Performance Characterization of the PentiumPro Processor

Dileep Bhandarkar and Jason Ding

HPCA 3
1997

Jason Ding
Senior Director, Performance Engineering, Cloud Computing Apps

PentiumPro

- 1995
- Superscalar processors starting to emerge
- In order superscalar processor
- Register renaming
- ROB and Reservation Stations
- Memory order buffer

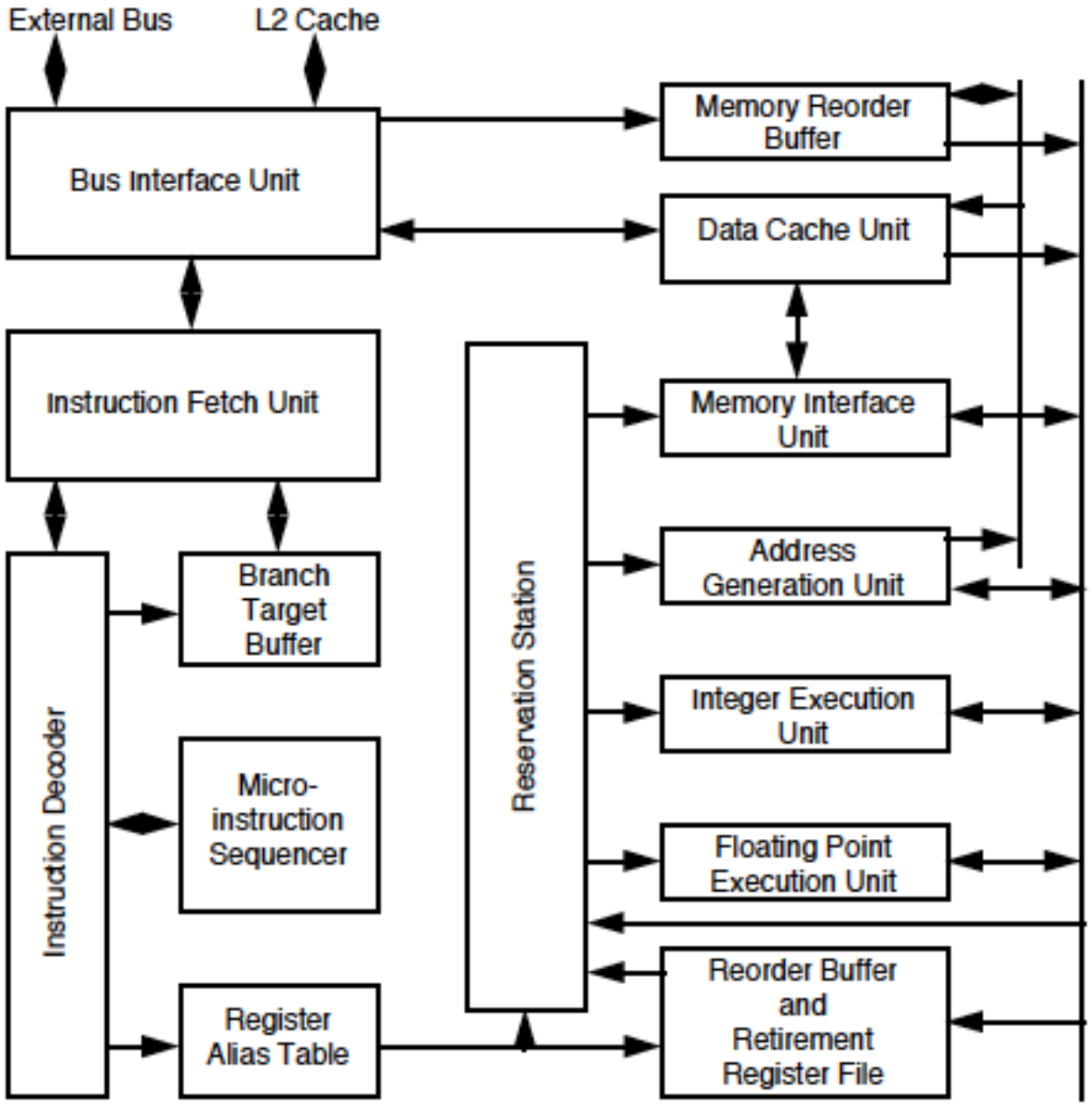


Figure 1 Pentium® Pro Processor Block Diagram

Pentium

- First superscalar from Intel
- In-order
- Two integer pipes (U-pipe and V-pipe)
- U- default pipeline to process all instr
- V- simpler pipeline for simple instrns only

- The following integer instructions are considered simple and may be paired:
- mov reg, reg/mem/imm
- mov mem, reg/imm
- alu reg, reg/mem/imm
- alu mem, reg/imm
- inc reg/mem
- dec reg/mem
- push reg/mem
- pop reg
- lea reg,mem
- jmp/call/jcc near
- nop

uops

- ADD [EBX+EAX], ECX
- Gets decomposed to

- MR1 <- EBX + EAX; effective address
- MR2 <- Mem[MR1]; load
- Flags, MR3 <- MR2 + ECX; add
- Mem[MR1] <- MR3; result store

Perf Monitoring on PentiumPro

- How many counters?
- 2
- Often an additional free-running cycle counter
- How are the counters accessed/controlled?
- Event select register controls what is counted
- Accessed using RDMSR and WRMSR instructions

Table 1. Pentium® Pro Processor Counter based Performance Metrics

Performance Metric	Numerator Event	Denominator Event
Data references per instruction	DATA_MEM_REFS	INST_RETIRED
L1 Dcache misses per instruction	DCU_LINES_IN	INST_RETIRED
L1 Icache misses per instruction	L2_IFETCH	INST_RETIRED
ITLB misses per instruction	ITLB_MISS	INST_RETIRED
Istalls cycles per instruction	IFU_MEM_STALL	INST_RETIRED
L1 cache misses per instruction	L2_RQSTS	INST_RETIRED
L2 cache misses per instruction	L2_LINES_IN	INST_RETIRED
L2 Miss ratio	L2_LINES_IN	L2_RQSTS
Memory transactions per instruction	BUS_TRAN_MEM	INST_RETIRED
FLOPS per instruction	FLOPS	INST_RETIRED
UOPS per instruction	UOPS_RETIRED	INST_RETIRED
Speculative execution factor	INST_DECODED	INST_RETIRED
Branch frequency	BR_INST_RETIRED	INST_RETIRED
Branch mispredict ratio	BR_MISS_PRED_RETIRED	BR_INST_RETIRED
Branch taken ratio	BR_TAKEN_RETIRED	BR_INST_RETIRED
BTB miss ratio	BTB_MISSES	BR_INST_DECODED
Branch Speculation factor	BR_INST_DECODED	BR_INST_RETIRED
Resource stalls per instruction	RESOURCE_STALLS	INST_RETIRED
Cycles per instruction	CPU_CLK_UNHALTED	INST_RETIRED

Key concepts to understand this paper well

- What is dynamic execution?
- What does dataflow order mean?
- What are reservation stations?
- What is ROB?
- What is MOB?
- What does register renaming mean?
- What is non-blocking cache?
- MSHR (Miss Status Holding Register)

More questions

- How long is the PPro pipeline?
- 8-3-3
- What is the issue width of PPro?
- 3
- How many RS entries?
- 20
- How many ROB entries?
- 40

What are instruction latencies/throughput of PPro?

- Add
- 1/1
- FPADD
- 3/1
- FPMUL
- 5/0.5
- IMUL
- 4/1
- LD
- 3 for L1 hit
- FDIV –
- 17/32/37 (single/double/extended)

More questions

- Which benchmarks (SPEC INT or FP) have better i-cache behavior?
- Which benchmarks (SPEC INT or FP) have better L1D-cache behavior?
- Int – 5-45; fp – 10 -140 (MPKI)
- Which benchmarks (SPEC INT or FP) have better L2-cache behavior?
- Int -1-15; fp – 10 -150 (MPKI)
- What is speculative execution factor?
- A program has low spec factor. Is that bad?

More questions

- What feature size was Ppro?
- 0.6 u
- Which benchmarks (SPEC INT or FP) have more TLB misses?
- INT – 0.01 to 0.9 MPKI
- FP – 0.01 to 0.05 MPKI
- How many memory transactions per K instrn?
- Int – 1 to 10
- Fp – 1 to 70

More questions

- In OOO processors, it is not always possible to find cause and effect relationships. Why?
- What metric is most correlated to CPI in this paper?
- L2 misses
- What is the average uop ratio for spec cpu programs?
- What is the average uop ratio for Excel?

Table 2. Basic Characteristics of Systems

Processor	Intel Pentium® Pro Processor	Intel Pentium® Processor
CPU Core Frequency	150 MHz	120 MHz
Bus Frequency	60 MHz	60 MHz
Data bus	64-bit	64-bit
Address bus	36-bit	32-bit
On-chip L1 cache	8 KB data, 8 KB instruction	8 KB data, 8 KB instruction
Off-chip L2 cache	4-way 256 KB	512 KB (Dell), 256 KB (Gateway)
L2 cache timing	4-1-1-1 @ 150 MHz CPU freq.	3-1-1-1 @ 60 MHz bus frequency
System Chip Set	82450GX/KX	82430FX
Memory timing (bus cycles)	14-1-1-1 (4-way interleaving) 14-2-2-2 (2-way interleaving) 14-4-4-4 (no interleaving)	13-3-3-3 (Fast Page Mode DRAM) 13-2-2-2 (EDO DRAM)
Basic Pipeline	14 stages	5 stages
Superscalar	3-way	2-way
Execution units	5	3
Branch prediction	4-way 512 entry BTB, 4-bit history, 2 level adaptive	4-way 256 entry BTB, 2-bit history
Execution model	Out of order	In order
Speculative Execution	Yes	No
McCalpin Streams Memory Bandwidth	140 MB/sec (4-way interleaving) 128 MB/sec (2-way interleaving) 97 MB/sec (no interleaving)	82 MB/sec (Gateway 2000 P120)
SYSmark/NT rating	497 (Digital Celebris* XL6150)	294 (Gateway 2000 P120)
SPECint95	6.08 (Intel Alder System)	3.53 (Dell Dimension XPS P120)
SPECfp95	5.42 (Intel Alder System)	2.92 (Dell Dimension XPS P120)

