ON THE COMPARISON OF DETERMINISTIC AND PROBABILISTIC WCET ESTIMATION TECHNIQUES

Jaume Abella[1,4], *Damien Hardy*[2], *Isabelle Puaut*[2], *Eduardo Quinones*[1,4], *Francisco J. Cazorla*[1,3,4]

1 Barcelona Supercomputing Center
Centro Nacional de Supercomputación

2 Université de Rennes 1

3 CSIC Consejo Superior de Investigaciones Científicas

4 PROXIMA

26th Euromicro Conference on Real-Time Systems (ECRTS14)
8-11 July Madrid, Spain

ECRTS 2014
8-11 July
Madrid, Spain

BACK IN ECRTS 2013 CONFERENCE

I have an argument about this WCET approach

I have a contra-argument

I have another argument

I argue on your argument

Well guys, let's investigate this and make a first comparison of these WCET techniques

Timing Validation

- Critical step in the design of a real-time system
  - WCET estimates derived per task (Unit of Scheduling)
  - Task WCET estimate given as an input to schedulability analysis
- Classification of existing WCET estimation techniques:
  - Deterministic Timing Analysis (DTA) vs Probabilistic Timing Analysis (PTA)
  - Static vs Measurement-Based

DTA and PTA

- Deterministic Timing Analysis (DTA)
  - Single WCET estimate
  - Designed primarily for deterministic HW/SW
- Probabilistic Timing Analysis (PTA)
  - Multiple WCETs with an associated probability (probabilistic WCET or pWCET)
  - HW/SW designs: deterministic and randomised
- DTA and PTA have their static and measurement-based variants

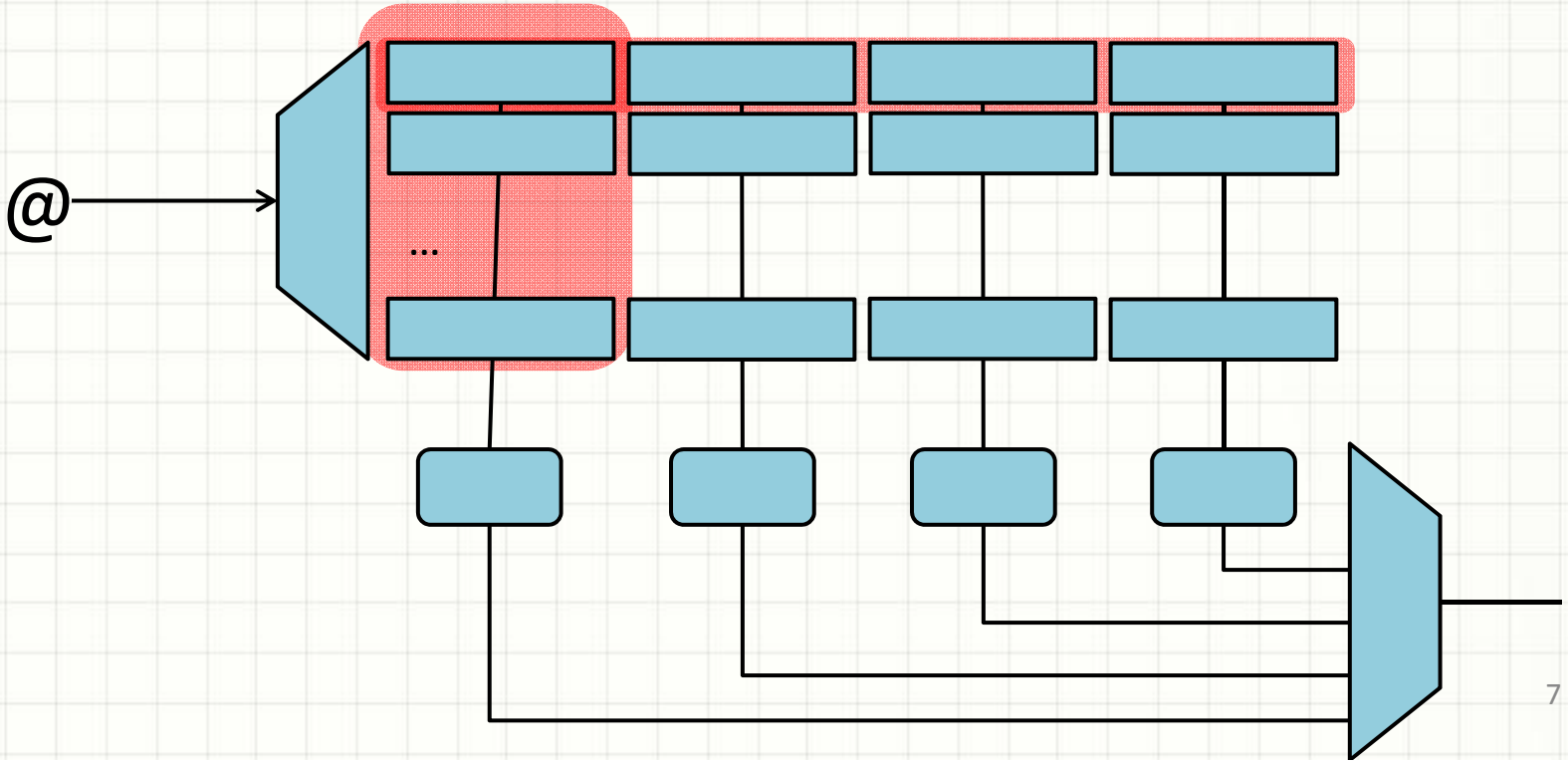
Variant	DTA	PTA
Static	SDTA	SPTA
Measurement based	MBDTA	MBPTA

Comparison

- Difficult so far:
  - PTA still in its infancy
  - Each approach (DTA and PTA) performs better on a different hardware design
- Our goal: Carry out the first comparison between DTA and PTA
  - No apocalyptic take-out message
    - 'This will never work'
  - Qualitative
    - Strengths and limitations of each technique
    - Sensitivity to different parameters
  - Quantitative
    - Common setup in which all methods are applicable

