Quality of Service Capabilities for Hard Real-Time Applications on Multi-core Processors

Real-Time Networks and Systems (RTNS)

16.-18.10.2013

Jan Nowotsch, Michael Paulitsch

University Augsburg

Advisor: Theo Ungerer

EADS Innovation Works

Advisor: Michael Paulitsch

Motivation and Background

Motivation

- Ongoing trend to deploy multi-core processors for real-time embedded systems yields
  - Increased system performance, enabling new applications
  - Increased power-efficiency
- Major issue: Timing analysability due to the inherent sharing of resources (e.g. main memory)
- System background
  - Mixed-criticality applications (multiple parallel hard real-time applications)
  - Requirement for certification
  - Commercial Off-The-Shelf (COTS) hardware platforms

- Extension of interference-sensitive Worst-Case Execution Time (isWCET) analysis approach towards increased average-case performance
- Reasoning
  - Generally: variations between worst-case and average-case execution characteristics
  - Multi-core: even increased variations due to variations in resource access latencies

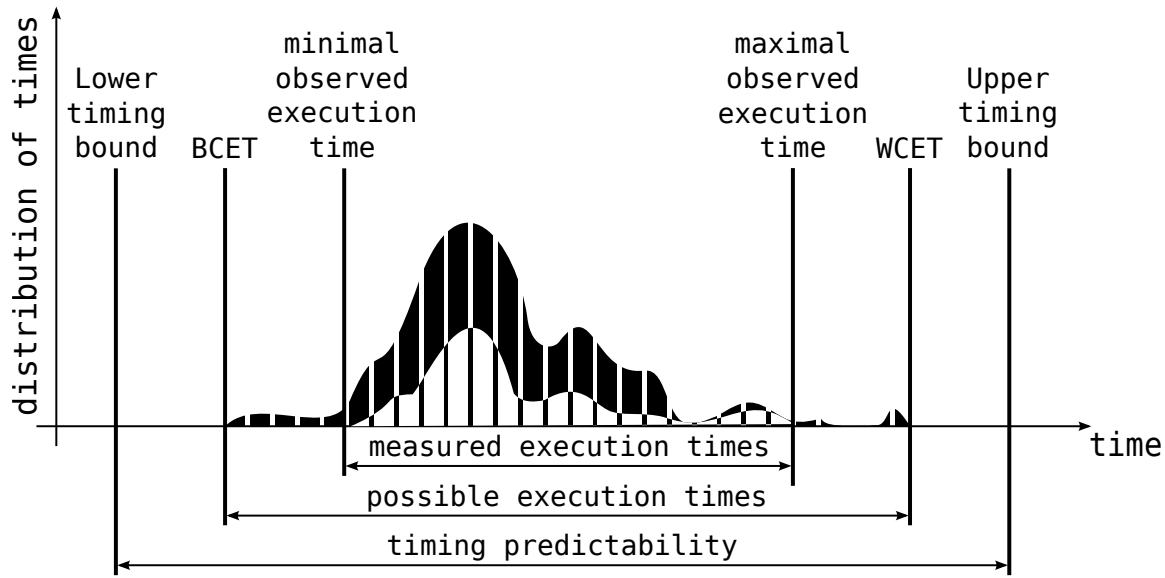
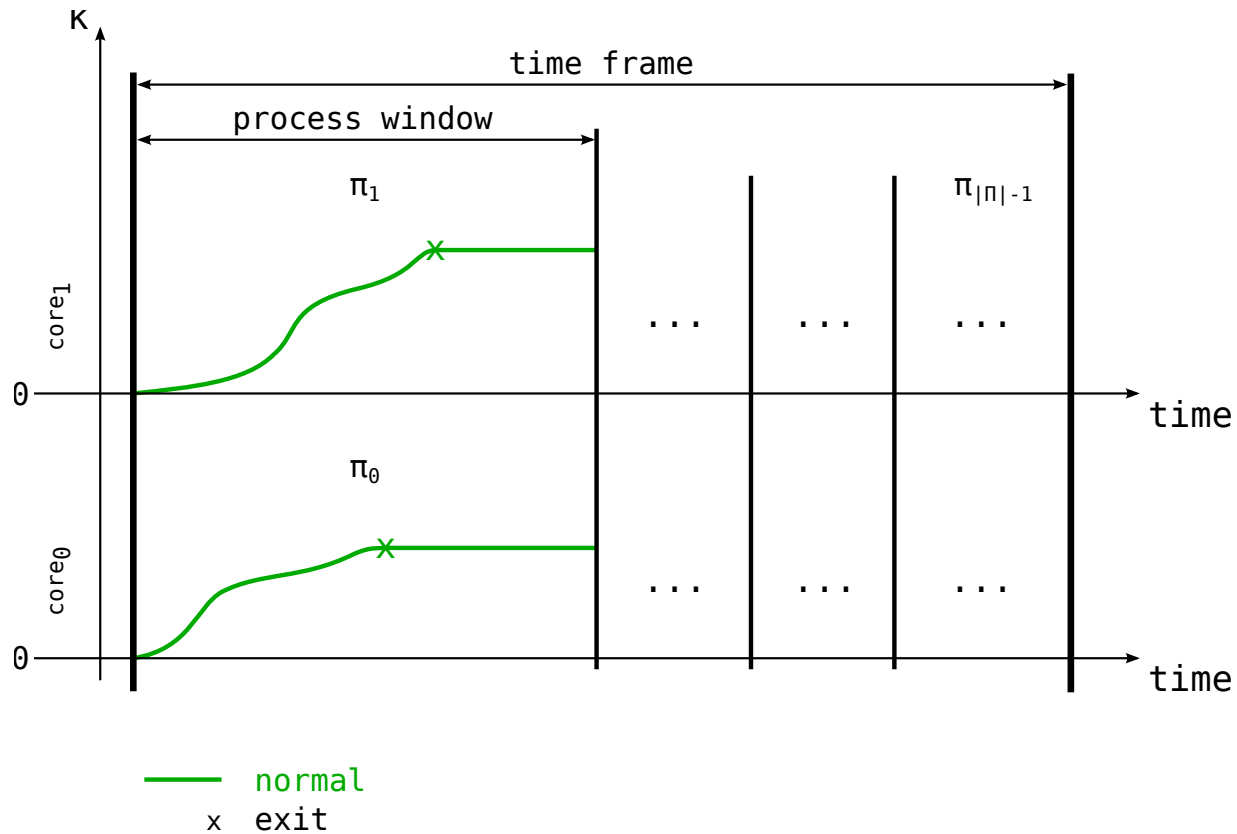


