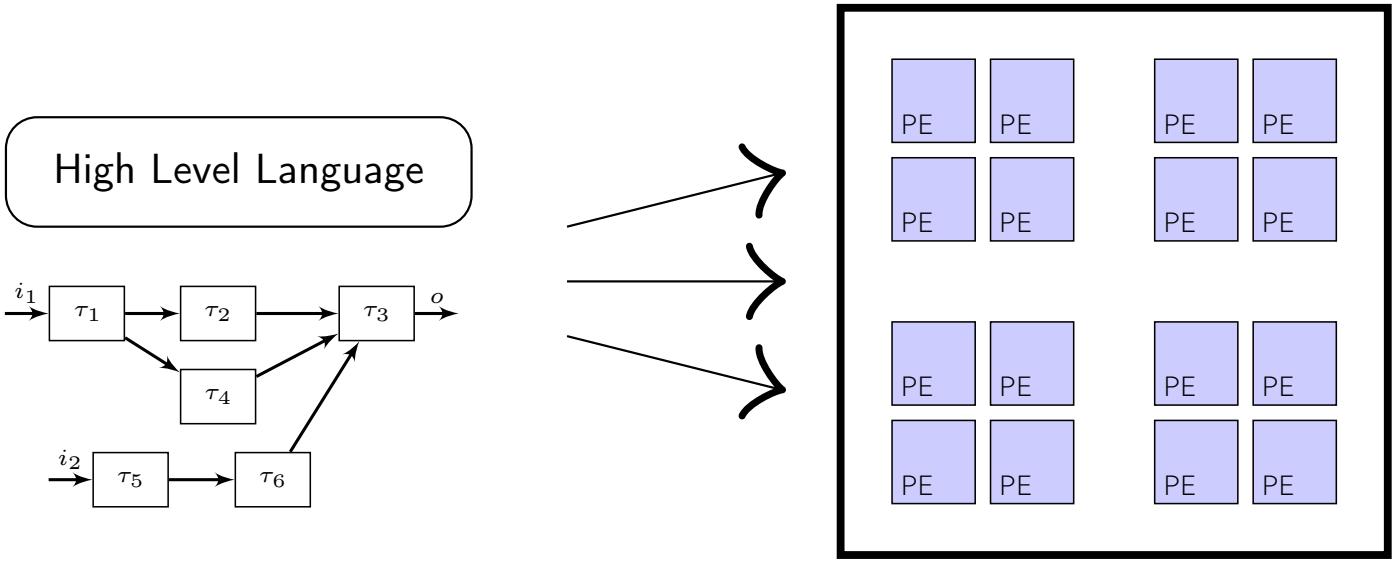
# Efficient Execution of Dependent Tasks on Many-Core Processors

Hamza Rihani, Claire Maiza, Matthieu Moy

Univ. Grenoble Alpes Verimag

RTSOPS 2016, July 5, 2016

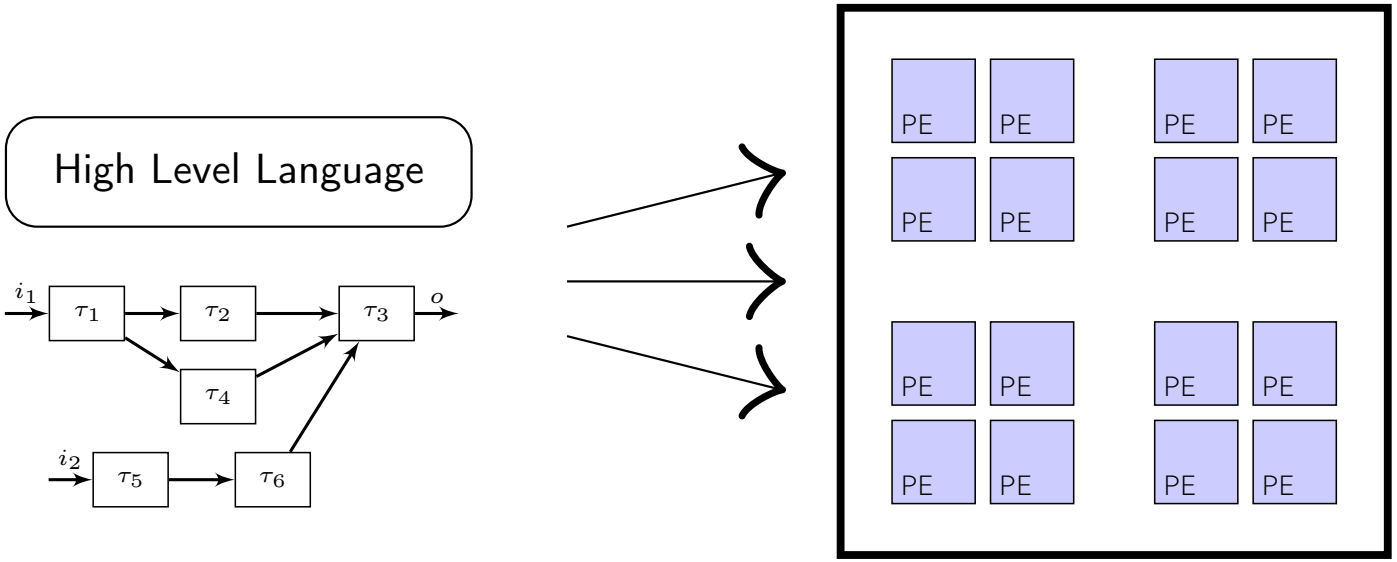

# Context




- Hard real-time systems
- Dependent tasks statically scheduled, on a many-core processor

