An analytical approach for calculating end-to-end response times in autonomous driving applications

Lukas Krawczyk, Mahmoud Bazzal, Ram Prasath Govindarajan, and Carsten Wolff

Institute for the Digital Transformation of Application and Living Domains

Dortmund University of Applied Sciences and Arts

44227 Dortmund, Germany

lukas.krawczyk@fh-dortmund.de

PANORAMA

The research leading to these results has received funding from the Federal Ministry for Education and Research (BMBF) under Grant 01IS18047D in the context of the ITEA3 EU-Project PANORAMA.

SPONSORED BY THE

Federal Ministry of Education and Research

ITEA3 - 17003

EUREKA

Agenda

- Introduction
- Data consistency
- Data propagation paths
- Analysis Approach
- Optimization
- Integrated Analysis and Optimization Results
- Conclusion and outlook

The research leading to these results has received funding from the Federal Ministry for Education and Research (BMBF) under Grant 01IS18047D in the context of the ITEA3 EU-Project PANORAMA.

SPONSORED BY THE
Federal Ministry of Education and Research

ITEA3 - 17003

EUREKA

Introduction

"Boosting Design Efficiency for Heterogeneous³ Systems"

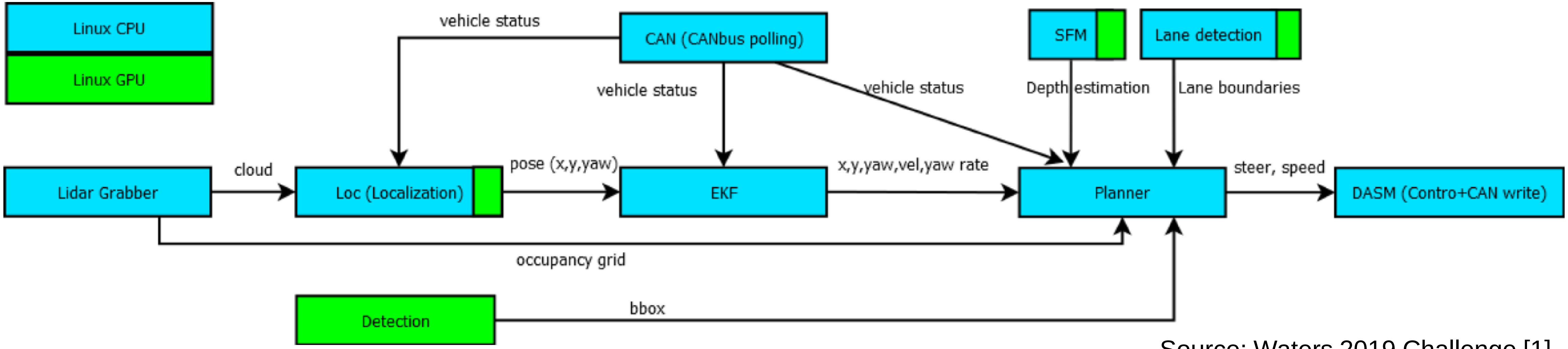
Programcall	ITEA 3 Call 4 17003
Title	Boosting Design Efficiency for Heterogeneous³ Systems
Period	Apr 2019 - Mar 2022
Status	**Running**
Domain	Services, Systems & Software Creation
Technology	Software
Effort	122 man-years
Costs	EUR 15.9 million
Project Leader	Jörg Tessmer (Bosch)
Partners	25
Countries (5)	Finland, Germany, Portugal, Sweden, Turkey https://itea3.org/project/panorama.html

20 YEARS ITEA 3 1998 - 2018

The research leading to these results has received funding from the Federal Ministry for Education and Research (BMBF) under Grant 01IS18047D in the context of the ITEA3 EU-Project PANORAMA.

