Technische Universität Braunschweig

Self-aware safety-critical systems-on-chip through an online system controller approach

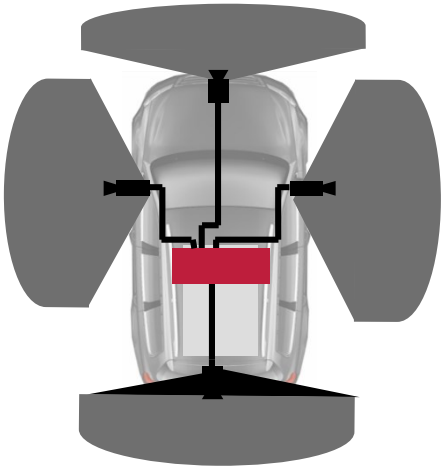
Thawra Kadeed and Rolf Ernst
TU Braunschweig, Germnay

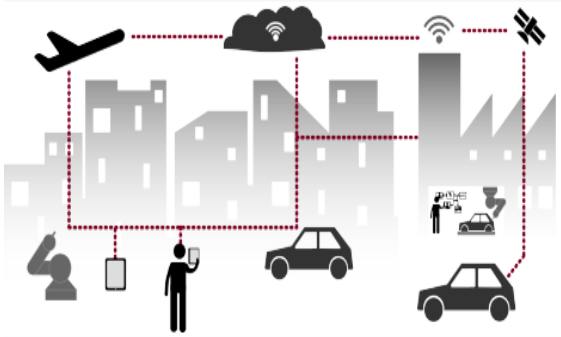
IWCR19 – Stuttgart - Germany

July 29th, 2019

Motivation

- many-core systems are reaching safety-critical real-time domains e.g.:
  - sensor fusion and recognition in highly automated driving
- complex multiprocessing architectures
  - mixed criticality
- variability versus QoS goals
- automated solutions are no longer sufficient
- service guarantees and reliable operation
- limited power and cost budget





Motivation – Safety Requirements

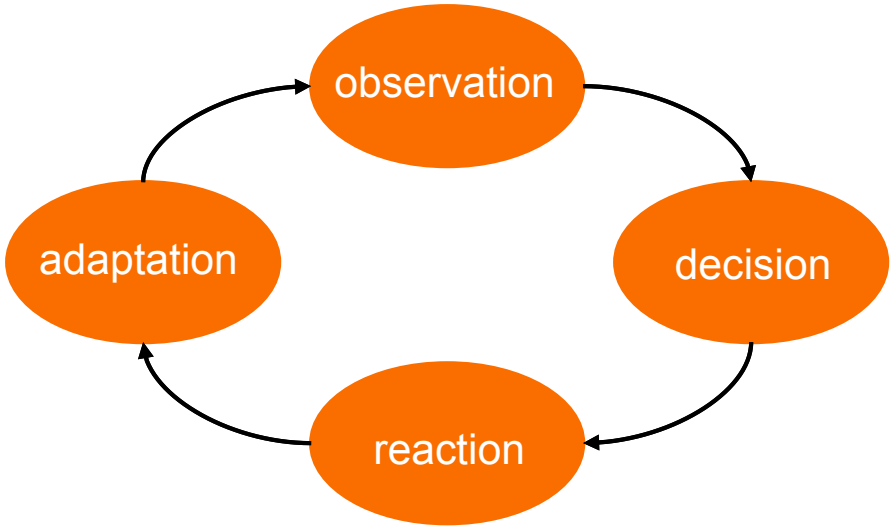
ISO 26262

- **Functional safety**
  - standards require functions designed according to highest relevant safety level → highly expensive
  - freedom from interference
    isolated cores, run independently from each other
- **Consequences:**
  - **Temporal guarantees**
    - App tasks/software modules are time critical, i.e., their deadlines must be met
    - worst-case behavior should be bounded and thus predicted
  - **Reliability**
    fail-operational: continuity of correct service
  - **Availability**
    readiness for correct service

Motivation – self awareness

Modern/future systems require self-awareness mechanisms to overcome the complexity challenge while adhering to strict safety requirements, and thus providing:

- system self-diagnosis
  - ability for observation
  - ability for error-detection
  - ability for error-reporting
- system self-organization
  - ability for error-handling
  - keep guaranteed service
- system self-adaptation
  - adoption of new operation conditions
  - adapts models to new system states

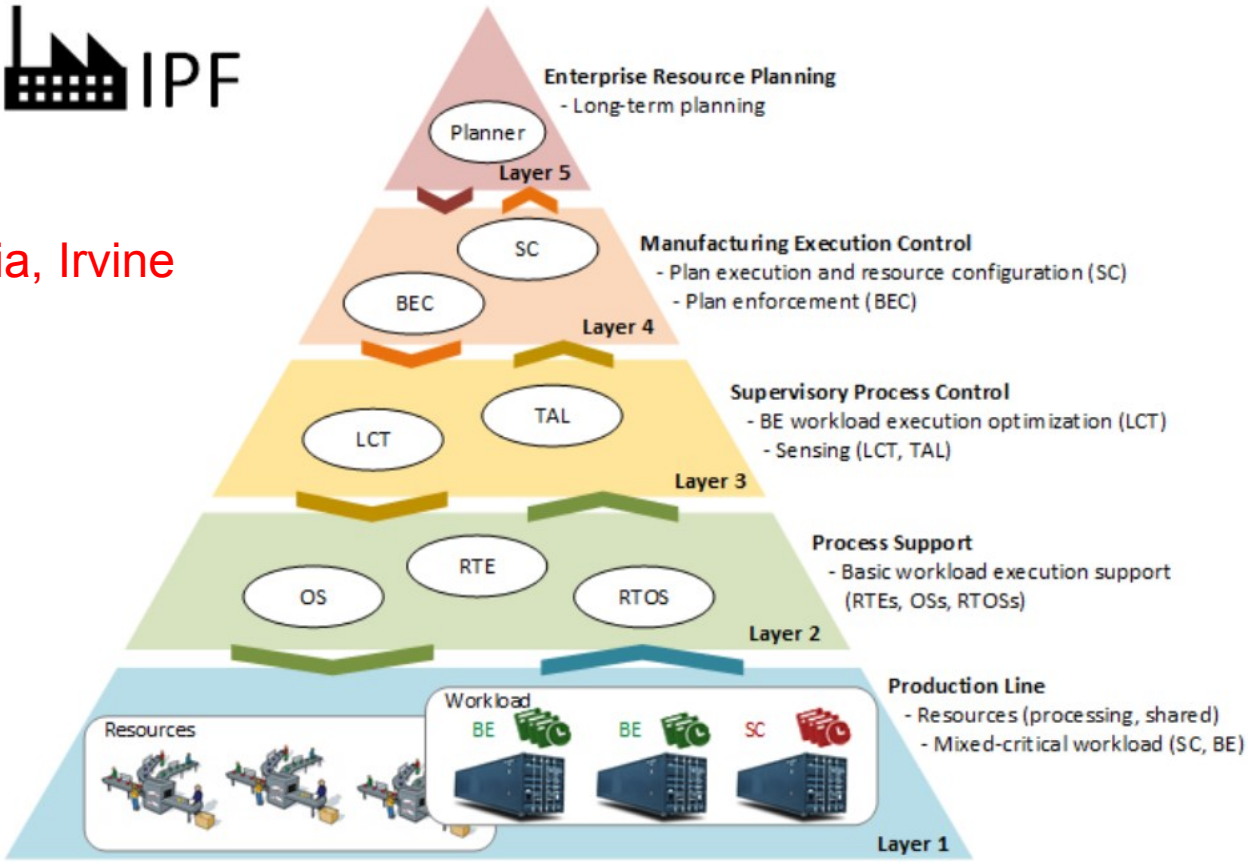


system self-awareness states

IPF (information processing factory) project

Joint project:

- University of California, Irvine
- TU Munich
- TU Braunschweig



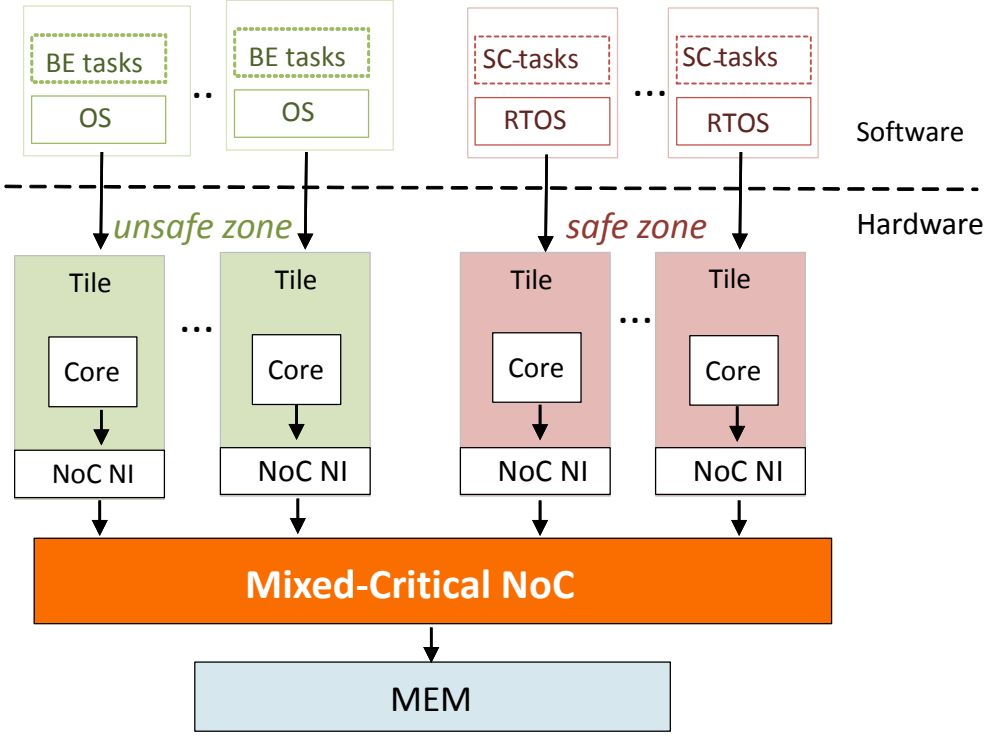
IPF's five-layer organization

published in arxiv [arxiv2019]

Outline

- Motivation
- Self awareness at different system layers
- Self-aware NoC through energy management
  - Safe and dynamic power-gating of real-time NoC routers
- Self-aware SoC through error handling
- Conclusion

Many-core NoC-based systems: as standards require



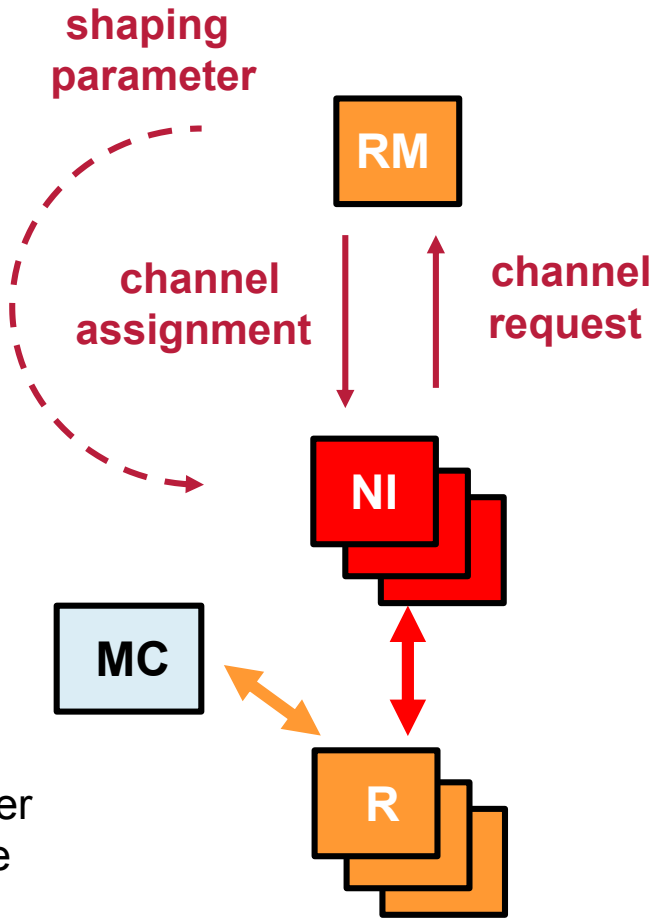
NoC-based mixed-critical system

NoC: network-on-chip
NI: network-interface
RTOS: real-time operating system
BE: best-effort
SC: safety critical

NoC Resource Manager (NoC-RM)

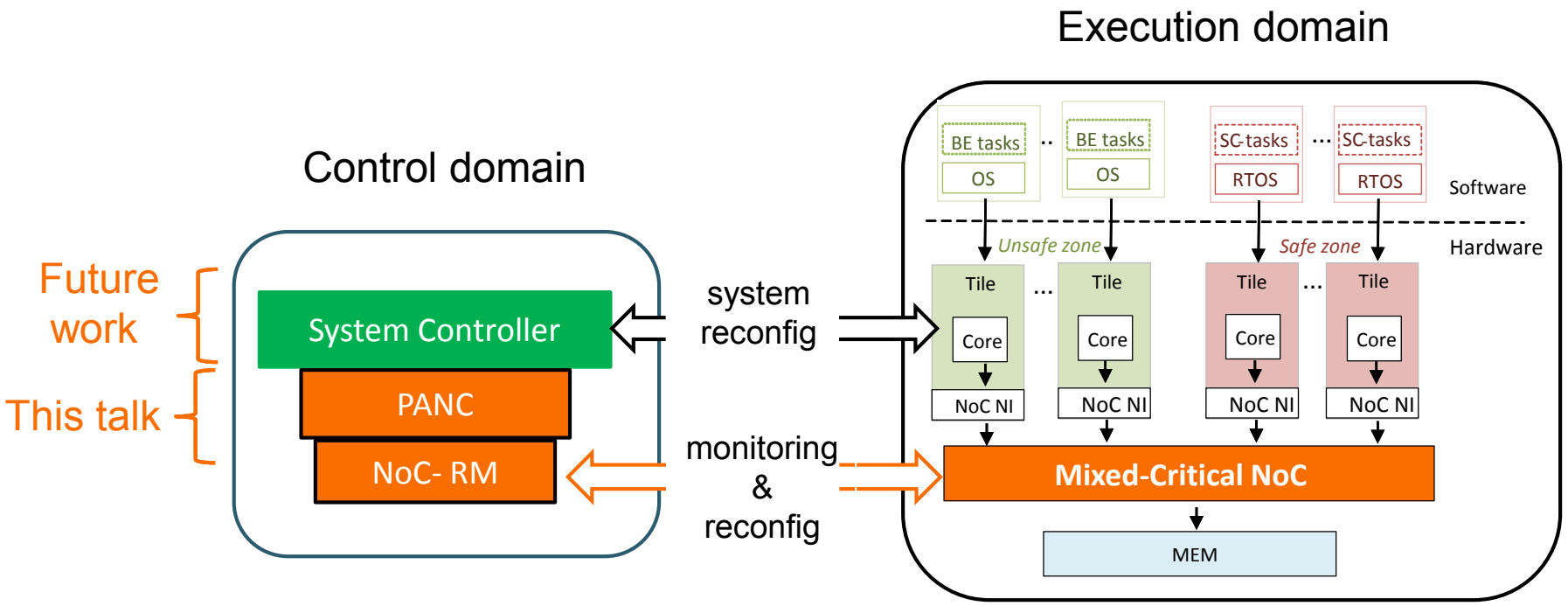
- **decouple admission control from flow control**

