DESIGN, AUTOMATION & TEST IN EUROPE
01 – 05 February 2021 · virtual conference
The European Event for Electronic System Design & Test

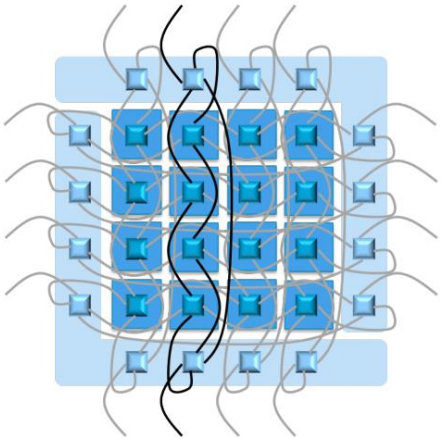
NoC Performance Model for Efficient Network Latency Estimation

Oumaima Matoussi

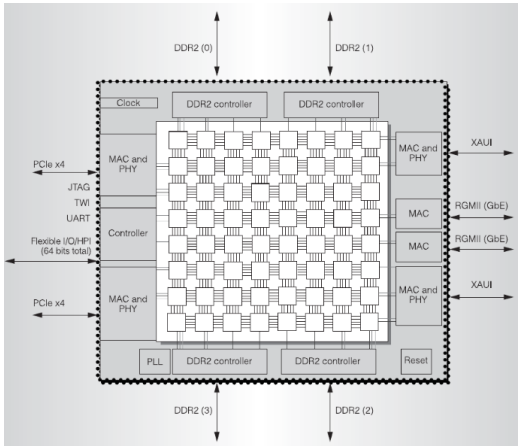
Université Paris-Saclay, CEA List, France

Introduction

- **NoCs are prevalent in many-core architectures**
  - **NoCs contribute to system performance and cost**
  - **Efficient and reliable NoC models are needed**



Kalray MPPA with 2D Torus [1]



Tilera Tile64 with 2D mesh [2]

Objective

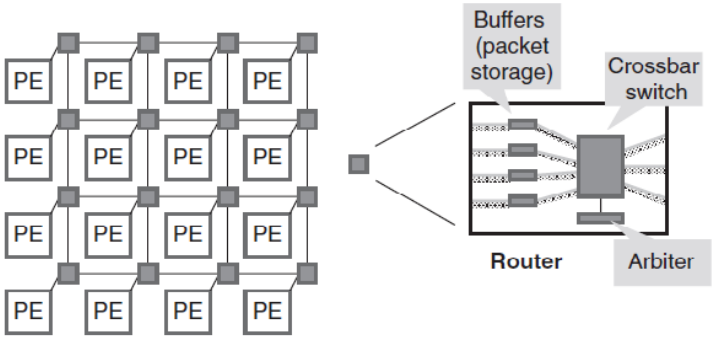
- **Fast and reliable NoC model for early performance estimation**
  - **Parametric** → **Design Space Exploration (DSE)**
  - **Tractable** → **Full System Simulation (FSS)**
  - **Scalable** → **Large scale MPSoC simulation**
  - **Realistic** → **Resource contention modeling**

Outline

- Introduction
- **NoC Background**
- Proposed NoC Model
- NoC Model Evaluation
- Conclusion

NoC Fundamentals

- **A NoC is mainly characterized by:**
  - **Topology**
  - **Routing algorithm**
  - **Flow control**



A mesh topology [3]

- **NoC performance is measured by packet latency**

$$Lat_{pkt} = Plat_{pkt} \times nbr\ hops + FT \times (L - 1) + W_{pkt}$$

NoC Modeling Approaches

Analytical approaches	Simulation approaches
Netwok Calculus [4] Queuing Theory [5] Real-time Analysis (WCTT) [6]	Garnet [7] BookSim [8] Noxim [9]
(+) Fast design space exploration (-) Not suitable for non deterministic traffic	(+) Flit-level granularity (-) Not suitable for large scale simulations

