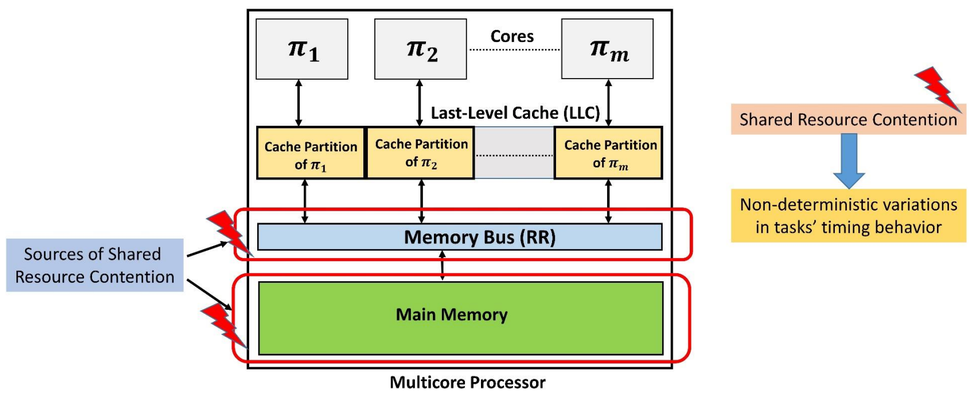
A Holistic Approach to WCRT Analysis for Multicore Systems

CISTER – Research Centre in Real-Time & Embedded Computing Systems

Jatin Arora, Syed Aftab Rashid, Cláudio Maia, Geoffrey Nelissen, Eduardo Tovar
{jatin;syara;crrm;emt}@isep.ipp.pt, g.r.r.j.p.nelissen@tue.nl

1. Motivation

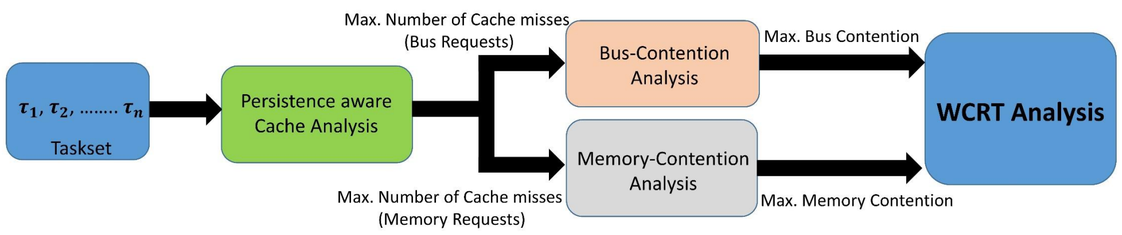
- **Shared resource contention** in multicore system may lead to non-deterministic variations in **WCET** (Worst-Case Execution Time)/ **WCRT** (Worst-Case Response Time).



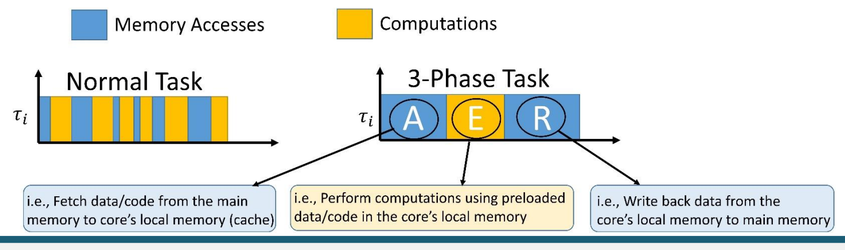
- Most existing works analyze each shared resource independently.
- Independent analysis of inter-core contention due to each shared resource, e.g., last-level cache (LLC), memory bus, main memory, etc., may lead to pessimistic results.
- **Holistic analysis** considering **interdependence** between shared resources may lead to tighter bounds on inter-core shared resource contention.

2. Contributions

- **Holistic approach** towards shared resource contention-aware WCRT analysis for multicore systems.