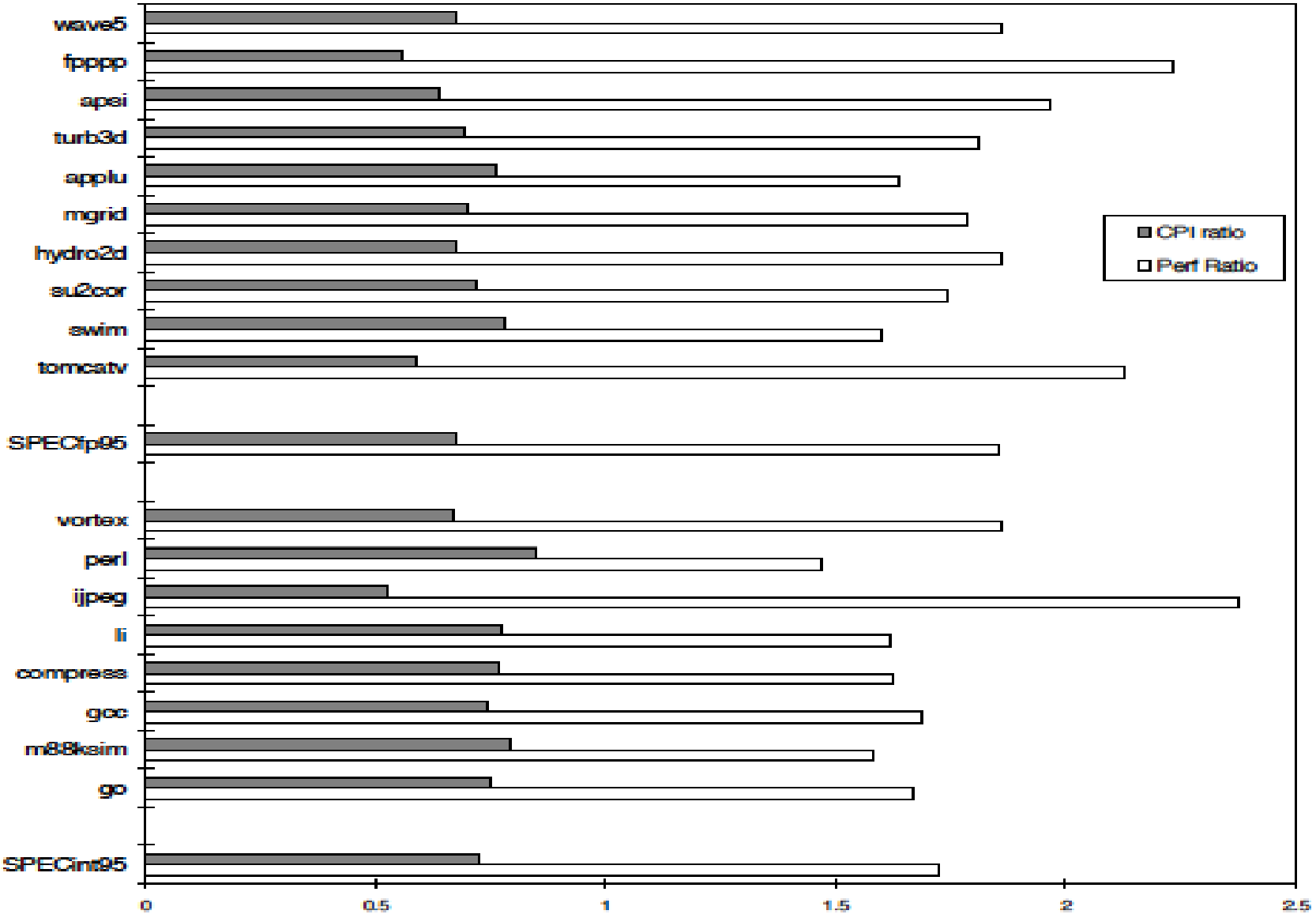


Figure 2 Performance Comparison of Pentium® and Pentium® Pro Processors on SPEC95

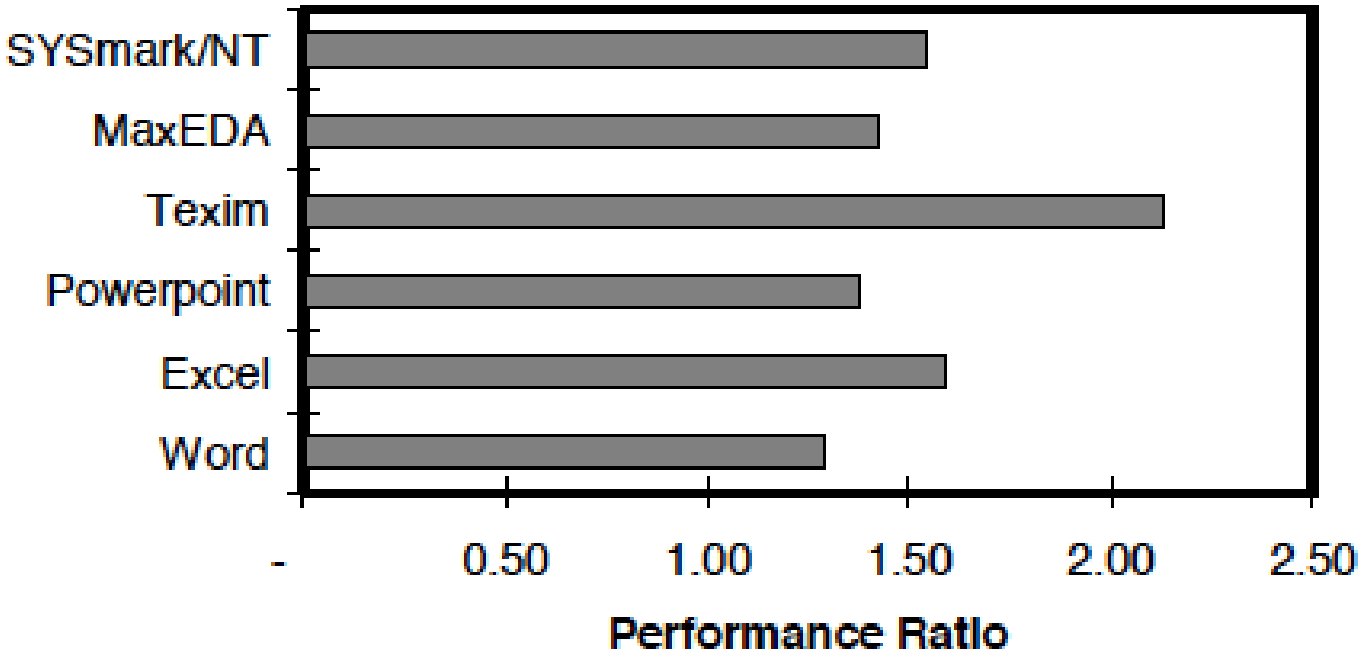
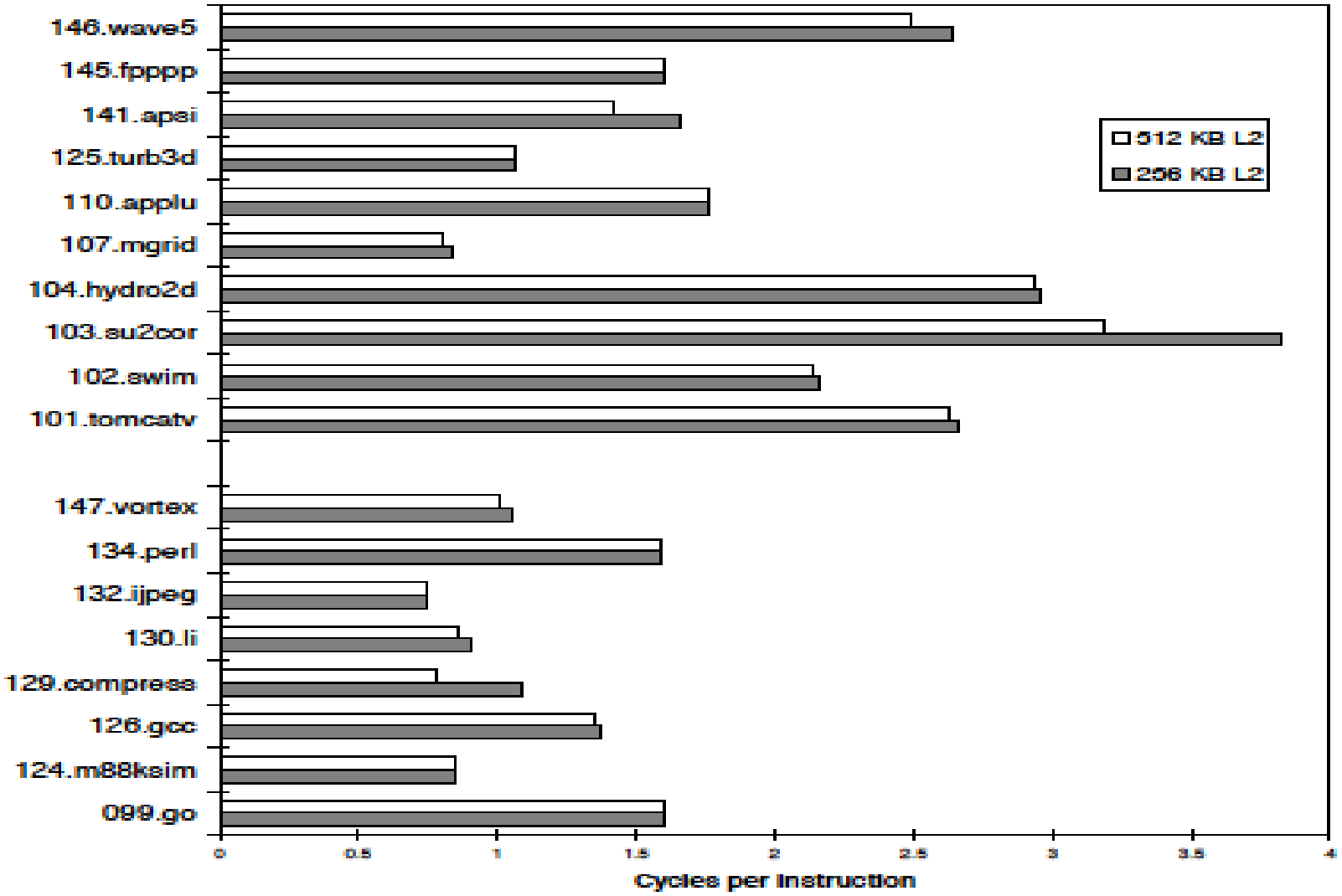


Figure 3 Performance Comparison of Pentium® and Pentium® Pro Processors on SYSmark/NT