→ Combine the advantages of analytical and simulation approaches in a hybrid NoC Model

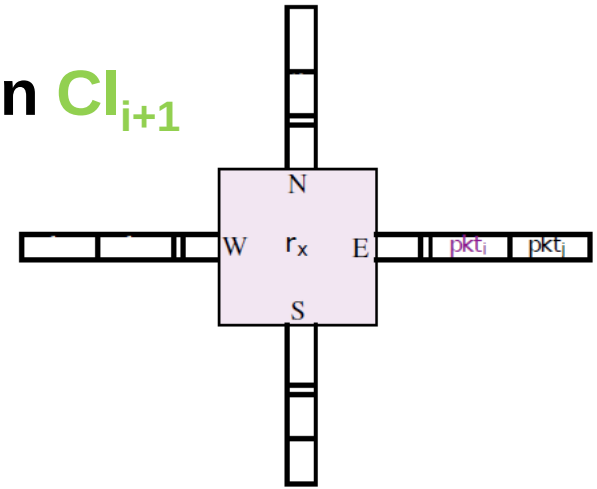
Outline

- Introduction
- NoC Background
- **Proposed NoC Model**
- NoC Model Evaluation
- Conclusion

Router Model and Packet Tracing

- **Information about pkt_k is collected in contention interval Ci_i**
  - **Route computation**
  - **Buffer update**
- **Network latency of pkt_k is computed in CI_{i+1}**

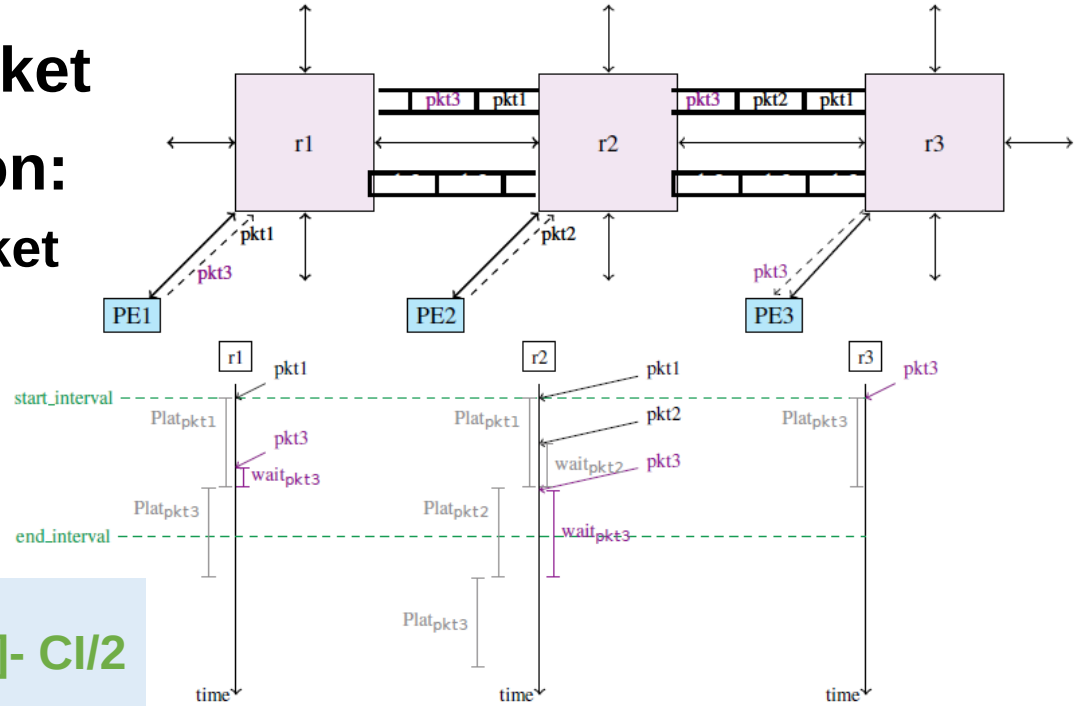
Interval_start=pkt_k.timestamp
Interval_end=interval_start + CI } CI bounds