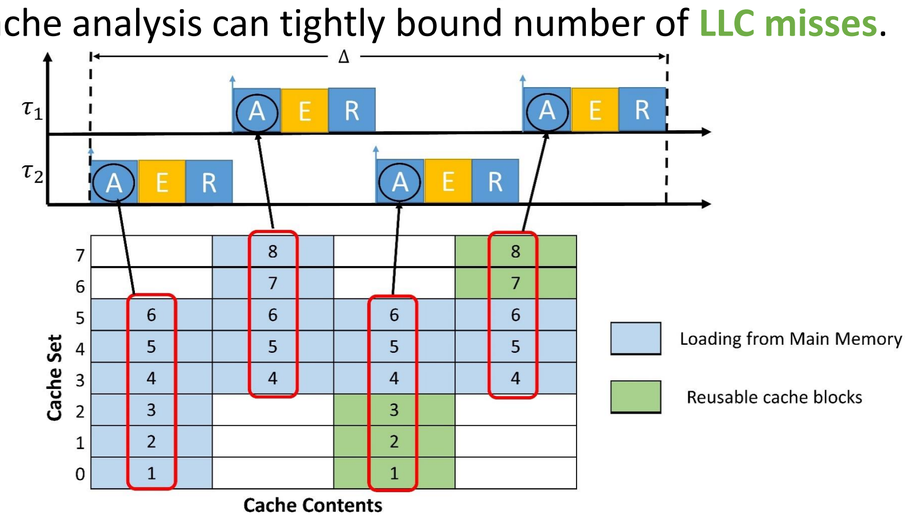


- **Cache Persistence**, i.e., cache reuse, aware cache analysis to bound LLC misses of tasks.
- Cache-aware bus contention analysis leading to tighter bounds on memory bus contention.
- Use of more predictable execution model, i.e., 3-phase model[1].



3. Cache Analysis

- Cache analysis can tightly bound number of **LLC misses**.



Total main memory fetches **without cache-persistence** during Δ = 10 (τ_1) +12 (τ_2)= 22
Total main memory fetches **with cache-persistence** during Δ= 8 (τ_1) +9 (τ_2)= 17

3.1 Bounding LLC Misses

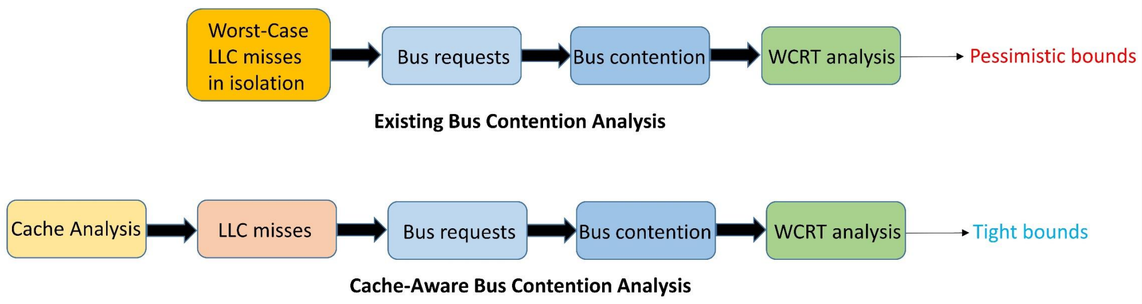
- The maximum number of memory requests made by the ***A-phases*** of τ_i during any time window Δ is given by:

$$\min\left(\left\lceil\frac{\Delta}{T_i}\right\rceil \times MD_i^A, |PCB_i| + \overline{MD_i^A} + \left(\left\lceil\frac{\Delta}{T_i}\right\rceil - 1\right) \times \left(\overline{MD_i^A} + |\rho_i|\right)\right)$$

- $\left\lceil\frac{\Delta}{T_i}\right\rceil$: Max number of jobs of τ_i during time window Δ
- MD_i^A: Maximum memory demand in isolation of the A-phase of τ_i
- $|PCB_i|$: Set of Persistent Cache blocks of τ_i
- $\overline{MD_i^A}$: Residual Memory Demand of the A-phase of τ_i
- $|\rho_i|$: CPRO suffered by the A-phase of τ_i due to HP tasks

4. Bus Contention Analysis

- State-of-the-art bus contention analysis for 3-phase tasks [2,3] does not account for cache persistence, leading to overestimation in the number of bus requests of tasks.
- Overestimation in the number of bus requests leads to pessimistic bounds on the bus contention.



5. Memory Contention

- Existing memory contention analysis [4] does not account for **cache reuse**.
- This leads to overestimation in the number of memory requests and eventually the **memory contention** suffered by the tasks.
- Memory contention analysis that considers the interdependence between caches, bus and main memory may lead to tighter WCRT bounds.

6. Ongoing Works

The following are the ongoing/future works:

- Cache-aware bus contention analysis for the 3-phase task model.
- Cache and bus-aware memory contention analysis.
- Memory contention analysis for 3-phase tasks considering different configurations.

References

[1] G. Durrieu et al., "Predictable Flight Management System Implementation on a Multicore Processor", ERTS'14, 2014.

[2] C. Maia et al., "Schedulability analysis for global fixed-priority scheduling of the 3-phase task model", 2017 IEEE RTCSA, 2017.

[3] J. Arora et al., "Schedulability analysis for 3-phase tasks with partitioned fixed-priority scheduling" Journal of Systems Architecture, 2022

[4] D. Casini et al., "A Holistic Memory Contention Analysis for Parallel Real-Time Tasks under Partitioned Scheduling" IEEE RTAS, 2020

Co-financed by: UIDP/04234/2020
FCT Fundação para a Ciência e a Tecnologia

CISTER Research Centre
ISEP, Polytechnic Institute of Porto
Rua Dr. Antº Bernardino de Almeida, 431
4249-015 Porto, Portugal
+351 228 340 502
www.cister-labs.pt
info@cister-labs.pt
facebook.com/cisterlabspt
isep Instituto Superior de Engenharia do Porto
P.PORTO