Schedulability Analysis of OpenMP Applications Under Heuristic Task-to-Thread Mapping

Presenter:

Mohammad Samadi

Co-authors:

Tiago Carvalho, Luis Miguel Pinho, Sara Royuela

Polytechnic Institute of Porto & INESC TEC, Porto, Portugal

Universitat Politècnica de Catalunya, Barcelona, Spain

Barcelona Supercomputing Center, Barcelona, Spain

January 2026

Introduction

- Modern embedded platforms are employed to provide high-performance capabilities to critical applications (e.g., automotive industry).
- Parallel programming models (e.g., OpenMP) can be used in these systems to effectively exploit the parallel nature of applications.
- Task-to-thread mapping process can be performed in these models through allocation and dispatching phases to improve their performance.
- Schedulability denotes whether tasks complete within their required deadlines under a scheduling algorithm, even in worst-case conditions.
- Application response time (and especially WCRT) can also be reduced by selecting an appropriate scheduler to improve schedulability.

Motivations

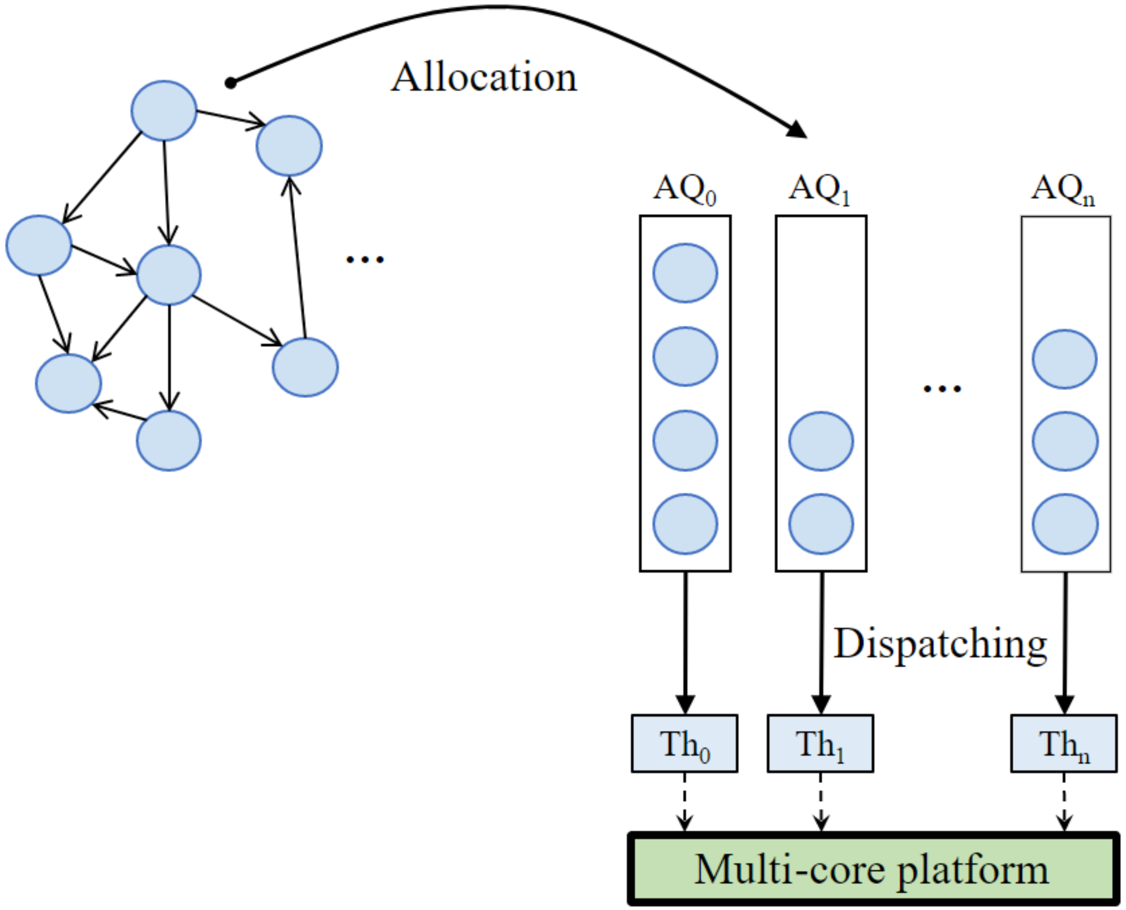
The main issues with current evaluations of static task-to-thread mapping and their impact on OpenMP applications are that:

- They do not consider recently presented task-to-thread mappings, which may improve the schedulability of the system compared to previous methods.
- Analyses do not consider large and different configurations, such as diverse applications with respect to their complexity, the number of threads used, and different possible deadlines.