Small Cache vs Bigger Cache



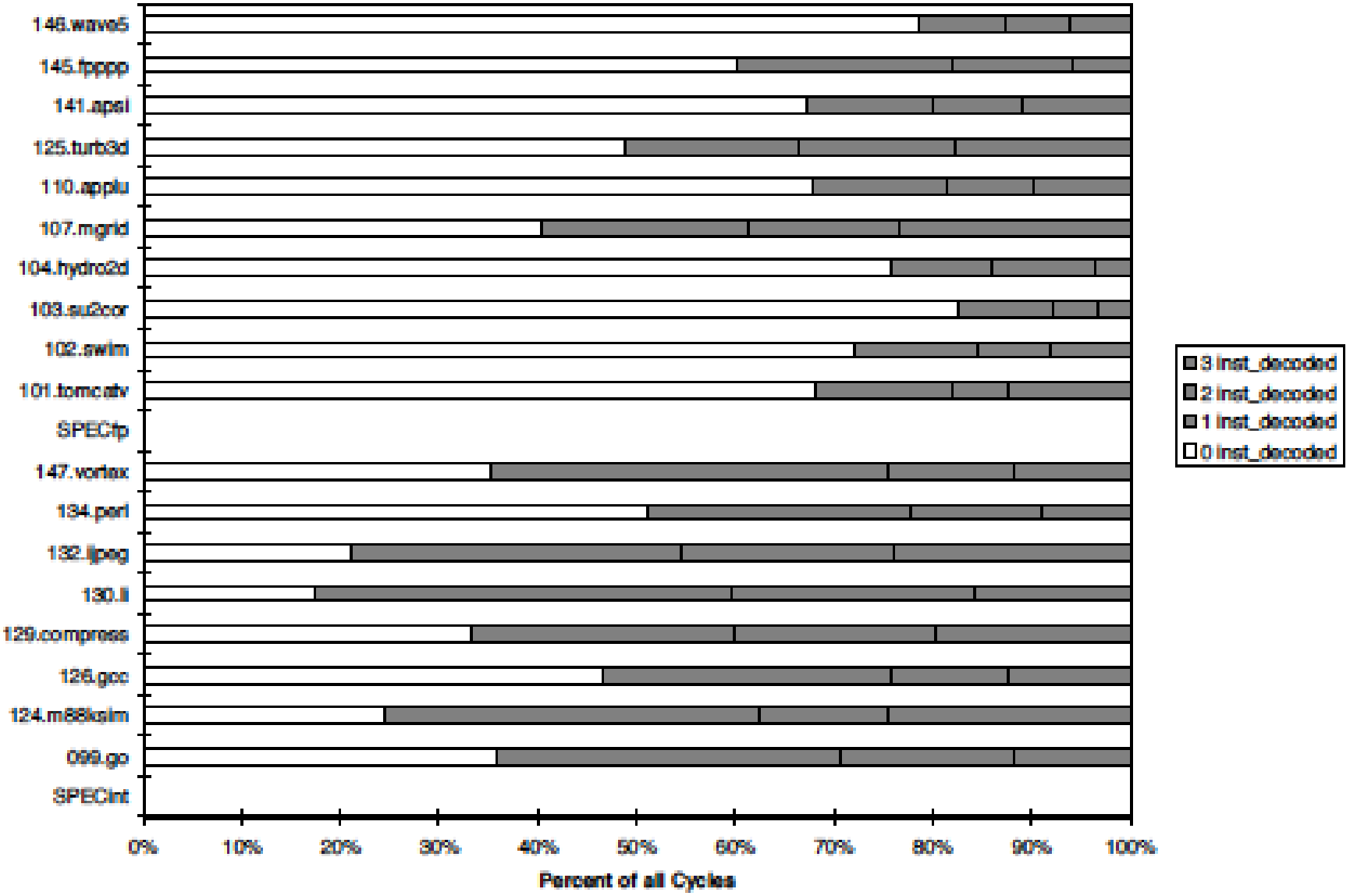
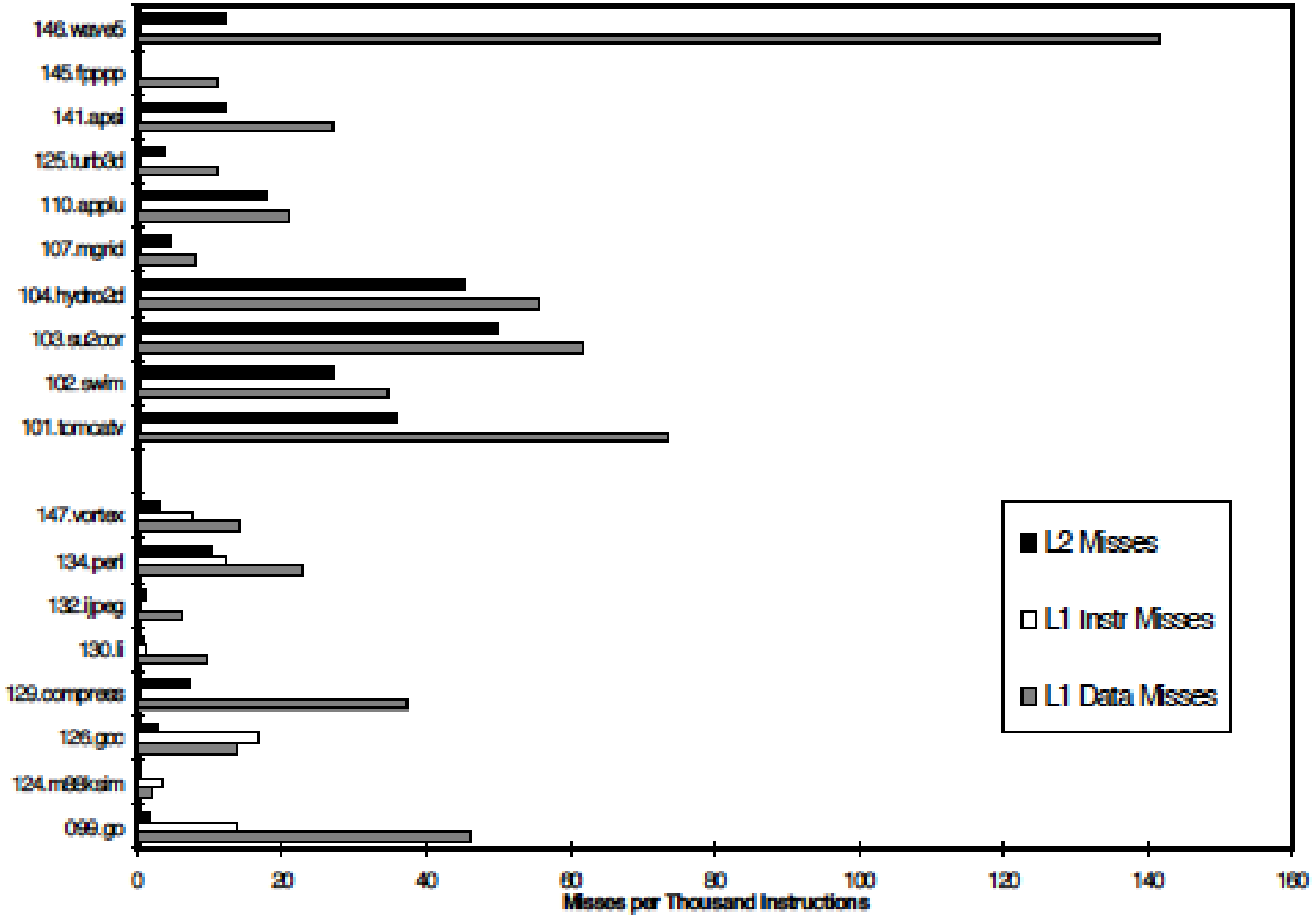


Figure 5 Instruction Decode Profile


146.wave5
145.fpppp
141.apsi
125.turb3d
110.applu
107.mgrid
104.hydro2d
103.su2cor
102.swim
101.tomcatv
147.vortex
134.perl
132.ijpeg
130.li
129.compress
126.gcc
124.m88ksim
099.go
L2 Misses
L1 Instr Misses
L1 Data Misses
0
20
40
60
80
100
120
140
160
Misses per Thousand Instructions

