

Providing Throughput Guarantees in Mixed-criticality Networks-on-Chip

IEEE SOCC 2017

Sebastian Tobuschat, Rolf Ernst

IDA, TU Braunschweig, Germany

08. September 2017, Munich, Germany

Networks-on-Chip (NoC)



- transmissions of different safety level share NoC resources
- standards require isolation
 - e.g. IEC 61508: "sufficient independence"
- consequences

- highest relevant safety level for all \rightarrow expensive
- or implement "sufficient independence"
 - \rightarrow Quality of Service mechanisms (QoS)
- main Challenge: QoS guarantees + high performance







Providing Quality of Service – Related Work

- static partitioning (e.g. TDMA):
 e.g. [Milberg2004], [Goossens2010], [Psarras2015]
 - typically reduced utilization
- prioritization:
 - e.g. [Bolotin2004], [Bjerregaard2005]
 - "criticality as priority": reduced performance for BE
- dual priority:
 - e.g. [Burns2014], [Indrusiak2015]
 - not accounting for NoC load; only for whole path
 - reduced exploitation of latency slack
- performance monitoring + dynamic prioritization, e.g.:
 - fluid meter [Tsai1015]: increased header and increased router
 - backsuction [Diemer2010]: no VC sharing, initial setup



Providing Quality of Service – Related Work

- static partitioning (e.g. TDMA): e.g. [Milberg2004], [Goossens2010], [Psarras2015]
 - typically reduced utilization
- prioritization: e.g. [Bolotin2004
 - "criticality as
- dual priority: e.g. [Burns2014]
 - not accounti
 - reduced exploration on latency slack
- performance monitoring + dynamic prioritization, e.g.:
 - fluid meter [Tsai1015]: increased header and increased router
 - backsuction [Diemer2010]: no VC sharing, initial setup



Goal: minimize negative performance impact of QoS mechanisms (on non-critical senders)

Idea: prioritize BE to exploit (latency) slack of critical applications



Outline

- Motivation
- Providing Quality of Service
- Throughput Guarantees
- Experimental Results
- Conclusion







- slack: difference between worst-case transmission time and deadline
- safety critical applications do not benefit from finishing before deadline or receiving more throughput than required
- but best effort applications benefit from low latency
- baseline approach:
 - two traffic classes: guaranteed throughput (GT) and best effort (BE)
 - prioritize BE over GT and limit interference BE induces to GT to exploit slack of GT



utility function of a firm/hard real-time task [Stankovic1998]



Limit BE Interference (1)



- GT injects with desired rate
- monitor progress of GT (locally in each buffer/router)
 - if too slow \rightarrow increase its priority (over BE)





Limit BE Interference (2)



- monitor fill level via threshold
 - if threshold reached, GT obtains priority over BE





Limit BE Interference (3)



- tag last flit of transmission
 - to obtain priority for end of transmission
 - needed as there might be no subsequent packets that fill up the buffer





Limit BE Interference (4)



- further extension
 - account for backlog at sender
 - if it was too slow in past, send packets with higher rate until average throughput is OK again
 - e.g. via size of sender buffer, bucket/burst size ...

- system design ensures that interference between GT senders does not violate requirements
 - otherwise system is not feasible
 - when GT has priority, it can only interfere with GT



Outline

- Motivation
- Providing Quality of Service
- Throughput Guarantees
- Experimental Results
- Conclusion



Worst-case Service



- based on [Tobuschat2017] compositional performance analysis
- local router analysis
 - worst-case waiting time at a port
 - maximum time until the port is ready to accept the q-th incoming flit
 - used to derive minimum accepted throughput $\hat{\beta}_p^-$ at each router port
 - break down into sum of different terms addressing different blocking factors
- for each stream
 - analyze routers along its path and propagate event models downstream
- formally analyze routers iteratively





$$\widehat{B}_{p}^{+}(q) \leq q * C + \max_{\theta \in \Theta^{k}} \left\{ \sum_{j \in \Theta} \left\{ B_{j}^{out} \left(\widehat{B}_{p}^{+}(q) - C, n \right) + \widehat{B}_{P(j),k+1}^{+}(q) \right\} \right\}$$



q : number of flits

 $\mathbf{\Theta}^k$: pos. mappings of k packets to output ports

- k : num. of packets q flits form
- n : number of flits in packet
- C: single flit transmission time



08. September 2017 | S. Tobuschat | IEEE SOCC 2017 | Providing Throughput Guarantees in Mixed-criticality Networks-on-Chip | Slide 13





For details and equations look into the paper

q : number of flits

 Θ^k : pos. mappings of k packets to output ports

- k : num. of packets q flits form
- n : number of flits in packet
- C: single flit transmission time



Technische Universität Braunschweig





n : number of flits in packet

C: single flit transmission time



Technische Universität Braunschweig

08. September 2017 | S. Tobuschat | IEEE SOCC 2017 | Providing Throughput Guarantees in Mixed-criticality Networks-on-Chip | Slide 15





