Improvements to Static Probabilistic Timing Analysis for Systems with Random Cache Replacement Policies

Rob Davis

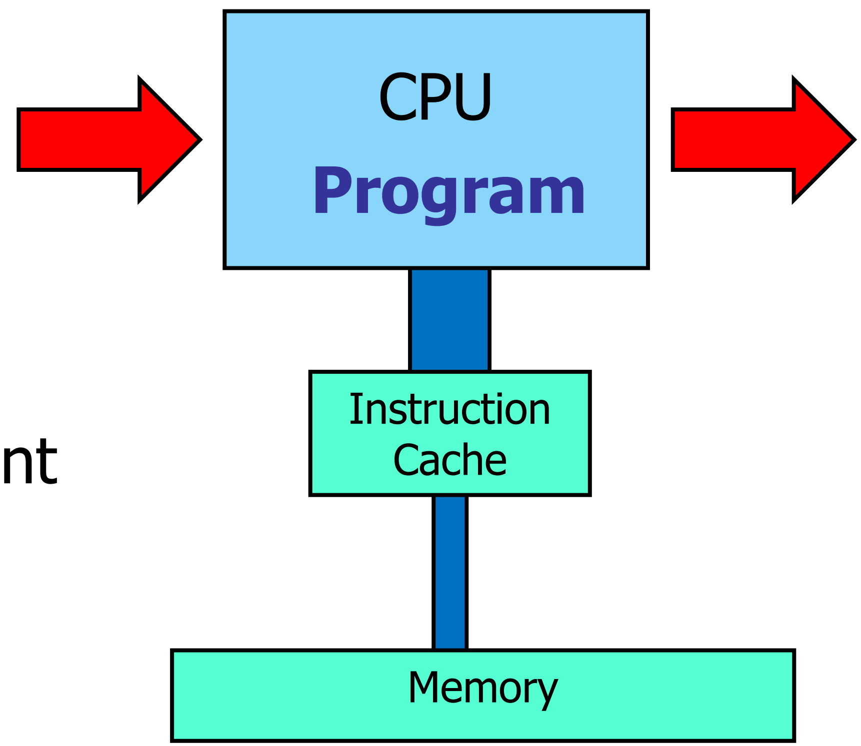
Real-Time Systems Research Group,
Department of Computer Science,
University of York, York, UK.

rob.davis@york.ac.uk

Static Probabilistic Timing Analysis (SPTA)

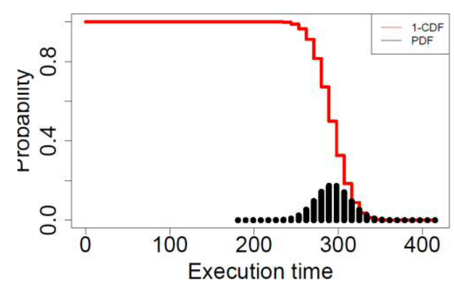
- Aim is to show that the probability of timing failure falls below some threshold e.g. 10^{-9} failures per hour: pWCET v. budget