A router model

Network Latency Estimation (1/2)

- **Queuing delay of a packet in a buffer B depends on:**
  - **Latency of previous packet**
  - **Duration of CI**
  - **Number of packets in B**



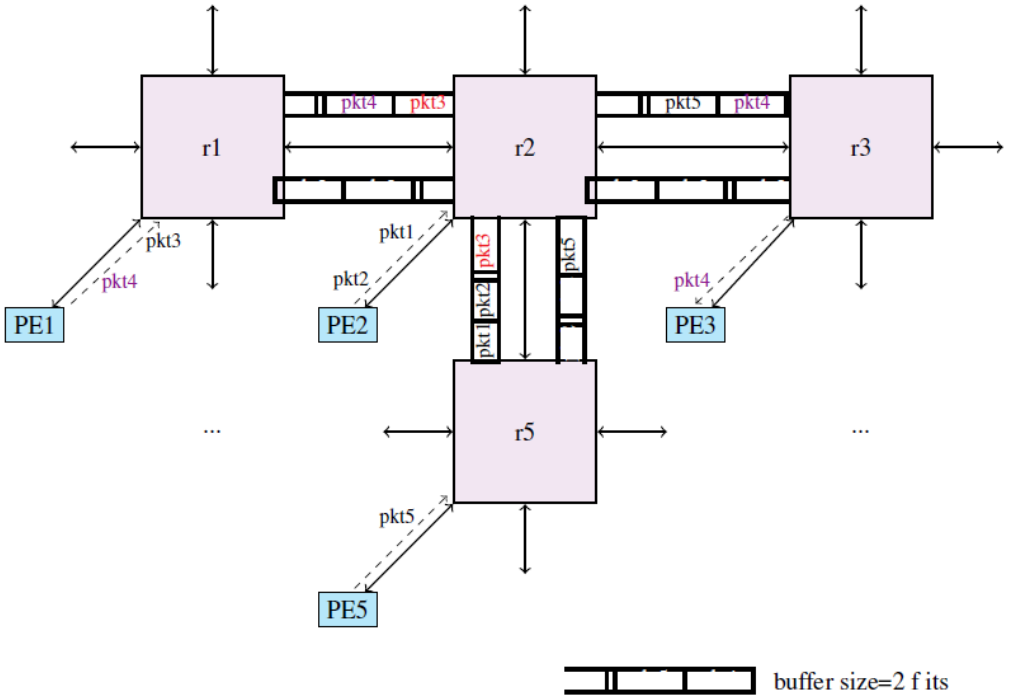
Link contention

$$Wait_{pkt3}[r1->r2] = 0 + Plat_{pkt1}[r1->r2] - CI/2$$

Network Latency Estimation (2/2)

- **Congestion delay is caused by:**
  - **Blocked HOL packet**
  - **Full destination Buffer**