SPONSORED BY THE Federal Ministry of Education and Research

Introduction



Source: Waters 2019 Challenge [1]

- Calculate applications end-to-end response time
  - Derive task chains for end-to-end paths
  - Develop integrated response time analysis approach
- Optimize the latency of the different task-chains
  - Our scope: Minimize the end-to-end response time

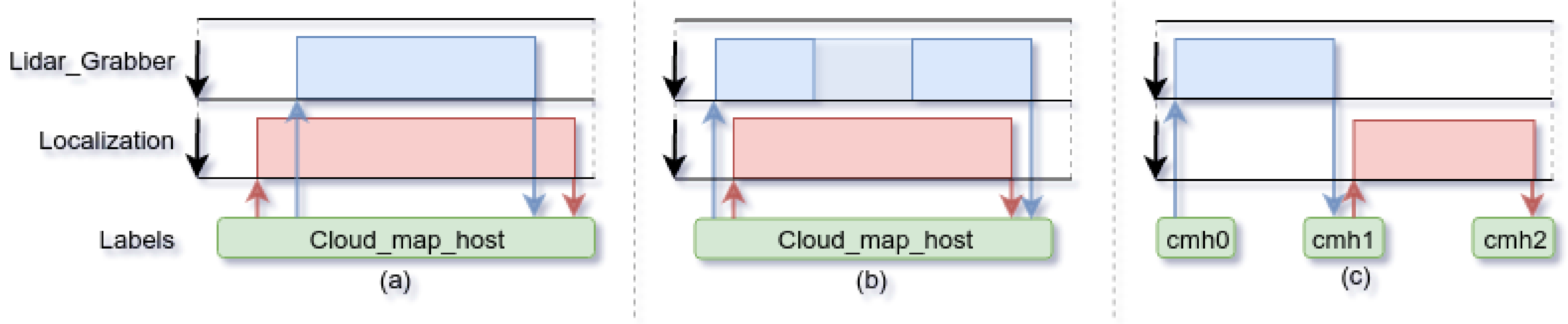
The research leading to these results has received funding from the Federal Ministry for Education and Research (BMBF) under Grant 01IS18047D in the context of the ITEA3 EU-Project PANORAMA.

SPONSORED BY THE Federal Ministry of Education and Research

ITEA3 - 17003

EUREKA

Data consistency



(a) Localization overwrites Lidar_Grabber

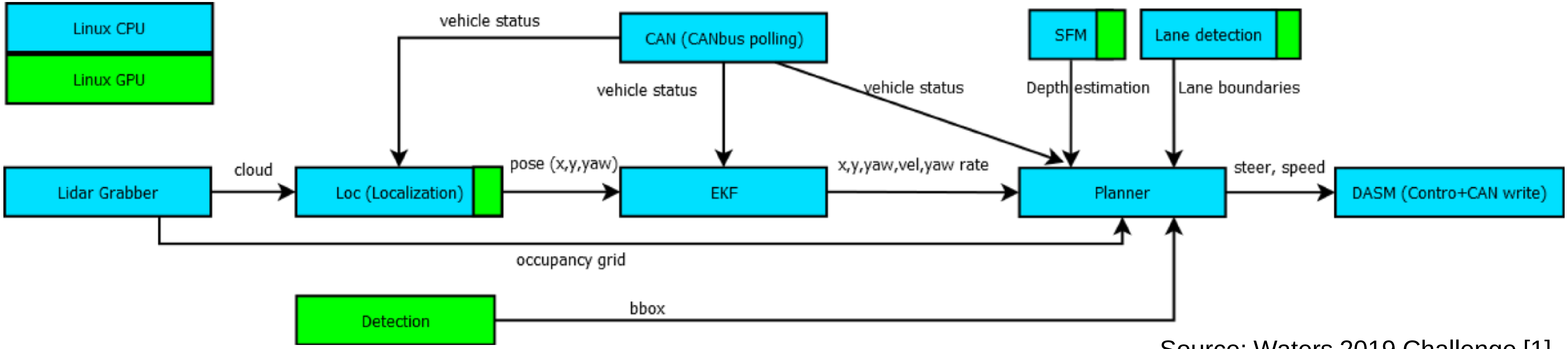
(b) Lidar_Grabber overwrites Localization

(c) Deterministic behaviour

- Higher memory consumption
- Increased latency compared to e.g. semaphore usage
- Correct behaviour can be realied at the cost of higher latency by an e.g. pipeline fashioned approach

The research leading to these results has received funding from the Federal Ministry for Education and Research (BMBF) under Grant 01IS18047D in the context of the ITEA3 EU-Project PANORAMA.
SPONSORED BY THE Federal Ministry of Education and Research
ITEA3 - 17003
EUREKA

Data Propagation Paths



Source: Waters 2019 Challenge [1]

- All **critical paths** from **sensor tasks** to **actuator tasks**
  - Lidar_Grabber → Loc → EKF → Planner → DASM
  - CAN → Loc → EKF → Planner → DASM
  - SFM → Planner → DASM
  - Lane_detection → Planner → DASM
  - Detection → Planner → DASM