Inputs

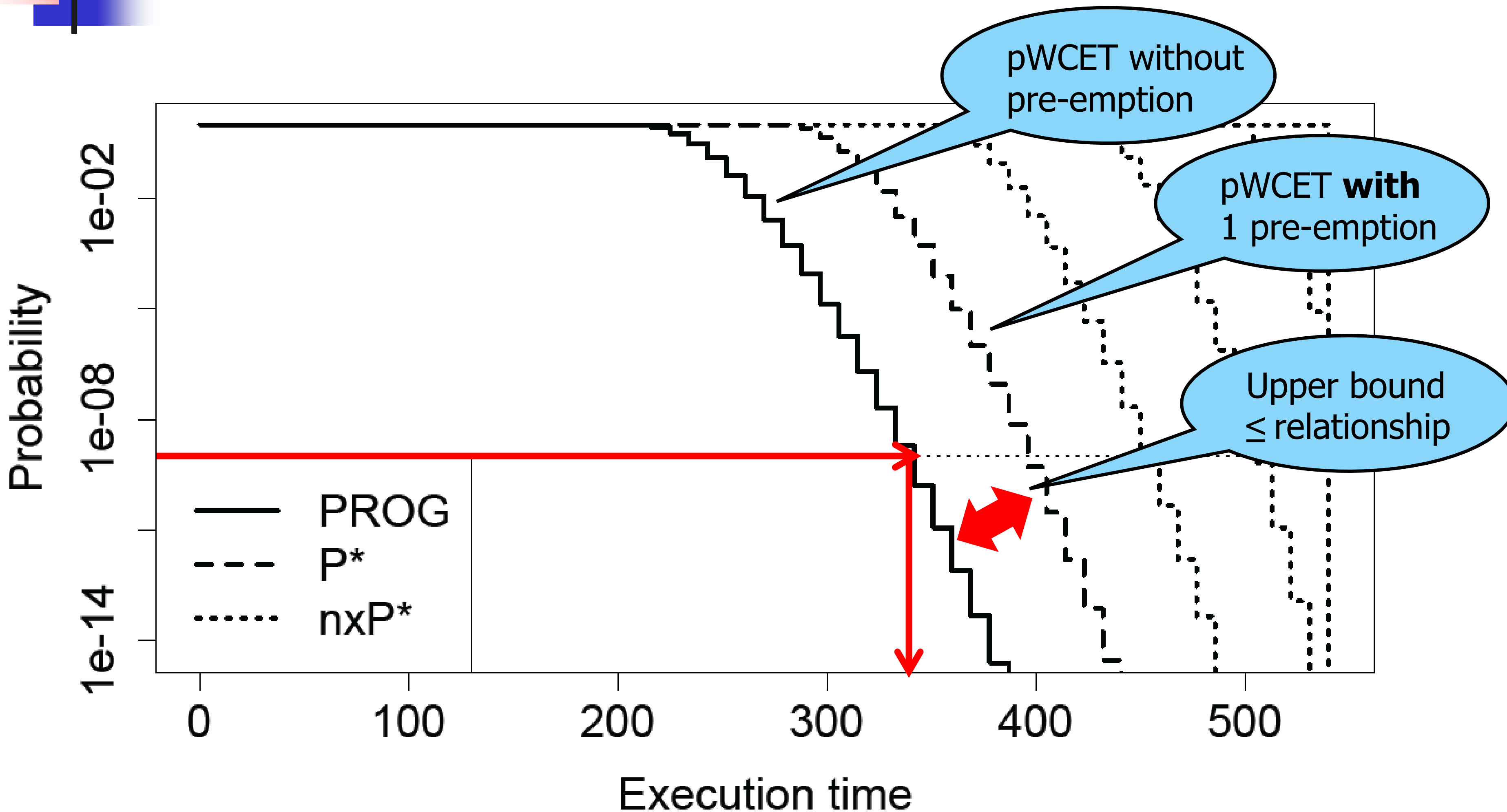


Probabilistic WCET (pWCET) distribution

Random replacement policy



pWCET distribution (1-CDF)



Simple model of execution

- Instructions are either:
  - Cache hit or cache miss
  - Misses take longer (H = 1 cycle, M = 10 cycles)
- Fully associative cache of N blocks
  - Memory blocks can be loaded into any block in cache
  - Each instruction resides in a memory block
- On a cache miss
  - Random choice of cache block to evict
  - Evict that block, load the requested block into the evicted location
- Probability of a cache hit:

$$P_{hit}(k) = \left(\frac{N-1}{N}\right)^k$$

(when $k < N$ otherwise 0)

  - k is **re-use distance** = number of intervening evictions since the memory block was last loaded into cache

Static Probabilistic Timing Analysis (for single path programs)