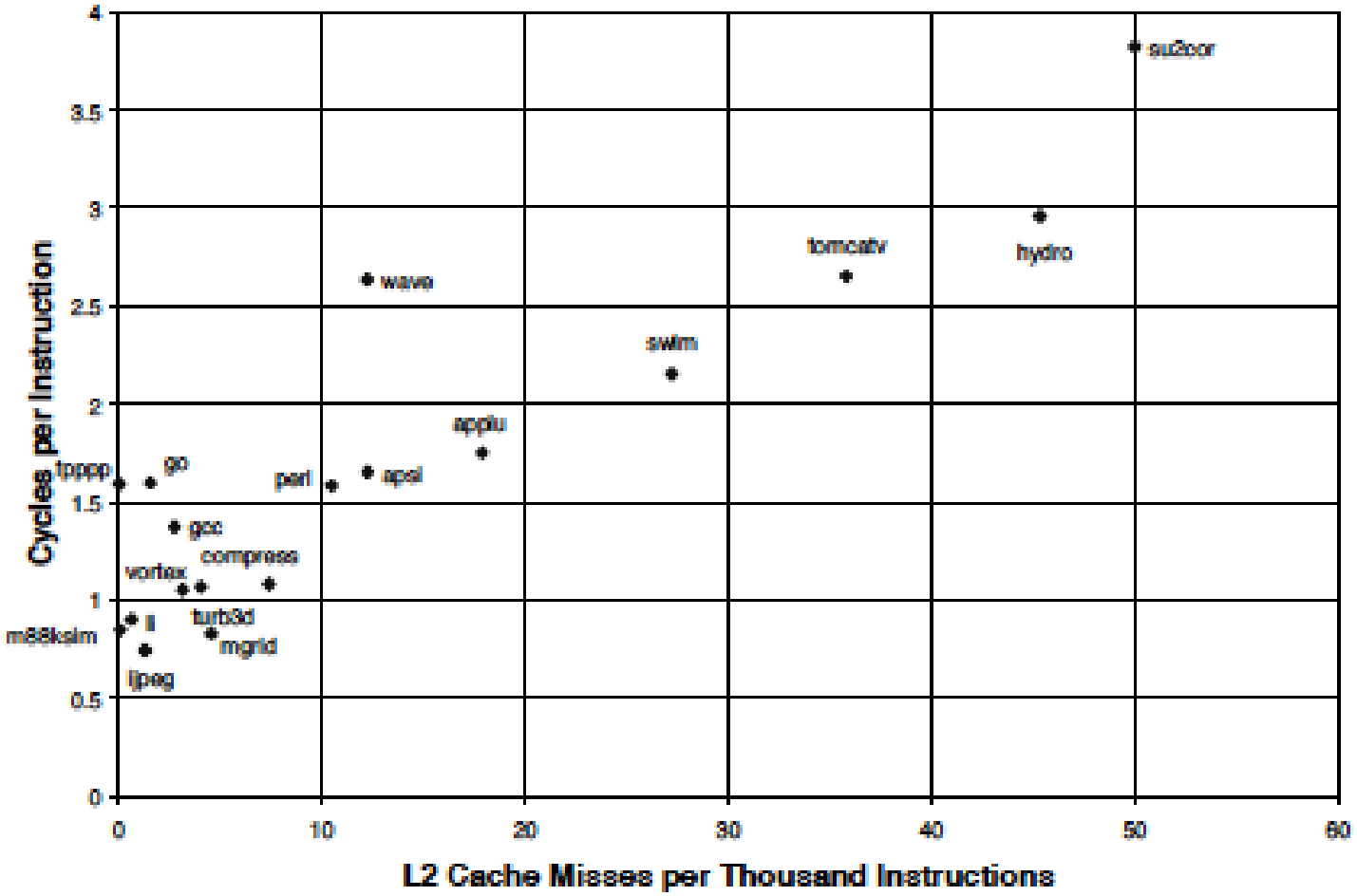


Figure 7 CPI versus L2 Cache Misses

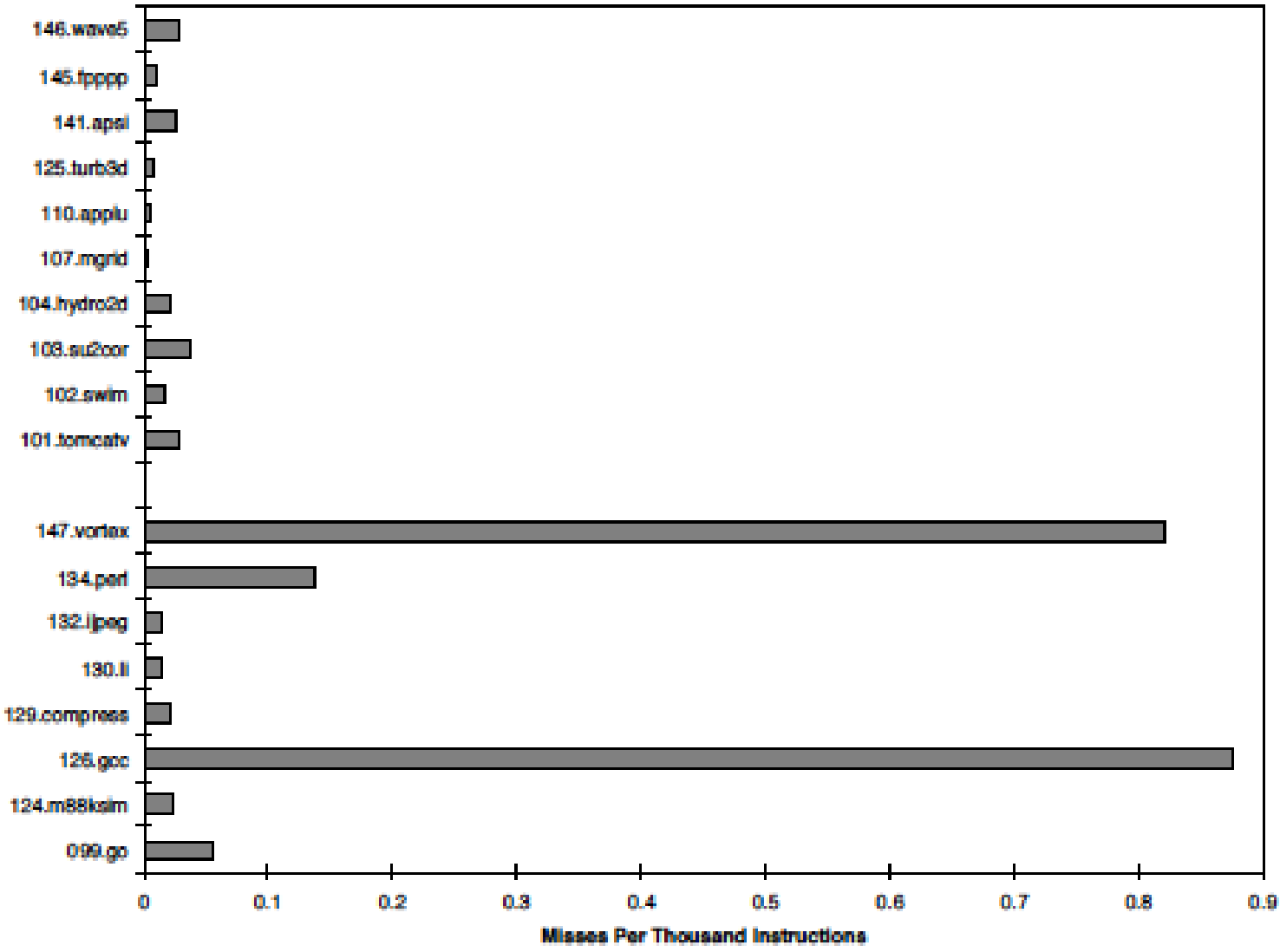


Figure 8 ITLB Statistics

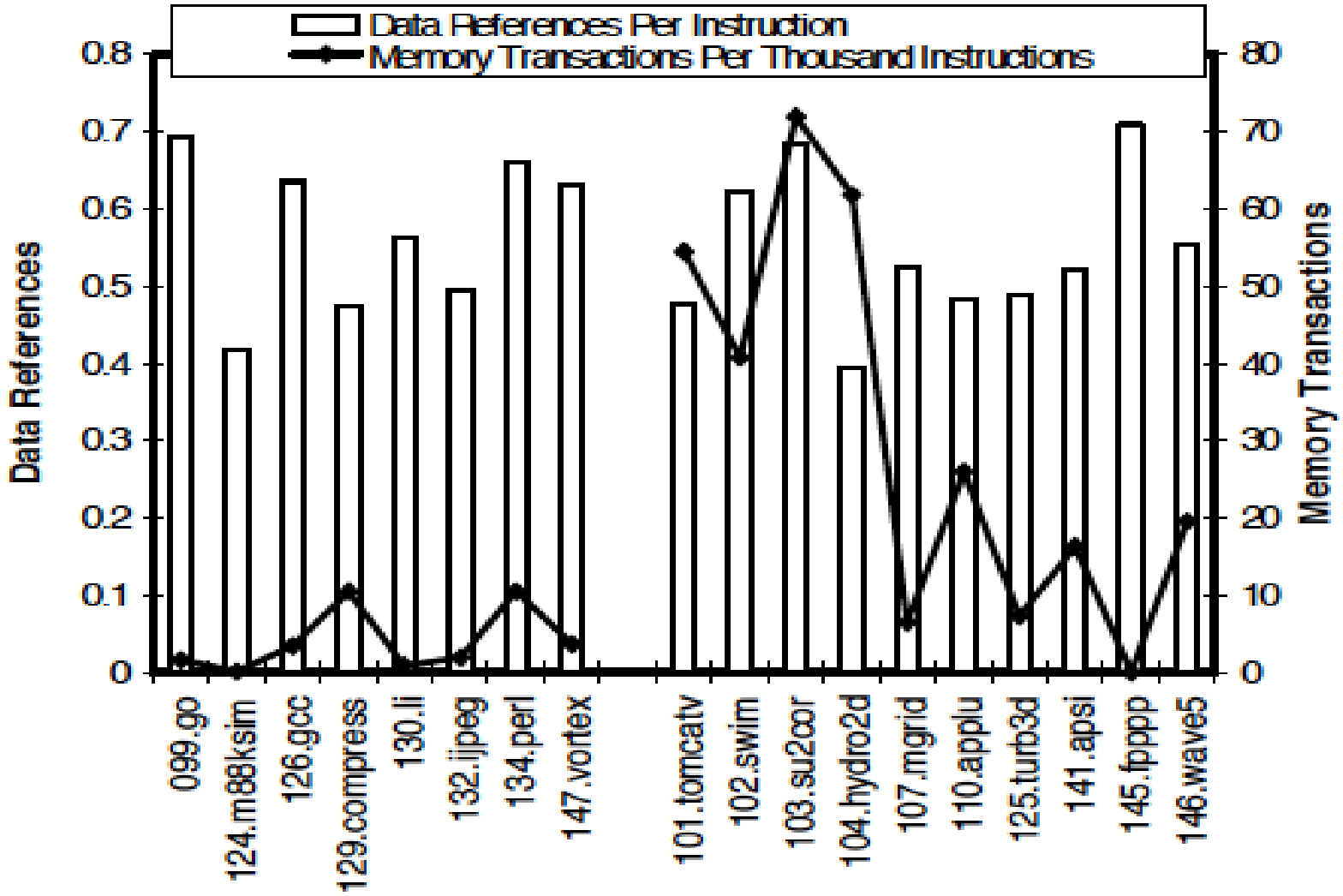
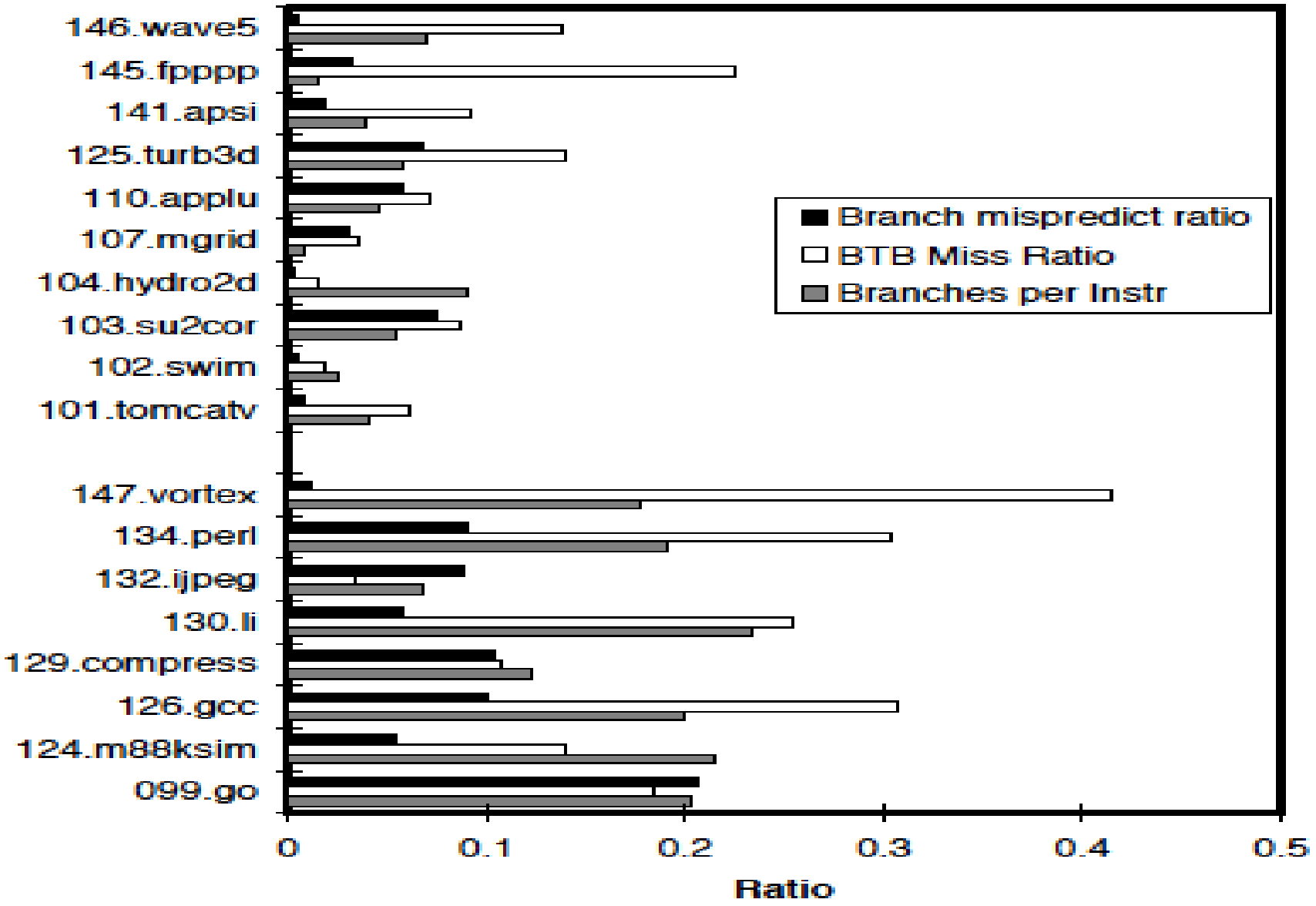