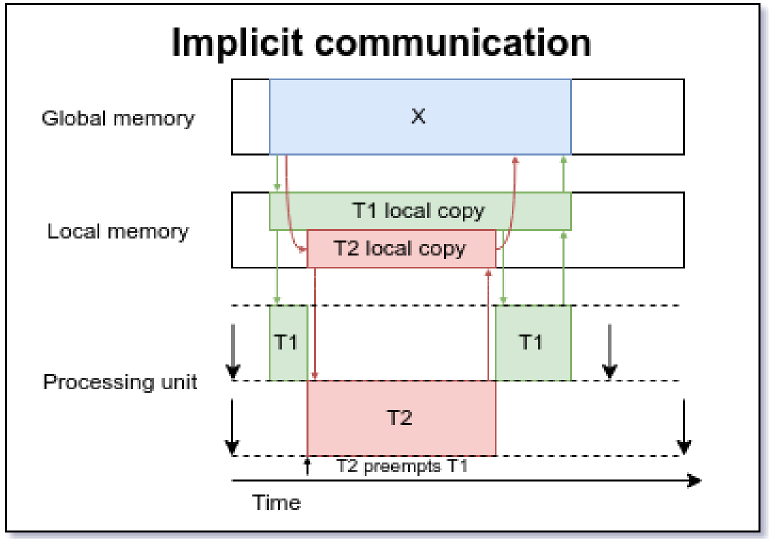
The research leading to these results has received funding from the Federal Ministry for Education and Research (BMBF) under Grant 01IS18047D in the context of the ITEA3 EU-Project PANORAMA.

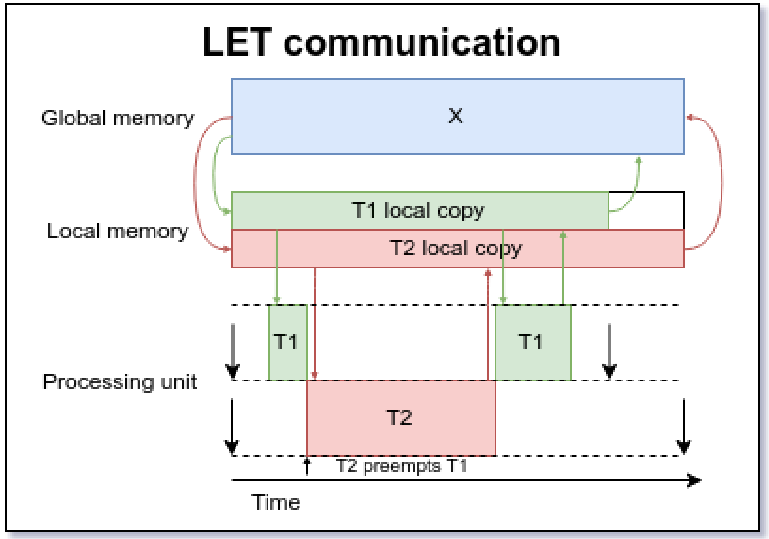
SPONSORED BY THE Federal Ministry of Education and Research

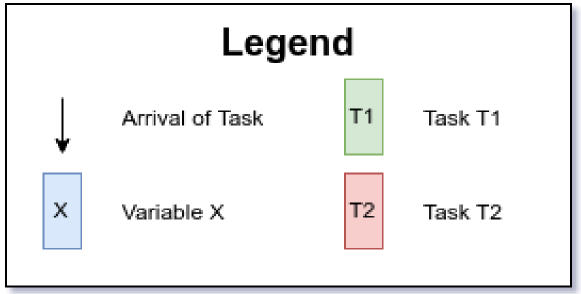
ITEA3 - 17003

EUREKA

Analysis







- Implicit communication
  - End-to-end response time can be optimized
  - E2E-RT by Kloda et al. [2]
- LET communication
  - Deterministic behaviour
  - Own implementation extending [2]

The research leading to these results has received funding from the Federal Ministry for Education and Research (BMBF) under Grant 01IS18047D in the context of the ITEA3 EU-Project PANORAMA.