- include Resource Manager:
  - controls resource assignment in NoC
  - RM, as a trusted unit, is allowed to reporogram the NIs.
  - **achieving functional safety in on-chip interconnects**



RM: resource manager
NI: network interface
R: router
MC: memory controller

Self awareness at multiple system layers



System self-awareness model

NoC-RM: [kostrzewa2017a] , [kostrzewa2017b]

RM: resource manager
PANC: power aware NoC controller

Outline

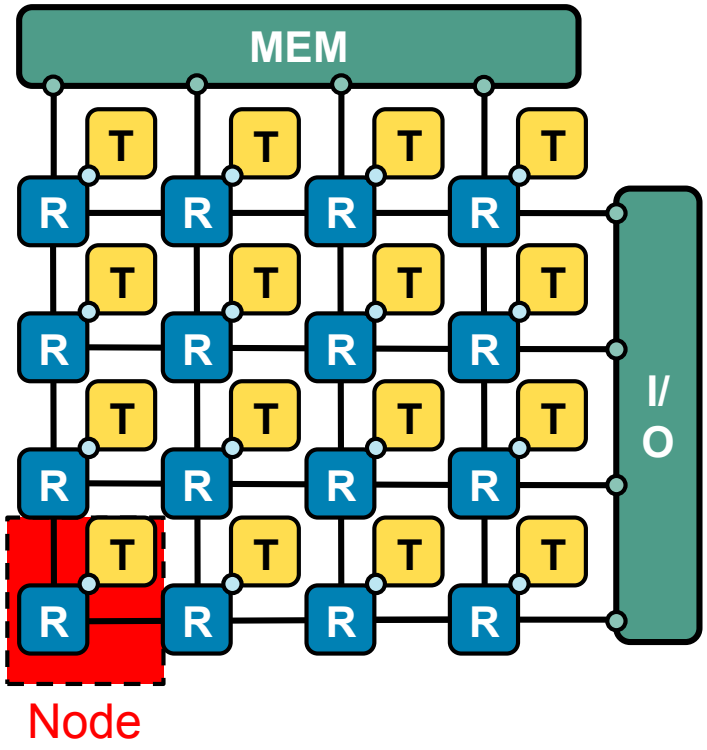
- Motivation
- Self awareness at different system layers
- Self-aware NoC through energy management
  - Safe and dynamic power-gating of real-time NoC routers
- Self-aware SoC through error handling
- Conclusion

Self-aware NoC through energy management

- NoC couples the execution of mixed-critical application and adopts complex architectures

Consequences:

- energy rapidly increase
- temperature runaway
- aging
- reliability reduction



NoC power-gating: State-of-the-art

Power-gating features leakage power-savings

- **Panthre**: (Power-aware NoC through Routing and Topology Reconfiguration) [Parikh2014]

  → Low power savings

- **Power Punch**: Towards Non-blocking power-gating of NoC Routers [Chen2015]

  → High performance but moderate power savings

- **Toot**: an efficient and scalable power-gating method for NoC routers [Farrokhbakht2016]

  → High power savings at the cost of performance

NoC power-gating: State-of-the-art

Power-gating features leakage power-savings

- **Panthre**: (Power-aware NoC through Routing and Topology Reconfiguration)
  [Parikh2014]

  ⟶ Low power savings

- **Pc**
  [C

- **T**
  [Farokhbakht2016]

  ⟶ High power savings at the cost of performance

Goal

Developing a global and efficient power savings approach for hard real-time NoCs, while providing temporal guarantees of real-time applications

Idea

PANC (Power-Aware NoC controller) approach

Outline

- Motivation
- Self awareness at different system layers
- Self-aware NoC through energy management
  - Safe and dynamic power-gating of real-time NoC routers
- Self-aware SoC through error handling
- Conclusion

Safe and dynamic power-gating of NoC routers

PANC through a control layer:

advantages:

- safely adjusts power dissipation
- isolated from the existing NoC data layer
- no complex modifications
- routers can be completely turned-off

comprised of:

PANC –

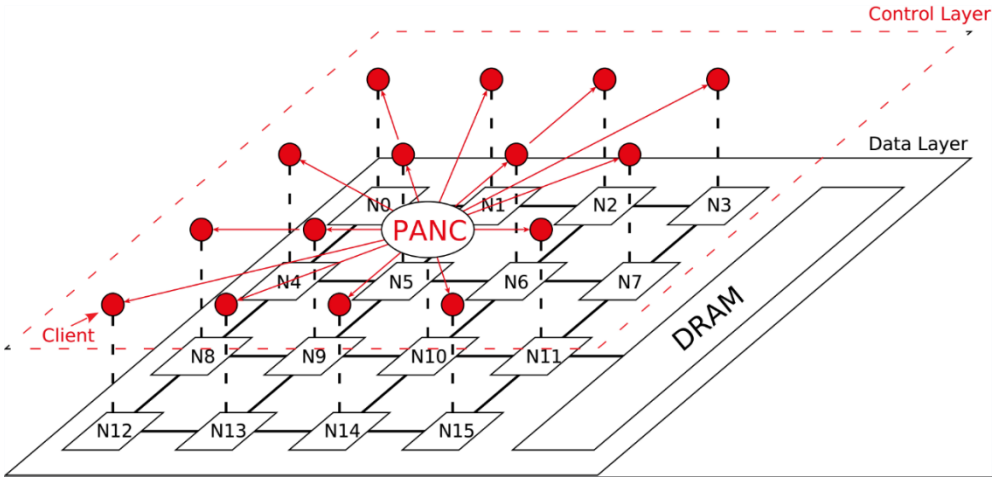
- trusted predictable unit
- turns on/off routers
- protocol-based synchronization

clients –

- local supervisors in nodes
- synchronize senders NoC access with PANC

Interconnect –

direct links



PANC within the control-layer (red color) for power-savings by monitoring traffic in data-layer

N: Node

Power-gating latency overhead

Timing Parameters

- **Wake-Up (WU) latency:** the number of cycles necessary to turn on a router
- **Break-Even Time (BET):** which determines the number of cycles the router should stay off before waking up to overcome WU power overhead