Contributions

- Applying a methodology for evaluating static task-to-thread mappings in OpenMP using an integration of our simulator and an existing schedulability analysis toolset.
- Analyzing schedulability of the recent heuristic task-to-thread mapping, compared to other relevant mappings, under different aspects, including small and large random graphs, based on partitioned mapping with the limited preemptive policy.
- Evaluating all the heuristic algorithms through analysis to identify their behavior in achieving acceptable schedulability.

Heuristic Task-to-Thread Mapping



The allocation heuristics to select threads	
MNTP	Minimum Number of Task-Parts
NT	Next Thread
MRIT	Most Recent Idle Time
MTET	Minimum Total Execution Time
MTRT	Maximum Total Response Time
TMCD	Total Multi-Criteria Decision based on MNTP, MRIT, and MTET

The dispatching heuristics to select tasks	
MET	Minimum Execution Time
MRT	Maximum Response Time
MCD	Multi-Criteria Decision based on MET and MRT

Methodology
(System Model)

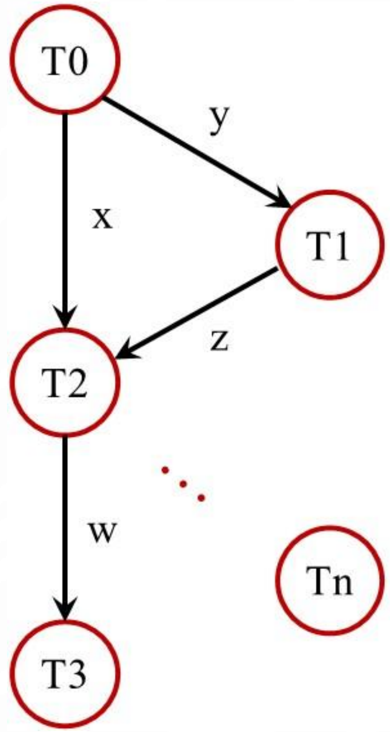
- A DAG is considered, where each node in the DAG is an OpenMP task (briefly task) and data dependencies can exist between tasks.
- The DAG includes a response time and a deadline.
- The schedulability of the tasking system is specified based on implicit and constrained deadlines.
- The homogeneous architecture is utilized for the mapping process, in which all tasks are executed by CPUs.
- The task-centric tasking model of OpenMP is considered.
- Tasks are executed in parallel using multiple threads (4 and 8 threads).

Methodology
(System Model)

OpenMP code

```
#pragma omp parallel num_threads (m)
#pragma omp single // Tmain
{
    #pragma omp task depend (out: x,y)
        T0();
    #pragma omp task depend (in: y; out: z)
        T1();
    #pragma omp task depend (in: x,z; out: w)
        T2();
    #pragma omp task depend (in: w)
        T3();
    ...
    #pragma omp task
        Tn();
}
```