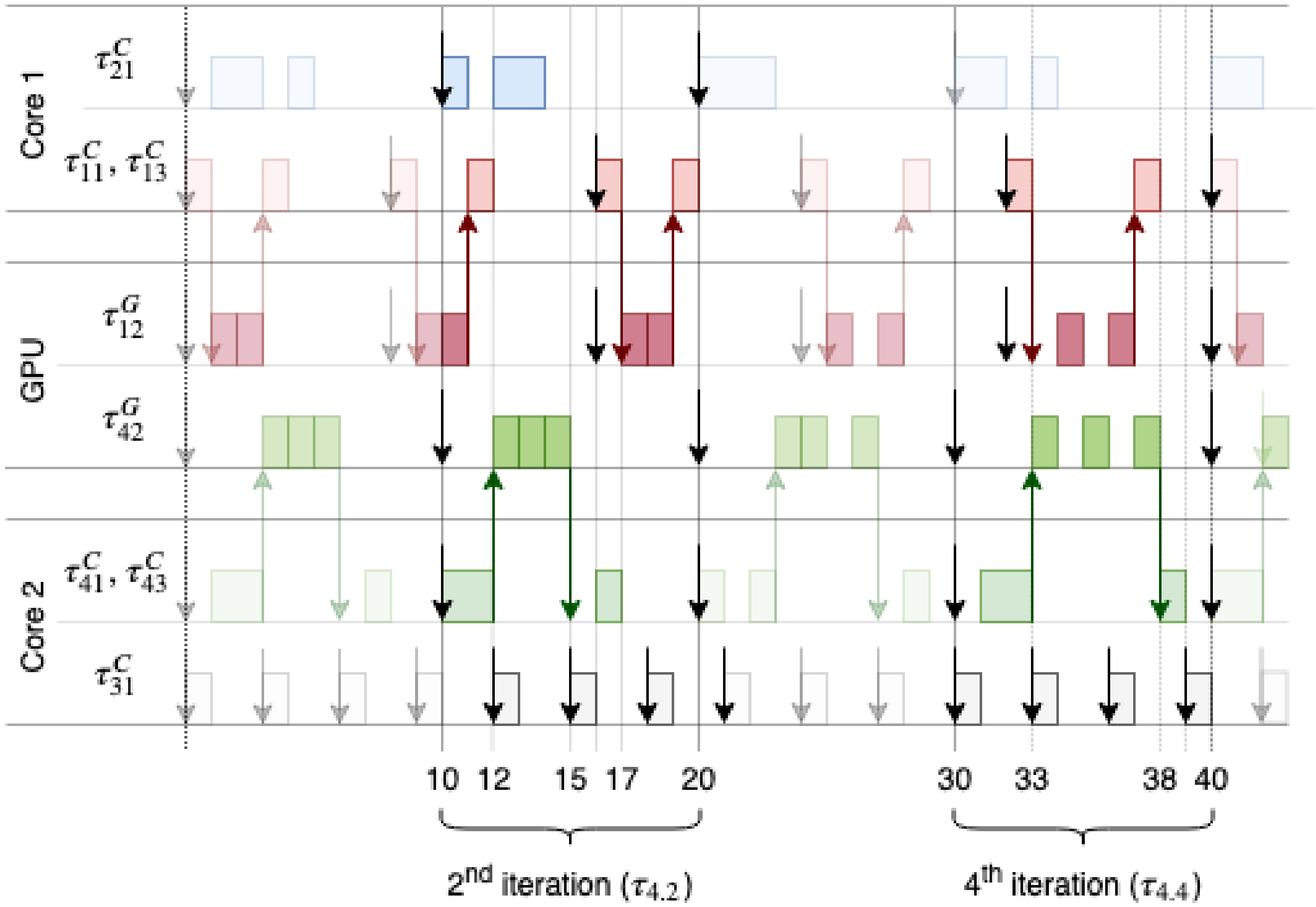
SPONSORED BY THE Federal Ministry of Education and Research

ITEA3 - 17003

EUREKA

Analysis

- Different scheduling strategies
  - Fixed priority preemptive (FPP) scheduling on CPUs
  - Weighted round-robin (WRR) scheduling on GPUs
  - Task suspension
- FPP: Palencia et al. [3]
- WRR: Racu et al. [4]



The research leading to these results has received funding from the Federal Ministry for Education and Research (BMBF) under Grant 01IS18047D in the context of the ITEA3 EU-Project PANORAMA.