Figure 9 Memory Reference Statistics


146.wave5
145.fpppp
141.apsi
125.turb3d
110.applu
107.mgrid
104.hydro2d
103.su2cor
102.swim
101.tomcatv
147.vortex
134.perl
132.ijpeg
130.li
129.compress
126.gcc
124.m88ksim
099.go
Branch mispredict ratio
BTB Miss Ratio
Branches per Instr
0
0.1
0.2
0.3
0.4
0.5
Ratio

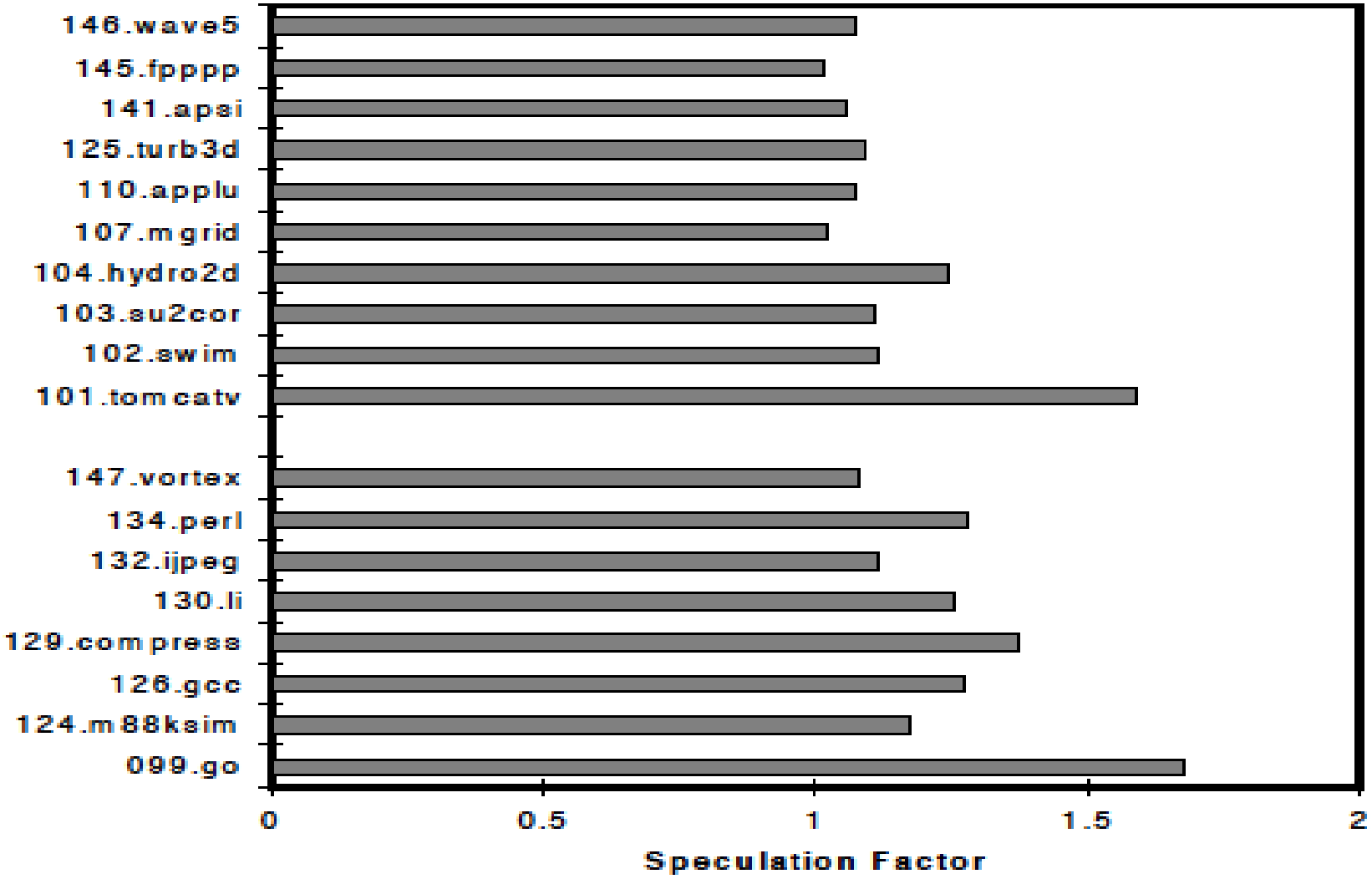


Figure 11 Speculation Factor

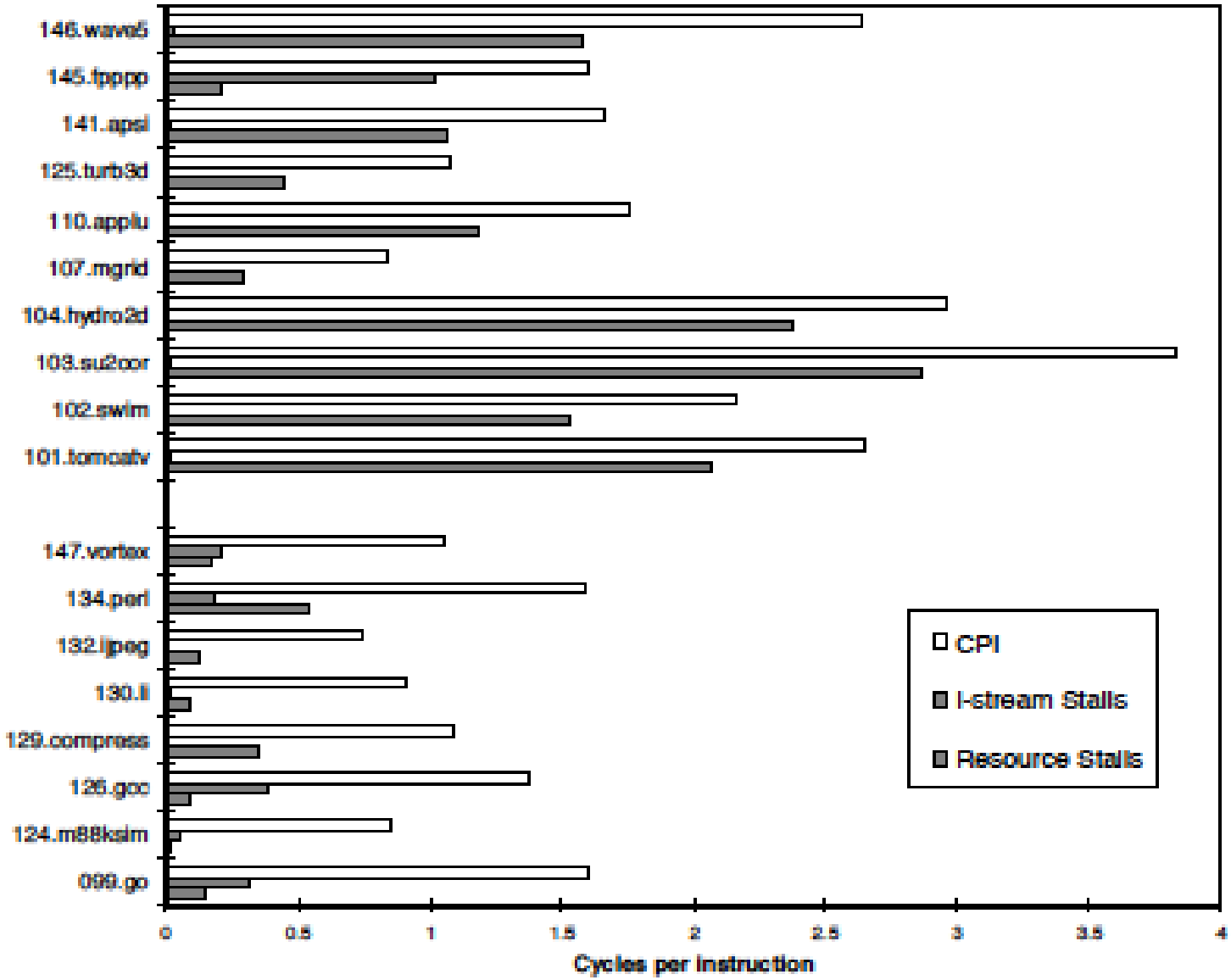
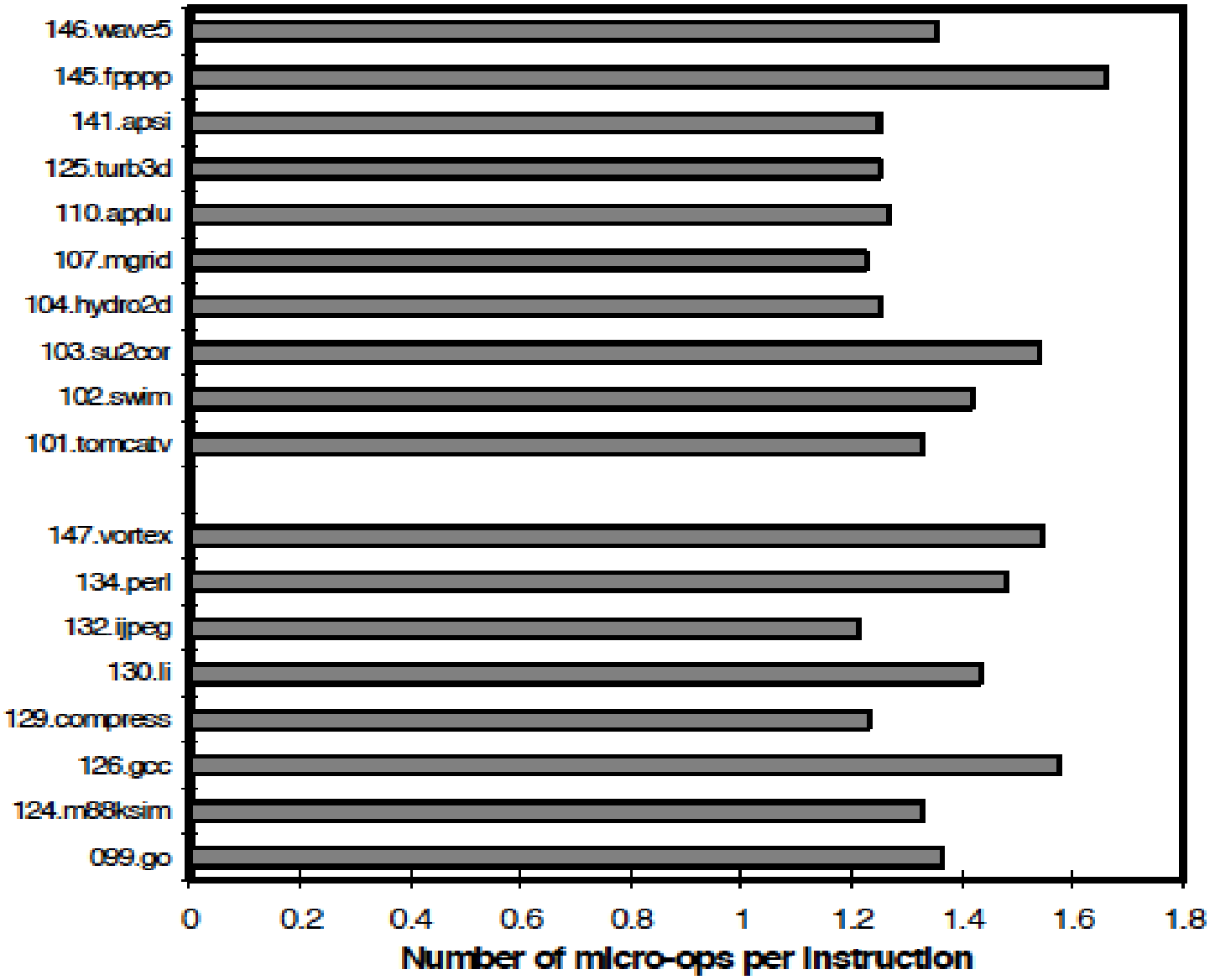