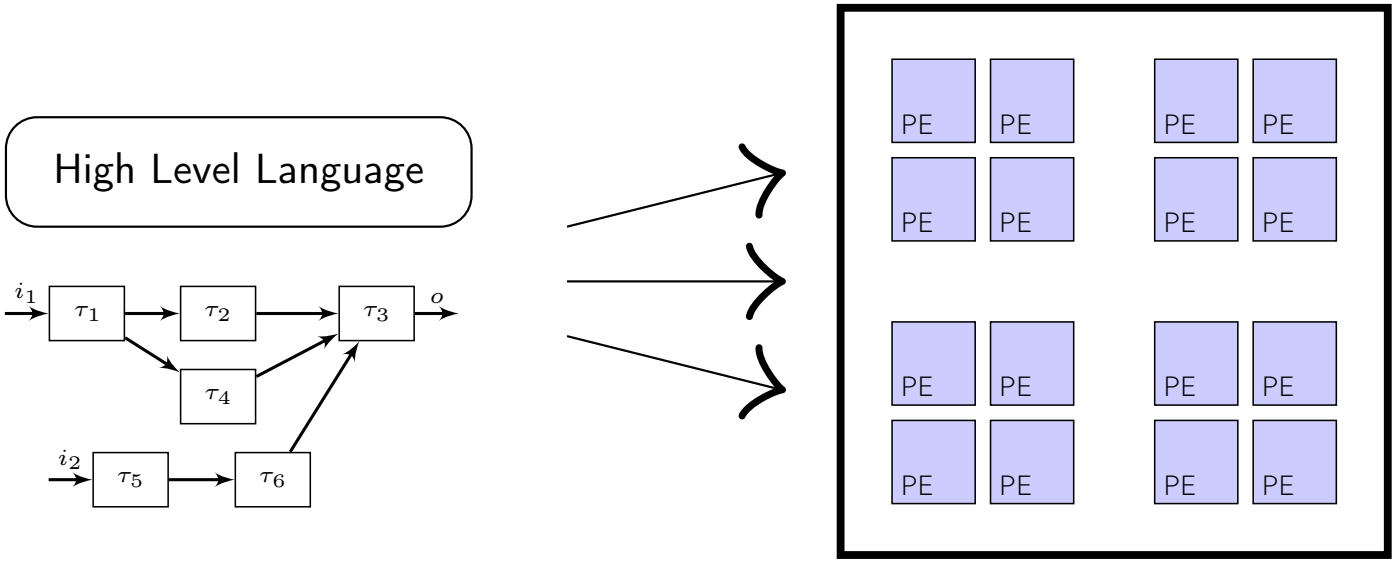
  ! Unpredictable delays due to shared resource interference

# Context




- Hard real-time systems
- Dependent tasks statically scheduled, on a many-core processor
  - ! Unpredictable delays due to shared resource interference
  - ✓ Use tightly estimated upper bounds on delays