Figure: Timing analysis notations, cf. [1].

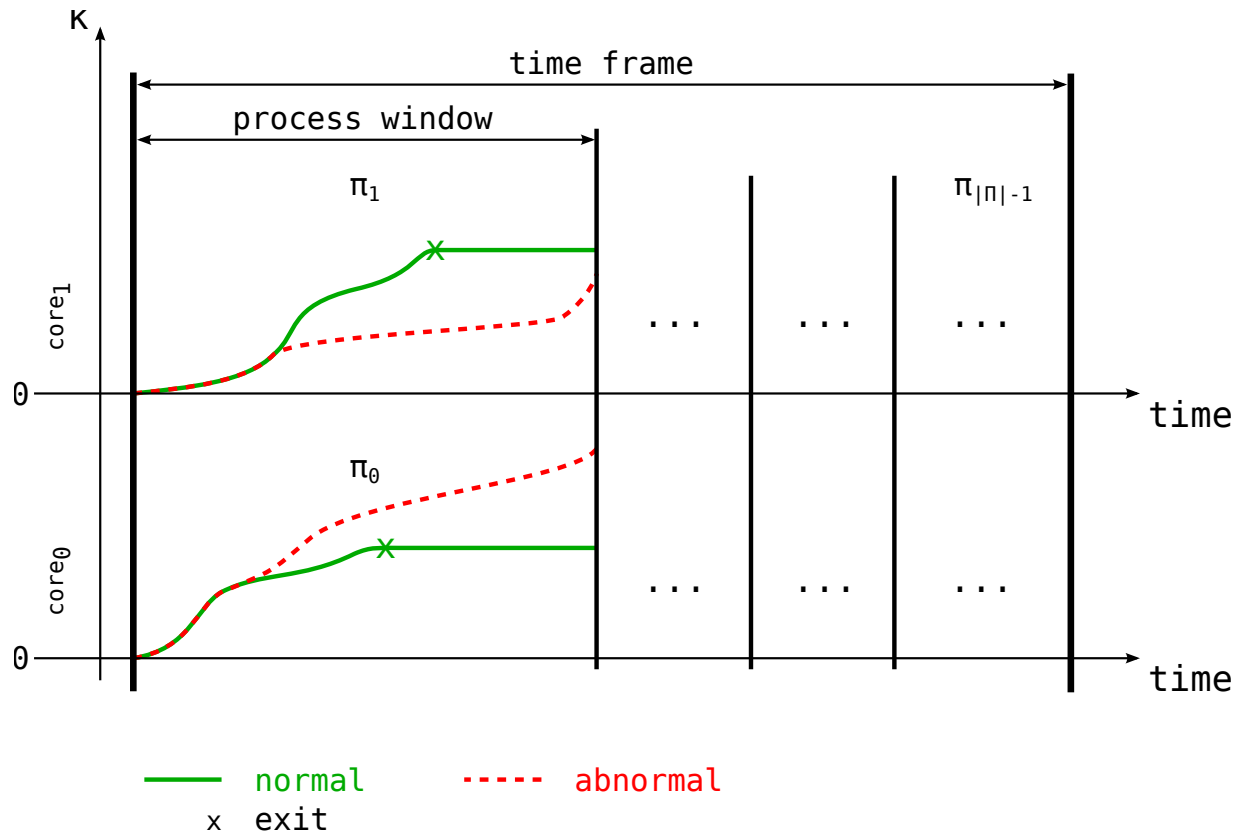
Contents

interference-sensitive WCET Analysis (isWCET)