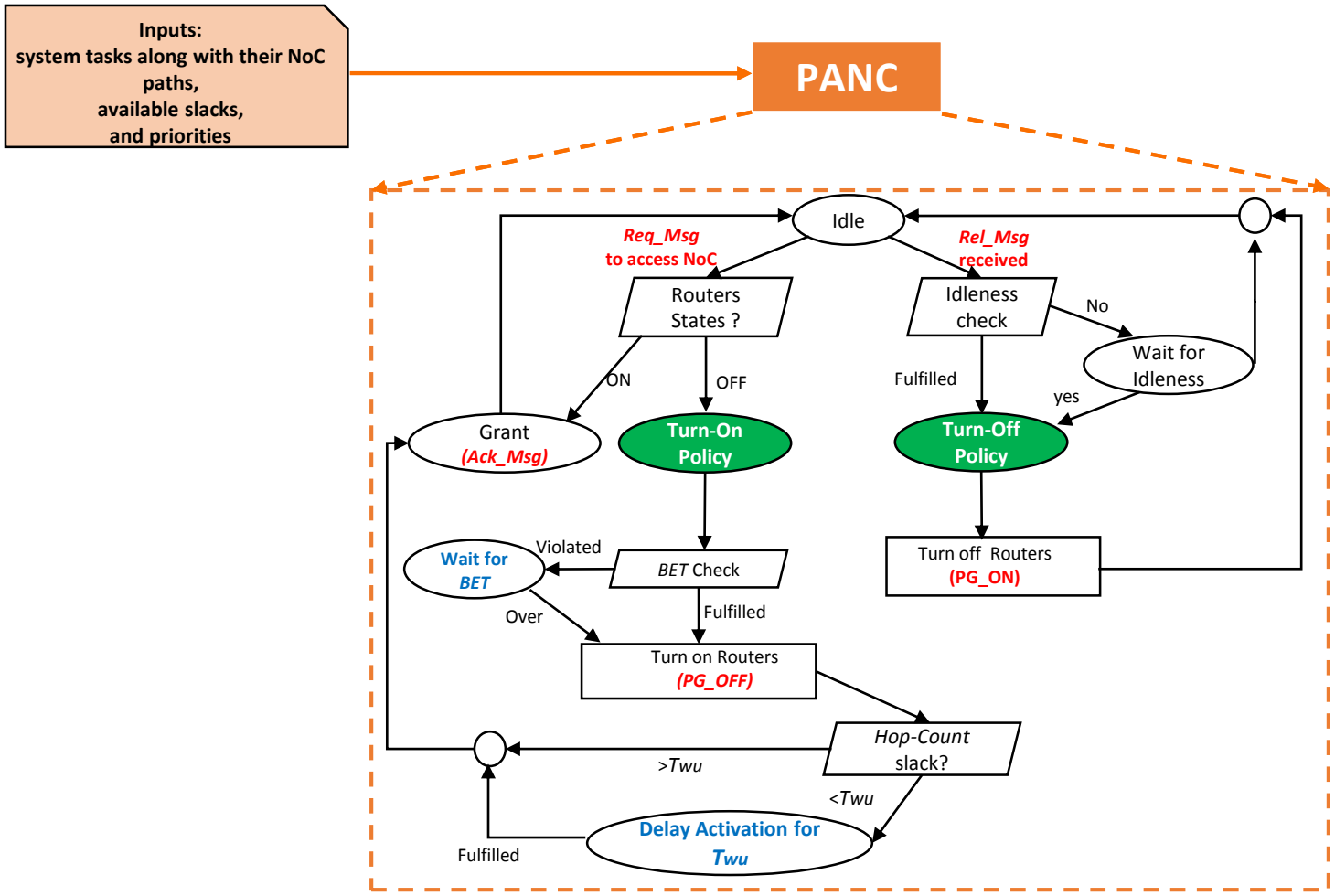
PANC must account to:

- **BET cycles:** to avoid power loss
- **safely switch between router states (on/off):** provides temporal guarantees

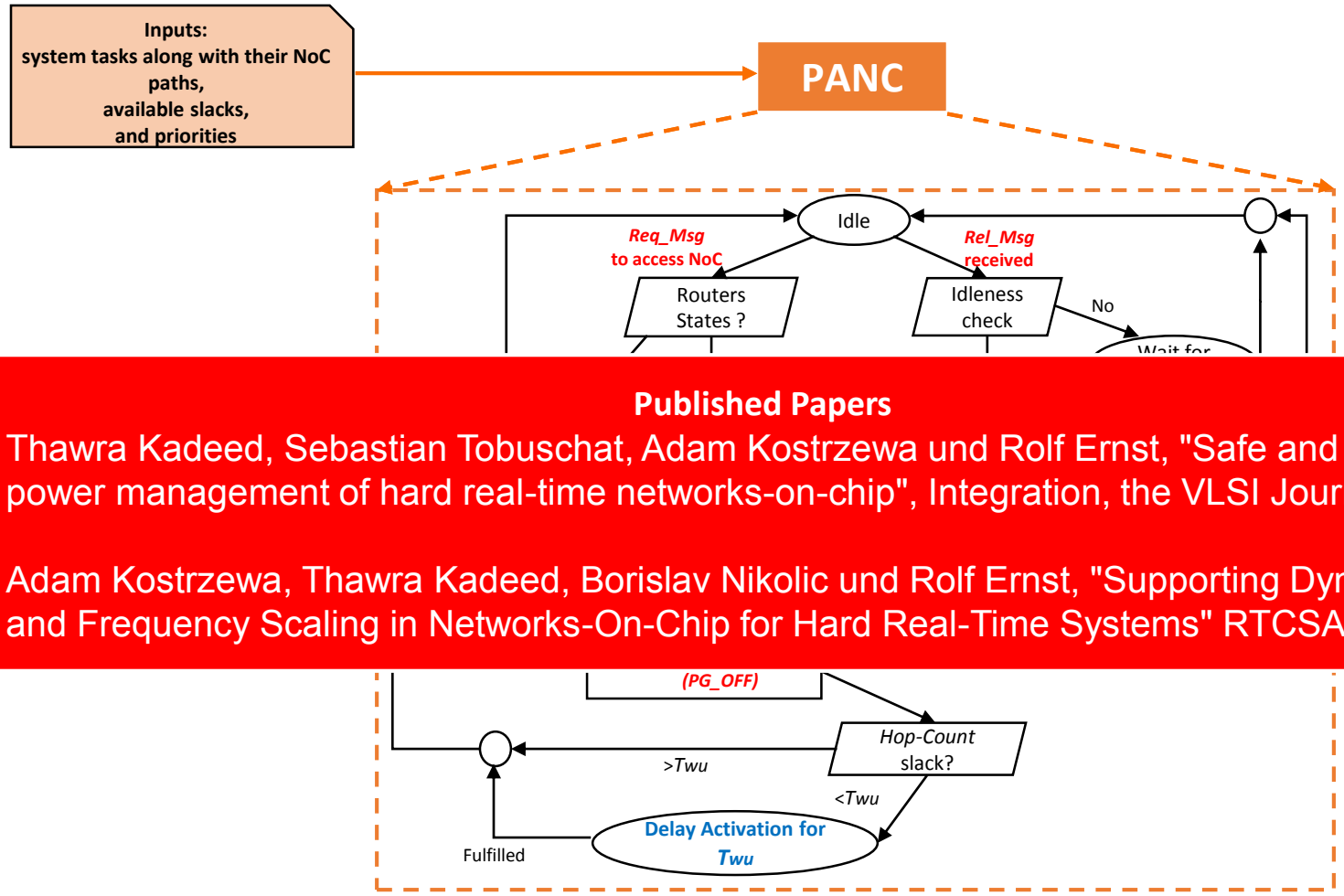
PANC exploits:

- **Task deadline slack**: $D_i - R_i$
- **Hop-count slack**: how far the respective router from the associated sender

PANC workflow



PANC workflow



Published Papers

- Thawra Kadeed, Sebastian Tobuschat, Adam Kostrzewa und Rolf Ernst, "Safe and efficient power management of hard real-time networks-on-chip", Integration, the VLSI Journal, 2018
- Adam Kostrzewa, Thawra Kadeed, Borislav Nikolic und Rolf Ernst, "Supporting Dynamic Voltage and Frequency Scaling in Networks-On-Chip for Hard Real-Time Systems" RTCSA, 2018

Temporal guarantees and predictability

- mechanism description → mathematical model
- calculate the worst-case response-time of a transmission including the approach overhead
- validate against the deadlines

$$R_i^{PANC} \leq D_i$$

- based on the extracted slacks:
  - PANC checks the available slacks and decides whether it is allowed to delay a task by BET cycles
- analysis framework: holistic analysis [Nikolic2019]
  - framework extension to get the latencies for communication protocol and power switches induced by our approach

PANC - Evaluation

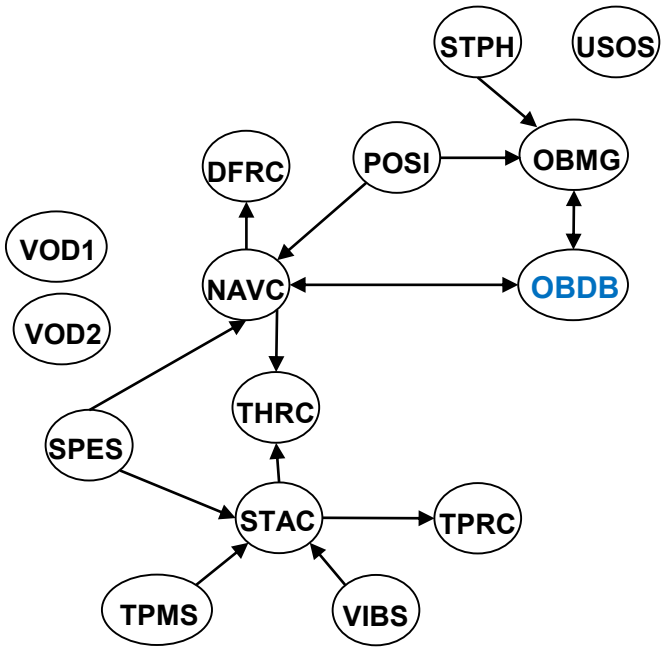
- **simulations:**
  - simulation using the OMNeT++ framework + HNOCs library
  - 2D mesh NoC (4x4)
  - four (virtual) channels

- **input data**

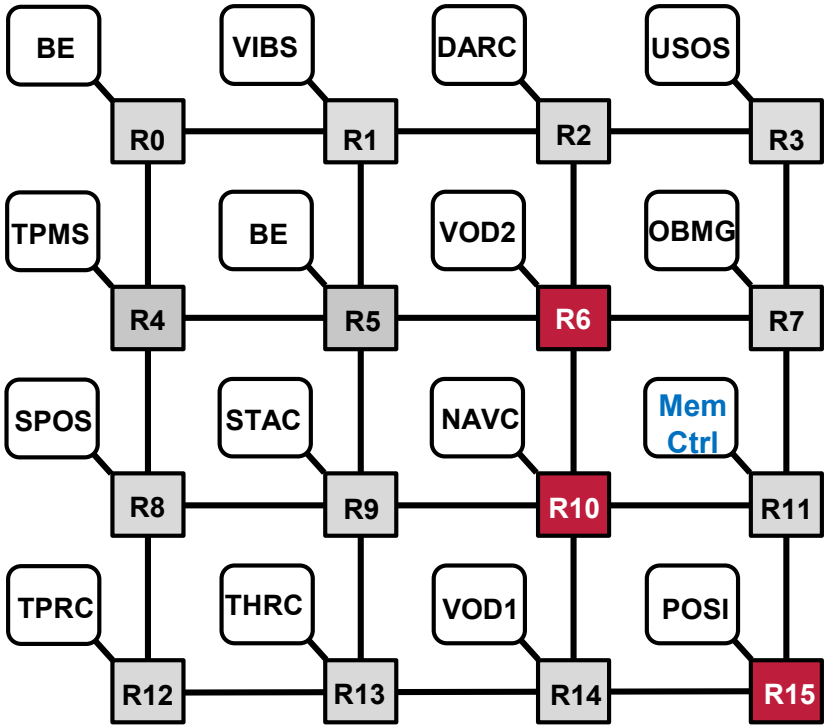
  footprints for the NoC power models were obtained with IDAMC platform [Tobuschat2012]
  - VHDL implementation at the gate level for routers
  - utilizing standard ASIC flow for 65nm process technology

- **AER task model is employed**
  - Acquisition, Execution, and Restitution (AER) task-model decouples task's execution phase from the communication one [Durrieu2014]
  - AER model increases the power-savings using PANC as a task needs to synchronize its NoC access once for the entire transmission

Use-case - Driver Assistance System



a) Use-case Graph



b) Use-case Mapping

- 17 communicating tasks extracted from [Shi2012]
  - obstacle detection, position sensors and ultrasonic sensors
- Tasks periods vary between 0.4ms and 1ms
- Tasks communication volume vary between 9kB and 0.3MB