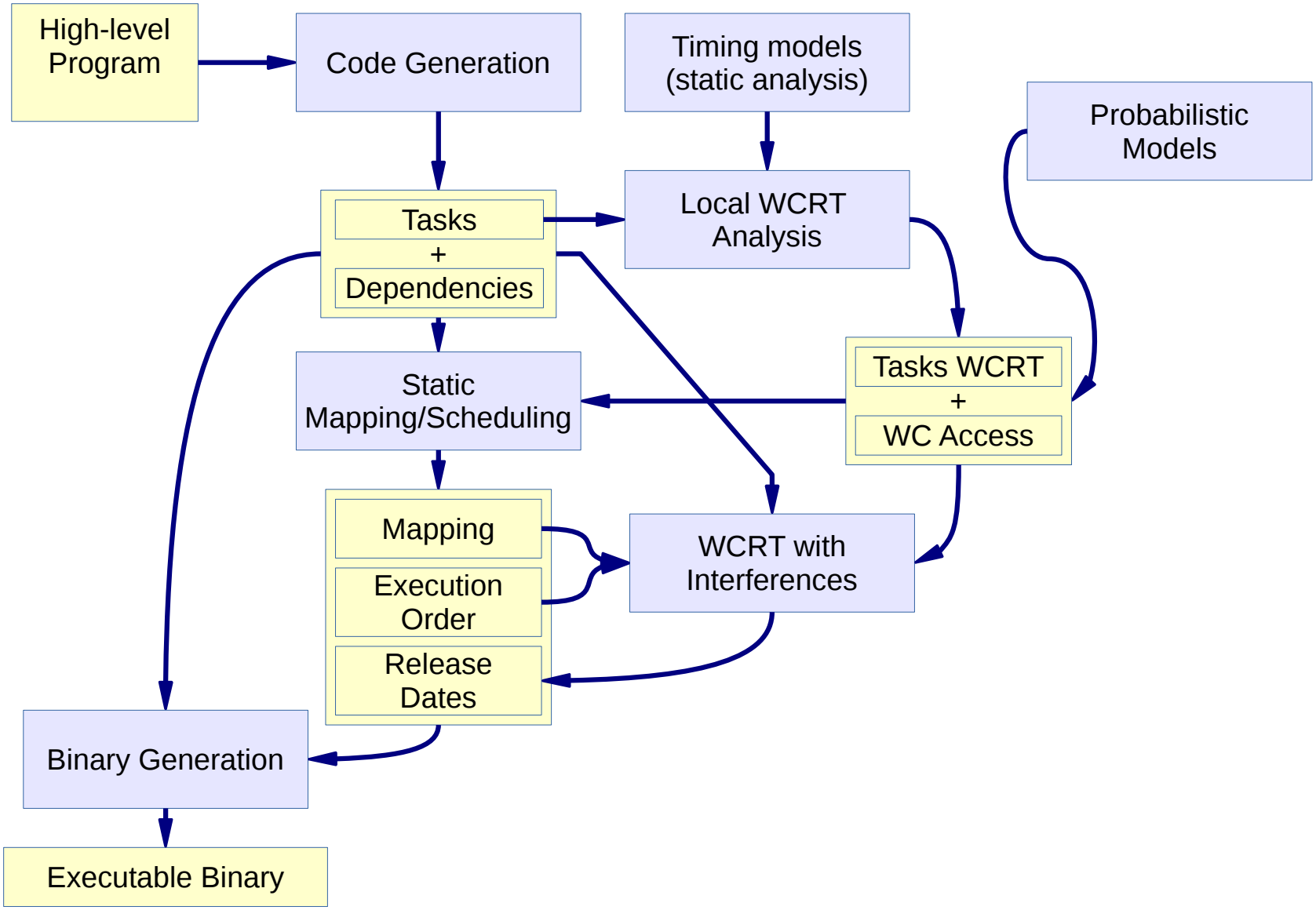
# Context



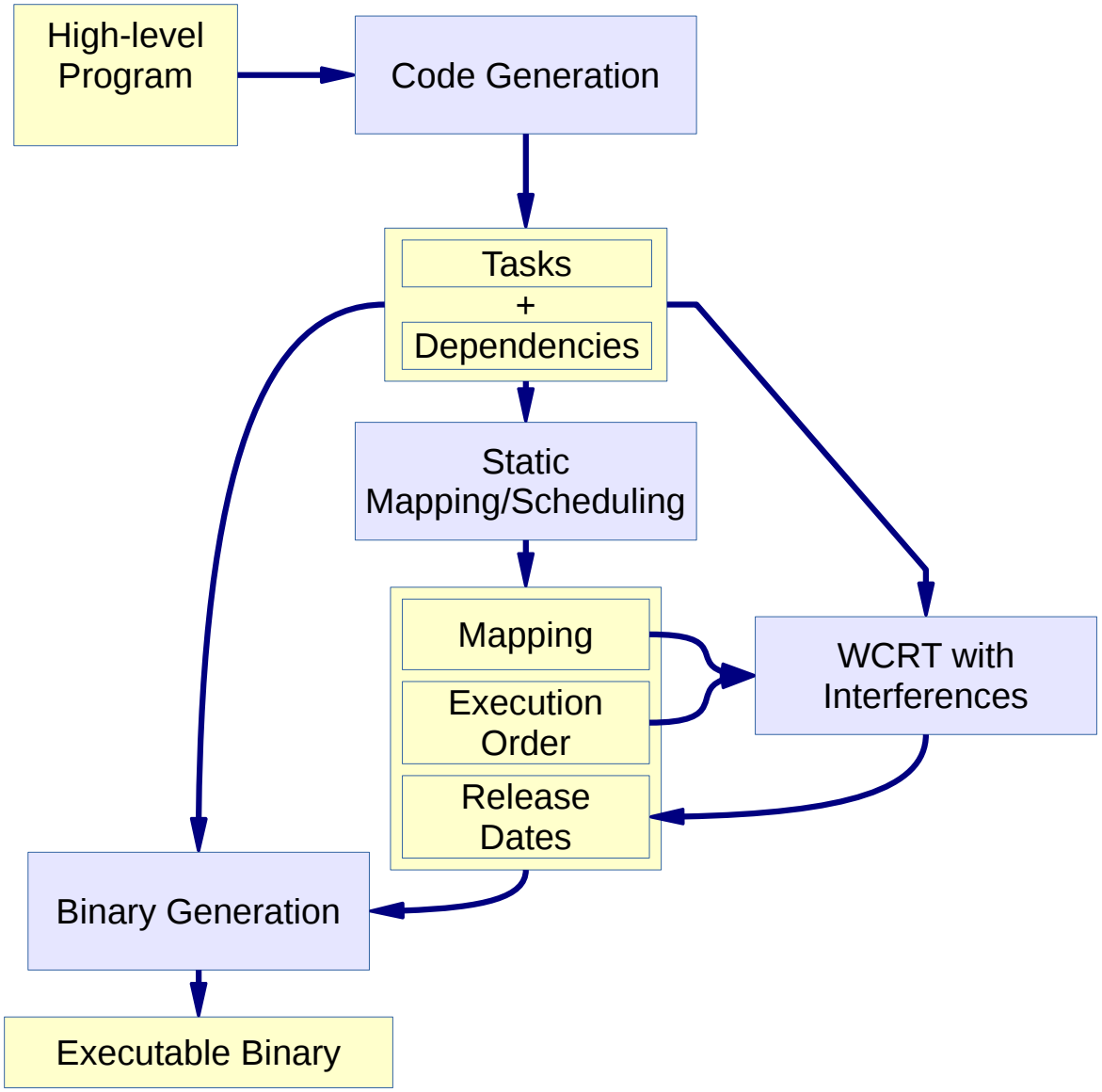

- Hard real-time systems
- Dependent tasks statically scheduled, on a many-core processor
  - ! Unpredictable delays due to shared resource interference
  - ✓ Use tightly estimated upper bounds on delays
  - ✓ Connect existing approaches for an optimally efficient execution