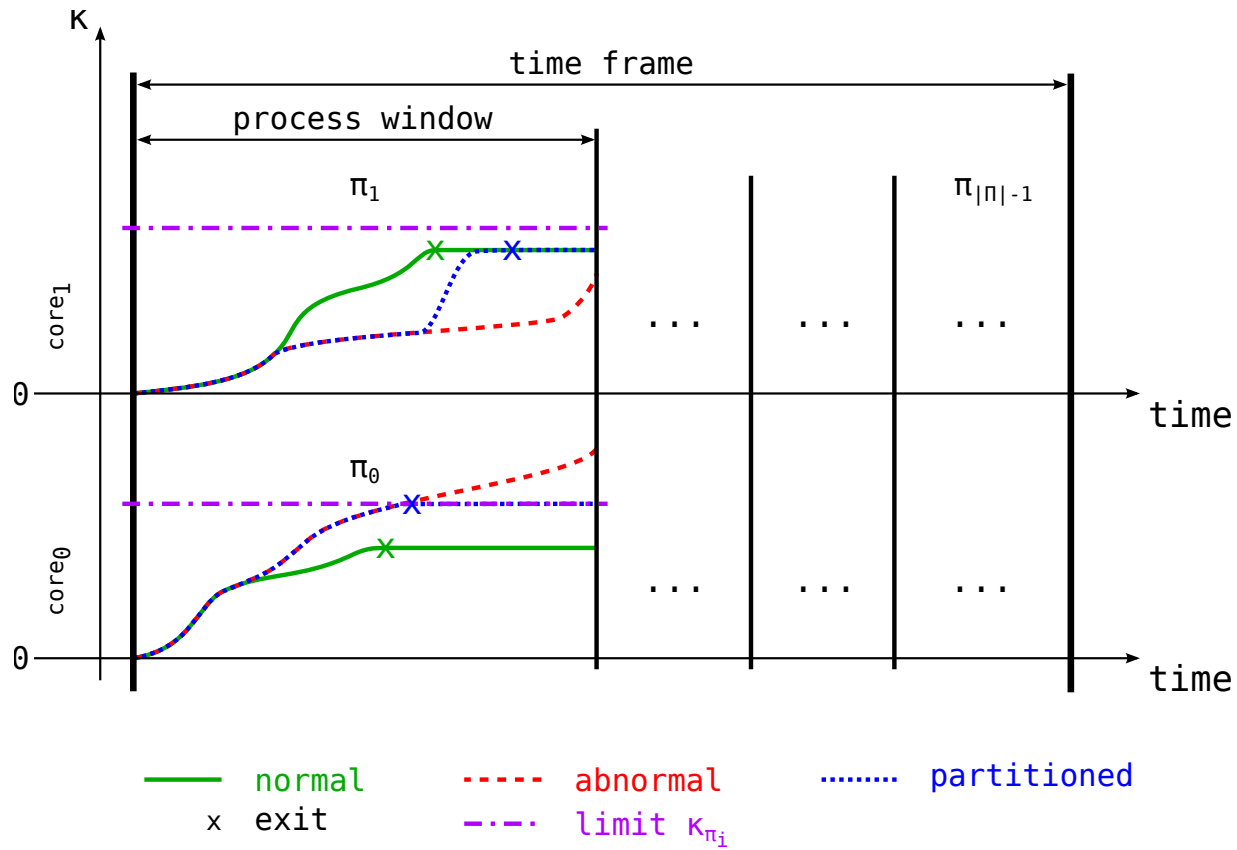
- Integrated approach of timing analysis and resource usage enforcement
- Concepts
  1. Access limitation
  2. Monitoring
  3. Suspension



- Integrated approach of timing analysis and resource usage enforcement
- Concepts
  1. Access limitation
  2. Monitoring
  3. Suspension



- Integrated approach of timing analysis and resource usage enforcement
- Concepts
  1. Access limitation
  2. Monitoring
  3. Suspension



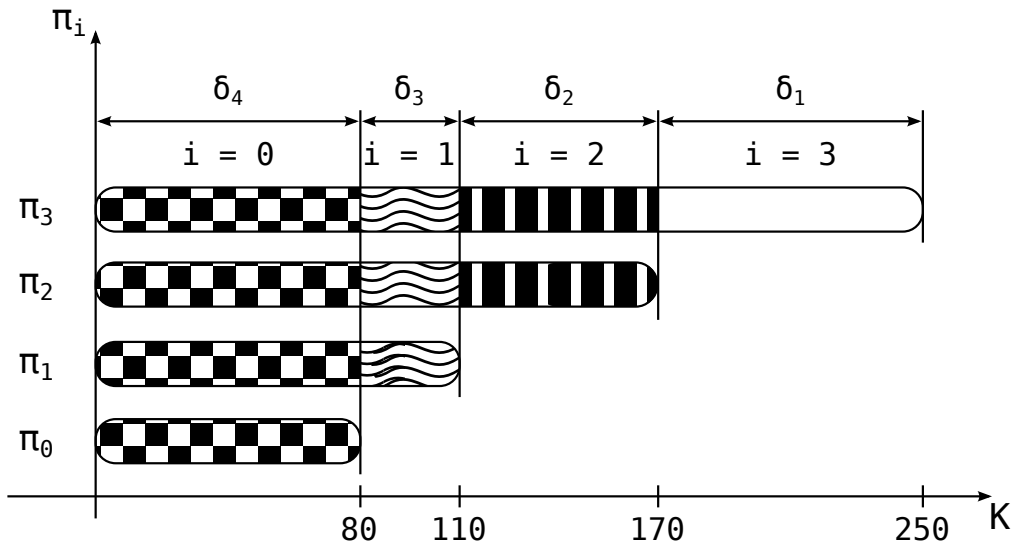
Idea

- Based on existing single-core analysis techniques
- Phases
  1. Single-core timing analysis
  2. Single-core resource usage analysis
  3. Computation of inter-process interference and isWCET