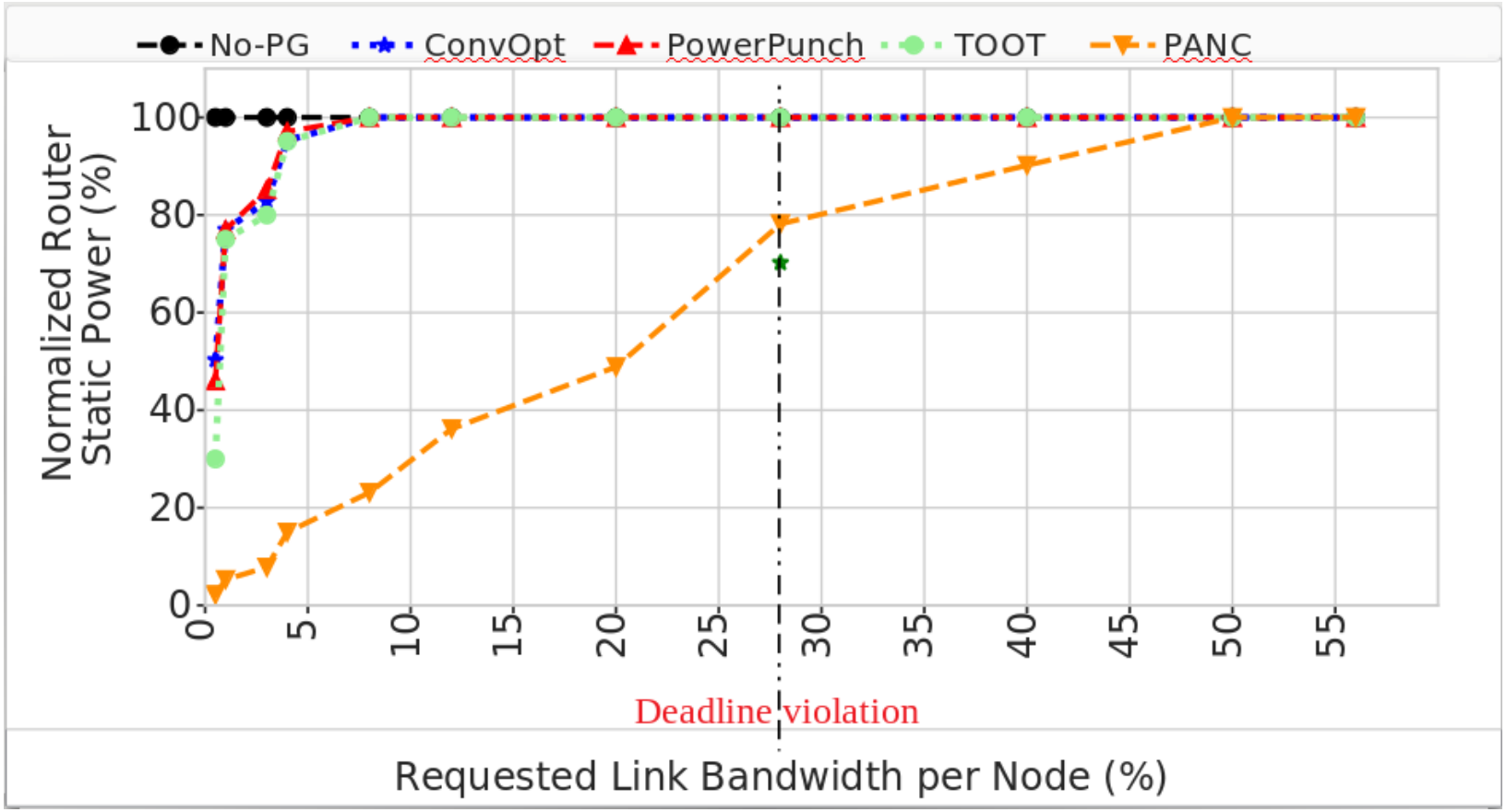
Experimental results

On average the network power savings using our approach is 82.7% compared with No-PG

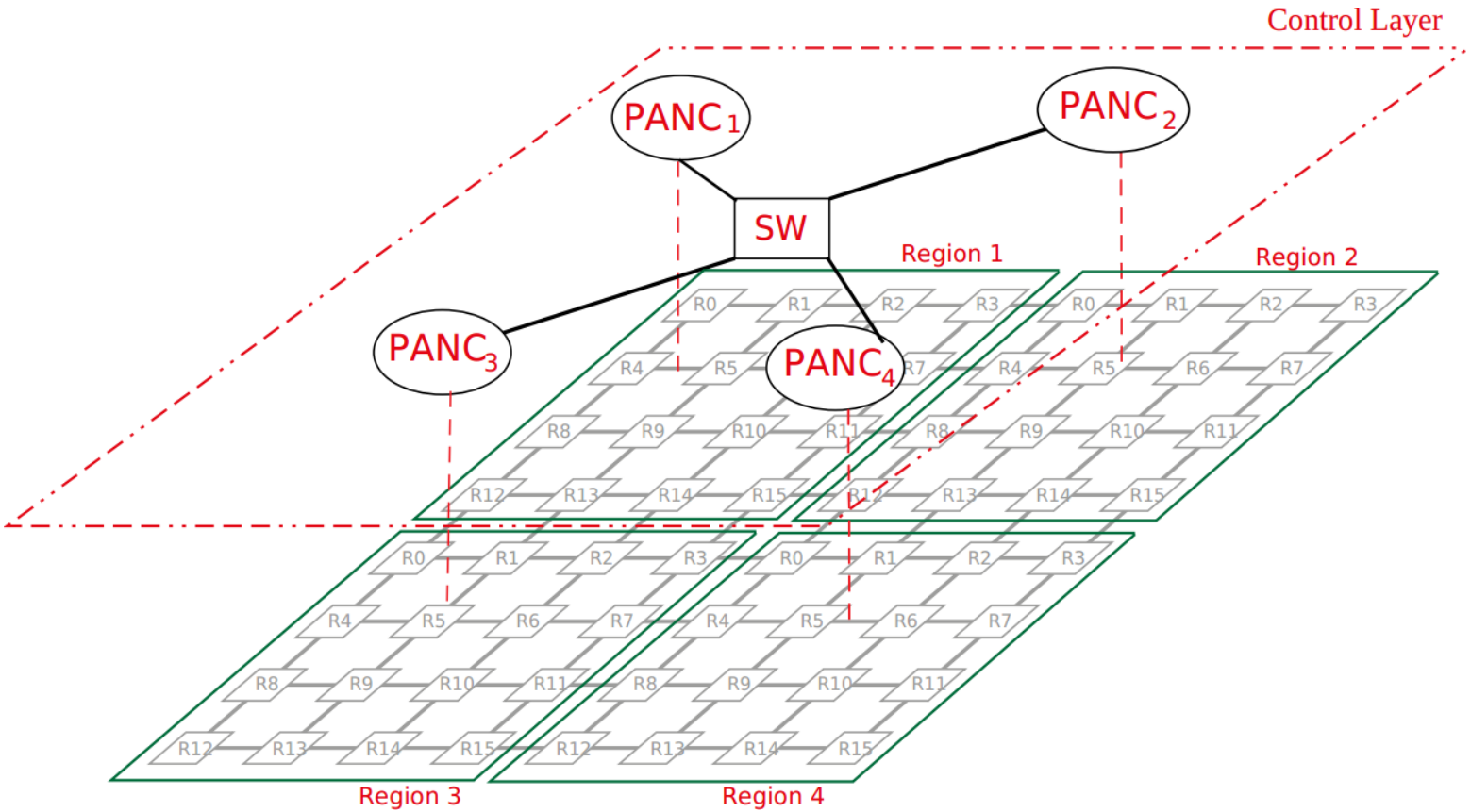
(y,x)	0	1	2	3
0	90.3%	91.08%	99.37%	78.67%
1	95.90%	92.59%	68.39%	77.85%
2	79.9%	79.08%	62.29%	70.5%
3	99.17	97.94%	73.30	66.82%

Leakage power-savings of individual routers in the 4×4 NoC, employing the vehicle use-case

Experimental results – random traffic



PANC evaluation under scalability (employing FADEC application)