- n : number of flits in packet
- C: single flit transmission time



Technische Universität Braunschweig

08. September 2017 | S. Tobuschat | IEEE SOCC 2017 | Providing Throughput Guarantees in Mixed-criticality Networks-on-Chip | Slide 16



$$B_i^{out}(\Delta t, q) \leq \Psi + \sum_{j \in Out_i} \left\{ C * \chi + \widehat{B}_{P(j),k+1}^+(\chi) \right\}$$



q : number of flits

 Θ^k : pos. mappings of k packets to output ports

- k : num. of packets q flits form
- n : number of flits in packet
- C: single flit transmission time







- n : number of flits in packet
- C: single flit transmission time



Technische Universität Braunschweig





- n : number of flits in packet
- C: single flit transmission time



Technische Universität Braunschweig

08. September 2017 | S. Tobuschat | IEEE SOCC 2017 | Providing Throughput Guarantees in Mixed-criticality Networks-on-Chip | Slide 19





n : number of flits in packet

C: single flit transmission time



Technische Universität Braunschweig

08. September 2017 | S. Tobuschat | IEEE SOCC 2017 | Providing Throughput Guarantees in Mixed-criticality Networks-on-Chip | Slide 20



$$\widehat{\boldsymbol{\beta}}_{p}^{-}(\Delta t) \geq \min\left\{\Delta t, \max\left\{\boldsymbol{m} \middle| \widehat{\boldsymbol{B}}_{p}^{+}\left(\max(\boldsymbol{0}, \boldsymbol{m} - \boldsymbol{Q}_{b})\right) < \Delta t\right\}\right\}$$

with $\Delta t \in \mathbb{N}^+$ and in multiple of flit transfer time

























Outline

- Motivation
- Providing Quality of Service
- Throughput Guarantees
- Experimental Results
- Conclusion



Evaluation (1)



- OMNeT++ framework + HNOCs library
- packet size: 4 flits
- buffer size: 4 packets (16 flits)
- router with 4 stage pipeline
- 3-4 VCs
- 8x8 mesh, XY-routing
- two sets of experiments:
 - synthetic workload: general properties
 - benchmark based: performance improvement



Evaluation (2)

ID

- two GT streams
 - node (0,1) to (6,4) requesting 30%
 - node (0,2) to (5,4) requesting 20%
- compared different approaches
 - RR2/RR3: round robin
 3VCs/4VCs (2-3 for BE, 1 for GT)
 - SP2/SP3: static priority 3VCs/4VCs (2-3 for BE, 1 for GT)
 - BS: backsuction
 4VCs (2 for BE, 2 for GT)
 - FP2/FP3: proposed 3VCs/4VCs (2-3 for BE, 1 for GT) threshold: 3 packets (75% buffer size)





Experiment 1



- synthetic workload: general properties
 - based on average link loads
- GT periodically sending burst of 4 packets
- increasing BE load
- check performance of BE
 - for highlighted BE node (0,3) to (4,4)
- random destinations for all other nodes (BE traffic)











Experiment 2

- benchmark based
 - traces from CHStone
 - extracted using Gem5: ARMv7, 32kB L1
 - accesses to network (e.g. memory access, communication, cache access)
 - for each trace on (0,3)
 - multiple (random) mappings of traces to other nodes
 - with random destinations for traffic
 - latency for highlighted BE node (0,3 to 4,4)
 - normalized to SP2 (over all runs)







Normalized BE Latency







Synthesis Results – NoC



- 2x2 NoC, 3-4 VCs, buffer size of 4 packets per VC
- Virtex-6 LX760 FPGA, Xilinx ISE 14.6, standard settings
 - RR3: round robin, 4 VCs (3 for BE, 1 for GT), not safe
 - SP3: static priority, 4 VCs (3 for BE, 1 for GT)
 - DP: flag to change priority of GT, 4 VCs (3 for BE, 1 for GT)
 - BS: backsuction, 4VCs (2 for BE, 2 for GT)
 - FP2/FP3: proposed, 3VCs/4Ccs (2-3 for BE, 1 for GT)

				+11.6%			
Unit	RR3	SP3	DP	BS	FP3	FP2	
#Registers	4368	4856	4868	4873	4875	3874	
#LUTs	5720	6301	6322	6338	6346	5132	
MHz	210	210	210	210	210	210	
			-	+0.7%		-19.0%	



Outline

- Motivation
- Providing Quality of Service
- Throughput Guarantees
- Experimental Results
- Conclusion



Conclusion



- run-time configurable, dynamic prioritization of GT to exploit slack of safety-critical applications
 - based on actual blocking through BE
 - prioritize BE over GT when possible
- → increased performance for BE
 - up to 16% lower average latency compared to static prioritization
 - less than 1% hardware overhead compared to static prioritization
- reduced hardware overhead compared to backsuction
 - while achieving similar performance benefits for BE
- future work:
 - dependency of/between buffer, threshold and packet sizes

Thank you for your attention. Questions?