# Outline

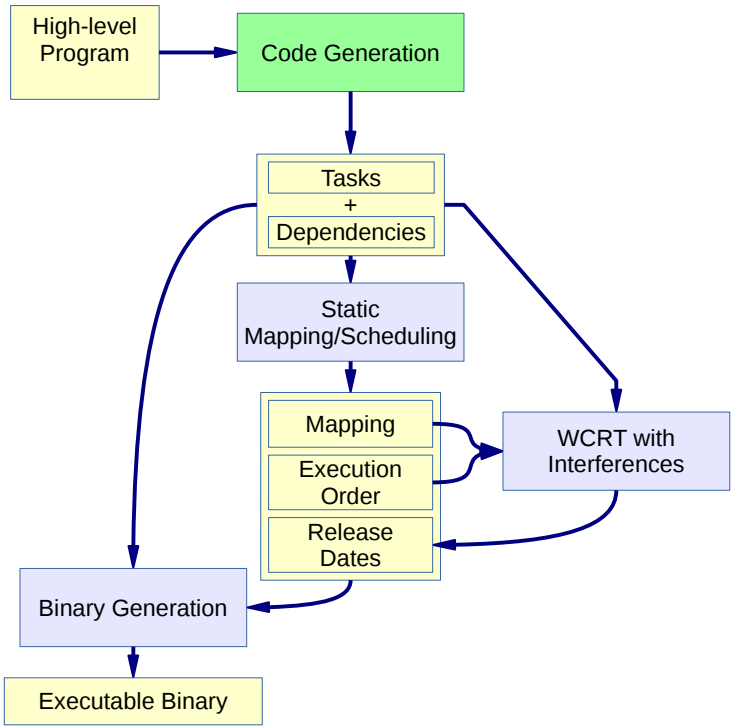
# Solved Problems

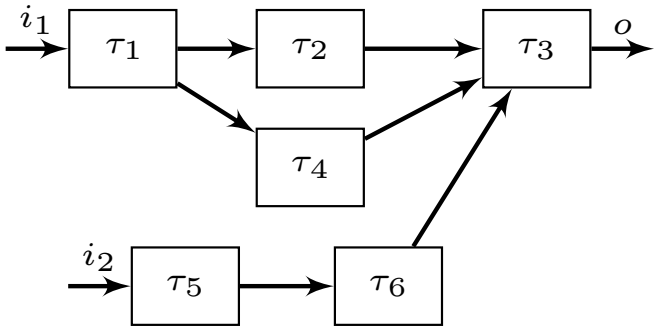


# Solved Problems



# Solved Problems: Code Generation

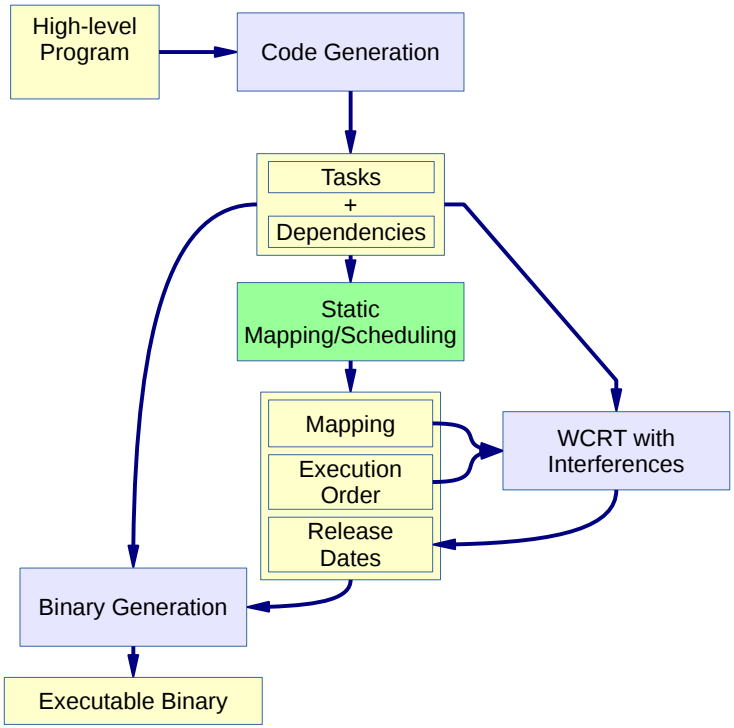


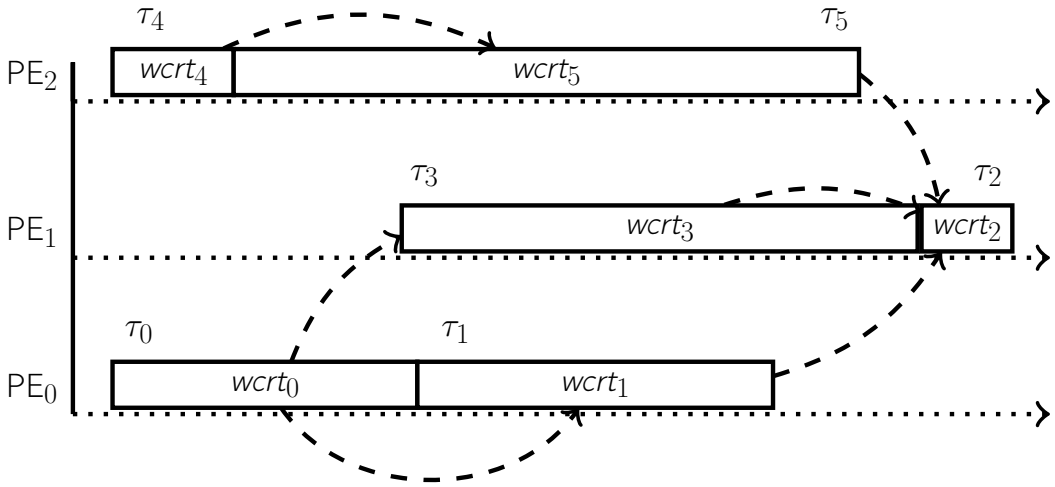




## Outputs

- Task binaries
- Task dependency graph
- Execution models: (Pellizzoni et al.[6])
  - Single phase execution
  - acquisition, execution, replication phases