Considered hardware

- Fixed execution latency but the instruct. cache
- Cache structured into sets and ways
  - Placement: Defines the possible set in cache
  - Replacement: Defines which block will be evicted

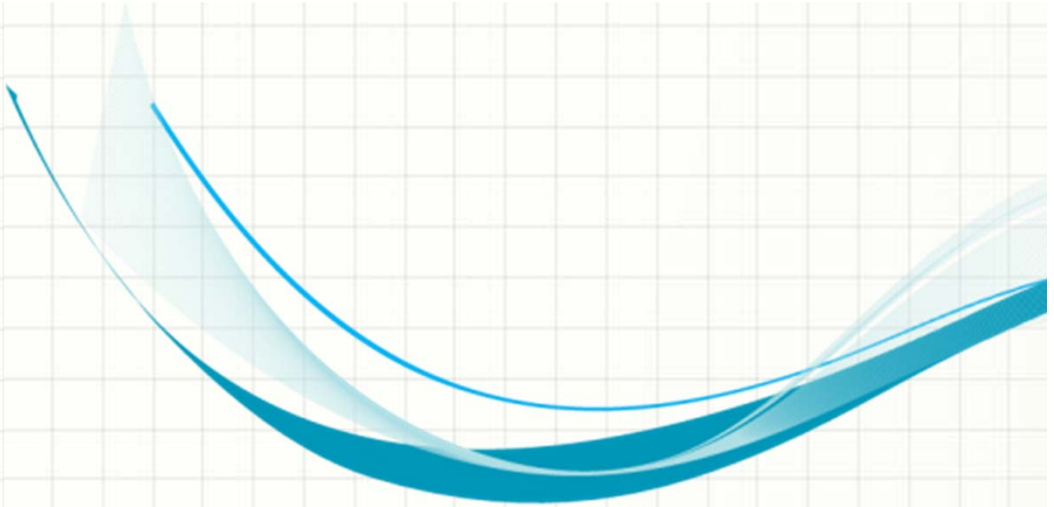


Considered hardware

- Cache structure: Fully-assoc., Set-Associative
- Placement: Deterministic (modulo) vs randomized
- Replacement policy: Deterministic (LRU) vs randomized (Evict on Miss, EoM)

Cache short names		Placement	
		Mod	Rand
Replacement	LRU	Time Det. (TD)	
	EoM	Partially Time Rand (pTR)	Time Rand (TR)

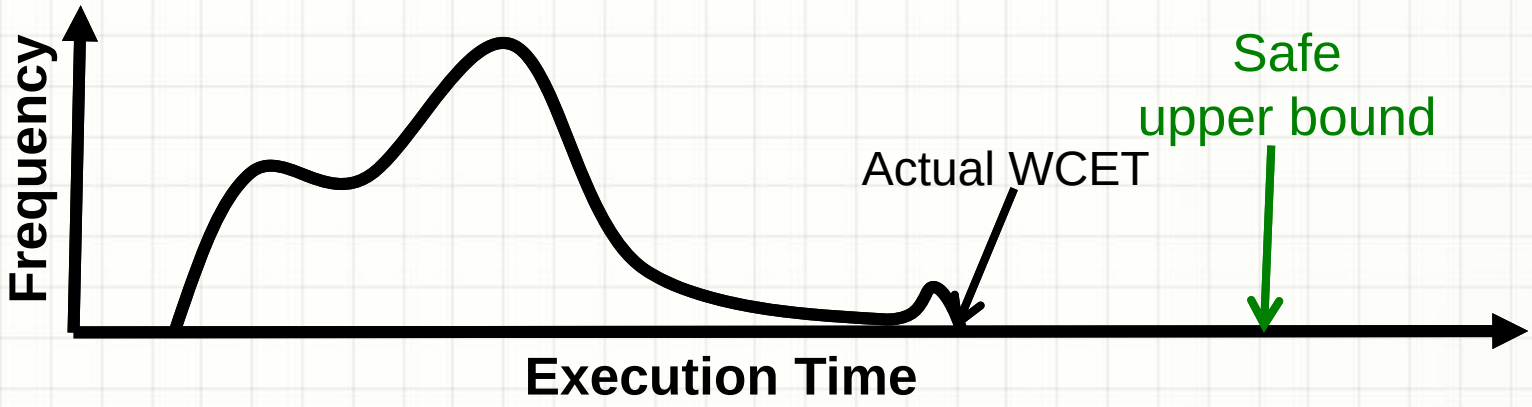
No quan-titative results



Deterministic Timing Analysis (DTA)

Static Deterministic Timing Analysis (SDTA)