SPONSORED BY THE
Federal Ministry of Education and Research

ITEA3 - 17003

EUREKA

Analysis – Task Model

- Tasks are described in terms of **transactions**, with:

  (Sub-Tasks (Runnables), Period, Priority)

$$\tau_i = (\{\tau_{i1}, \ldots, \tau_{i|\tau_i|}\}, P_i, \pi_i)$$

- Sub-Task on **CPU**, with:

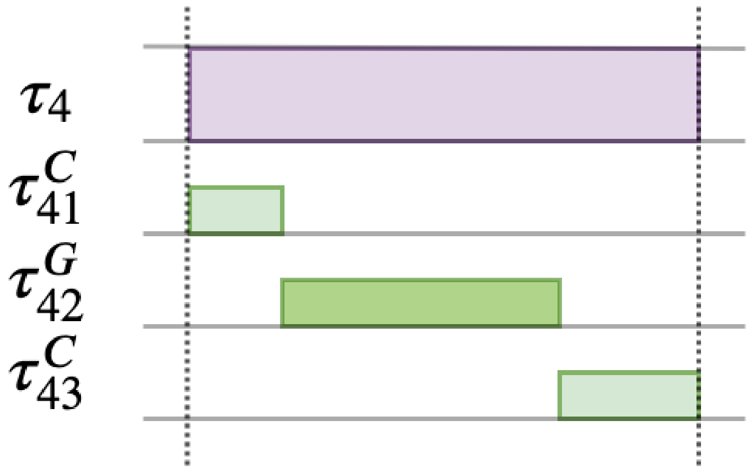
  (Execution Time, Offset, Jitter)

$$\tau_{ij}^C = (C_{ij,\rho}, O_{ij}, J_{ij})$$

- Sub-Task on **GPU**, with:

  (Execution Time, Offset, Jitter, Time-Slice)

$$\tau_{ij}^G = (C_{ij,\rho}, O_{ij}, J_{ij}, \phi_{ij})$$



The research leading to these results has received funding from the Federal Ministry for Education and Research (BMBF) under Grant 01IS18047D in the context of the ITEA3 EU-Project PANORAMA.

SPONSORED BY THE
Federal Ministry of Education and Research

ITEA3 - 17003

EUREKA

Fachhochschule Dortmund
University of Applied Sciences and Arts
IDiAL Institute for the Digital Transformation of Application and Living Domains

Analysis – Data transfer times

PANORAMA

- Number of **label accesses**

$$\lambda_{ij} = \sum_{l \in \mathcal{L}_{ij}} \left\lceil \frac{size(l)}{size(cacheline)} \right\rceil$$

- Memory **access times**

	Best	Worst
A57	20 ns	220 ns
Denver	8 ns	38 ns
GPU	3 ns	6 ns

- Total **work** for a sub-task

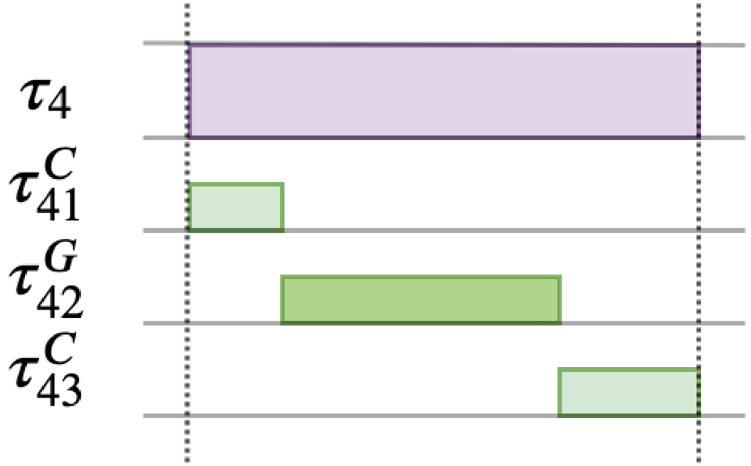
$$\mathcal{W}_{ij} = \mathcal{C}_{ij,\rho} + \lambda_{ij} \cdot \mathcal{A}_{\rho}$$

- Sub-task‘s best case response time

$$\mathcal{R}^{+}_{ij} = \sum_{k=1\ldots j} \mathcal{W}^{+}_{ik}$$

- Task‘s worst case response time

$$\mathcal{R}^{-}_{i} = \mathcal{R}^{-}_{i|\tau_i|}$$



The research leading to these results has received funding from the Federal Ministry for Education and Research (BMBF) under Grant 01IS18047D in the context of the ITEA3 EU-Project PANORAMA.
SPONSORED BY THE Federal Ministry of Education and Research
ITEA3 - 17003
EUREKA