# Solved Problems: Task Mapping/Scheduling



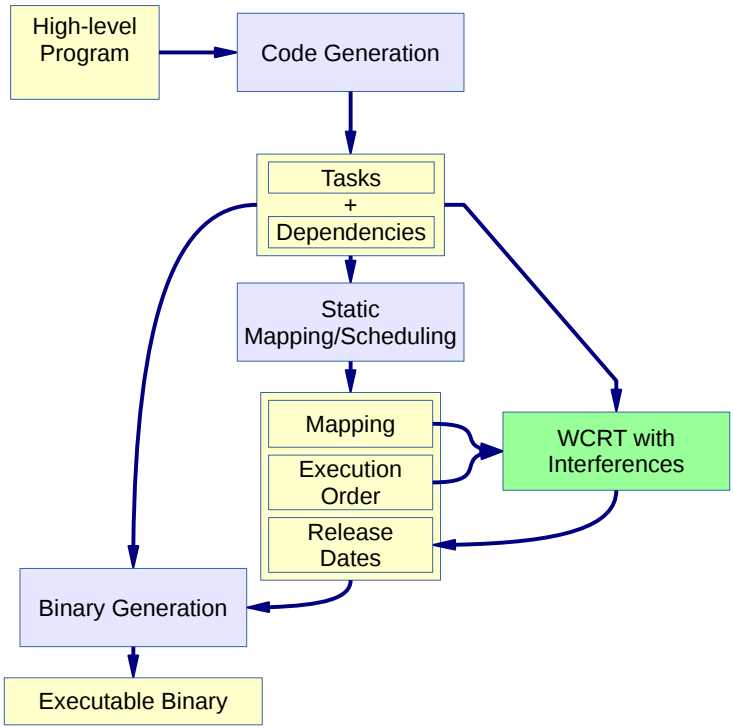



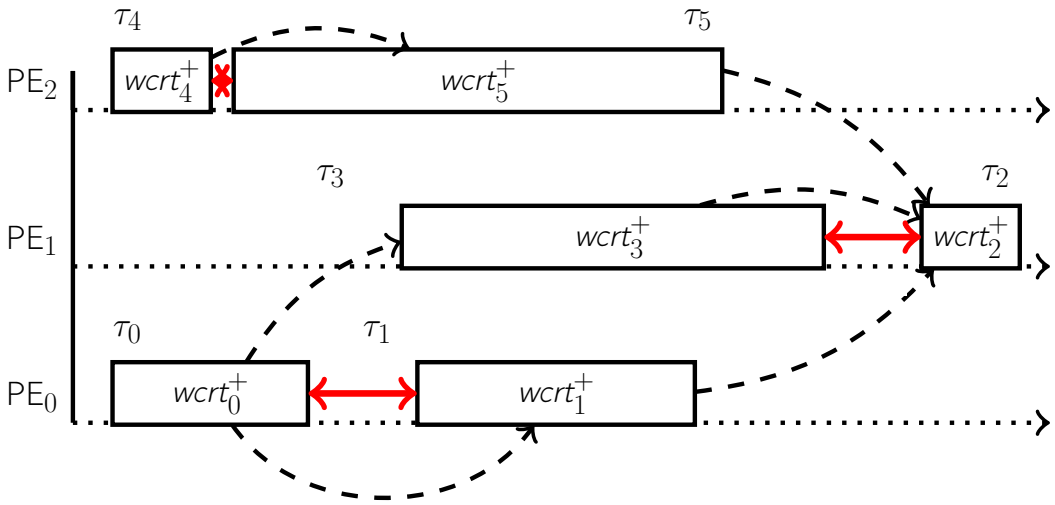

$(*)wcrt_x$: safe WCRT

- Respect the dependency constraints
- Optimize the overall response time

Puffitsch et al. 2013 [7],
Giannopoulou et al. 2013 [4],
Walter et al. 2015 [8]