- Sequence of instructions represented by their memory blocks

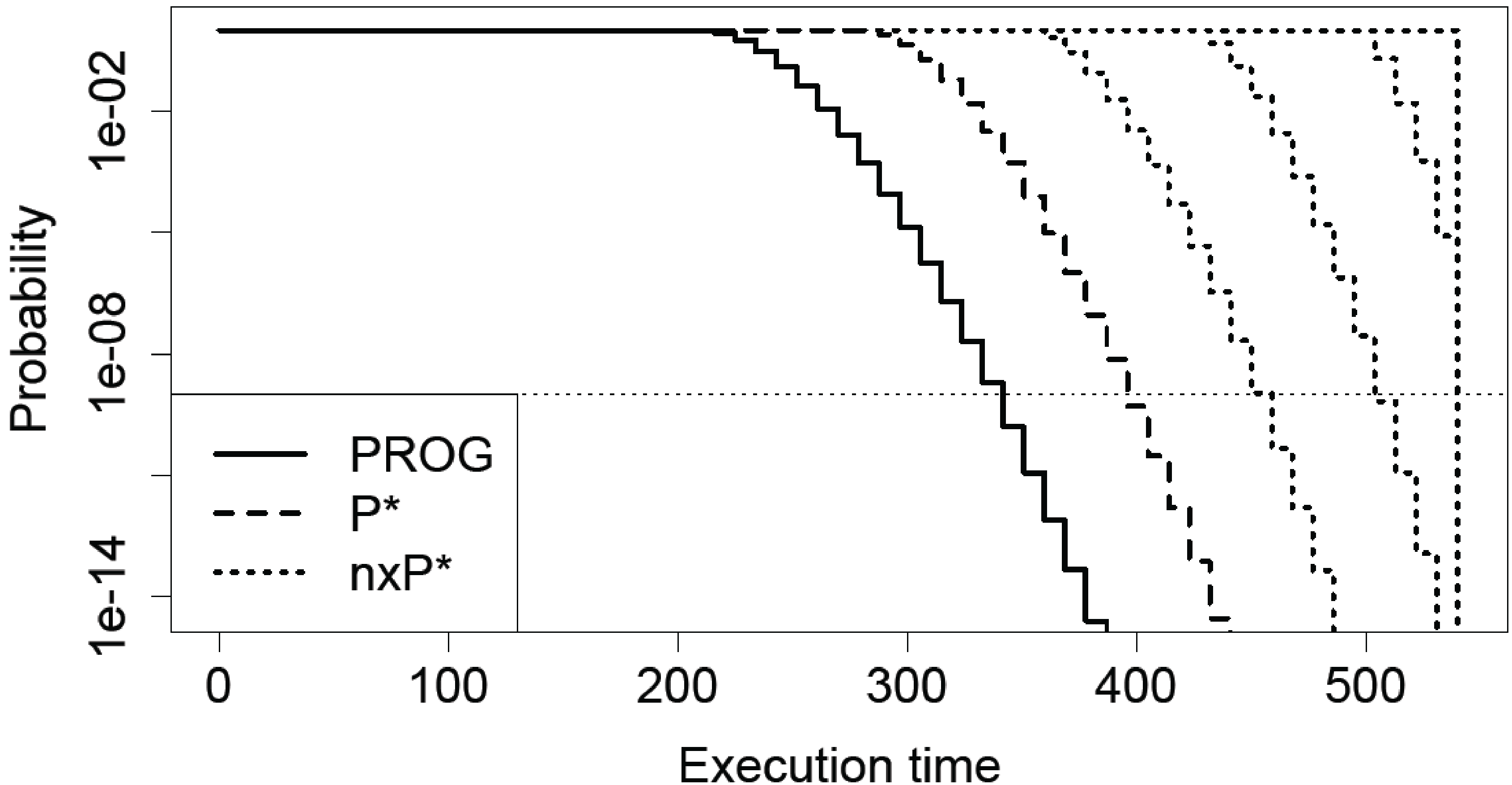
  $a, b, a^1, c, d, b^3, c^2, d^2, a^5$

- Get a probability distribution (pWCET) for each instruction
  - Depends only on re-use distance *k*
  - Possible to model instructions as independent, hence we can convolve distributions for instructions to get a pWCET distribution for a sequence of instructions

    E.g. two instructions with P_{hit} = 0.8 and 0.7

$$\begin{pmatrix} 1 & 10 \\ 0.8 & 0.2 \end{pmatrix} \otimes \begin{pmatrix} 1 & 10 \\ 0.7 & 0.3 \end{pmatrix} = \begin{pmatrix} 2 & 11 & 20 \\ 0.56 & 0.38 & 0.06 \end{pmatrix}$$

pWCET distribution (1-CDF)



SPTA has some pessimism

- Sequence of instructions represented by their memory blocks
  $a, b, a^1, c, d, b^3, c^2, d^2, a^5$

- Consider the a^5
  - 5 because of the intervening instructions c, d, b^3, c^2, d^2
  - c, d, are definitely misses
  - b^3, c^2, d^2 considered as misses when analysing a^5

Pessimistic because the probability that b^3, c^2, d^2 are all misses is already $< 7.1 \times 10^{-7}$ (with $N = 256$)

How can we obtain a tighter pWCET that is still correct (not optimistic)?