Figure 12 Stall Cycles


146.wave5
145.fpppp
141.apsi
125.turb3d
110.applu
107.mgrid
104.hydro2d
103.su2cor
102.swim
101.tomcatv
147.vortex
134.perl
132.ijpeg
130.li
129.compress
126.gcc
124.m88ksim
099.go
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
Number of micro-ops per Instruction

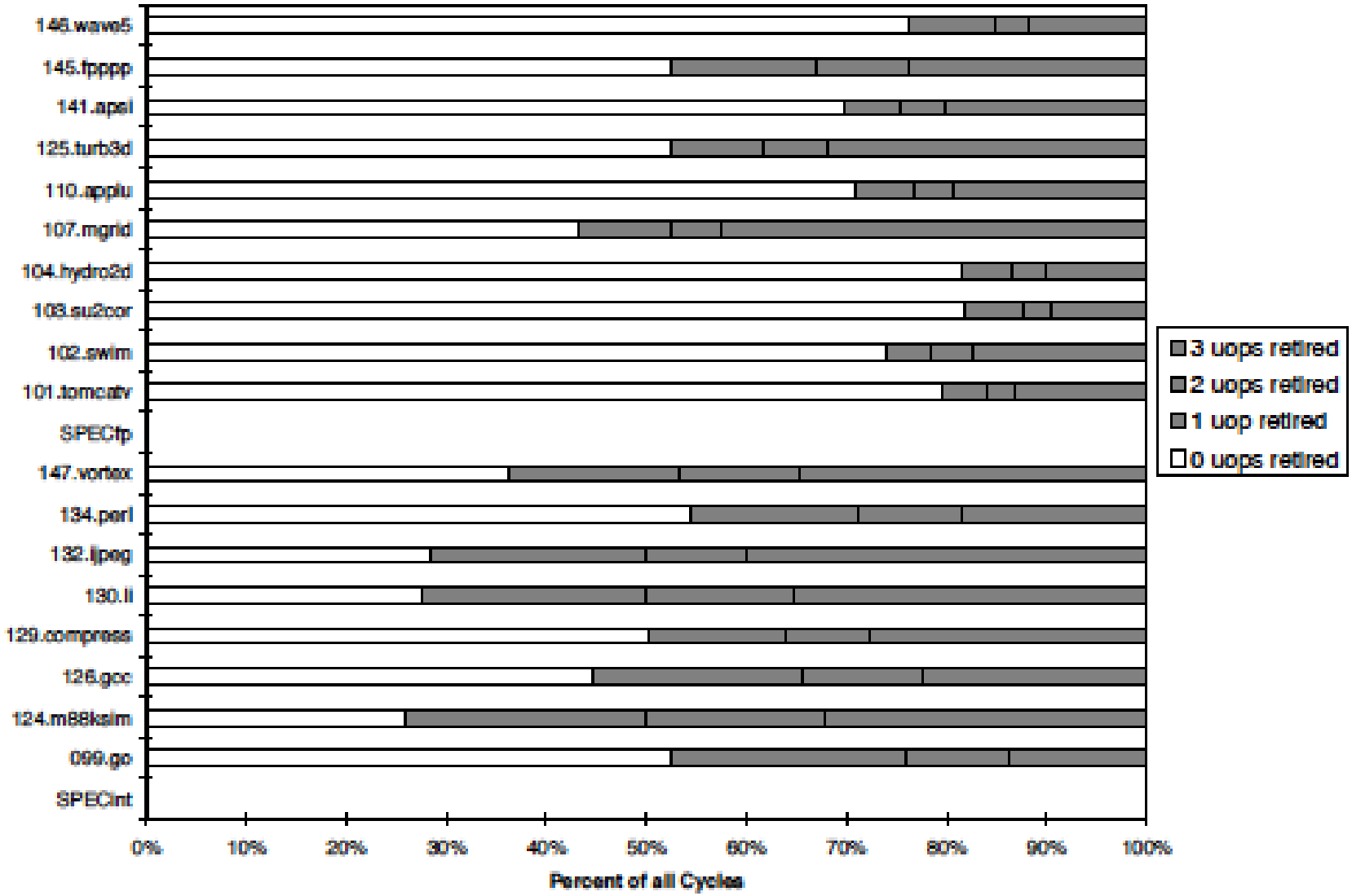