References



- [Bjerregaard2005]T. Bjerregaard and J. Sparsoe, "Scheduling discipline for latency and bandwidth guarantees in asynchronous network-on-chip," in Asynchronous Circuits and Systems, 2005. ASYNC 2005. Proceedings. 11th IEEE International Symposium on, pp. 34–43, March 2005.
- [Bolotin2004] E. Bolotin, I. Cidon, R. Ginosar, and A. Kolodny, "QNoC: QoS architecture and design process for network on chip," J. Syst. Archit., vol. 50, pp. 105–128, Feb. 2004.
- [Burns2014] A. Burns, J. Harbin, and L. Indrusiak, "A wormhole NoC protocol for mixed criticality systems," in Real-Time Systems Symposium (RTSS), 2014 IEEE, pp. 184–195, Dec. 2014.
- [Diemer2010] J. Diemer and R. Ernst, "Back suction: Service guarantees for latency-sensitive on-chip networks," in Networks-on-Chip (NOCS), 2010 Fourth ACM/IEEE International Symposium on, 2010, pp. 155–162.
- [Goossens2010] K. Goossens and A. Hansson, "The aethereal network on chip after ten years: Goals, evolution, lessons, and future," in Proceedings of the 47th Design Automation Conference, DAC '10, (New York, NY, USA), pp. 306–311, ACM, 2010.
- [Indrusiak2015] L. Indrusiak, J. Harbin, and A. Burns, "Average and worst-case latency improvements in mixed-criticality wormhole networks-on-chip," in Real-Time Systems (ECRTS), 2015 27th Euromicro Conference on, pp. 47–56, July 2015.
- [Milberg2004] M. Millberg, E. Nilsson, R. Thid, and A. Jantsch, "Guaranteed bandwidth using looped containers in temporally disjoint networks within the nostrum network on chip," in Design, Automation and Test in Europe Conference and Exhibition, 2004. Proceedings, vol. 2, pp. 890–895 Vol.2, Feb 2004.
- [Psarras2015] A. Psarras, I. Seitanidis, C. Nicopoulos, and G. Dimitrakopoulos, "Phasenoc: Tdm scheduling at the virtual-channel level for efficient network traffic isolation," in Proceedings of the 2015 Design, Automation & Test in Europe Conference & Exhibition, DATE '15, (San Jose, CA, USA), pp. 1090–1095, EDA Consortium, 2015.
- [Rambo2015] E. A. Rambo and R. Ernst, "Worst-case communication time analysis of networks-on-chip with shared virtual channels," in Proceedings of the 2015 Design, Automation & Test in Europe Conference & Exhibition, DATE '15, (San Jose, CA, USA), pp. 537–542, EDA Consortium, 2015.
- [Tsai2015] W. C. Tsai, H. E. Lin, Y. C. Lan, S. J. Chen, and Y. H. Hu, "A novel flow fluidity meter for BiNoC bandwidth resource allocation," in 2015 28th IEEE International System-on-Chip Conference (SOCC), Sept 2015,pp. 281–286.





Backup Slides



IEC 61508-2: 2010



"7.4.2.3 Where an E/E/PE safety-related system is to implement both safety and non-safety functions, then all the hardware and software shall be treated as safety-related unless it can be shown that the implementation of the safety and non-safety functions is **sufficiently independent** (i.e. that the failure of any non-safety-related functions does not cause a dangerous failure of the safety-related functions).

NOTE 1: Sufficient independence of implementation is established by showing that the **probability of a dependent failure between the non-safety and safetyrelated parts is sufficiently low** in comparison with the highest safety integrity level associated with the safety functions involved.

NOTE 2: Caution should be exercised if non-safety functions and safety functions are implemented in the same E/E/PE safety-related system. While this is allowed in the standard, it may lead to greater complexity and increase the difficulty in carrying out E/E/PE system safety lifecycle activities (for example design, validation, functional safety assessment and maintenance)."



Experiment 1



- synthetic workload: general properties
 - based on average link loads
- check for achieved throughput for GT
 - GT periodically sending burst of 4 packets
 - increasing BE load
- check performance of BE
 - for highlighted BE node (0,3) to (4,4)
- random destinations for all other nodes





Observed Achieved Throughput







Compositional Performance Analysis for NoCs

- based on analysis from [Rambo2015]
- analysis performed iteratively
- step 1: local analysis (at each router)
 - compute worst-case latency R⁺_i of flits based on critical instant (busy window)
 - derive output event models
- step 2: global analysis
 - propagate event models downstream
 - go to step 1 if any event model has changed
 - otherwise, terminate







CPA Approach



- worst-case end-to-end latency relies on response times R⁺ from local analyses
- for each stream
 - analyze routers along its path and propagate event models downstream
- formally analyze routers iteratively





Mapping NoC Domain to Processor Resource Model

IDA

- output ports → processing resources
- input ports → shared resources with mutually exclusive access
- traffic stream → chain of tasks mapped to resources
- flit transmission → task execution
- flit arrival → task activation
 - input and output event models





Complex Activation Patterns



- variety of activation patterns used in practice
 e.g. periodic + spontaneous, dual cyclic, on change
- timing verification can consider them through use of minimum distance functions
 - i.e. specification of the minimum distance between any n consecutive events
 - derived from specification or ratelimiter