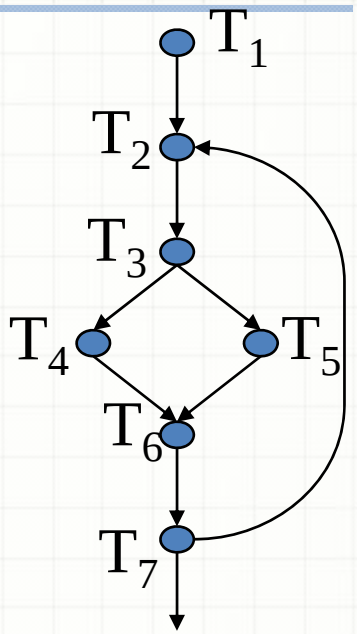
- SDTA provides safe WCET estimates



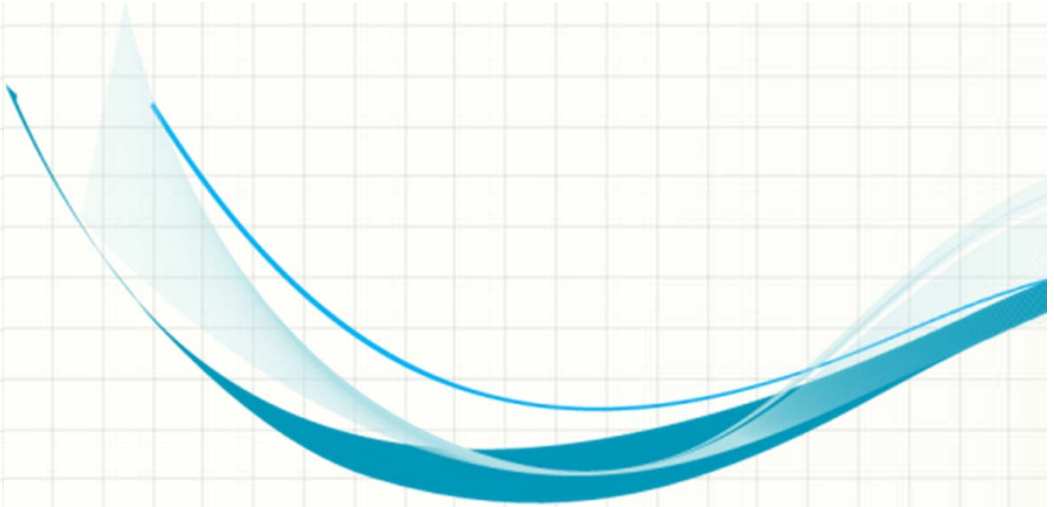
- Needs: Information about the HW/SW
- Approach: Analysis divided into 2 steps
  - Low-level analysis: modeling of hardware timing
    - Caches, pipeline, predictors…
  - High-level analysis: longest execution path computation

Static Deterministic Timing Analysis (SDTA)

- High-level analysis Longest execution path computation
  - Implicit Path Enumeration Technique (IPET)
    - ILP formulation of the WCET calculation problem
    - Linear programming solver



- Low level analysis: Cache Analysis
  - Based on Abstract Interpretation
  - Determines guaranteed hit
    - Cache Hit Miss Classification: Always hit, First miss
  - Defined for different replacement policies
    - LRU, EoM…

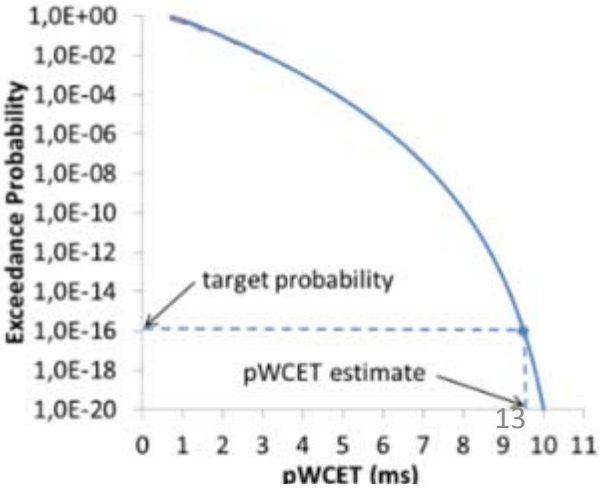


Probabilistic Timing Analysis (PTA)

Probabilistic Timing Analysis (PTA)

- Different WCETs with associated probabilities
- Approach: introduce randomization in the time behavior of HW and SW and apply probabilistic and statistical techniques
- Most techniques presented here from:

PROARTIS PROXIMA

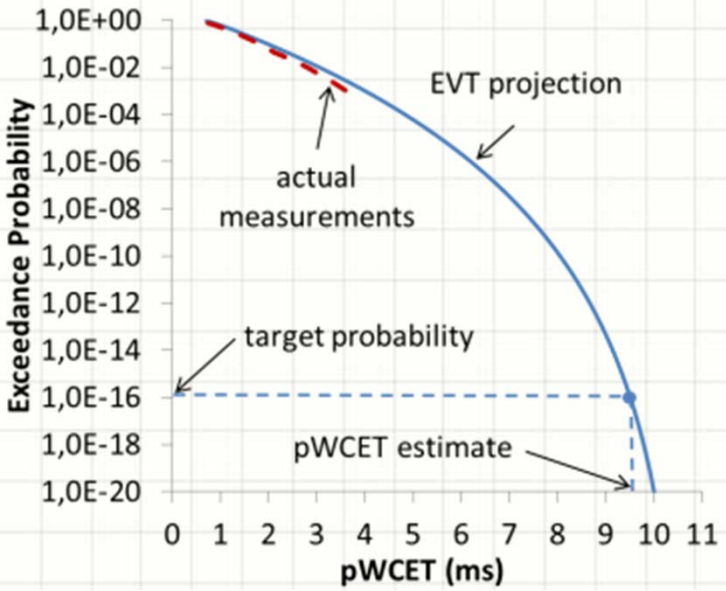


Static PTA (SPTA)[6][2][9]

- Execution Time profile (ETP)
  - ETP = $<(L_1, L_2, \ldots, L_n),(P_1, P_2, \ldots P_n)>$
- Along each path instructions combined using convolution $\otimes$
  - Inherits $\otimes$ requirements
- Require probabilities in the ETP
  - To be computed or upper-bounded
  - Probability (or its bound) should be independent from history of execution
  - Example: ETP of instruction using probability of hit from [9]: if k>=N, 0, else $((N-1)/N)^k$

Measurement-based PTA (MBPTA)

- Works with end-to-end runs
  - Uses Extreme Value theory (EVT)
  - Inherits EVT reqs (i.i.d) and also has its own [1]
- Probabilities must exist, not to be computed
  - Approximation expression to show probabilistic behavior of cache hit/misses [16][17]



Qualitative comparison

Suitability: cache designs for SDTA and SPTA

- Time Deterministic caches: Works for SDTA
- Time Randomized caches: Pessimistic for SDTA. OK for MBPTA. SPTA not shown to work yet
- pTR caches (partially Time Randomised)
  - Dependence among addresses and cache lines
  - SPTA:
    - Requires addresses (or alignments) of accesses
    - As much information as SDTA → defies PTA goal [JR]

Policy		SDTA	SPTA	MBPTA
Placement	Replacement			
TD	TD	yes	no	no
TD	TR	pessimistic	yes	yes
TR	TR	very pessimistic	no	yes

[JR] Randomized Caches Considered Harmful in Hard Real-Time Systems . Jan Reineke . LITES.