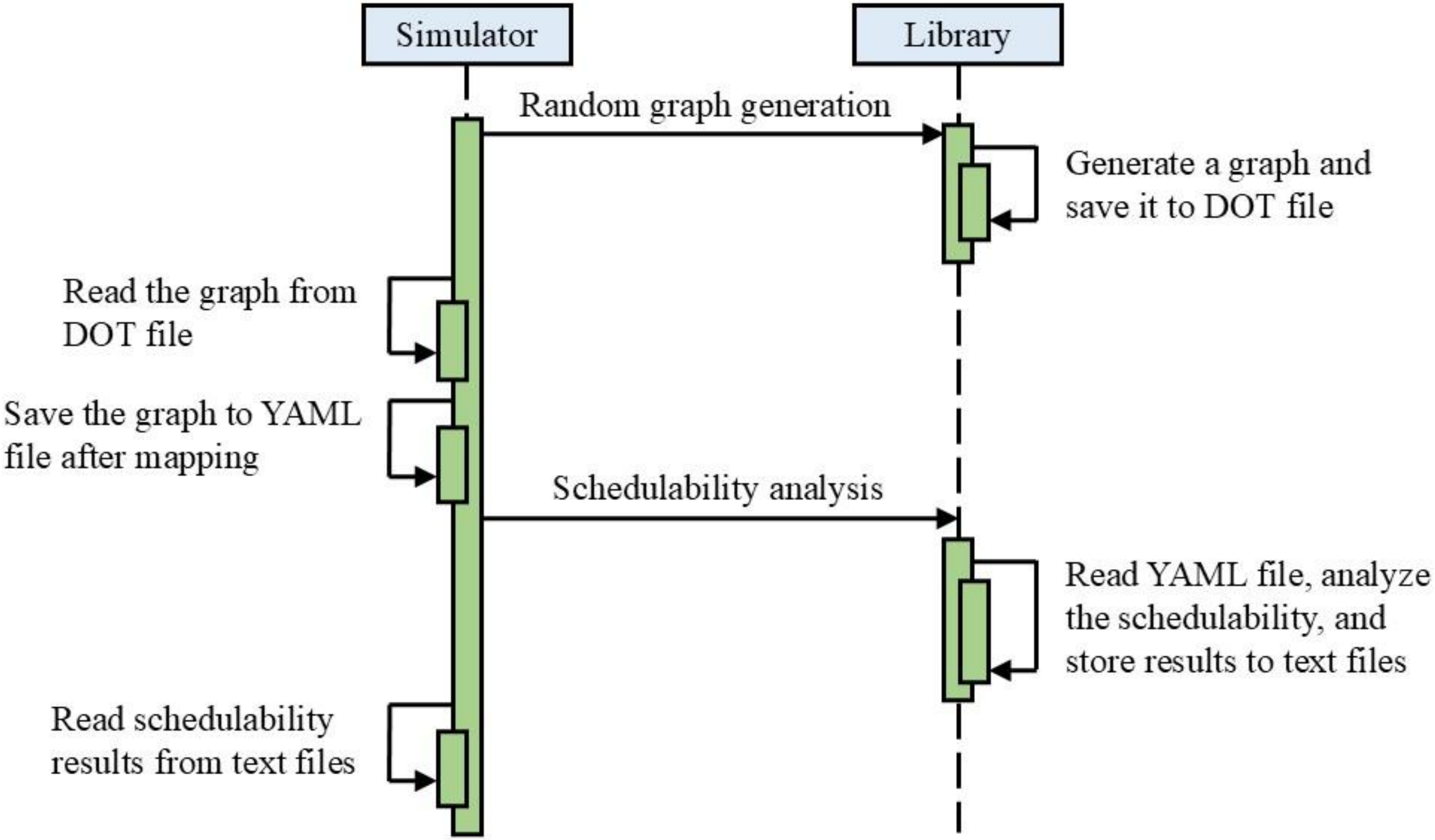
OpenMP-DAG



Methodology
(Tools)

- A DAG-scheduling library is integrated in our developed simulator.
- Two types of random graphs are used:
  - Regular graphs, with a number of tasks in the range of [1, 50] and exposing limited parallelism.
  - Wide graphs, with a number of tasks in the range of [51, 100] and exposing extensive parallelism.
- A new graph is generated in each iteration to perform the analysis under different conditions across all iterations.
- The model focuses primarily on one graph of parallel computation (single DAGs).
- The analysis is done with limited preemptive policy based on partitioned scheduling.

Methodology
(Integration)



Performance Evaluation
(Setup)

- The evaluation process is performed based on schedulability and difference in worst-case response time (WCRT) obtained using the mapping methods.
- The schedulability is specified based on implicit and constrained deadlines, with a deadline equal to the period in implicit deadlines, and smaller than or equal to the period in constrained deadlines.
- The analysis is conducted based on utilization of 1, 2, 4, and 8, where the utilization of the graphs is calculated as follows.

$$Utilization = \frac{Vol}{T}$$

Performance Evaluation
(Setup)

- The schedulability given with each method is determined as follows.

$$Schedulability = \frac{\sum_{i=1}^{itr} SchedVal_i}{itr} * 100$$

- The difference in WCRT is determined as below.

$$Diff = 1 - \frac{WCRT_{WFPA}}{WCRT_{meth}} * 100$$

- The cases considered in the analysis are as following.