Single-core Analysis

- Determine timing behaviour - Worst-Case Execution Time (WCET) ($\tau_{s,i}$)
- Determine shared resource usage - Worst-Case number of shared Resource Accesses (WCRA) (κ_{π_i})
- Can be implemented with either measurement-based, static or hybrid techniques

isWCET



$$\begin{aligned}\tau_{id} = \quad & \delta_4 * 80 \\ & +\delta_3 * 30 \\ & +\delta_2 * 60 \\ & +\delta_1 * 80\end{aligned}$$

$$\tau_{is}(\pi_x) = \overbrace{\tau_{s,x}}^{\text{single-core bound}} + \underbrace{\delta_{|\Pi_{||}|} \cdot \kappa_{\pi_0} + \sum_{i=1}^{x} \left(\delta_{|\Pi_{||}|-i} \cdot (\kappa_{\pi_i} - \kappa_{\pi_{i-1}}) \right)}_{\text{interference delay}(\tau_{id})}$$

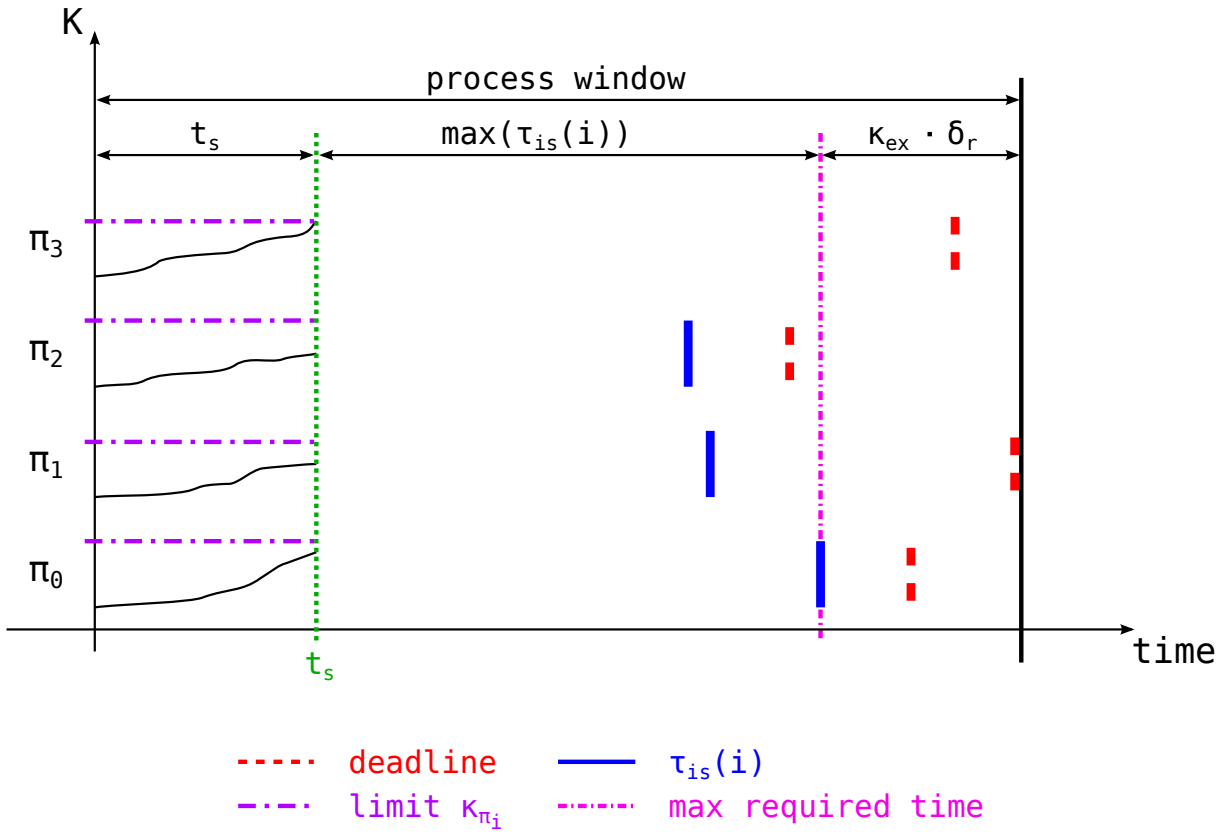
Average-Case Performance Extension

Motivation

- Overestimation of WCET and resource usage
- Additional overestimation due to access latency variations
  $\rightarrow$ high deviation between worst-case and average-case executions
  $\rightarrow$ idle resources in practice
  $\rightarrow$ reduced system utilisation

Idea

- Extend isWCET monitoring approach
- Basic isWCET: suspend process on limit violation
- Extension: dynamic adjustment of limits, based on actual system progress



$$\kappa_{ex} = \frac{\text{process window} - t_s - \max_{i=0}^{|\Pi_{||}|-1}(\tau_{is}(\pi_i))}{\delta_r}$$