Other elements of comparison

- Multipath programs: affect analysis
  - SDTA: safely supported
  - MBPTA: path coverage issue → PUB (next talk ☺)
  - SPTA: Some first methods already [9]
- Sensitivity to the lack of information
  - Get addresses (alignments) of accesses is complex
  - Impact of lack of addresses/alignments
    - SDTA (TD cache): sensitive
    - MBPTA (TR caches): insensitive
    - SPTA (pTR caches): sensitive

Other challenges

- SPTA for random placement
  - Random placement not modelled yet [JR]
  - Mod. placement: assume accesses go to the same set
- SPTA tighter hit/miss
  - Baseline upperbound formula [9]
    - Approximation formulas in [16][17] not meant for SPTA
- Trustworthiness of MBPTA for random placement [JR] → HoG technique (last talk)

[JR] Randomized Caches Considered Harmful in Hard Real-Time Systems . Jan Reineke . LITES.

Quantitative comparison

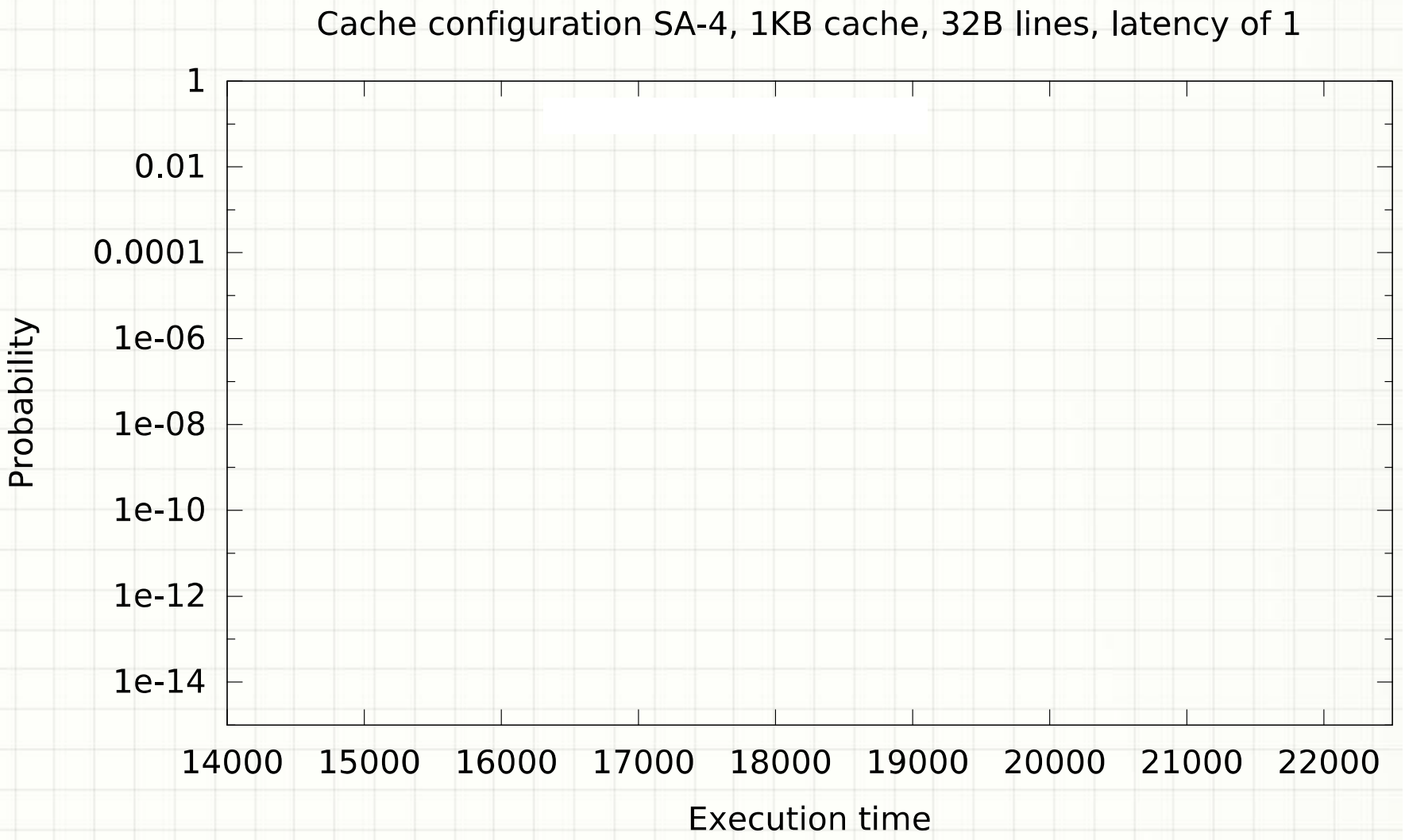
Experimental setup

- Common denominator supported by all methods
  - Programs: Single path code & addresses are known
    - Mälardalen WCET benchmarks
    - Autobench benchmarks
  - Common ISA (MIPS) + compiler toolchain (gnu)
- Instruction cache design
  - 1KB cache – 32B lines
  - Structure
    - 4-way Set-Associative (SA-4), latency: 1/70 cycles
    - Fully Associative (FA), latency 2/70 cycles
  - Placement: deterministic (modulo)
  - Replacement
    - Deterministic (LRU)
    - Random (EoM)