Case #	Size of graph	Number of tasks	Deadline policy
1	Small	1-50	Implicit
2	Large	51-100	Implicit
3	Small	1-50	Constrained
4	Large	51-100	Constrained

Performance Evaluation
(Results)

	Case 1 with 4 threads	Case 2 with 4 threads	Case 2 with 8 threads	Case 3 with 4 threads	Case 4 with 4 threads	Case 4 with 8 threads
BFS	100%	93%	100%	35%	4%	10%
LPT	100%	91%	100%	34%	2%	17%
SPT	100%	88%	100%	38%	2%	15%
LNSNL	100%	91%	100%	37%	5%	15%
MNTP-MET	100%	99%	100%	27%	3%	24%
MNTP-MRT	100%	100%	100%	28%	4%	25%
MNTP-MCD	100%	100%	100%	28%	4%	25%
MTET-MET	100%	100%	100%	32%	6%	25%
MTET-MRT	100%	100%	100%	33%	5%	25%
MTET-MCD	100%	100%	100%	33%	5%	25%
TMCD-MET	100%	99%	100%	32%	6%	25%
TMCD-MRT	100%	100%	100%	34%	2%	24%
TMCD-MCD	100%	100%	100%	34%	2%	24%
WFPA	100%	87%	100%	37%	1%	19%

Performance Evaluation
(Results)

In the comparison results:

- Negative difference (started with '-') means that the WCRT obtained with a mapping method is reduced compared to the baseline (i.e., WFPA).
- Positive difference means that the WCRT achieved using the mapping method is increased compared to the baseline.
- '0' means that there is no difference between the performance of the mapping method and the baseline.

Performance Evaluation
(Results)

Case 1 with 4 Threads

	1	2	4	8
BFS	3%	0%	10%	1%
LPT	3%	0%	8%	3%
SPT	3%	0%	10%	4%
LNSNL	-1%	0%	10%	4%
MNTP-MET	3%	-6%	5%	0%
MNTP-MRT	-2%	-9%	3%	-3%
MNTP-MCD	-2%	-9%	3%	-3%
MTET-MET	1%	-5%	0%	-2%
MTET-MRT	-12%	-4%	1%	-2%
MTET-MCD	-12%	-4%	1%	-2%
TMCD-MET	1%	-5%	-2%	-5%
TMCD-MRT	-8%	-4%	3%	1%
TMCD-MCD	-8%	-4%	3%	1%
WFPA	Baseline			

Utilization

Evaluation

The schedulability results reveal no significant difference in mapping performance.

The WCRT results indicate that the WCRT obtained with these heuristics is lower than that reported in the other methods.

MTET-X and TMCD-X perform better than MNTP-X, indicating that some metrics, such as task execution time, can efficiently reduce WCRT.

Performance Evaluation
(Results)

Case 2 with 4 threads

	1	2	4	8
BFS	-13%	9%	7%	53%
LPT	40%	26%	39%	-27%
SPT	8%	32%	3%	24%
LNSNL	33%	31%	-24%	23%
MNTP-MET	-60%	-41%	-85%	-93%
MNTP-MRT	-59%	-35%	-57%	-86%
MNTP-MCD	-59%	-35%	-57%	-86%
MTET-MET	-73%	-37%	-84%	-91%
MTET-MRT	-69%	-36%	-95%	-93%
MTET-MCD	-69%	-36%	-95%	-93%
TMCD-MET	-73%	-35%	-84%	-80%
TMCD-MRT	-69%	-36%	-91%	-96%
TMCD-MCD	-69%	-36%	-91%	-96%
WFPA	Baseline			

Utilization

Evaluation

All heuristics achieve higher efficiency than the other methods, as most graphs are schedulable with the heuristics when the utilization is set to 1.

The WCRT given with these heuristics is lower than that obtained with the other methods.

Efficiency of the heuristics in large DAGs is generally better than that in small DAGs, as the dispatching phase is better suited to cases where queues contain many tasks and there are multiple options for selecting them.

Performance Evaluation
(Results)

Case 2 with 8 threads

	1	2	4	8
BFS	0%	6%	5%	65%
LPT	36%	34%	62%	-2%
SPT	6%	4%	-2%	24%
LNSNL	-7%	33%	-2%	1%
MNTP-MET	-18%	-1%	-15%	-17%
MNTP-MRT	-10%	-11%	-14%	-10%
MNTP-MCD	-10%	-11%	-14%	-10%
MTET-MET	-13%	-11%	-15%	-17%
MTET-MRT	-26%	-8%	-14%	-11%
MTET-MCD	-26%	-8%	-14%	-11%
TMCD-MET	-21%	-11%	-15%	-15%
TMCD-MRT	-26%	-8%	-14%	-11%
TMCD-MCD	-26%	-8%	-14%	-11%
WFPA	Baseline			

Utilization

Evaluation

The schedulability obtained using all the methods is the same, since all graphs are schedulable.

The schedulability rate improves by increasing the number of threads, as the heuristics can better exploit the capabilities of the possible mappings.

The difference between the WCRT obtained with these heuristics and that obtained with the preliminary methods is noticeable.