# Solved Problems: WCRT Analysis

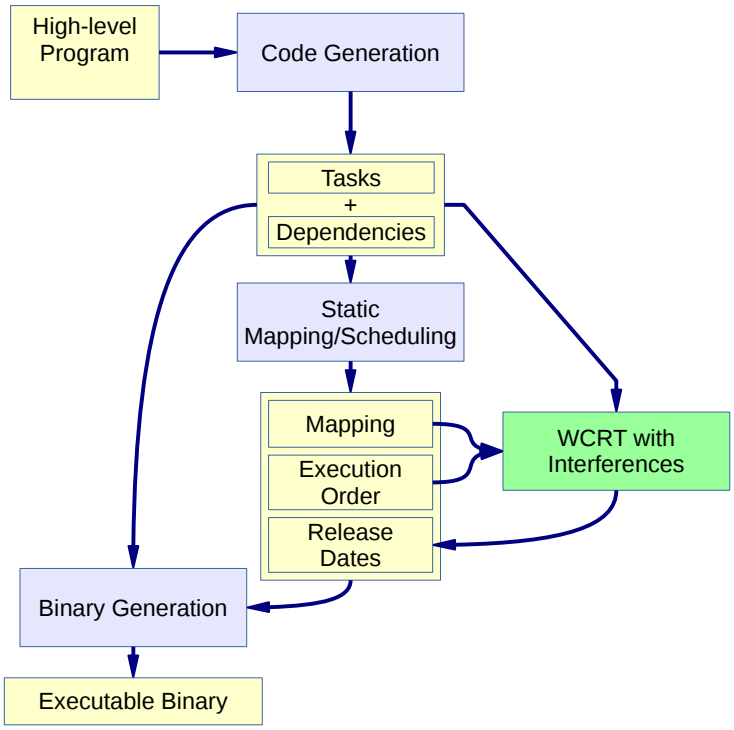


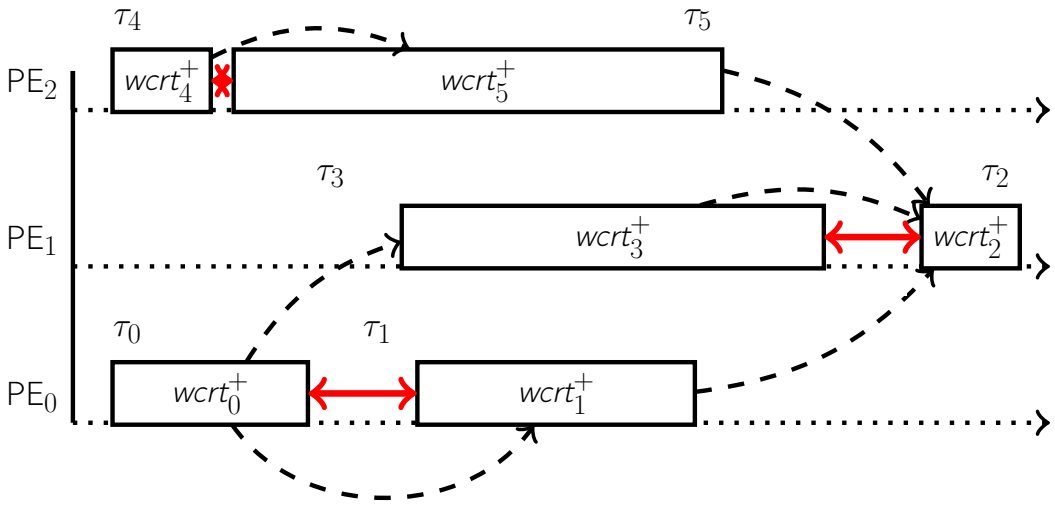




$(*)wcrt_x^+$: refined WCRT

- Take the interference into account
- Update the release times

# Solved Problems: WCRT Analysis



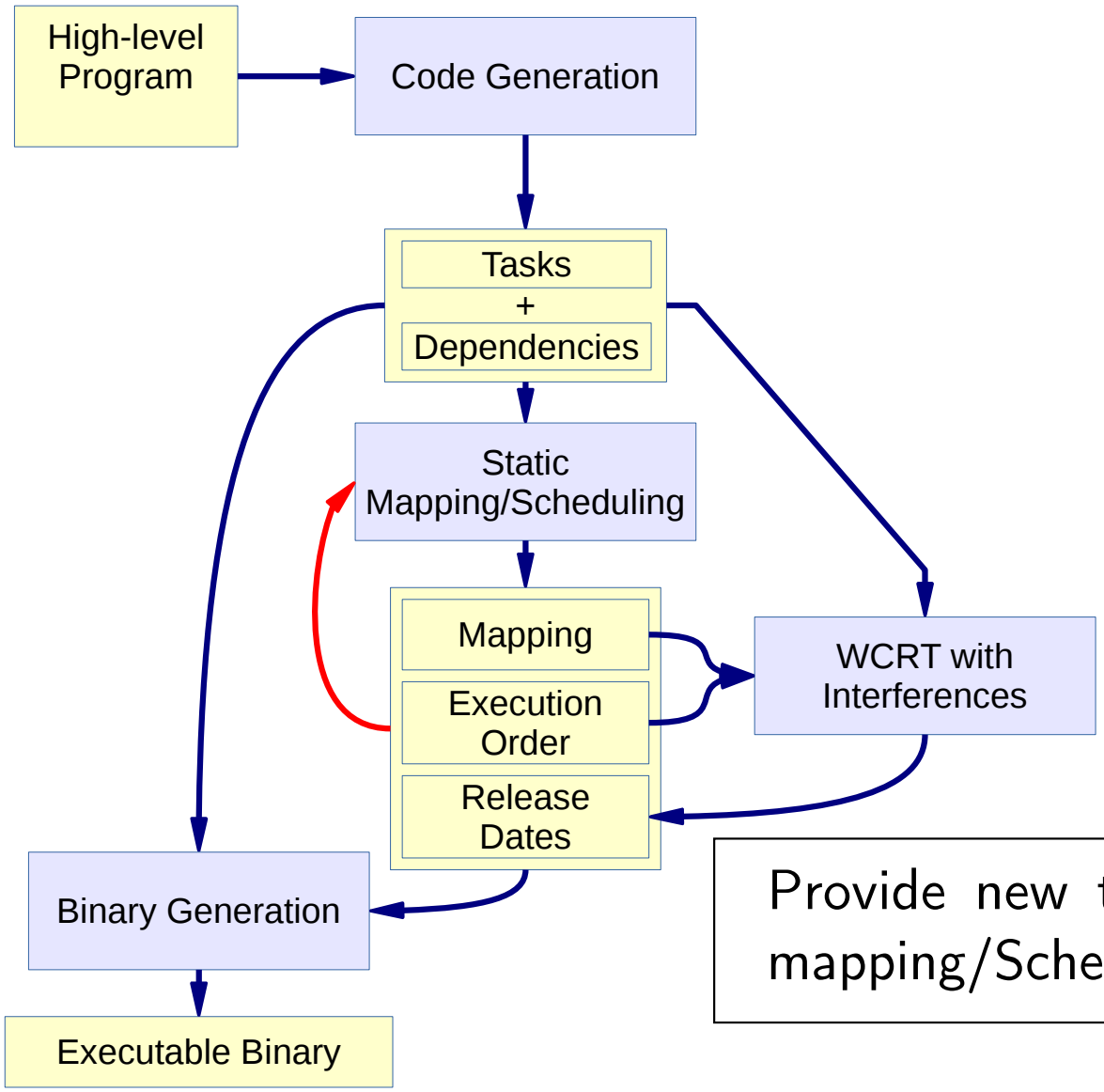




$(*)wcrt_x^+$: refined WCRT

- Take the interference into account
- Update the release times

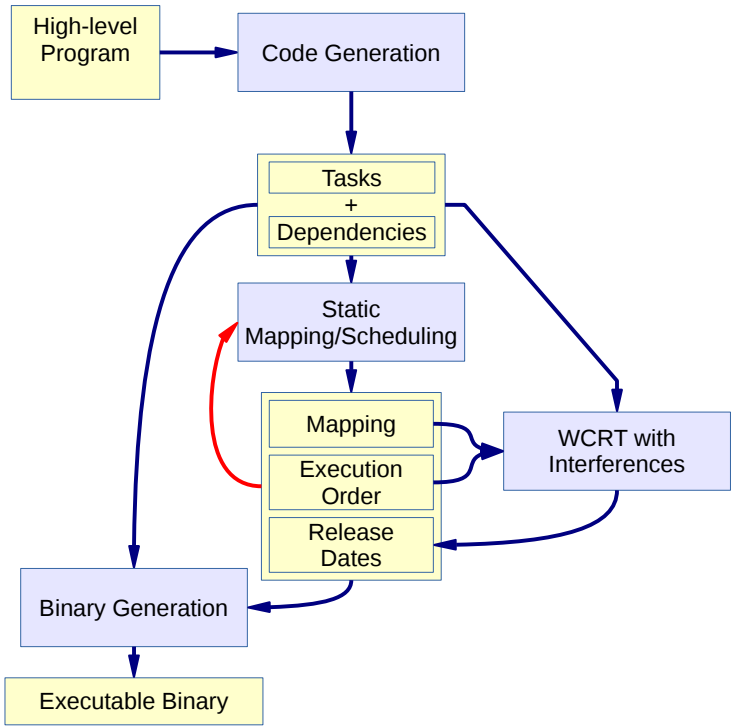
**The overall response time may not be optimal**