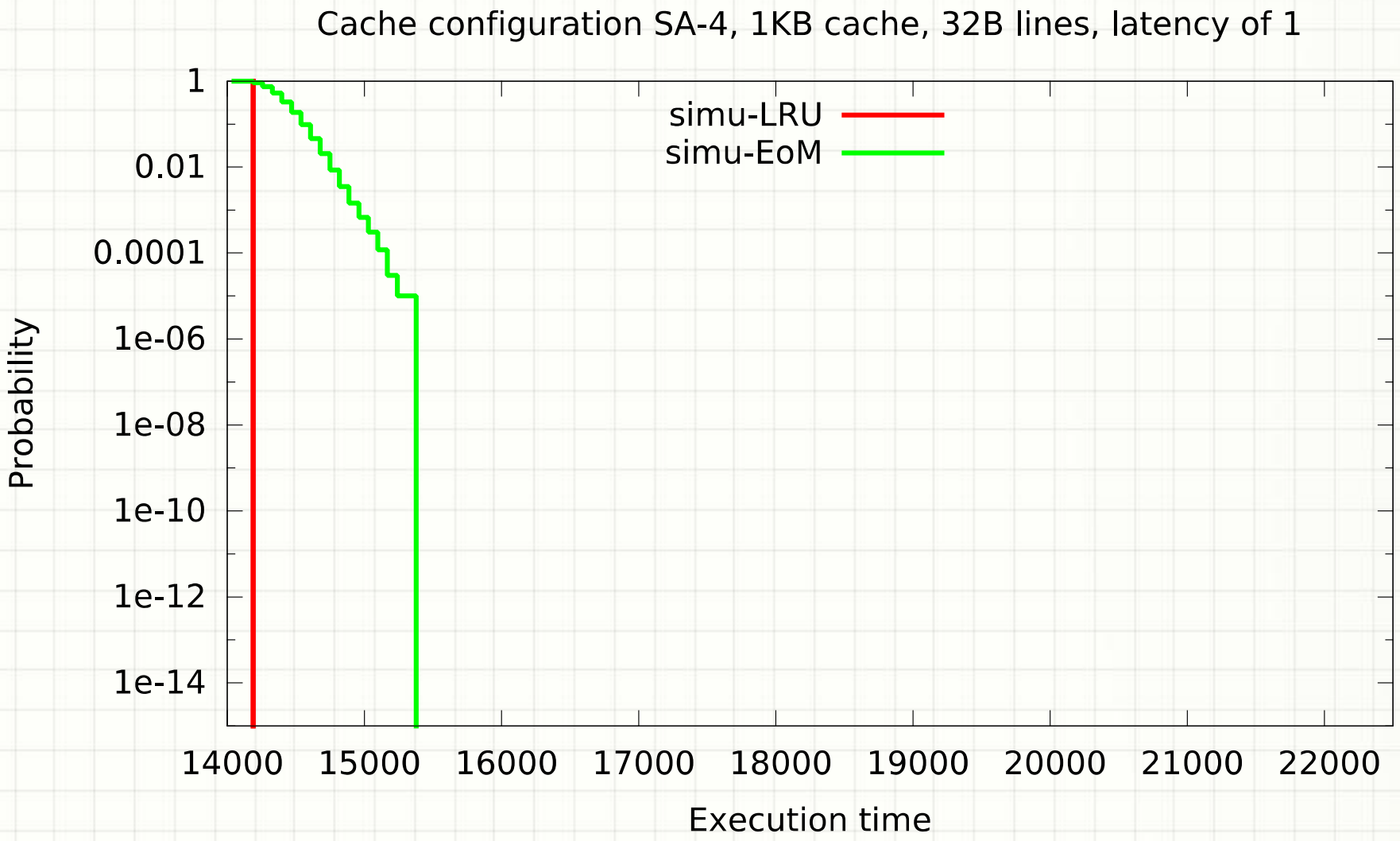
WCET estimation techniques

- SDTA
  - Heptane WCET estimation tool
    - AI-based analysis of cache + IPET
- SPTA
  - Formula from [9] to derive ETPs on traces
  - Convolutions of ETPs
- MBPTA
  - End-to-end measurements + EVT
  - Statistical test to check i.i.d property

pWCET distributions (SA-4, minver)



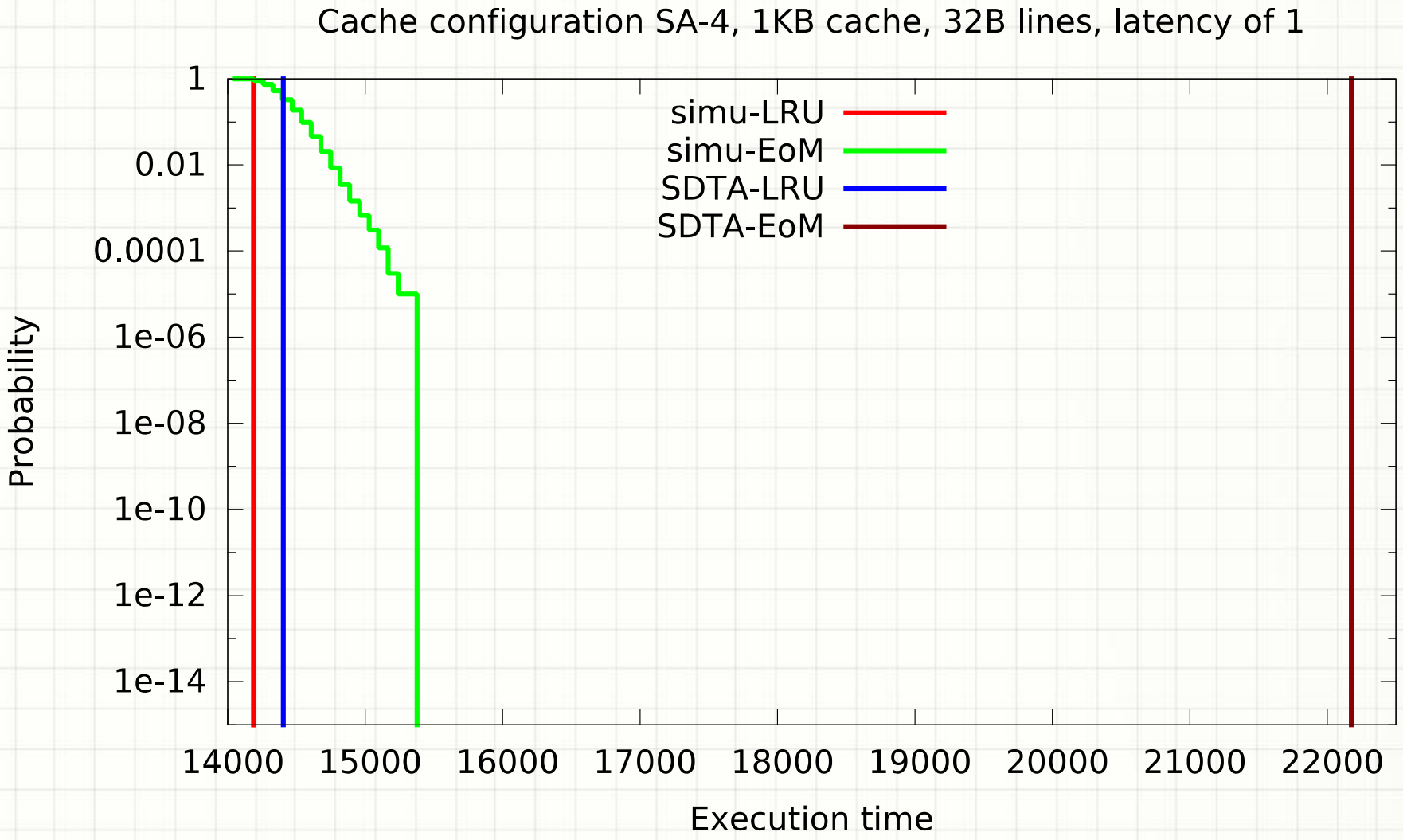
pWCET distributions (SA-4, minver)



