Cache Performance Analysis using gem5

Harsh Vishwakarma : 2022205

August 2025

Abstract

This report presents a cache performance analysis using gem5 simulations across varying L2 and L3 cache configurations. Here I evaluate miss rates and simulation ticks for both TimingSimpleCPU and O3CPU models, and visualize trends to identify optimal architectural trade-offs.

1 Introduction

Efficient cache design is critical for modern processor performance. This study uses gem5 to simulate different L2 and L3 cache configurations under two CPU models: TimingSimpleCPU and O3CPU. The goal is to understand how cache parameters affect miss rates and execution time.

2 Methodology

Simulations were run using the gem5 framework with the `qsort_small.elf` benchmark. Cache sweeps varied size and associativity for both L2 and L3 caches. Performance metrics were extracted from `stats.txt` and compiled into CSV summaries.

3 Configuration Overview

Table 1: Sample Cache Configurations

CPU Model	**Cache Setup**
TimingSimpleCPU	L2: 128kB, 4-way; L3: 1MB, 16-way
O3CPU	L2: 256kB, 8-way; L3: 2MB, 8-way

4 Simulation Results

4.1 Cache Sweep Summary

Table 2: Cache Sweep Results for TimingSimpleCPU and O3CPU

CPU Type	L2 Size	L2 Assoc	L3 Size	L3 Assoc	SimTicks	L2 Miss Rate	L3 Miss Rate
timing	128kB	4	1MB	8	58563352000	0.645072	0.437514
timing	128kB	4	1MB	16	58578874000	0.645072	0.441642
timing	128kB	4	2MB	8	57286842000	0.645072	0.231503
timing	128kB	4	2MB	16	57286842000	0.645072	0.231503
timing	128kB	8	1MB	8	58563352000	0.645904	0.436950
timing	128kB	8	1MB	16	58580268000	0.645904	0.441074
timing	128kB	8	2MB	8	57294368000	0.645904	0.231205
timing	128kB	8	2MB	16	57294368000	0.645904	0.231205
timing	256kB	4	1MB	8	57846969000	0.523496	0.539135
timing	256kB	4	1MB	16	57866325000	0.523496	0.544124
timing	256kB	4	2MB	8	56570893000	0.523496	0.285267
timing	256kB	4	2MB	16	56570893000	0.523496	0.285267
timing	256kB	8	1MB	8	57862503000	0.526866	0.535687
timing	256kB	8	1MB	16	57879390000	0.526866	0.540658
timing	256kB	8	2MB	8	56590279000	0.526866	0.283442
timing	256kB	8	2MB	16	56590279000	0.526866	0.283442
timing	512kB	4	1MB	8	57099506000	0.397425	0.702125
timing	512kB	4	1MB	16	57146303000	0.397425	0.716140
timing	512kB	4	2MB	8	55840426000	0.397425	0.375759
timing	512kB	4	2MB	16	55840426000	0.397425	0.375759
timing	512kB	8	1MB	8	57077362000	0.397940	0.696557
timing	512kB	8	1MB	16	57139770000	0.397940	0.712532
timing	512kB	8	2MB	8	55845200000	0.397940	0.375273
timing	512kB	8	2MB	16	55845200000	0.397940	0.375273
o3	128kB	4	1MB	8	10569818000	0.648201	0.433015
o3	128kB	4	1MB	16	10566328000	0.648207	0.437074
o3	128kB	4	2MB	8	9798428000	0.648061	0.227034
o3	128kB	4	2MB	16	9798428000	0.648061	0.227034
o3	128kB	8	1MB	8	10572681000	0.649276	0.432251
o3	128kB	8	1MB	16	10570625000	0.649245	0.436393
o3	128kB	8	2MB	8	9803998000	0.649126	0.226637
o3	128kB	8	2MB	16	9803998000	0.649126	0.226637
o3	256kB	4	1MB	8	10291128000	0.525640	0.534562
o3	256kB	4	1MB	16	10291521000	0.525615	0.539358
o3	256kB	4	2MB	8	9524179000	0.525232	0.280678
o3	256kB	4	2MB	16	9524179000	0.525232	0.280678
o3	256kB	8	1MB	8	10295975000	0.528942	0.531102
o3	256kB	8	1MB	16	10299307000	0.529014	0.535820
o3	256kB	8	2MB	8	9536305000	0.528583	0.278905
o3	256kB	8	2MB	16	9536305000	0.528583	0.278905
o3	512kB	4	1MB	8	10051328000	0.399798	0.696029
o3	512kB	4	1MB	16	10074639000	0.399783	0.708776
o3	512kB	4	2MB	8	9306164000	0.399415	0.370458
o3	512kB	4	2MB	16	9306164000	0.399415	0.370458
o3	512kB	8	1MB	8	10025415000	0.400299	0.689975
o3	512kB	8	1MB	16	10060609000	0.400298	0.705770
o3	512kB	8	2MB	8	9308646000	0.400032	0.369866
o3	512kB	8	2MB	16	9308646000	0.400032	0.369866