r... number of active processes + 1

Evaluation

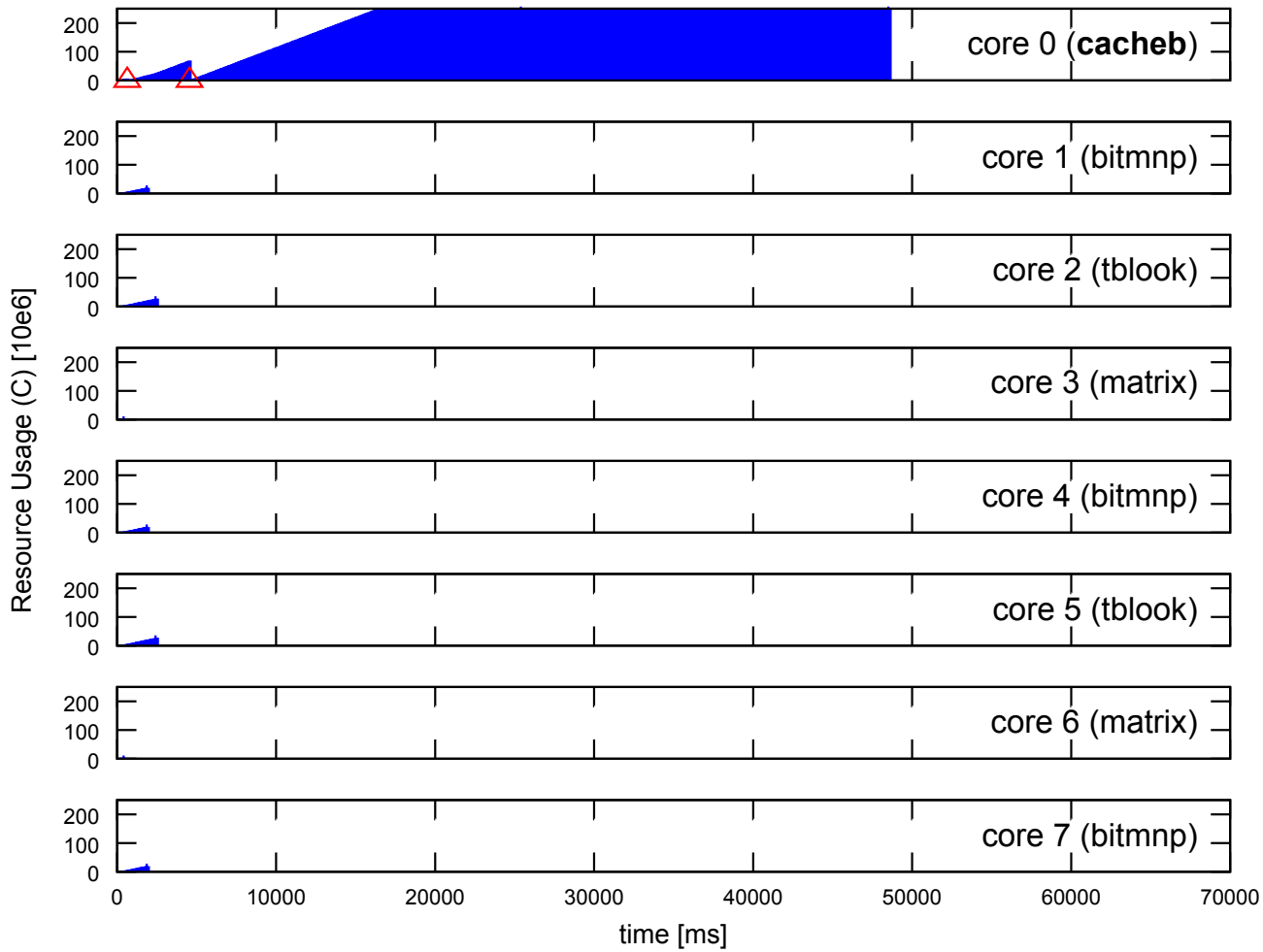
- Benchmarks: EEMBC Autobench
- Benchmark characterisation

bmark	*msens*	$a/t[1/us]$	locality	application
cacheb	5.74	50.7	low	data loading
iirflt	4.37	38.0	low	audio processing
rspeed	3.52	50.5	low	control
a2time	1.99	36.9	medium	control
bitmnp	1.09	3.2	high	display
tblook	1.28	7.2	high	image
matrix	1.44	9.1	high	optimisation
aifftr	3.61	30.5	medium	image

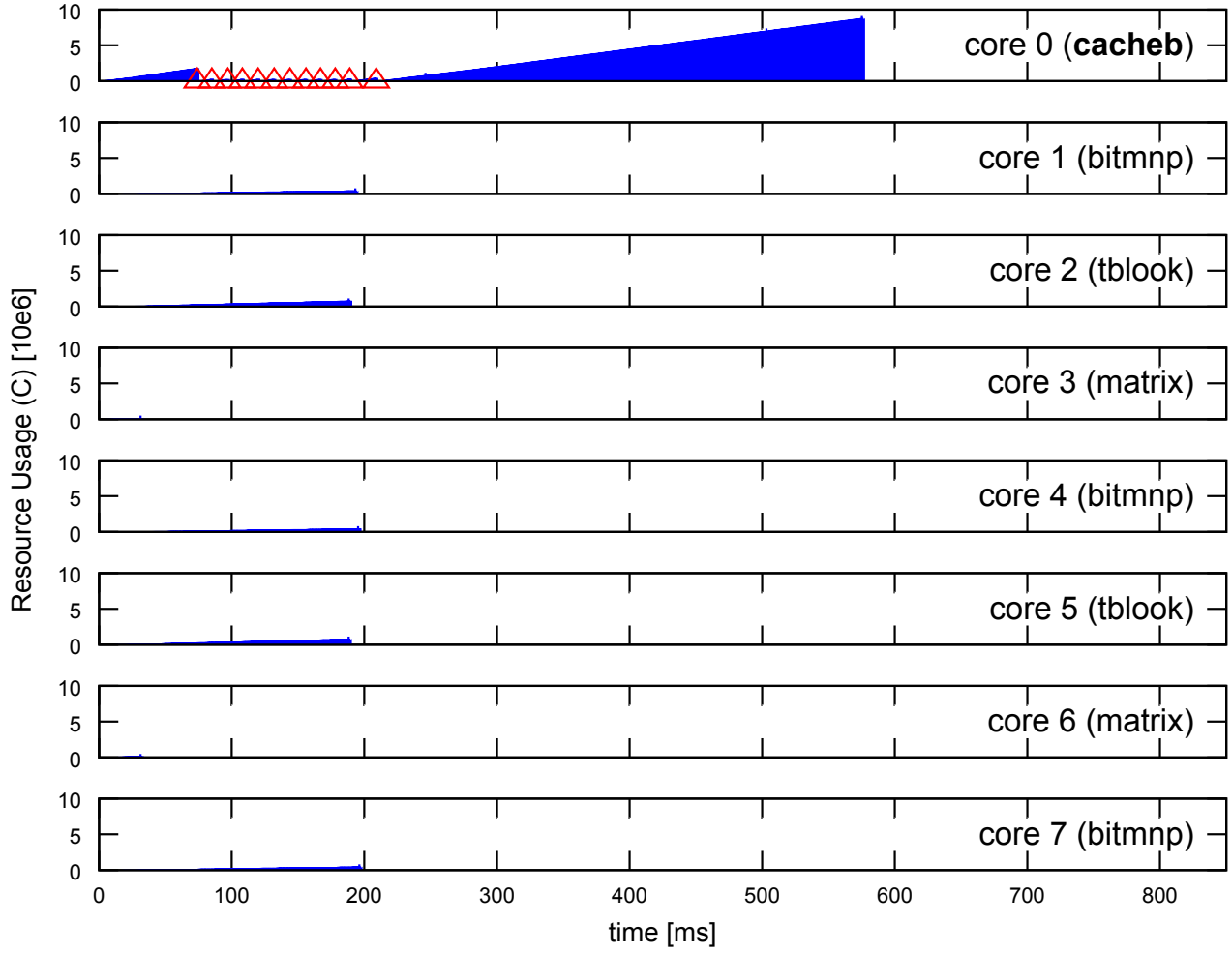
- Configurations
  1. low progress (lowp): only a single extendable application