Figure 14 Micro-operations retirement profile

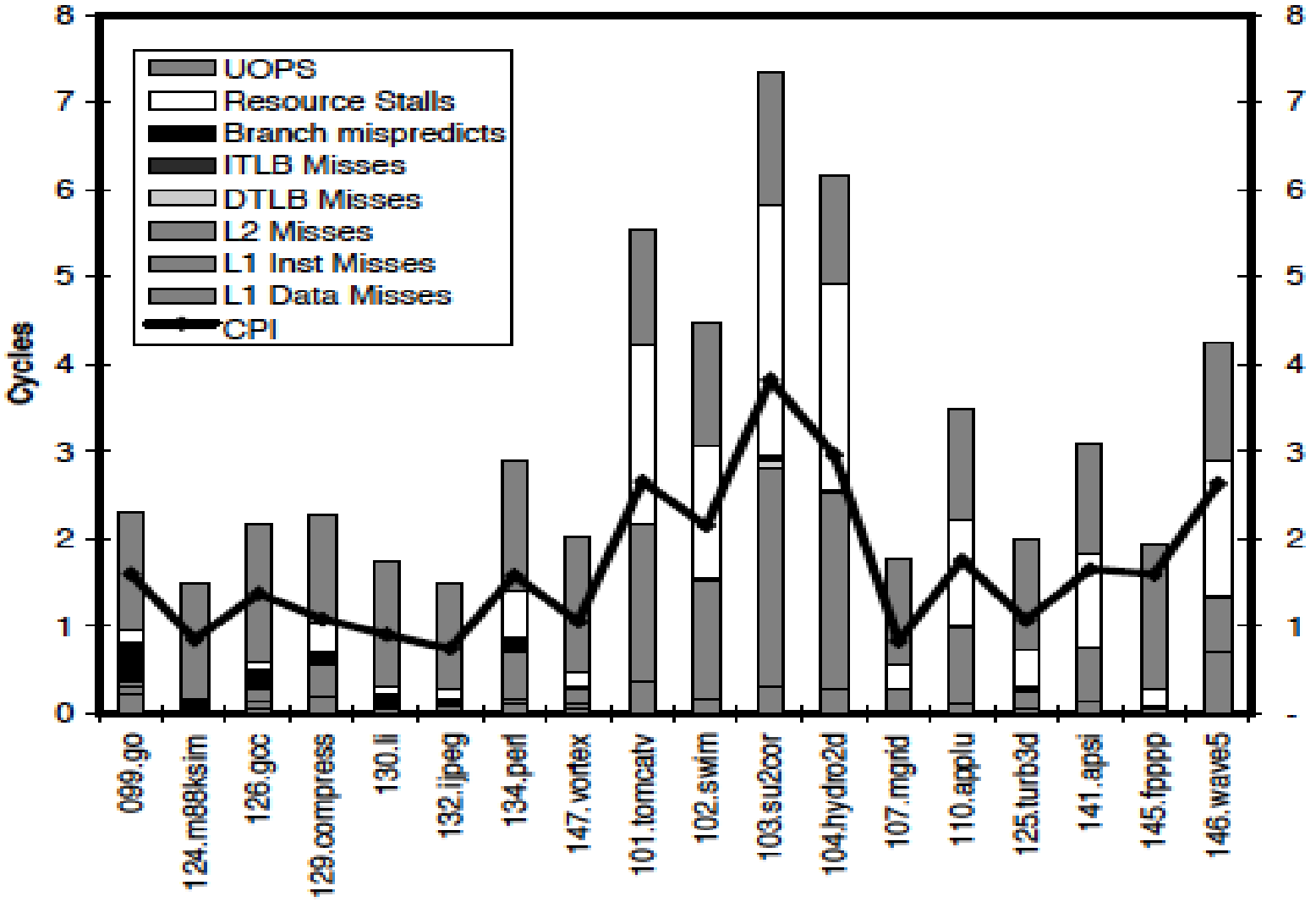


Figure 15 CPI vs. Latency Components

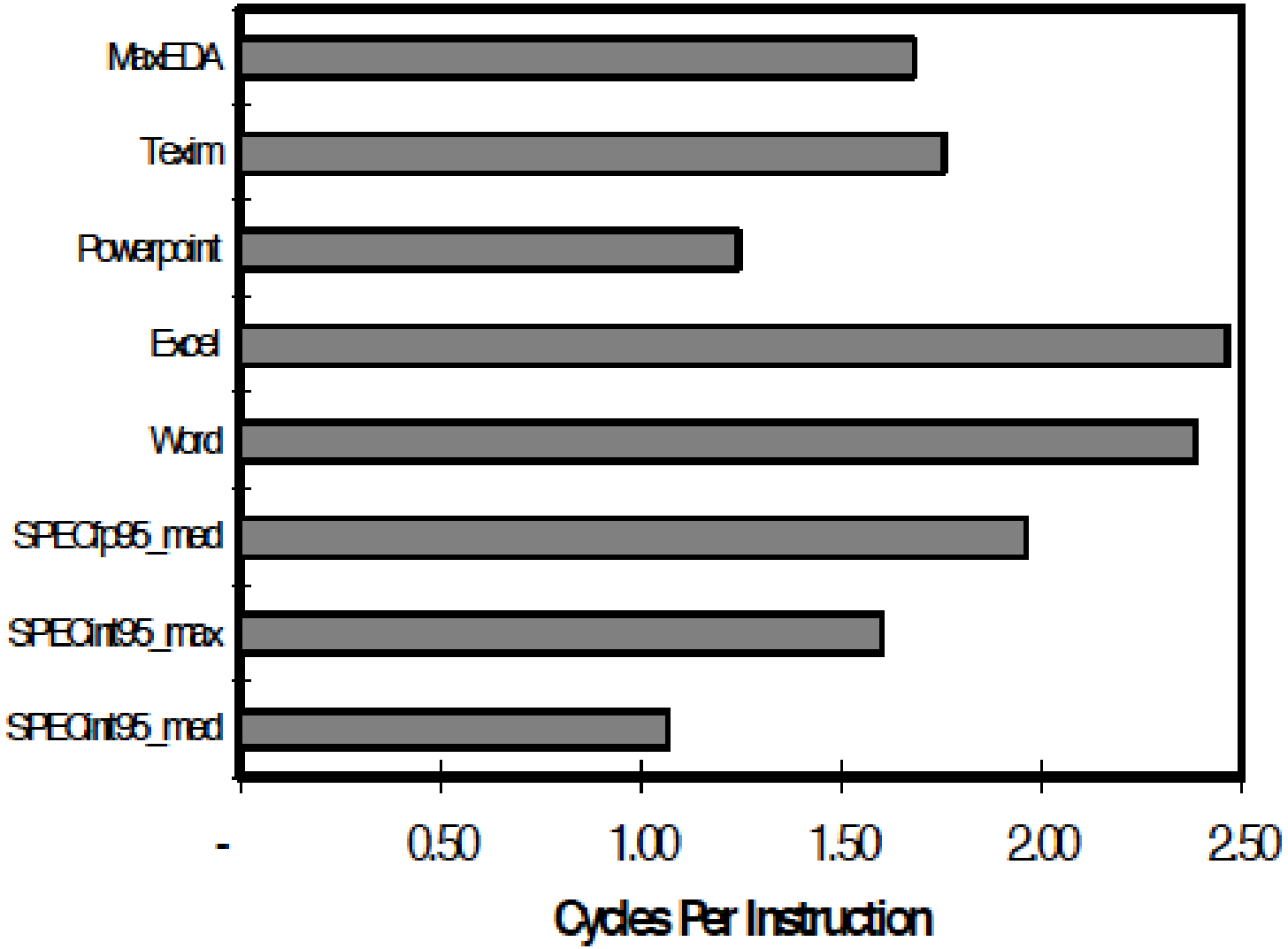


Figure 16 CPI for SYSmark/NT

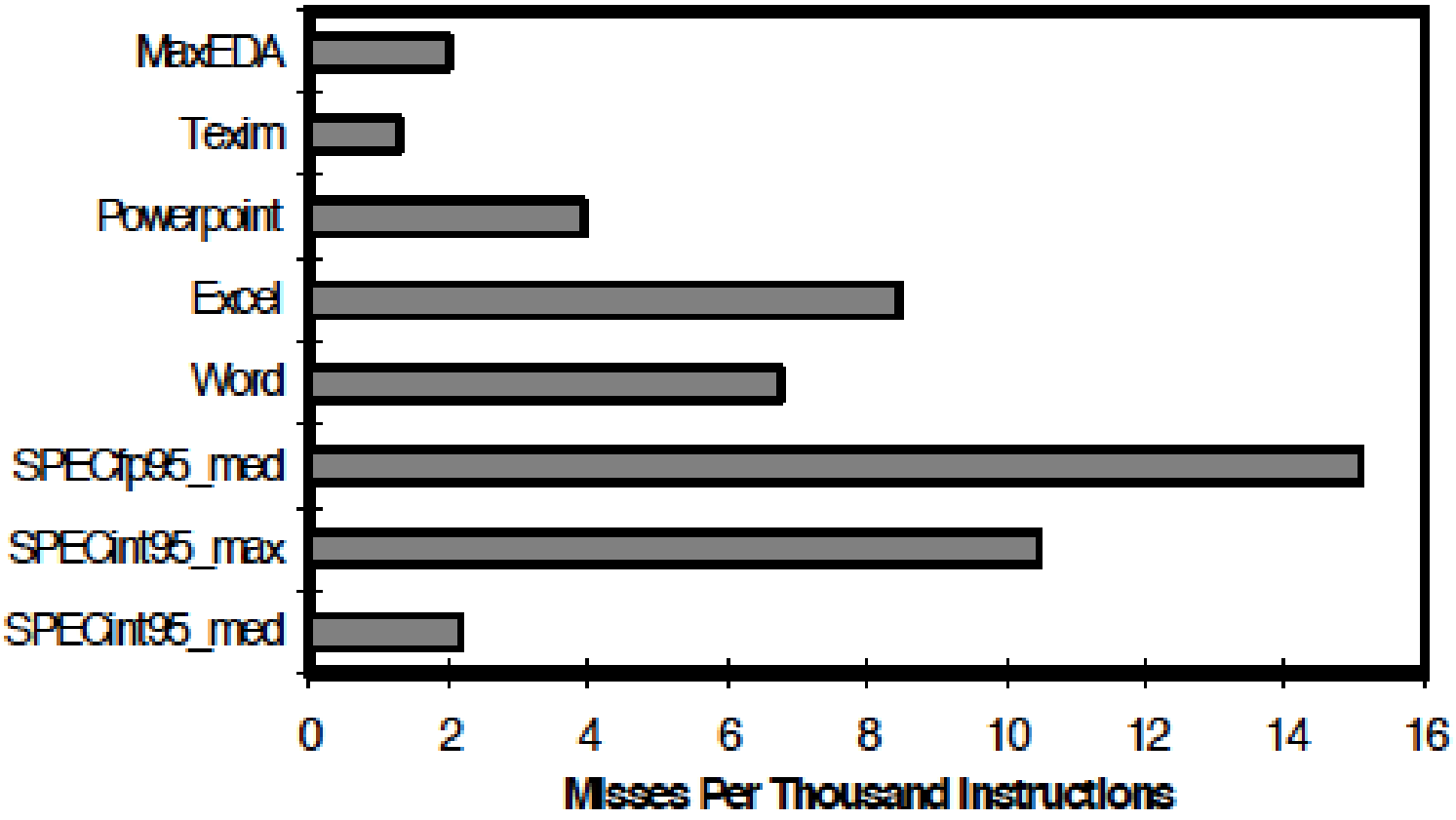


Figure 17 SYSmark/NT L2 Cache Misses