4.2 Miss Rate Plots

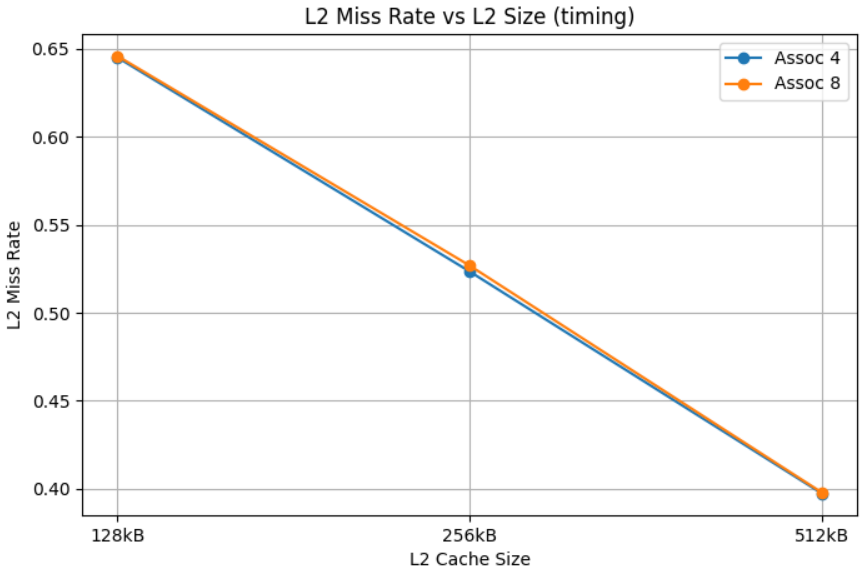

Figure 1: L2 Miss Rate - TimingSimpleCPU

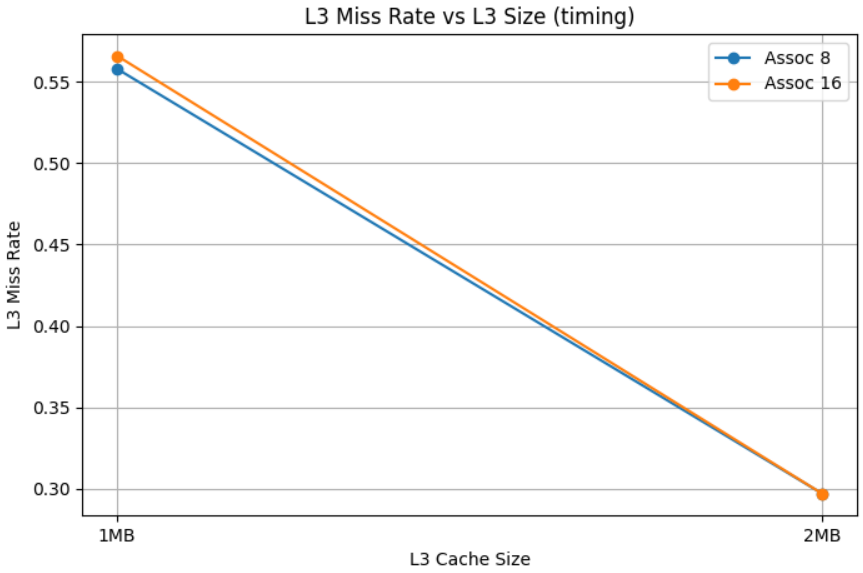

Figure 2: L3 Miss Rate - TimingSimpleCPU

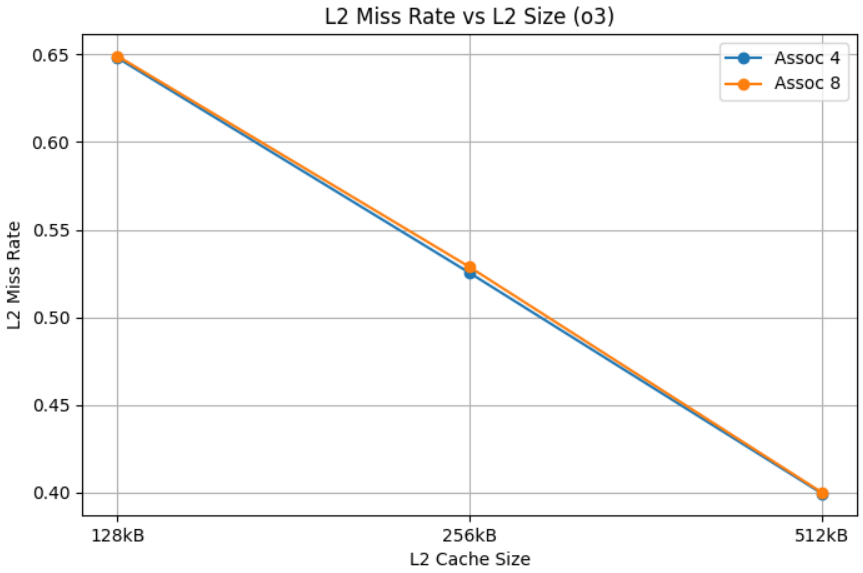

Figure 3: L2 Miss Rate - O3CPU

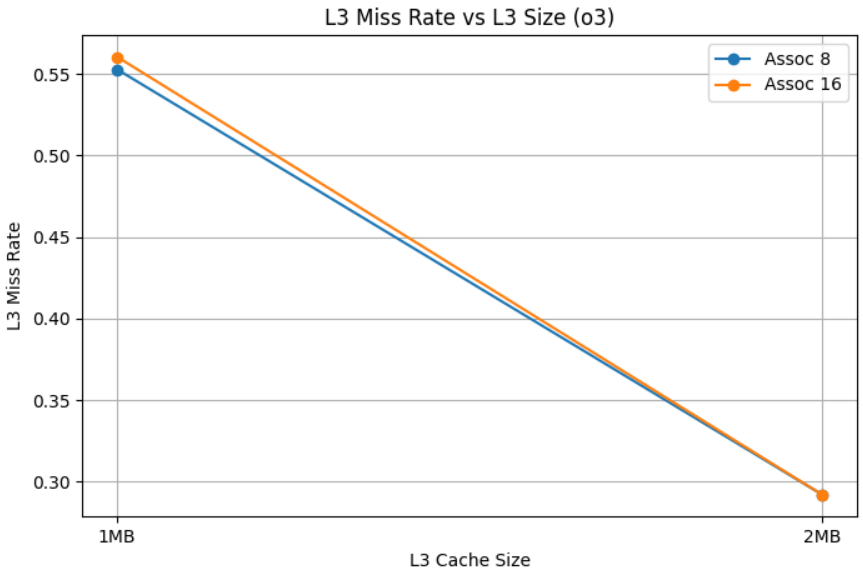

Figure 4: L3 Miss Rate - O3CPU

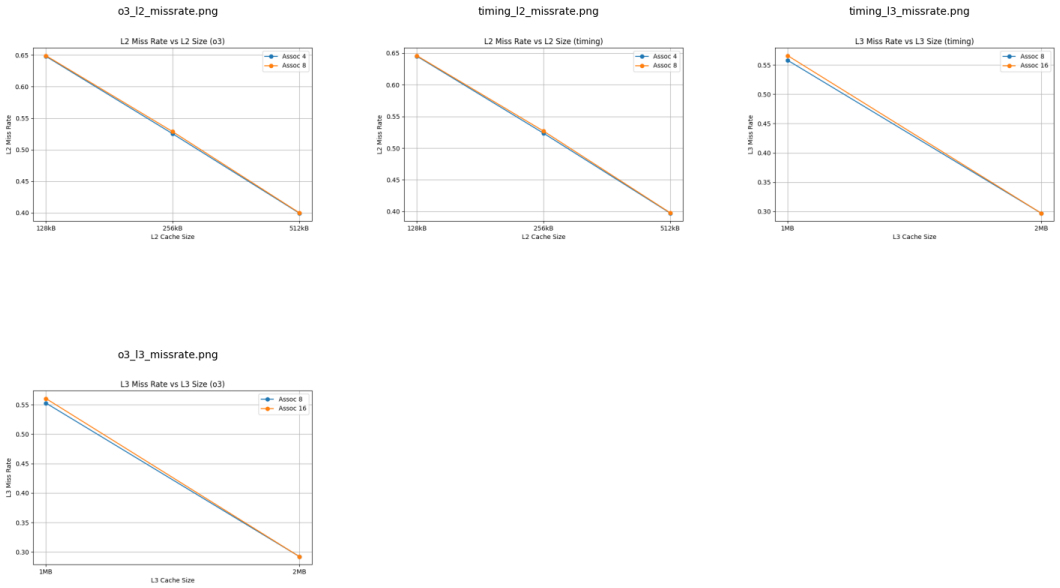

Figure 5: Combined Miss Rate Grid

5 Simulation Ticks Summary

Table below reports the observed simulation ticks for each CPU model across selected cache configurations. O3CPU consistently outperforms TimingSimpleCPU due to its out-of-order execution capabilities.

Cache Config (L2/L3)	**TimingSimpleCPU**	**O3CPU**
128kB / 1MB	58.56B	10.57B
256kB / 2MB	56.57B	9.53B
512kB / 2MB	55.84B	9.30B

Table 3: Simulation ticks for qsort_small.elf under different CPU models

6 Analysis

The TimingSimpleCPU shows higher miss rates and longer simulation ticks compared to O3CPU, especially under smaller cache configurations. Increasing associativity generally improves miss rates, but with diminishing returns beyond 8-way. L3 cache size has a more pronounced effect on O3CPU performance.

7 Conclusion

Cache configuration significantly impacts performance in gem5 simulations.

- Larger caches reduce miss rates (expecially for L3)
- Associativity helps (Only up to a point like till 8)
- O3CPU is more efficient
- TimingSimpleCPU is useful for baseline analysis