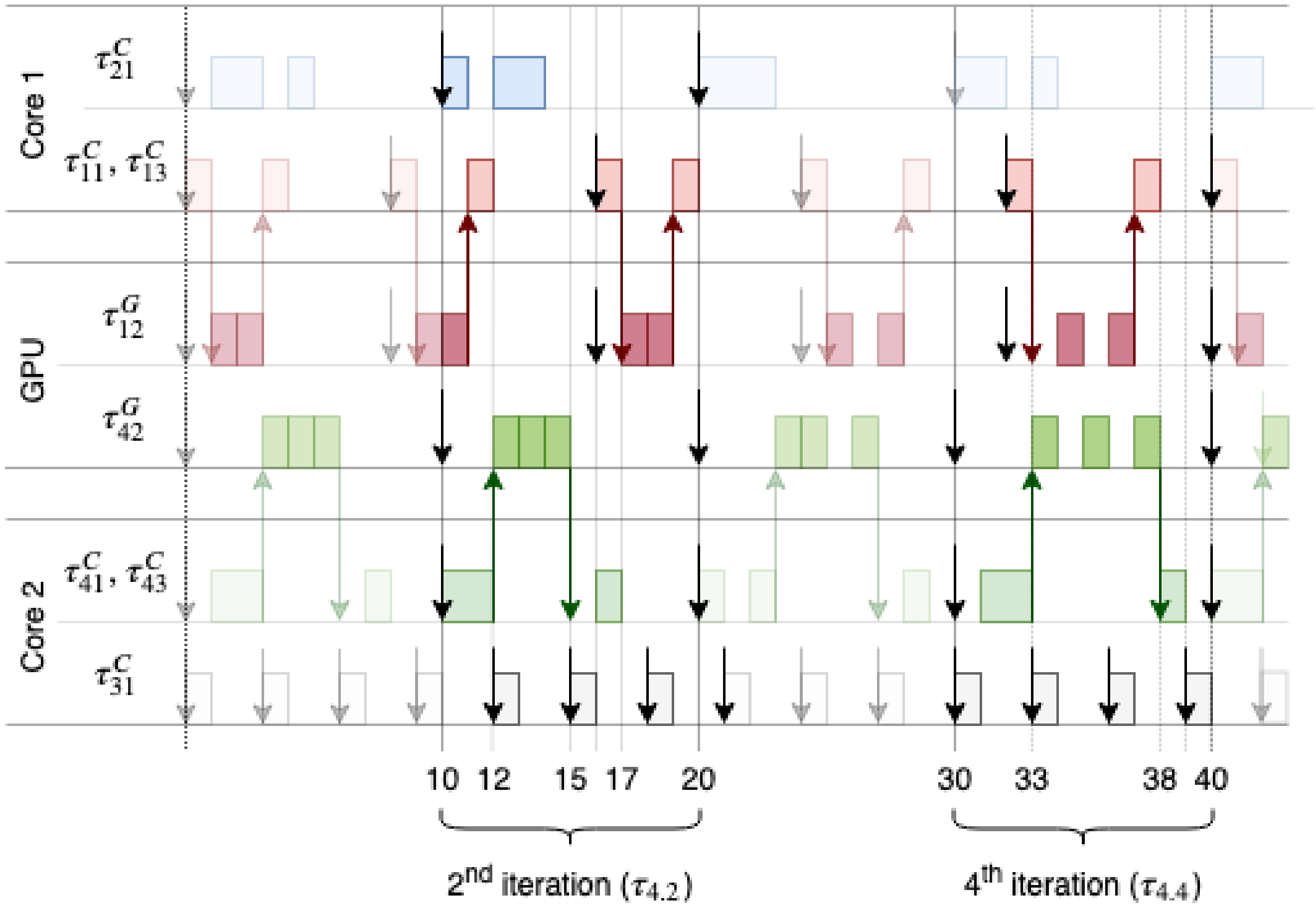
Analysis

- Integration of both approaches as iterative strategy
  - Update the **offset** of the **successor,** set it to the **best case response time** of its **predecessor**

$$O_{ij} = R^{+}_{ij-1}$$

  - Update the **jitter**, set it to the difference between **worst case response time** and **offset (BCRT)**

$$J_{ij} = R^{-}_{ij-1} - R^{+}_{ij-1}$$



The research leading to these results has received funding from the Federal Ministry for Education and Research (BMBF) under Grant 01IS18047D in the context of the ITEA3 EU-Project PANORAMA.