Performance Evaluation
(Results)

Case 3 with 4 threads

	1	2	4	8
BFS	-1%	0%	11%	-2%
LPT	0%	0%	-2%	1%
SPT	0%	0%	0%	0%
LNSNL	0%	0%	0%	1%
MNTP-MET	-1%	-3%	-7%	0%
MNTP-MRT	-2%	-5%	-2%	-3%
MNTP-MCD	-2%	-5%	-2%	-3%
MTET-MET	-1%	-9%	-7%	-3%
MTET-MRT	-4%	-5%	-2%	-5%
MTET-MCD	-4%	-5%	-2%	-5%
TMCD-MET	-3%	-9%	-4%	-2%
TMCD-MRT	-4%	-10%	-3%	2%
TMCD-MCD	-4%	-10%	-3%	2%
WFPA	Baseline			

Utilization

Evaluation

The schedulability of the heuristics is slightly lower than that of some previous methods. But MTET-X and TMCD-X provide very similar results.

All these heuristics perform more efficiently in minimizing WCRT than the other methods.

The results for small graphs with implicit deadline policy indicate that the heuristics perform more efficiently under this policy, since the system (including a few tasks) misses its deadline under the constrained deadline policy.

Performance Evaluation

(Results)

Case 4 with 4 threads

	1	2	4	8
BFS	5%	7%	-2%	38%
LPT	4%	17%	1%	-14%
SPT	-21%	9%	35%	55%
LNSNL	70%	17%	-40%	26%
MNTP-MET	-67%	-34%	-95%	-78%
MNTP-MRT	-78%	-48%	-103%	-77%
MNTP-MCD	-78%	-48%	-103%	-77%
MTET-MET	-89%	-47%	-99%	-78%
MTET-MRT	-87%	-47%	-98%	-77%
MTET-MCD	-87%	-47%	-98%	-77%
TMCD-MET	-94%	-41%	-68%	-65%
TMCD-MRT	-21%	-49%	-99%	-79%
TMCD-MCD	-21%	-49%	-99%	-79%
WFPA	Baseline			

Utilization

Evaluation

The schedulability obtained with all the methods is low, and the performance of the heuristics, especially MTET-X, is slightly better than some other methods, including LPT, SPT, and WFPA.

The heuristics perform efficiently in large DAGs, so a shorter deadline, as in the constrained deadline system, has little effect on their efficiency.

The WCRT obtained with these heuristics is lower than that obtained with the other methods, especially with the graph utilization of 4.

Performance Evaluation

(Results)

Case 4 with 8 threads

	1	2	4	8
BFS	67%	-6%	-50%	30%
LPT	23%	-22%	-53%	47%
SPT	-1%	-2%	-53%	-5%
LNSNL	0%	29%	-45%	-5%
MNTP-MET	-4%	-35%	-61%	-11%
MNTP-MRT	-4%	-30%	-61%	-14%
MNTP-MCD	-4%	-30%	-61%	-14%
MTET-MET	-4%	-29%	-61%	-14%
MTET-MRT	-4%	-24%	-61%	-14%
MTET-MCD	-4%	-24%	-61%	-14%
TMCD-MET	-4%	-29%	-61%	-14%
TMCD-MRT	-4%	-35%	-61%	-14%
TMCD-MCD	-4%	-35%	-61%	-14%
WFPA	Baseline			

Utilization

Evaluation

The schedulability of the graphs obtained with the heuristics, compared to the other methods, especially BFS, SPT, and LNSNL, is noticeable.

The difference in the WCRT achieved with these heuristics and the previous ones is very considerable, especially in the utilization of 4.

The heuristics perform more efficiently with 8 threads, as large DAGs, including many tasks, allow for a more suitable allocation in the ready queues when the number of threads is high.

Conclusion

- The schedulability obtained with most heuristics from the heuristic mapping for large DAGs with implicit deadlines, running using 4 threads, and with constrained deadlines, especially in cases running using 8 threads, is higher than the previous methods.
- Most graphs generated based on the utilization of 2, 4, and 8 are not schedulable with any of the methods.
- The WCRT is minimized with most heuristics, compared to the others, especially in the cases for large DAGs with 4 threads.
- The schedulability of DAGs generated based on implicit deadlines is higher than that generated based on constrained deadlines, in all the cases.

Future Works

- Schedulability analysis of real-world programs, especially those used in artificial intelligence applications with enormous workloads, will be conducted under different application and system configurations.
- Schedulability analysis will be performed based on other policies, such as global or semi-partitioned approaches.

Thank you for your attention!

For more information:
Please contact me at mohammad.samadi@upc.edu.