$$HOL_{pkt4}[r2->r3] = wait_{pkt2}[r2->r5]$$



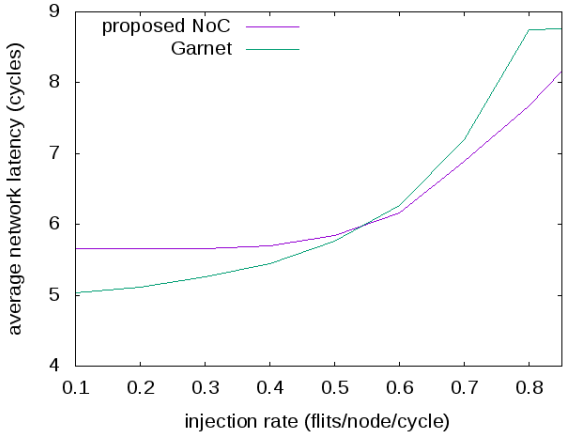
HOL blocking

Outline

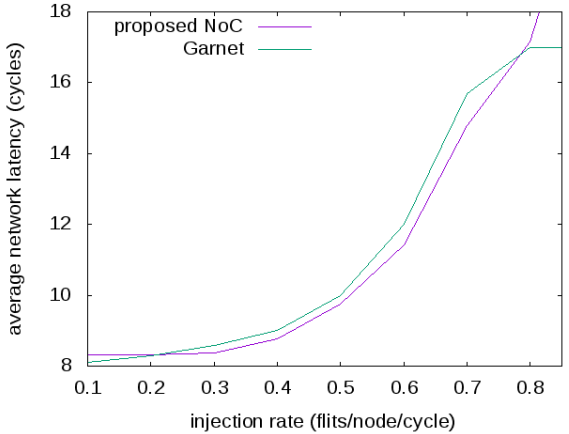
- Introduction
- NoC Background
- Proposed NoC Model
- **NoC Model Evaluation**
- Conclusion

Standalone Mode (1/2)

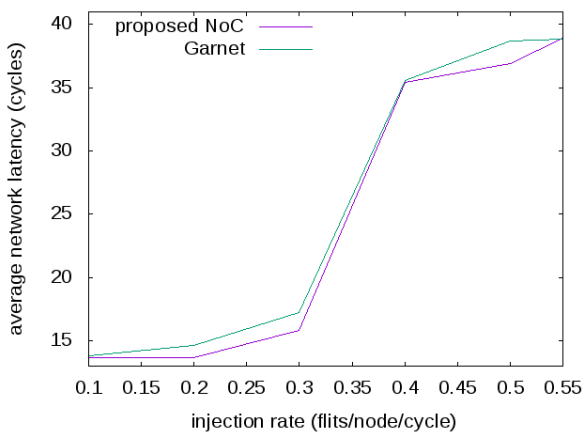
- **Impact of mesh size on average network latency under uniform-random traffic**



2*2 mesh



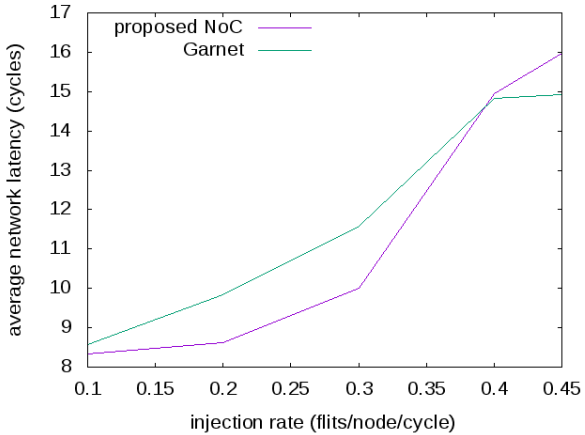
4*4 mesh



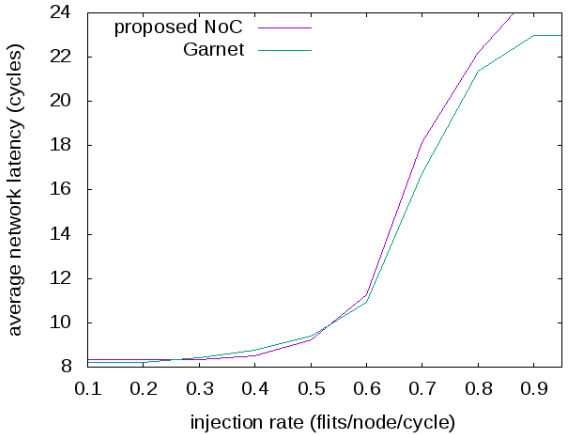
8*8 mesh

Standalone Mode (2/2)