     (**cacheb**, 3 x bitmnp, 2 x tblook, 2 x matrix) - **bold** = extension enabled
  2. real: potential real task set

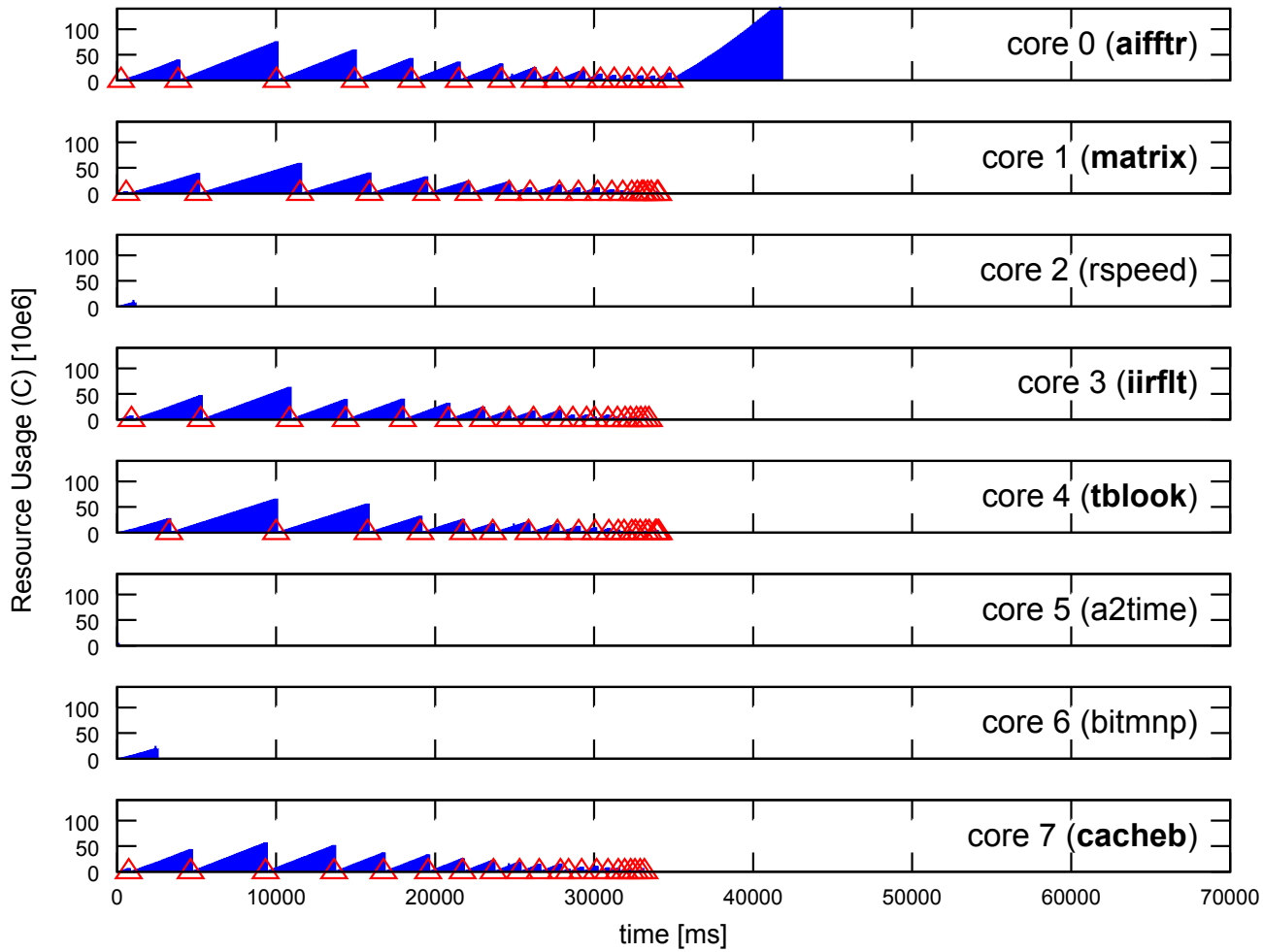
     (**aifft, matrix, iirflt, tblook, cacheb**, rspeed, a2time, bitmnp)
  3. static analysis
  4. measurement-based analysis