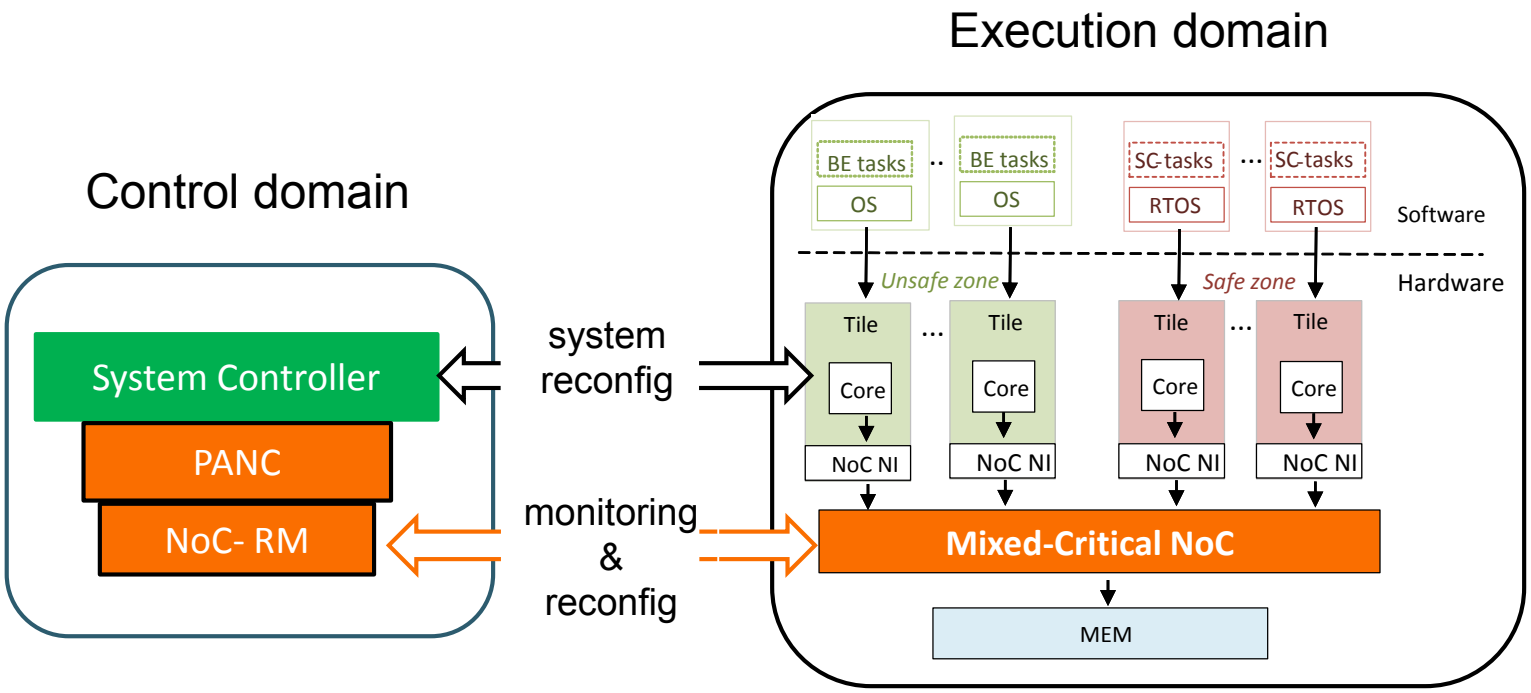
- the experimental results (8x8) indicate that PANCs save up to 79:13% of the static power compared with No-PG NoC after accounting for the power overhead of the control layer.

Outline

- Motivation
- Self awareness at different system layers
- Self-aware NoC through energy management
  - Safe and dynamic Power-gating of real-time NoC routers
- Self-aware SoC through error handling
- Conclusion

New idea
New challenge
Future work

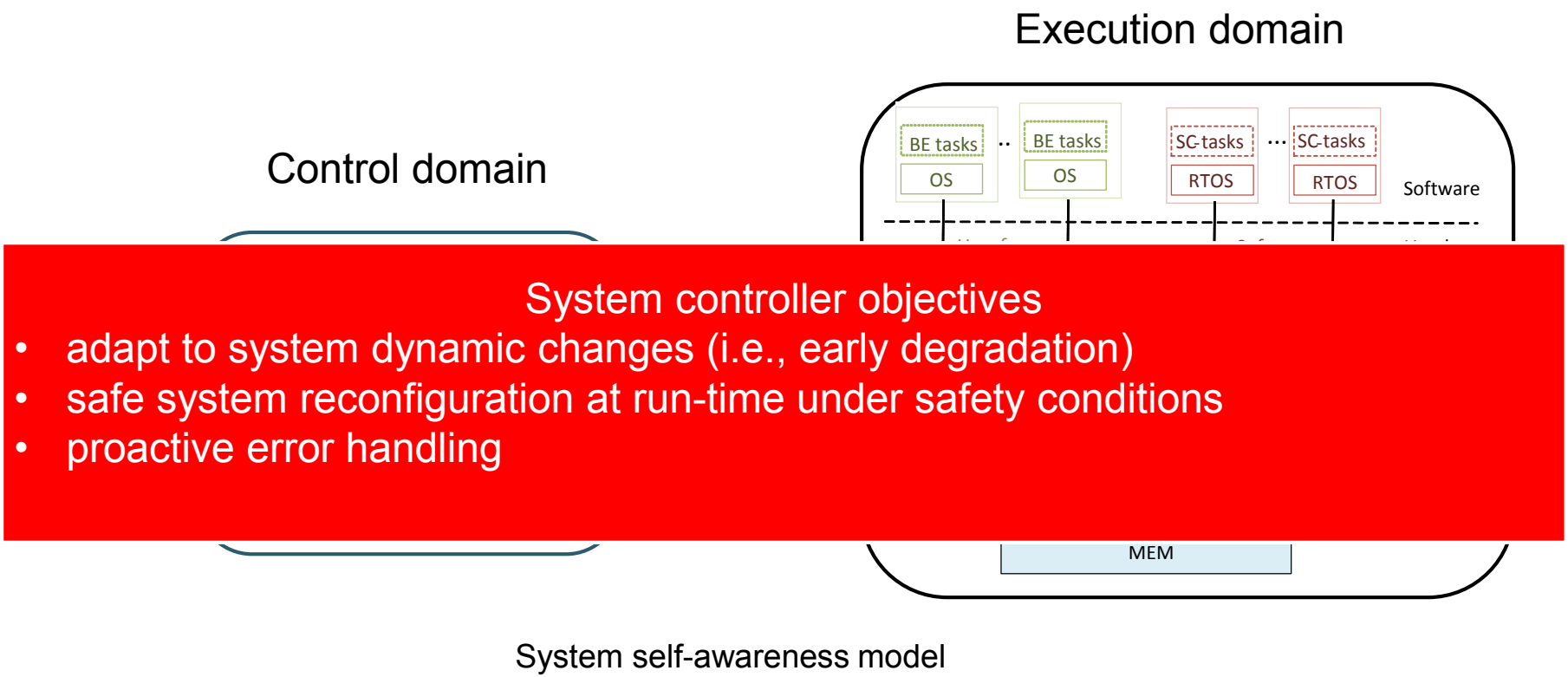
Self awareness at system-level via the system controller



System self-awareness model

New idea
New challenge
Future work

Self awareness at system-level via the system controller



System controller objectives

- adapt to system dynamic changes (i.e., early degradation)
- safe system reconfiguration at run-time under safety conditions
- proactive error handling