Reducing the pessimism

- *"A Cache Design for Probabilistic Real-time Systems"*, DATE 2013 [5]

$$P_{hit} = \left(\frac{N-1}{N}\right)^{\sum P_{miss}}$$

Sum of probabilities of cache misses of intervening instructions

- But is it correct?

  consider a, b, a^1, b^1 with $N=2$

  for b^1 $\quad P_{hit} = \left(\frac{1}{2}\right)^{1/2} = 1/\sqrt{2}$

  Irrational value for a probability ?

Counter example: Analysis from [5]

Consider a, b, a^1, b^1 with $N=4$

- Distributions for a, b, a^1 $\begin{pmatrix} 10 \\ 1 \end{pmatrix}$ $\begin{pmatrix} 10 \\ 1 \end{pmatrix}$ $\begin{pmatrix} 1 & 10 \\ 0.75 & 0.25 \end{pmatrix}$
- For b^1 $\left(\frac{3}{4}\right)^{0.25} = 0.9306$ according to [5]
- Hence

$$\begin{pmatrix} 10 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 10 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 10 \\ 0.75 & 0.25 \end{pmatrix} \otimes \begin{pmatrix} 1 & 10 \\ 0.9306 & 0.0694 \end{pmatrix} = \begin{pmatrix} 22 & 31 & 40 \\ 0.69795 & 0.2847 & 0.01735 \end{pmatrix}$$

Counter example: Precise analysis

Consider a, b, a^1, b^1 with N =4. Two cases:

- Case 0: a^1 is a hit (probability of occurrence = 0.75)
  - Given that a^1 is a hit then b^1 is guaranteed to also be a hit

Partial pWCET = $\begin{pmatrix} 10 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 10 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0.75 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 22 \\ 0.75 \end{pmatrix}$

- Case 1: a^1 is a miss (probability of occurrence = 0.25)
  - Given that a^1 is a miss then b^1 has P_{hit} = 0.75

Partial pWCET = $\begin{pmatrix} 10 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 10 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 10 \\ 0.25 \end{pmatrix} \otimes \begin{pmatrix} 1 & 10 \\ 0.75 & 0.25 \end{pmatrix} = \begin{pmatrix} 31 & 40 \\ 0.1875 & 0.0625 \end{pmatrix}$

- Overall pWCET = $\begin{pmatrix} 22 & 31 & 40 \\ 0.75 & 0.1875 & 0.0625 \end{pmatrix}$

This is precise – we covered all possibilities

Counter example: comparison

Consider a, b, a^1, b^1 with $N = 4$.

- Precise analysis:

$$\begin{pmatrix} 22 & 31 & 40 \\ 0.75 & 0.1875 & 0.0625 \end{pmatrix}$$

Exact but exponential complexity

- Analysis from [5]:

$$\begin{pmatrix} 22 & 31 & 40 \\ 0.69795 & 0.2847 & 0.01735 \end{pmatrix}$$

OPTIMISTIC

- Simple analysis from [3]:

$$\begin{pmatrix} 22 & 31 & 40 \\ 0.5625 & 0.375 & 0.0625 \end{pmatrix}$$

Pessimistic but Ok

Open Problem: Can we tighten the pWCET (1-CDF) found by SPTA?

- Sequence of instructions represented by their memory blocks

  $a, b, a^1, c, d, b^3, c^2, d^2, a^5$

  with re-use distances
- Probability of a hit for a single instruction (for $k < N$)

$$P_{hit}(k) = \left(\frac{N-1}{N} \right)^k$$

- Convolve pWCET distributions for individual instructions to get overall pWCET distribution for the sequence
- Existing analysis is simple but somewhat pessimistic as intervening instructions are not necessarily certain to be misses
- Precise analysis is exponential in complexity

Can we find a tighter upper bound on the pWCET that can be computed efficiently?

Extent of the pessimism