLB Hit: 649228 Thread id: 0 THREAD DTLB Miss: 248521 Thread id: 0 THREAD instruction: 4951261 Thread id: 1 THREAD IPC: 0.206077 THREAD D-cache miss: 62164 Thread id: 1 THREAD D-cache hit: 1552954 Thread id: 1 THREAD data hazard: 90491 Thread id: 1 THREAD control hazard : 924487 Thread id: 1 THREAD Store-load forwarding: 20 Thread id: 1 THREAD Branch Predictor Mispredictions: 49544 Thread id: 1 THREAD Branch Predictor OK predictions: 874943 Thread id: 1 THREAD DTLB Hit: 1000744 Thread id: 1 THREAD DTLB Miss: 306955 Thread id: 1	Thus, weighted speedup calculated is $WS_{23} = \left(\frac{0.139629}{0.337157} + \frac{0.206077}{0.345431} \right) / 2$ Therefore, $WS_{23} = 0.505357893$

2.0 Performance Implications of Memory Intensive Traces

First, we will classify the traces into Memory Intensive and Non-memory Intensive traces. A trace can be said to be memory intensive if it spends more time in memory than CPU. Time spent by the application in accessing memory is proportional to total number of DRAM accesses made by the trace. We will calculate number of cycles spent in memory as follows –

Number of cycles spent in memory

$$= (\text{\#of Row Buffer Hits} * \text{ROW_HIT_LATENCY}) \\ + (\text{\#of Row Buffer Misses} * \text{ROW_MISS_LATENCY})$$

	Trace 1	Trace 2	Trace 3
DRAM row hits	2089	6107	6119
DRAM row misses	940	1110	1106
Total DRAM accesses cycles	396900	832700	833100
Total cycles	7861921	3354758	4951261
% Time spent in memory	5.04838449 %	24.82146 %	16.82602 %
Memory Intensive?	No	Yes	Yes

Thus, it can be seen that Trace 1 spends only 5% time in memory, whereas, Traces 2 and 3 spend 25% and 17% time in memory respectively. Thus, Trace 1 is non-memory-intensive trace whereas Traces 2 and 3 are memory intensive traces.

Performance Implications

When a memory intensive trace is run with a non-memory intensive one, we can expect that non-memory intensive trace will progress more than the memory intensive trace. This is due to fact that when memory intensive trace is blocked on memory access, the execution unit can execute instructions from the other trace. This can be seen from the fact that Thread IPC for both Trace 2 and 3 is less when they are run with Trace 1, but it improves when they are run together.

Thus, it can be inferred that when we need to maximize the throughput, a proper mix of memory and non-memory intensive applications should be run with MT. On the other hand, when we need to reduce the latency of memory intensive applications, they should be run with other memory intensive applications.

3.0 Branch prediction with MT

We will calculate branch prediction accuracy for all the three traces running alone and with MT enabled.

	Correct predictions	Mis-predictions	Branch pred accuracy (%)
Trace 1 alone	1270105	50706	96.160995
Trace 2 alone	620242	36617	94.4254399
Trace 3 alone	873466	51021	94.4811555
Benchmark 1 : Total	1683471	89107	94.973028
Benchmark 1 : Trace 1	1067367	48352	95.6662923
Benchmark 1 : Trace 2	616104	40755	93.7954721
Benchmark 2 : Total	1658694	98810	94.3778222
Benchmark 2 : Trace 1	792224	40793	95.1029811
Benchmark 2 : Trace 3	866470	58017	93.7244115
Benchmark 3 : Total	1494462	86884	94.5056933
Benchmark 3 : Trace 2	619519	37340	94.3153706
Benchmark 3 : Trace 3	874943	49544	94.6409198

It can be seen that branch prediction accuracy decreases when MT is enabled. This is due to the fact that the PHT table in our branch predictor is shared between all the threads. As a result, the states of 2bC in PHT table get affected due to interference of other threads, causing the branch predictor accuracy to decrease.

An interesting result is seen in Benchmark 3 where the branch prediction accuracy of trace 3 increases when multi-threading is enabled. Thus, trace 2 is affecting the branch predictor in positive way for trace 3. It can be inferred that we can not actually predict if the branch predictor accuracy will improve or degrade by enabling MT.

4.0 Reducing Simulation Errors

To reduce the simulation errors, following measures could be taken –

1. We are running the simulator till the longest instruction trace exits. This leads to problems as the shorter trace could get completed very quickly and yet we are considering total cycles of larger trace to calculate thread IPC of both the threads. We should be keeping separate cycle count per thread to get around this problem.
2. The above problem can also be solved by stopping the simulator when any one of the running threads completes the execution.
3. Another problem is that the simulator that we are running is not very close to the actual hardware. Firstly, the TLB is blocking. Secondly, when memory stage is blocked due to TLB miss or the op is waiting to access d-cache for number of cycles that equal d-cache latency, the execution stage is also blocked. As we have multiple threads available, the execution stage can execute the instructions from other threads while the previous thread is blocked in the memory.

These steps could be followed to minimize the simulation errors in the benchmarks.