SPONSORED BY THE
Federal Ministry of Education and Research

ITEA3 - 17003

EUREKA

Optimization

- Genetic Algorithm Implementation based on Jenetics (Java)
- Already integrated into App4MC (OpenMapping)
- Degrees of freedom (DoF)
  - Allocation (Task to Processing Unit)
  - Allocation (Offloadable sub-task to Processing Unit)
  - Time Slice (Sub-Task on GPU only)

The research leading to these results has received funding from the Federal Ministry for Education and Research (BMBF) under Grant 01IS18047D in the context of the ITEA3 EU-Project PANORAMA.
SPONSORED BY THE
Federal Ministry of Education and Research
ITEA3 - 17003
EUREKA

Integrated Analysis and Optimization Results

- Similar end-to-end latency for LET and implicit communication
- Response times close to the task‘s period
- Runtime: 287 seconds
  - Reason: Audsleys priority assignment algorithm

Name	P	π	C^-	$\lambda \cdot \mathcal{A}^-$	R^-	ϕ
Core 0 (Denver)						
Planner	12	9	11.2	0.8	12.0	–
Core 1 (Denver)						
SFM*	33	6	6.7	3.6	31.5	–
Lane_detection	66	2	42.2	1.2	53.6	–
Core 2 (A57)						
CANbus_polling	10	5	0.6	0.0	0.6	–
EKF	15	1	4.8	0	5.4	–
Core 3 (A57)						
Localization	400	4	387.4	5.2	392.6	–
Core 4 (A57)						
Lidar_Grabber	33	8	13.7	12.0	25.7	–
Detection*	200	7	4.7	1.8	198.0	–
Core 5 (A57)						
OS_Overhead	100	0	50	0.0	79.9	–
DASM	5	3	1.9	0.0	1.9	–
GP10B (GPU)						
Detection	200	–	116.0	0.5	170.5	7.0
SFM	33	–	7.9	0.4	15.2	11.6

Task Chain	LET end-to-end	Implicit end-to-end
σ_1	886	859.9
σ_2	865	836.9
σ_3	67	59.9
σ_4	100	71.9
σ_5	230	221.9

The research leading to these results has received funding from the Federal Ministry for Education and Research (BMBF) under Grant 01IS18047D in the context of the ITEA3 EU-Project PANORAMA.

SPONSORED BY THE
Federal Ministry of Education and Research

ITEA3 - 17003

EUREKA

Conclusion and outlook

- Analysis of end-to-end response time of a given application following an **implicit** and **LET** communication paradigm

- Accounting all mandatory delays:
  - Data transfer time for **copy engine** (GPU <-> CPU)
  - Data transfer time between **CPU and shared main memory**
  - **Synchronous** and **asynchronous** offloading
  - Application of given memory contention approach

- Response time analysis for coupled task sets scheduled on an **heterogeneous architecture** consisting of processing units with **fixed priority preemptive** (CPU) and **weighted round robin** (GPU) scheduling

- **Minimization** of the applications maximum **end-to-end response** time among all task chains for a implicit communication paradigm

The research leading to these results has received funding from the Federal Ministry for Education and Research (BMBF) under Grant 01IS18047D in the context of the ITEA3 EU-Project PANORAMA.

SPONSORED BY THE
Federal Ministry of Education and Research

ITEA3 - 17003

EUREKA

Conclusion and outlook

- Simplification of the model was required (transitive labels, planner task)
- Cooperative scheduling (FPFP optimistic assumption)
- Scalability
- Fully integrated approach
- Comparison with practical demonstration results

The research leading to these results has received funding from the Federal Ministry for Education and Research (BMBF) under Grant 01IS18047D in the context of the ITEA3 EU-Project PANORAMA.

SPONSORED BY THE
Federal Ministry of Education and Research

ITEA3 - 17003

EUREKA

References

[1] Arne Hamann, Dakshina Dasari, and Falk Wurst. WATERS Industrial Challenge, 2019.