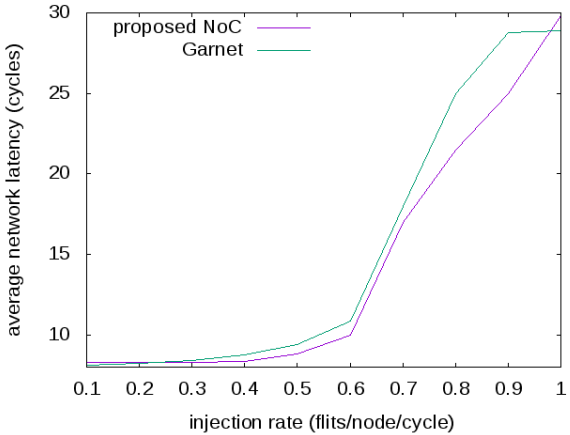
- **Impact of VC variation on average network latency of 4*4 mesh**



1 VC



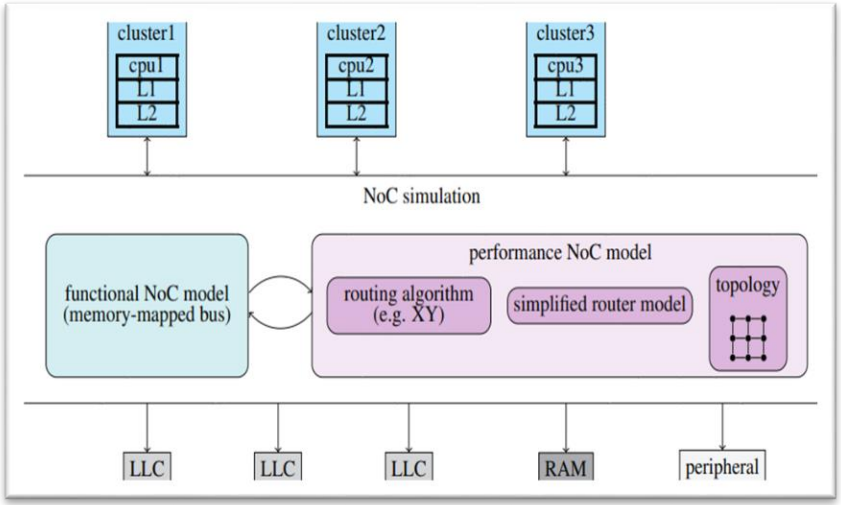
3 VCs



4 VCs

Full System Simulation Mode

- **Integration of NoC model in VPSim** [10]



NoC model in an FSS environment

	swaptions	radiosity	barnes
slowdown	1,6	2,5	1,5

Slowdown of VPSim in MIPS

Conclusion

- **A hybrid NoC model:**
  - **Abstract router model**
  - **Analytical formulae for latency computation**
    - **link contention**
    - **buffer congestion**
- **Suitable for Full System Simulation**

14⨯ Speedup
<= 17% error

References

1. B. D. de Dinechin and A. Graillat, “Network-on-chip service guarantees on the kalray mppa-256 bostan processor”, AISTECS .
2. Tile-gx36 processor, Available: https://www.mellanox.com/related-docs/prod multi core/PB TILEGx36.pdf, [Online].
3. S.Pasricha, N.Dutt, “On-Chip Communication Architectures – System On Chip Interconnect“.
4. F. Giroudot and A. Mifdaoui, “Tightness and computation assessment of worst-case delay bounds in wormhole networks-on-chip”, RNS.
5. N. Nikitin and J. Cortadella, “A performance analytical model for network-on-chip with constant service time routers”, ICCAD.
6. L. Abdallah, M. Jan, “Wormhole networks properties and their use for optimizing worst case delay analysis of many-cores”, SIES.
7. N. Agarwal, T. Krishna, “Garnet: A detailed on-chip network model inside a full-system simulator”.
8. Nan Jiang, D. U. Becker, “A detailed and flexible cycleaccurate network-on-chip simulator“, ISPASS.
9. V. Catania, A. Mineo, “Noxim: An open, extensible and cycle-accurate network on chip simulator”, ASAP.
10. A. Charif, G.Busnot, R. Mameesh, T. Sassolas, “Fast virtual prototyping for embedded computing systems design and exploration”, RAPIDO.

Acknowledgements

This work is partially funded by **H2020 European Processor Initiative** (grant agreement No 826647).