simu-LRU: 1 run (single path)

simu-EoM: 100K runs

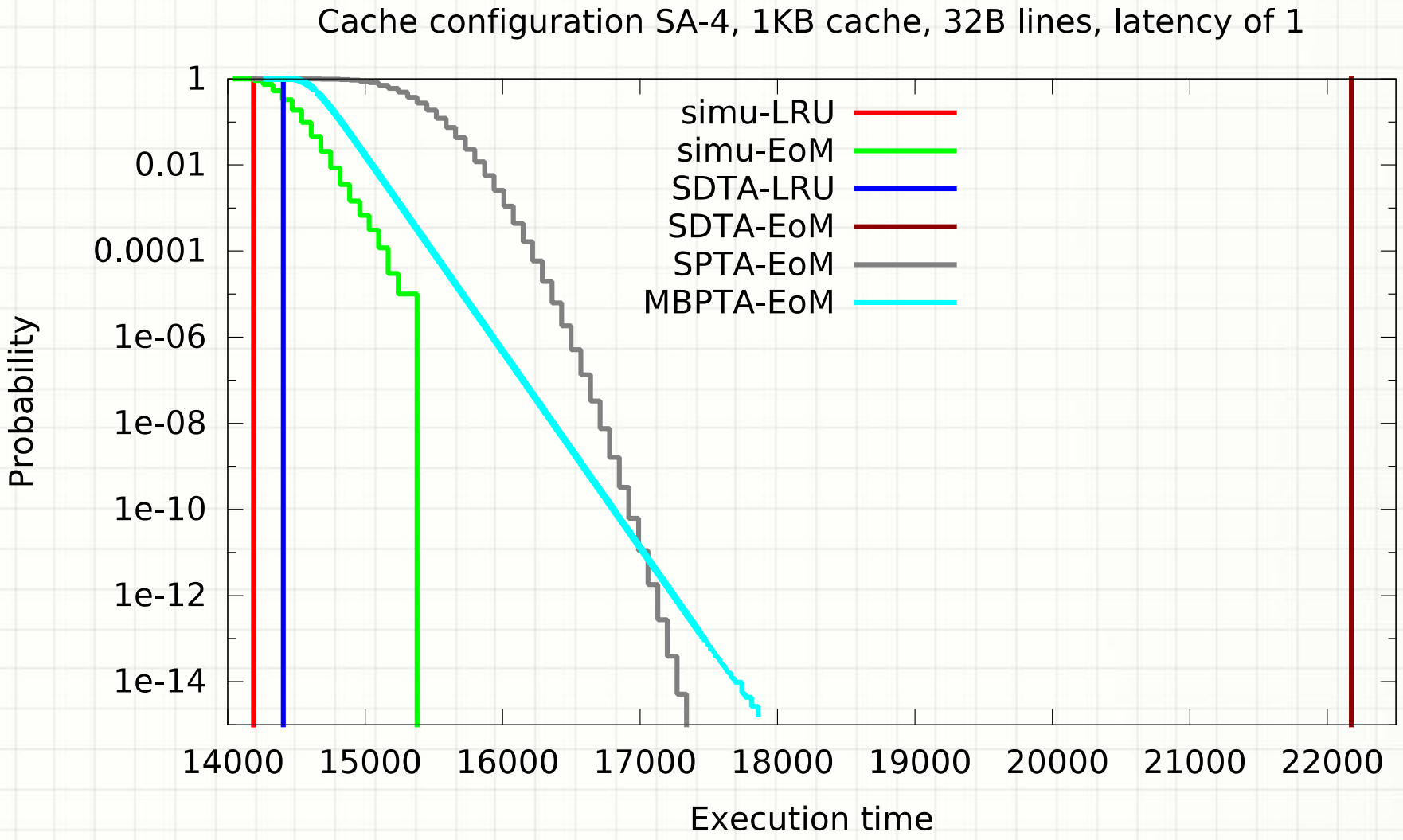
pWCET distributions (SA-4, minver)