# Toward a Solution



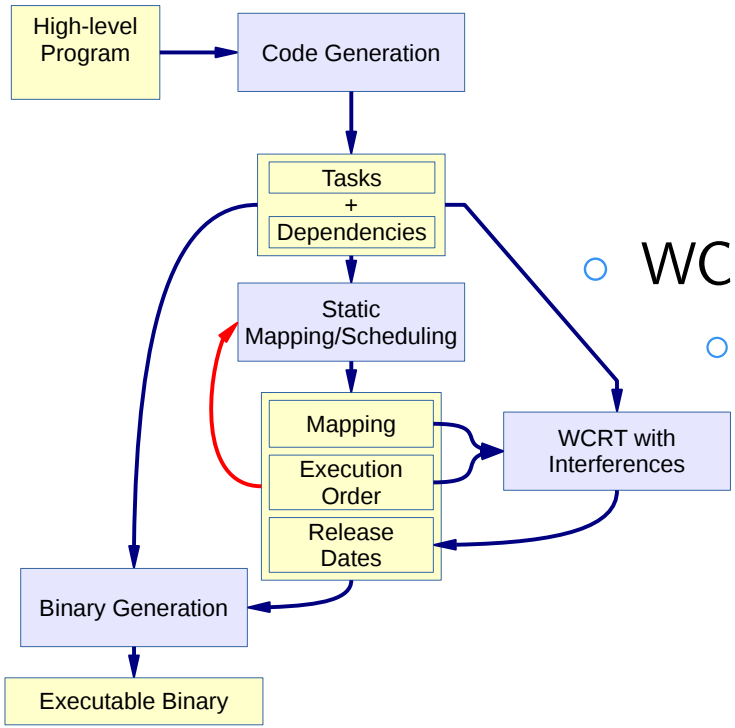

Provide new timing information to the mapping/Scheduling analysis

# Toward a Solution



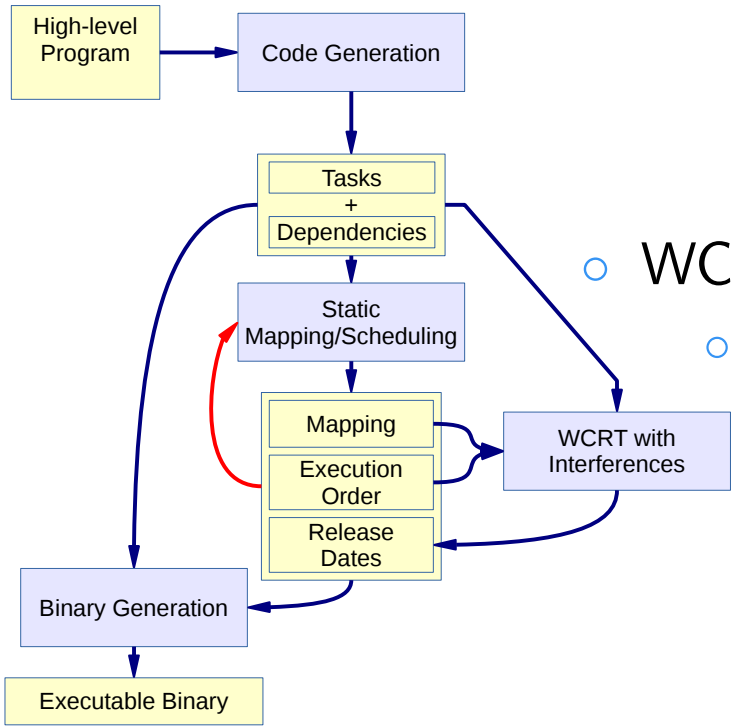

- Mapping/Scheduling:
  - Taking into account new timing information
    - Co-schedule communications and computations (Melani et al. 2015 [5])
    - Clustering non-interfering tasks (Choi et al. 2016 [2])

# Toward a Solution




- Mapping/Scheduling:
  - Taking into account new timing information
    - Co-schedule communications and computations (Melani et al. 2015 [5])
    - Clustering non-interfering tasks (Choi et al. 2016 [2])
- WCRT Analysis:
  - Trade-off: run-time/ pessimism
    Altmeyer et al. 2015 [1], Dasari et al. 2015[3]

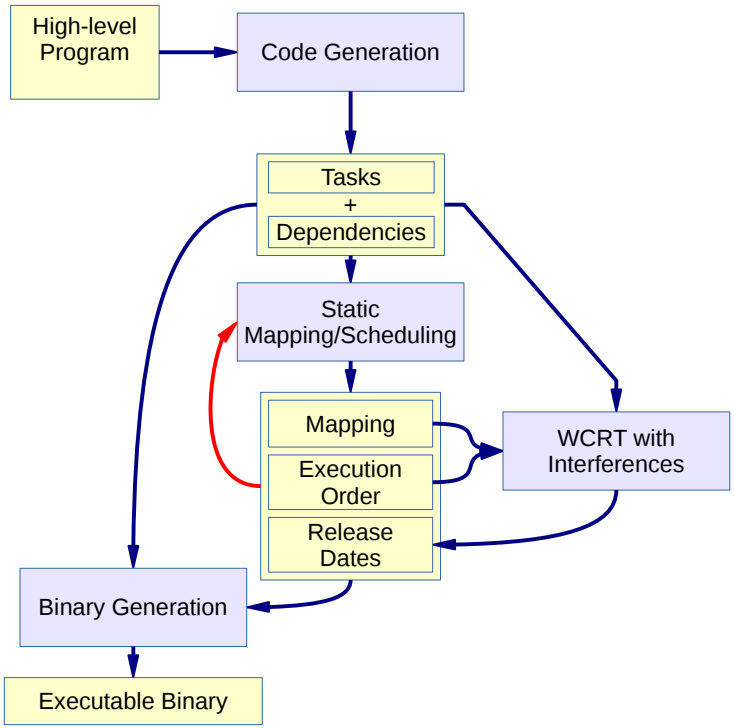
# Toward a Solution




- Mapping/Scheduling:
  - Taking into account new timing information
    - Co-schedule communications and computations (Melani et al. 2015 [5])
    - Clustering non-interfering tasks (Choi et al. 2016 [2])
- WCRT Analysis:
  - Trade-off: run-time/ pessimism
    Altmeyer et al. 2015 [1], Dasari et al. 2015[3]