core 0 (**cacheb**)
core 1 (bitmnp)
core 2 (tblook)
core 3 (matrix)
core 4 (bitmnp)
core 5 (tblook)
core 6 (matrix)
core 7 (bitmnp)
200
100
0
Resource Usage (C) [10e6]
0
10000
20000
30000
40000
50000
60000
70000
time [ms]

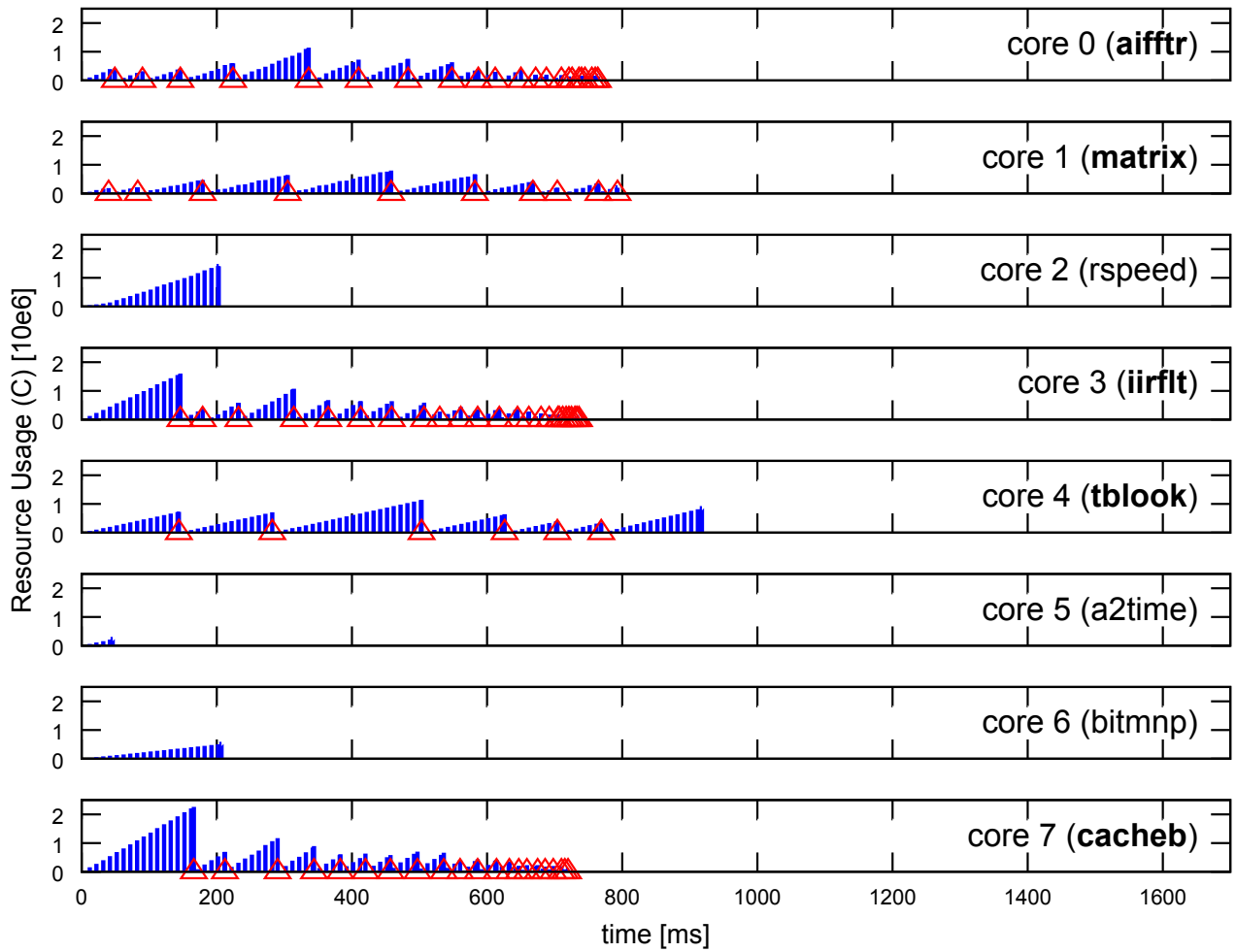
Results - low progress, measurement-based analysis



Results - real, static analysis



Results - real, measurement-based analysis



scenario	process window [ms]	system utilisation basic [%]	system utilisation ext. [%]	system utilisation $\frac{\text{ext.}}{\text{basic}}$	max. core utilisation basic [%]	max. core utilisation ext. [%]	max. core utilisation $\frac{\text{ext.}}{\text{basic}}$
lowp, aiT	48630	3.2	15.5	4.9	1.3	99.9	76.4
lowp, meas	580	23.7	34.5	1.5	12.8	99.1	7.8
real, aiT	41813	2.8	53.8	19.1	0.6	99.8	171.1
real, meas	922	13.6	59.3	4.4	15.6	99.5	6.4

scenario	max. resource accesses basic [10^6]	max. resource accesses ext. [10^6]	max. resource accesses $\frac{\text{ext.}}{\text{basic}}$
lowp, aiT	9.5	998.1	105.1
lowp, meas	2.0	12.3	6.3
real, aiT	9.5	389.7	41.0
real, meas	0.2	4.0	19.6