SDTA-LRU: close to simu-LRU

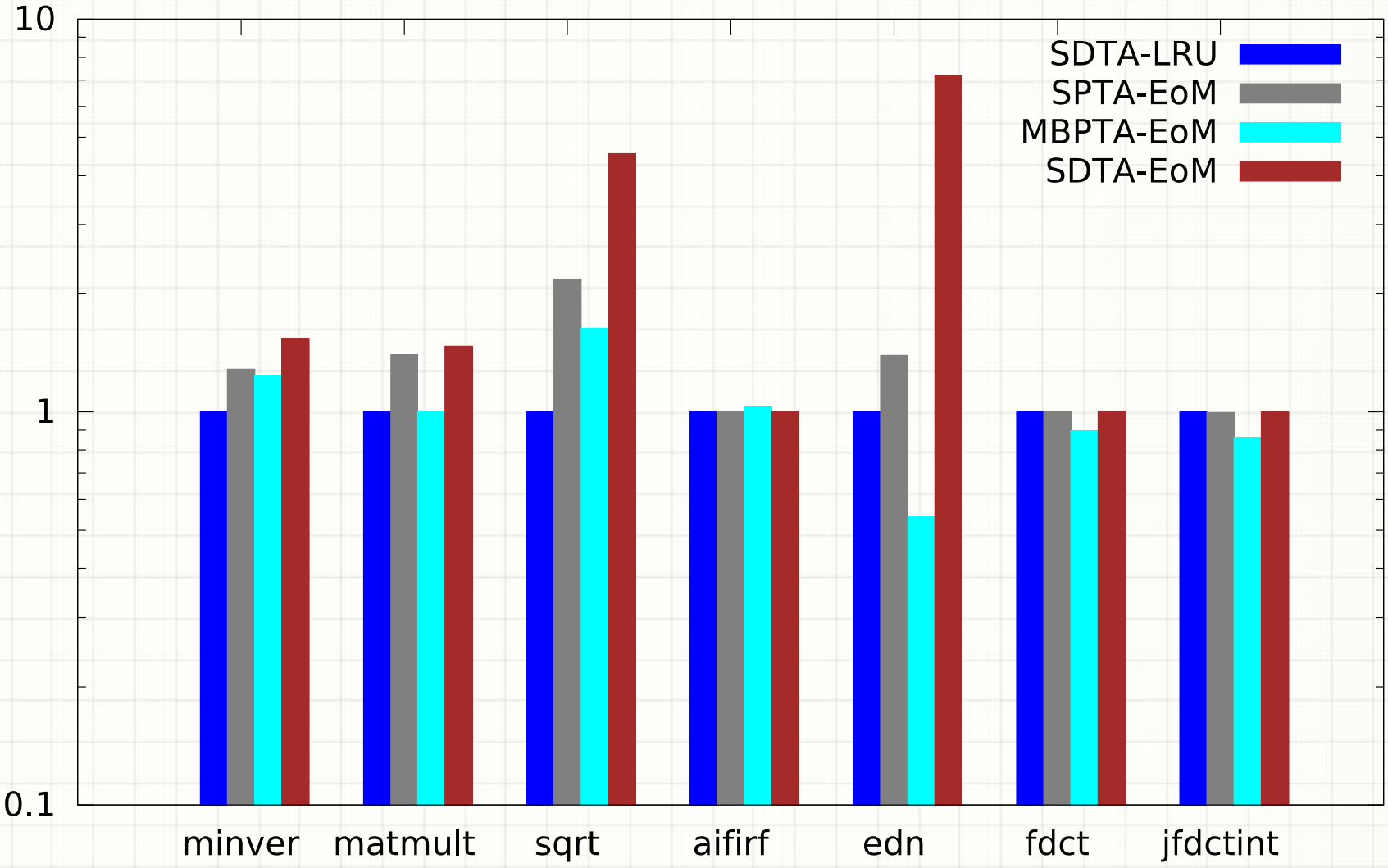
SDTA-EoM: pessimistic estimation (1/4 of the cache)

pWCET distributions (SA-4, minver)

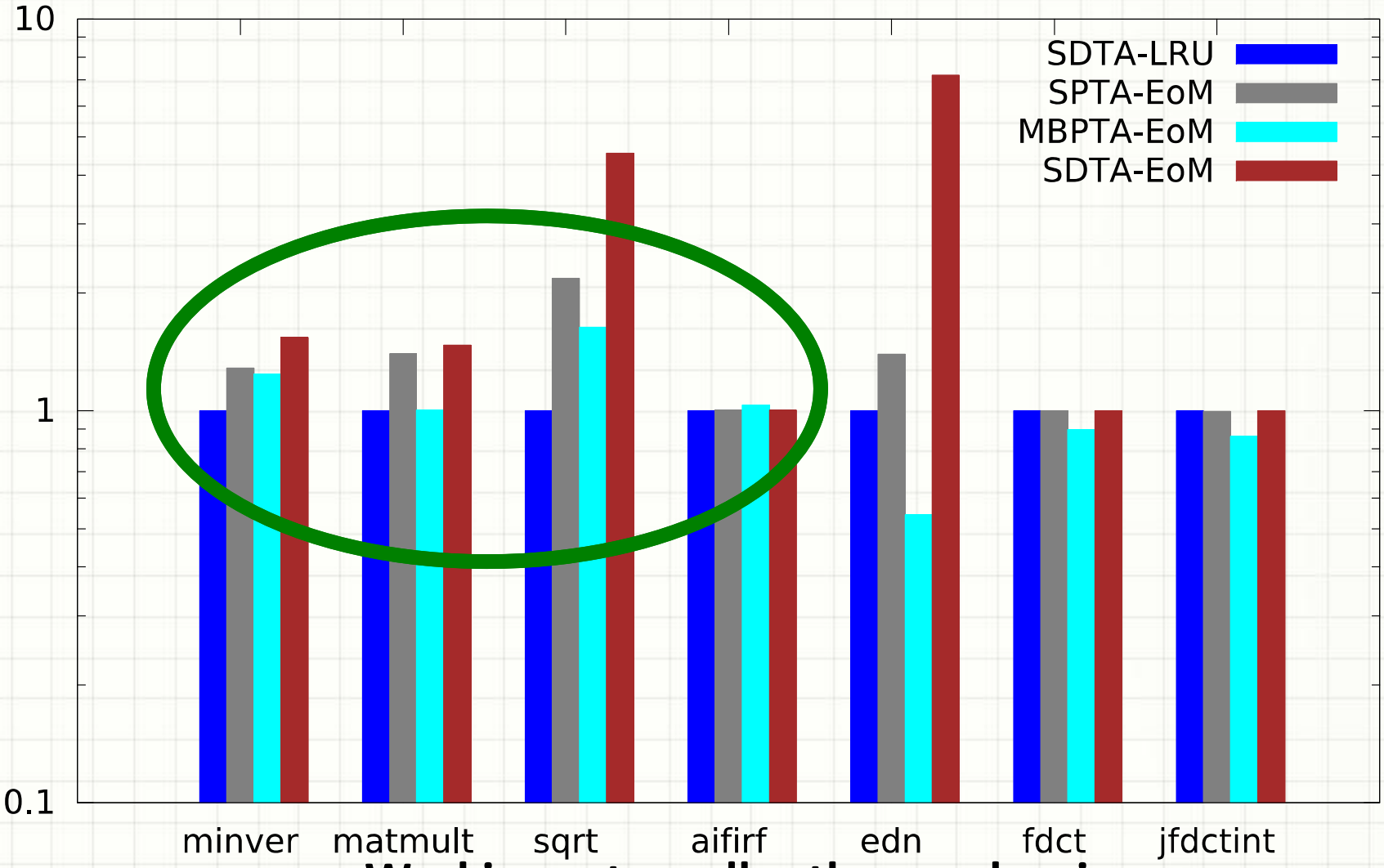


SPTA-EoM and MBPTA-EoM: strictly higher than simu-EoM & pWCET estimates between SDTA-LRU and SDTA-EoM