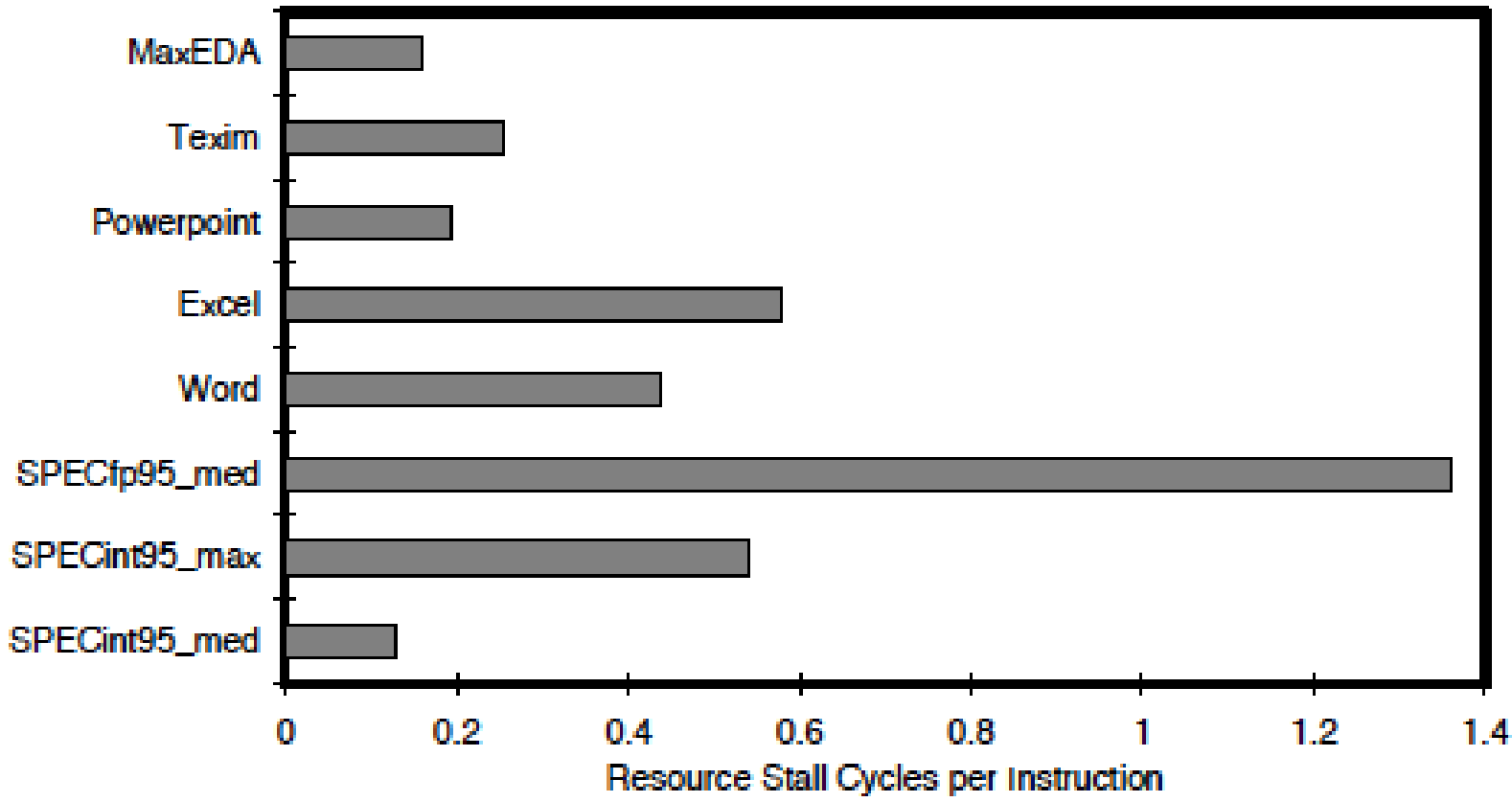


Figure 18 SYSmark/NT Resource Stalls

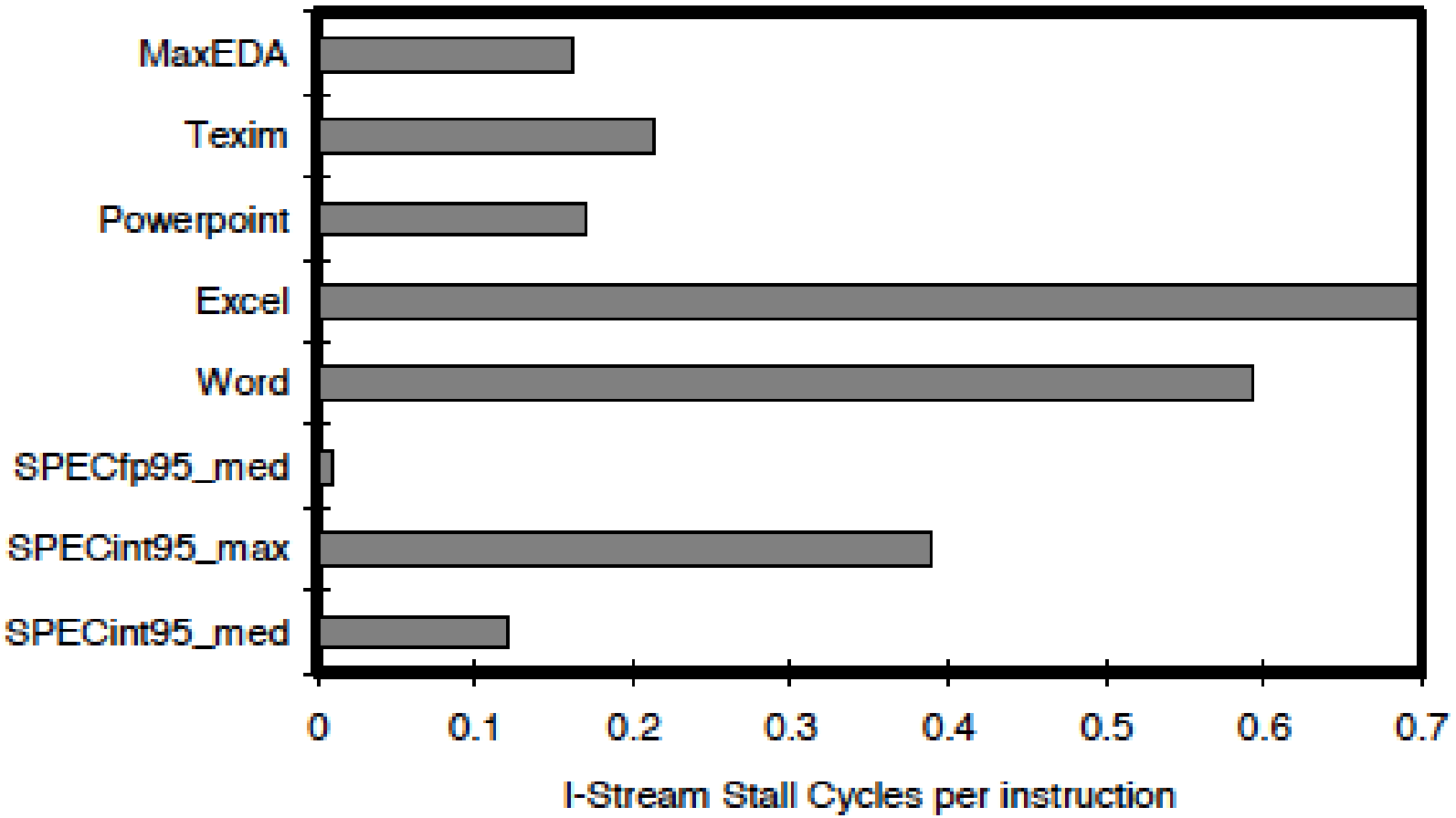


Figure 19 SYSmark/NT Instruction Stalls

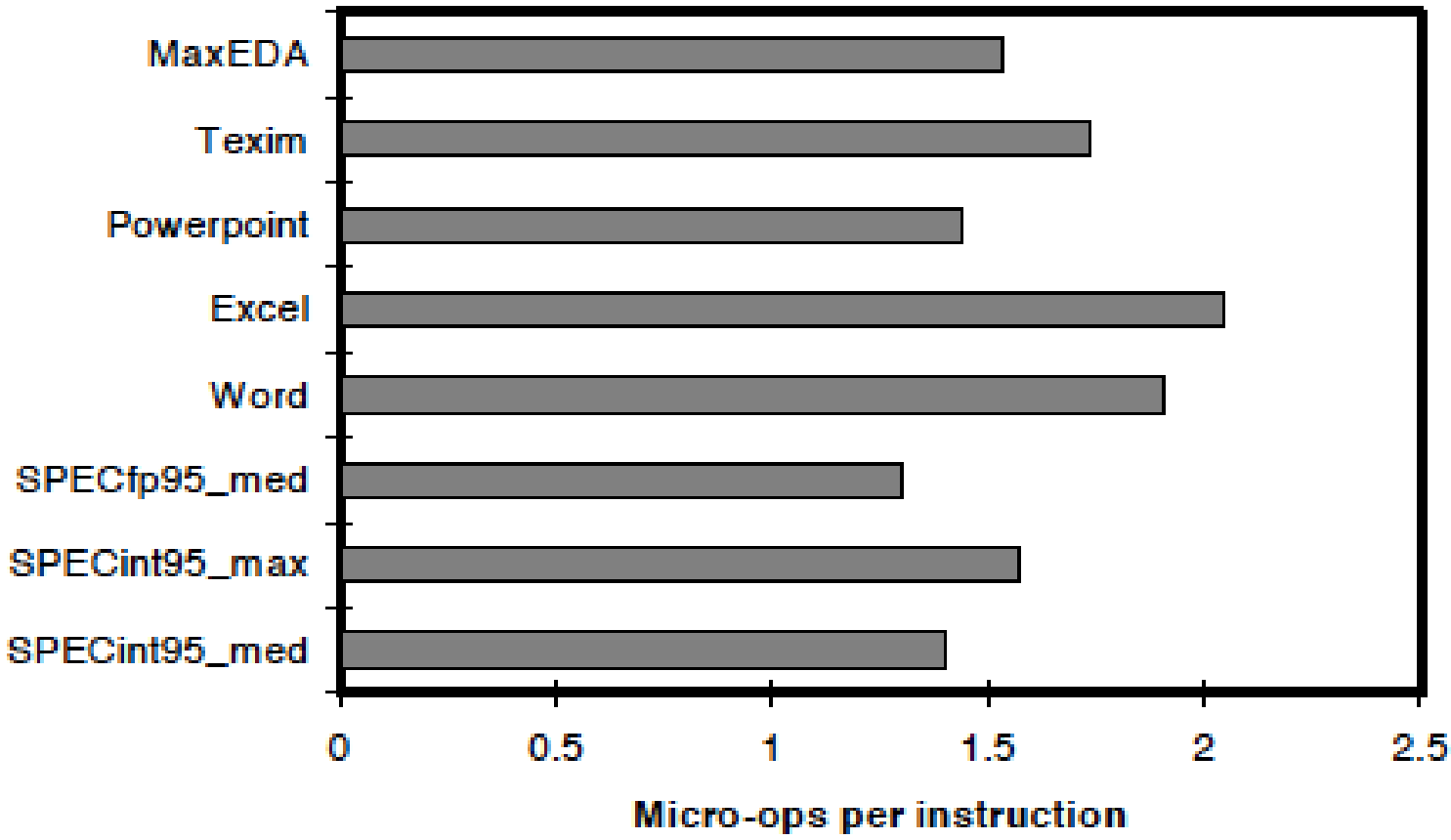


Figure 20 SYSmark/NT Micro-ops Per Instruction

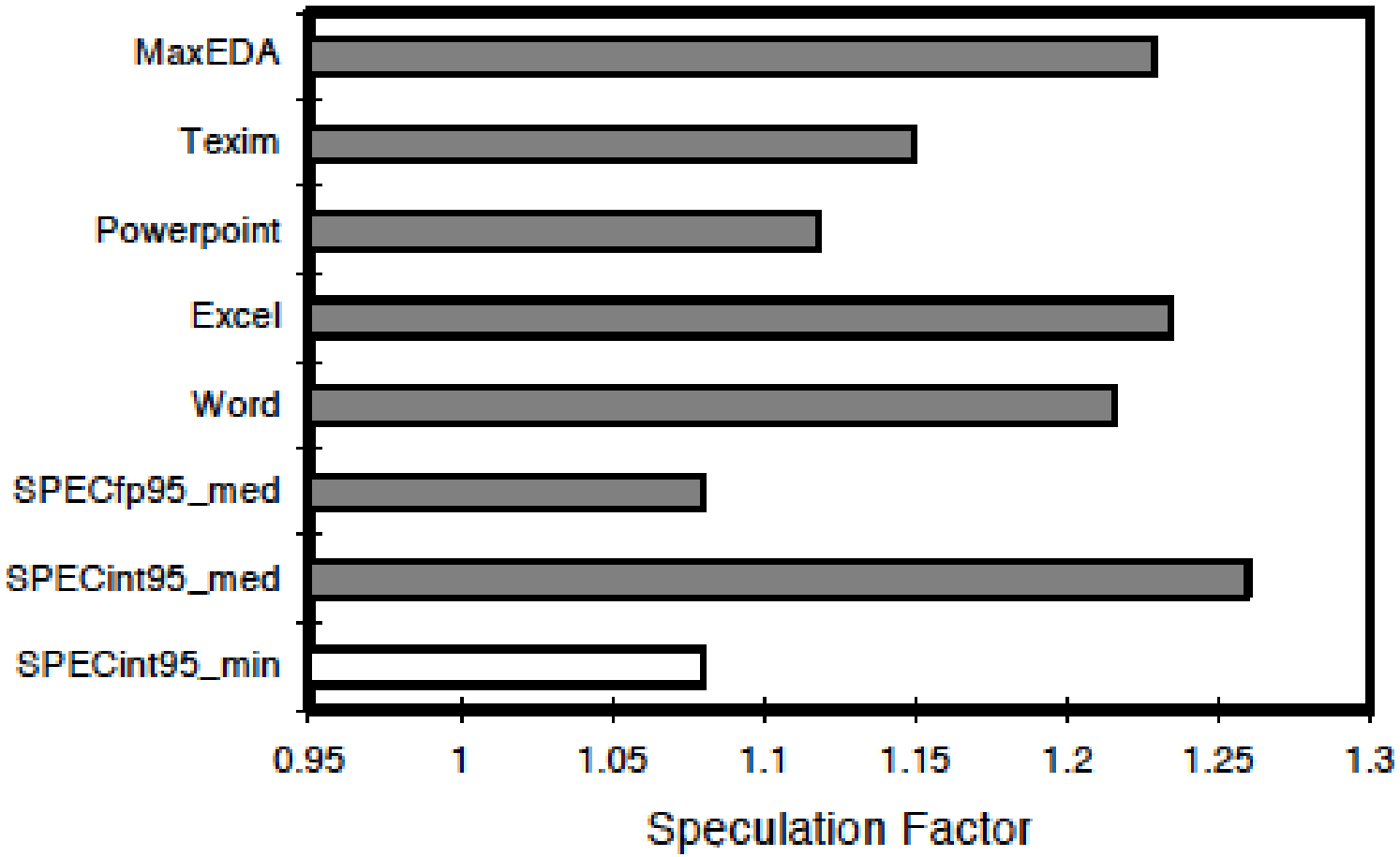


Figure 21 SYSmark/NT Speculation Factor