[2] Tomasz Kloda, Antoine Bertout, and Yves Sorel. Latency analysis for data chains of real-time periodic tasks. In 23rd IEEE International Conference on Emerging Technologies and Factory Automation, ETFA 2018, Torino, Italy, September 4-7, 2018, pages 360–367. IEEE, 2018.

[3] J.C. Palencia and M. Gonzalez Harbour. Schedulability analysis for tasks with static and dynamic offsets. 2002.

[4] Razvan Racu, Li Li, Rafik Henia, Arne Hamann, and Rolf Ernst. Improved response time analysis of tasks scheduled under preemptive round-robin. In Proceedings of the 5th International Conference on Hardware/Software Codesign and System Synthesis, CODES+ISSS 2007, Salzburg, Austria, September 30 - October 3, 2007, pages 179–184. ACM, 2007.

The research leading to these results has received funding from the Federal Ministry for Education and Research (BMBF) under Grant 01IS18047D in the context of the ITEA3 EU-Project PANORAMA.

SPONSORED BY THE Federal Ministry of Education and Research

ITEA3 - 17003

EUREKA

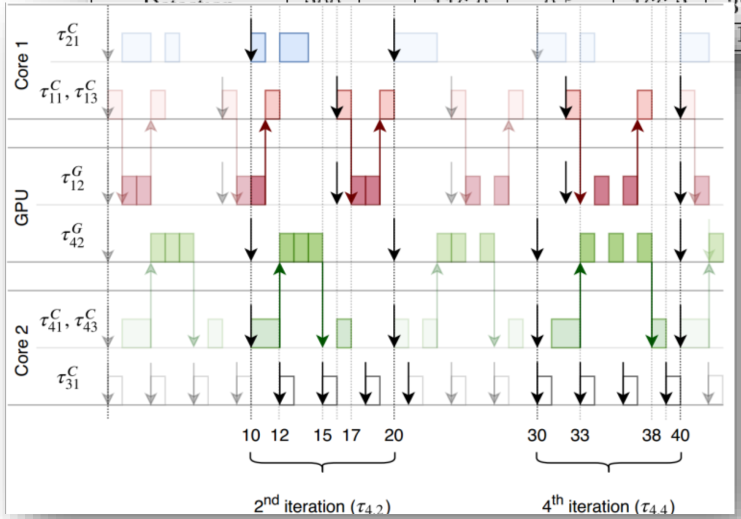
Fachhochschule Dortmund
University of Applied Sciences and Arts
IDiAL Institute for the Digital Transformation of Application and Living Domains

Thank you for your attention

PANORAMA

- Analysis of End-to-End latencies of a given application following an **implicit** and **LET** communication paradigm
- Response time analysis for coupled task sets scheduled on an **heterogeneous architecture** consisting of processing units with **fixed priority preemptive** (CPU) and **weighted round robin** (GPU) scheduling
- **Minimization** of the applications maximum **end-to-end response time** among all task chains for a implicit communication paradigm

Name	P	π	C^-	$\lambda \cdot \mathcal{A}^-$	R^-	ϕ
Core 0 (Denver)						
Planner	12	9	11.2	0.8	12.0	–
Core 1 (Denver)						
SFM*	33	6	6.7	3.6	31.5	–
Lane_detection	66	2	42.2	1.2	53.6	–
Core 2 (A57)						
CANbus_polling	10	5	0.6	0.0	0.6	–
EKF	15	1	4.8	0	5.4	–
Core 3 (A57)						
Localization	400	4	387.4	5.2	392.6	–
Core 4 (A57)						
Lidar_Grabber	33	8	13.7	12.0	25.7	–
Detection*	200	7	4.7	1.8	198.0	–
Core 5 (A57)						
OS_Overhead	100	0	50	0.0	79.9	–
DASM	5	3	1.9	0.0	1.9	–
GP10B (GPU)						
Detection	200	[illegible]	116.0	0.5	166.2	375
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.6



Questions?

Lukas Krawczyk, Mahmoud Bazzal, Ram Prasath Govindarajan, and Carsten Wolff

Institute for the Digital Transformation of Application and Living Domains
Dortmund University of Applied Sciences and Arts
44227 Dortmund, Germany
lukas.krawczyk@fh-dortmund.de

The research leading to these results has received funding from the Federal Ministry for Education and Research (BMBF) under Grant 01IS18047D in the context of the ITEA3 EU-Project PANORAMA.

SPONSORED BY THE Federal Ministry of Education and Research

ITEA3 - 17003

EUREKA