Comparison at fixed probability (10^{-15}, SA-4)



Comparison at fixed probability (10^{-15}, SA-4)

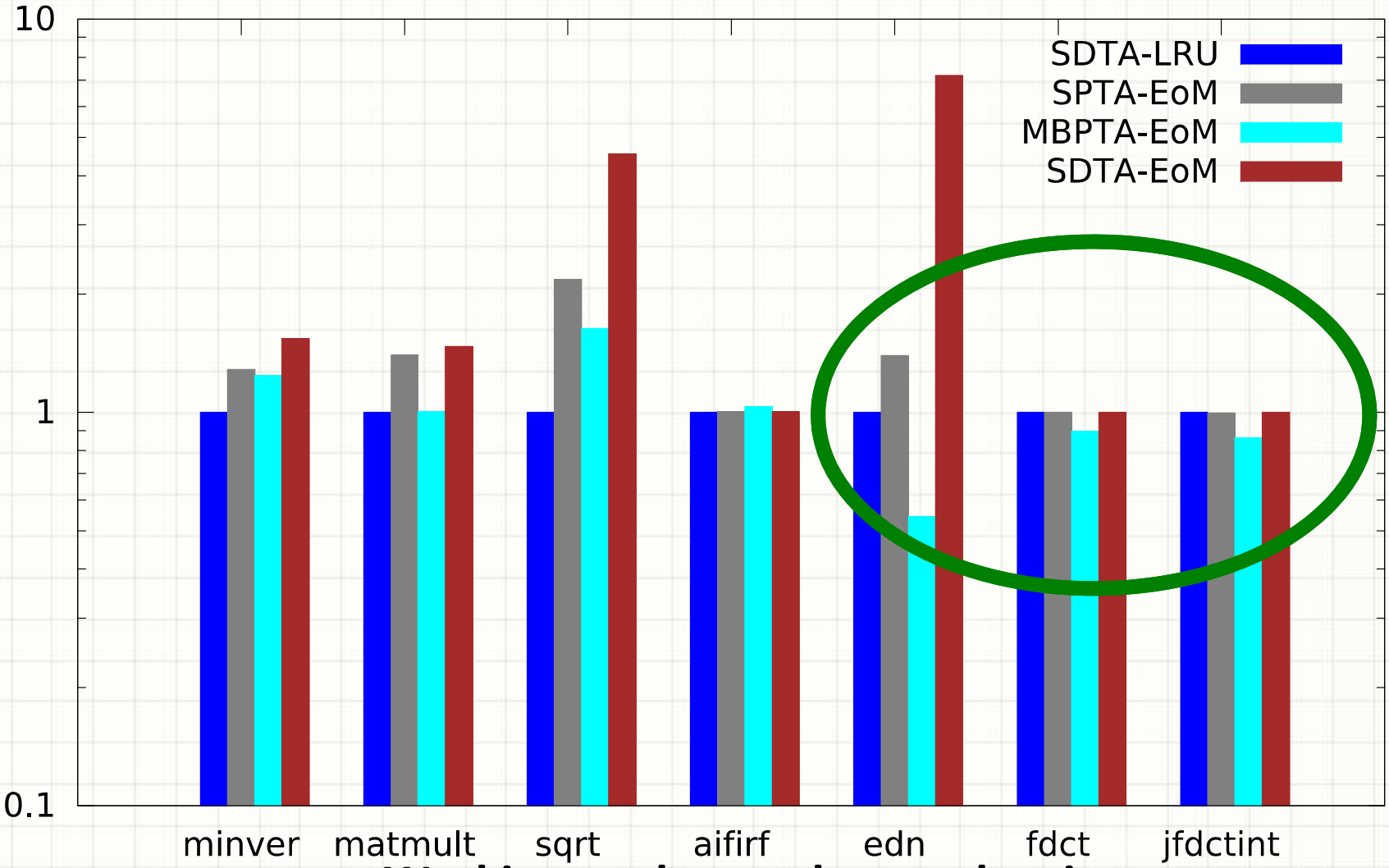


Working set smaller than cache size

LRU: only cold misses

EoM: non zero probability of replacements to occur

Comparison at fixed probability (10^{-15}, SA-4)



Working set larger than cache size

LRU: conflict misses

EoM: non-null probability of survivability

Conclusions

Conclusions & Future work

- We proposed a first comparison between deterministic and probabilistic methods
  - Qualitative & Quantitative studies to identify the HW & SW for which each method perform best
- Next step
  - Multi-path programs
  - Data caches
  - Sensitivity to the lack of address information
- Challenges
  - SPTA tighter hit/miss probability estimation
  - SPTA for random placement
  - …

ON THE COMPARISON OF DETERMINISTIC AND PROBABILISTIC WCET ESTIMATION TECHNIQUES

Jaume Abella[1,4], *Damien Hardy*[2], *Isabelle Puaut*[2], *Eduardo Quinones*[1,4], *Francisco J. Cazorla*[1,3,4]

1 Barcelona Supercomputing Center
Centro Nacional de Supercomputación

2 Université de Rennes 1

3 CSIC Consejo Superior de Investigaciones Científicas

4 PROXIMA

26th Euromicro Conference on Real-Time Systems (ECRTS14)
8-11 July Madrid, Spain

ECRTS 2014
8-11 July
Madrid, Spain