Fixed-point search algorithms

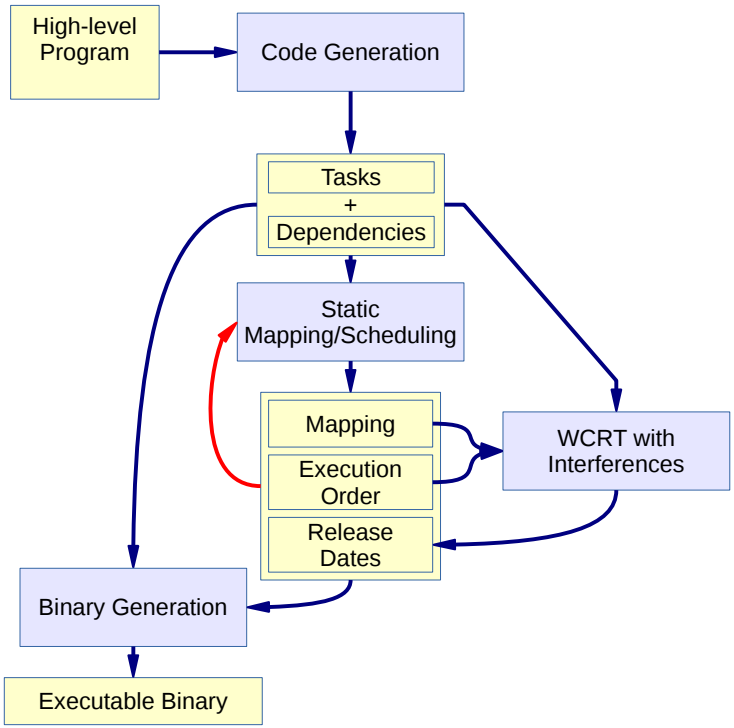
# The Open Problem




Iterate until an optimal solution is found

**What about convergence?**

# The Open Problem



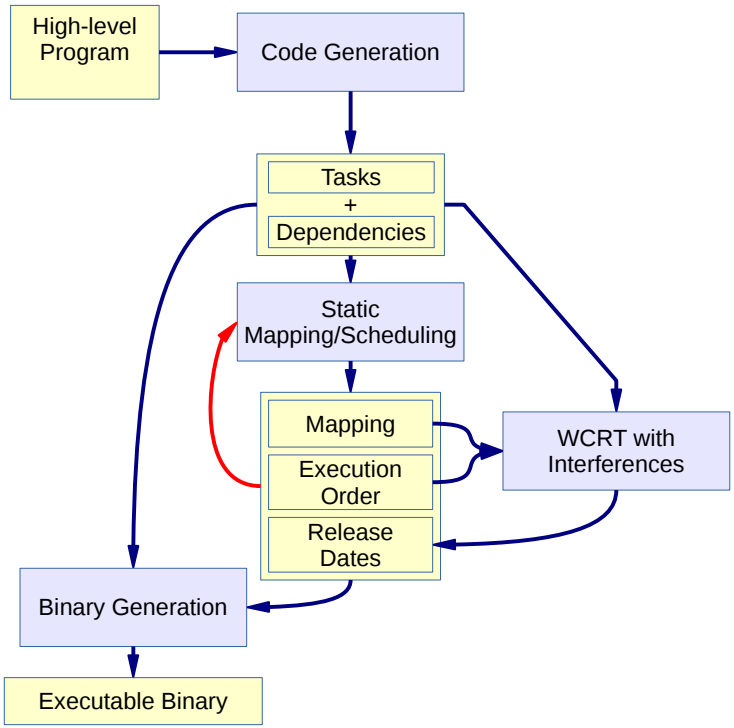

Iterate until an optimal solution is found

**What about convergence?**

Suboptimal:

- Compute several solutions, choose the best one
- **How many iterations?**

# The Open Problem




Iterate until an optimal solution is found

**What about convergence?**

Suboptimal:

- Compute several solutions, choose the best one
- **How many iterations?**

Multi/Many-core processors are a game changer in the interaction between WCRT analysis and task mapping/scheduling

# Efficient Execution of Dependent Tasks on Many-Core Processors

Hamza Rihani, Claire Maiza, Matthieu Moy

Univ. Grenoble Alpes
Verimag


# References I


S. Altmeyer, R. I. Davis, L. Indrusiak, C. Maiza, V. Nelis, and J. Reineke.

A generic and compositional framework for multicore response time analysis.

In *Proceedings of the 23rd International Conference on Real Time and Networks Systems*, RTNS '15, pages 129–138. ACM, 2015.

J. Choi, D. Kang, and S. Ha.

Conservative modeling of shared resource contention for dependent tasks in partitioned multi-core systems.

In *2016 Design, Automation Test in Europe Conference Exhibition (DATE)*, pages 181–186.

D. Dasari, V. Nelis, and B. Akesson.

A framework for memory contention analysis in multi-core platforms.

*Real-Time Systems*, pages 1–51, 2015.

# References II


G. Giannopoulou, N. Stoimenov, P. Huang, and L. Thiele.

Scheduling of mixed-criticality applications on resource-sharing multicore systems.

In *Embedded Software (EMSOFT), 2013 Proceedings of the International Conference on*, pages 1–15, Sept 2013.

A. Melani, M. Bertogna, V. Bonifaci, A. Marchetti-Spaccamela, and G. Buttazzo.

Memory-processor co-scheduling in fixed priority systems.

In *23rd ACM International Conference on Real-Time Networks and Systems (RTNS)*, Lille, France, November, 2015.

R. Pellizzoni, A. Schranzhofer, J.-J. Chen, M. Caccamo, and L. Thiele.

Worst case delay analysis for memory interference in multicore systems.

In *Proceedings of the Conference on Design, Automation and Test in Europe*, DATE '10, pages 741–746.

# References III


W. Puffitsch, E. Noulard, and C. Pagetti.

Mapping a multi-rate synchronous language to a many-core processor.

In *Real-Time and Embedded Technology and Applications Symposium (RTAS), 2013 IEEE 19th*, pages 293–302.

J. Walter and W. Nebel.

Energy–aware mapping and scheduling of large–scale macro data–flow applications.

In *1st International Workshop on Investigating Dataflow in Embedded Computing Architecture*, 2015.