Summary

Summary

- Presented an extension to the isWCET approach to increase the average-case utilisation
- Idea: dynamic adjustment of resource usage bounds depending on actual system progress
- Reasoning: great deviations between worst-case and average-case performance
- Evaluation: EEMBC Autobench
- Results:
  - Core utilisation increased up to 99% for all setups
  - System utilisation increase highly depends task set characteristics
- Conclusion:
  - Assumption could be proven true
  - Extension drastically increases core and system utilisation
  - Benefit depends on task set

Thank you for your attention

[1] R. Wilhelm, J. Engblom, A. Ermedahl, N. Holsti, S. Thesing, D. Whalley, G. Bernat, C. Ferdinand, R. Heckmann, T. Mitra, F. Mueller, I. Puaut, P. Puschner, and J. Staschulat, "The worst-case execution time problem - overview of methods and survey of tools," *ACM Transactions on Embedded Computing Systems*, 2008.

bmark	$\tau_{s,x}$ [ms]	κ_{π_i} [10^6]	$t_{\overline{ext}}$ [ms]	t_{ext} [ms]	$u_{\overline{ext}}$ [%]	u_{ext} [%]	$\frac{u_{ext}}{u_{\overline{ext}}}$
cacheb	389	9.5	636	48566	1.3	99.9	76.4
bitmnp	2393	53.8	1926	-	4.0	-	-
tblook	2371	56.8	2481	-	5.1	-	-
matrix	4707	99.9	465	-	1.0	-	-
bitmnp	2393	53.8	1925	-	4.0	-	-
tblook	2371	56.8	2474	-	5.1	-	-
matrix	4707	99.9	477	-	1.0	-	-
bitmnp	2393	53.8	1926	-	4.0	-	-
				system utilisation	3.2	15.5	4.9
				process window [ms]	48630		

bmark	accesses [10^6]		
	static	extend	factor
cacheb	9.5	998.1	105.1

bmark	$\tau_{s,x}$ [ms]	κ_{π_i} [10^6]	$t_{\overline{ext}}$ [ms]	t_{ext} [ms]	$u_{\overline{ext}}$ [%]	u_{ext} [%]	$\frac{u_{ext}}{u_{\overline{ext}}}$
cacheb	38	2.0	74	574	12.8	99.1	7.8
bitmnp	170	0.5	193	-	33.4	-	-
tblook	126	0.9	188	-	32.5	-	-
matrix	22	0.2	31	-	5.5	-	-
bitmnp	170	0.5	195	-	33.7	-	-
tblook	126	0.9	188	-	32.5	-	-
matrix	22	0.2	31	-	5.5	-	-
bitmnp	170	0.5	196	-	33.9	-	-
				system utilisation	23.7	34.5	1.5
				process window [ms]	580		

bmark	accesses [10^6]		
	static	extend	factor
cacheb	2.0	12.3	6.3

bmark	$\tau_{s,x}$ [ms]	κ_{π_i} [10^6]	$t_{\overline{ext}}$ [ms]	t_{ext} [ms]	$u_{\overline{ext}}$ [%]	u_{ext} [%]	$\frac{u_{ext}}{u_{\overline{ext}}}$
aifftr	7193	190.0	244	41746	0.6	99.8	171.1
matrix	4707	99.9	561	34016	1.3	81.4	60.6
rspeed	862	19.3	1088	-	2.6	-	-
iirflt	516	13.5	896	33431	2.1	80.0	37.3
tblook	2371	56.8	3267	34050	7.8	81.4	10.4
a2time	151	3.2	167	-	0.4	-	-
bitmnp	2393	53.8	2472	-	5.9	-	-
cacheb	389	9.5	729	33150	1.7	79.3	45.5
				system utilisation	2.8	53.8	19.1
				process window [ms]	41813		

bmark	accesses [10^6]		
	static	extend	factor
cacheb	9.5	389.7	41.0
iirflt	13.5	369.3	27.4
tblook	56.8	305.5	5.4
matrix	99.9	311.8	3.1
aifftr	190.0	577.1	3.0

bmark	$\tau_{s,x}$ [ms]	κ_{π_i} [10^6]	$t_{\overline{ext}}$ [ms]	t_{ext} [ms]	$u_{\overline{ext}}$ [%]	u_{ext} [%]	$\frac{u_{ext}}{u_{\overline{ext}}}$
aifftr	14	0.4	49	762	5.3	82.7	15.6
matrix	22	0.2	40	791	4.3	85.9	19.8
rspeed	75	3.8	202	-	22.0	-	-
iirflt	61	2.4	146	724	15.8	78.6	5.0
tblook	126	0.9	144	917	15.6	99.5	6.4
a2time	10	0.4	45	-	5.0	-	-
bitmnp	170	0.5	206	-	22.4	-	-
cacheb	38	2.3	166	718	18.0	77.9	4.3
				system utilisation	13.6	59.3	4.4
				process window [ms]	922		

bmark	accesses [10^6]		
	static	extend	factor
cacheb	2.3	9.0	3.9
iirflt	2.4	7.8	3.3
tblook	0.9	4.1	4.5
matrix	0.2	4.0	19.6
aifftr	0.4	7.4	16.8