System self-awareness model

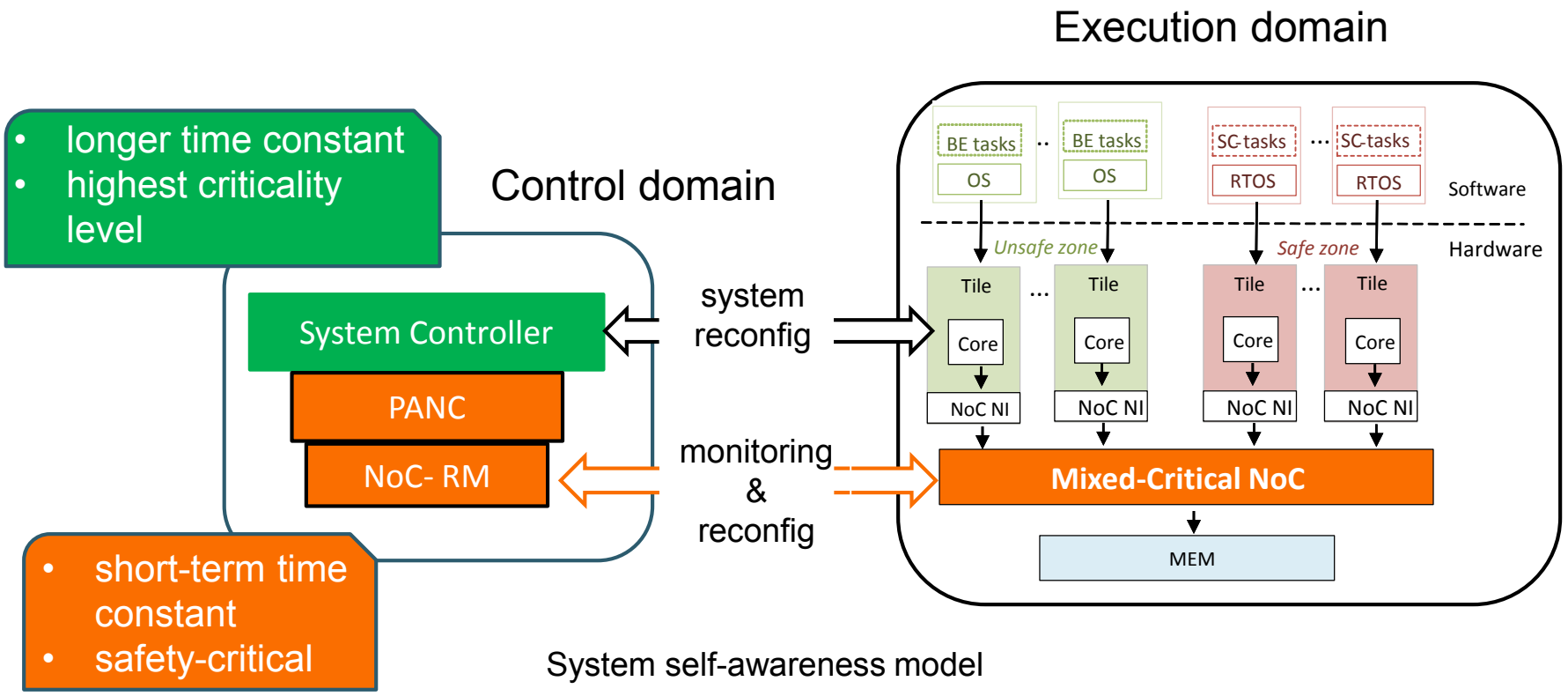
Proactive vs reactive error handling

- reactive solutions (e.g., modular redundancy: DMR/TMR): the primary limitation is handling an error once it occurs, jeopardizing the system safety during the recovery process

- proactive solutions advantages:
  - increased system safety
    - decreases the probability of failure occurrences
  - increased lifecycle of dependable system operation
    - system requires fewer reactions to immediate errors
  - increased flexibility
    - more time available to proactively handle an impending permanent error

New idea
New challenge
Future work

Self awareness at system-level via the system controller



System self-awareness model

Conclusion

- the current/future complexity of embedded system requires new methods to provide system lifecycle management
- new method for safe power management in NoCs for real-time systems has been introduced employing PANC
- extending the concept to cover error-handling at system-level employing the system controller
- safe proactive error handling that adheres to safety conditions, results in system safety, reliability, and availability increase
- safety standards require both NoC power controller and the system controller to be at highest level of criticality

Thanks for your attention
Happy with your questions!

References

- **[arxiv2019]** Eberle A. Rambo, Bryan Donyanavard, Minjun Seo, Florian Maurer, Thawra Kadeed, Caio B. de Melo, Biswadip Maity, Anmol Surhonne, Andreas Herkersdorf, Fadi Kurdahi, Nikil Dutt, and Rolf Ernst. 2019. The Information Processing Factory:Organization, Terminology, and Definitions. arXiv:1907.01578

- **[kostrzewa2017a]** Adam Kostrzewa, Sebastian Tobuschat, Leonardo Ecco, and Rolf Ernst, "Adaptive load distribution in mixed-critical Networks-on-Chip" in 22nd Asia and South Pacific Design Automation Conference (ASP-DAC), 2017

- **[kostrzewa2017b]** Adam Kostrzewa, Sebastian Tobuschat, and Rolf Ernst, "Self-Aware Network-On-Chip Control in Real-Time Systems", IEEE Design & Test, 2017

- **[Parikh2014]** R. Parikh, R. Das, and V. Bertacco, “Power-aware nocs through routing and topology reconfiguration,” in esign Automation Conference (DAC), 2014 51st ACM/EDAC/IEEE, 2014, pp. 1–6

- **[Chen2015]** L. Chen, D. Zhu, M. Pedram, and T. M. Pinkston, “Power punch: Towards non-blocking power-gating of noc routers,” in High Performance Computer Architecture (HPCA), 2015 IEEE 21st International Symposium on, 2015, pp. 378–389`

- **[Farrokhbakht2016]** H. Farrokhbakht, M. Taram, B. Khaleghi, and S. Hessabi, “Toot: an efficient and scalable power-gating method for noc routers,” in Networks-on-Chip (NOCS), 2016 Tenth IEEE/ACM International Symposium on, 2016, pp. 1–8

References

- **[Kadeed2018]** Thawra Kadeed, Sebastian Tobuschat, Adam Kostrzewa und Rolf Ernst, "Safe and efficient power management of hard real-time networks-on-chip", Integration, the VLSI Journal, 2018

- **[Kostrzewa2018]** Adam Kostrzewa, Thawra Kadeed, Borislav Nikolic und Rolf Ernst, "Supporting Dynamic Voltage and Frequency Scaling in Networks-On-Chip for Hard Real-Time Systems" RTCSA, 2018

- **[Nikolic2019]** Borislav Nikolic, Sebastian Tobuschat, Leandro Soares Indrusiak, Rolf Ernst, Alan Burns: Real-time analysis of priority-preemptive NoCs with arbitrary buffer sizes and router delays. Real-Time Systems 55(1): 63-105 (2019)
- **[Durrieu2014]** Durrieu, G., Faugere, M., Girbal, S., P´erez, D. G., Pagetti, C., Puffitsch, W., 2014. Predictable flight management system implementation on a multicore rocessor. In: Embedded Real Time Software (ERTS'14)
- **[Shi2012]** Shi, Z., Burns, A., Indrusiak, L. S., 2012. Schedulability analysis for real time on-chip communication with wormhole switching. Innovations in embedded and real-time systems engineering for communication, 198
- **[Tobuschat2013]:** Tobuschat, S., Axer, P., Ernst, R., Diemer, J., 2013. Idamc: A noc for mixed criticality systems. In: Embedded and Real-Time Computing Systems and Applications (RTCSA),19th International Conference on. IEEE, pp. 149{156.