



# A Fine-Grained Link-Level Fault-Tolerant Mechanism For NoCs

Arseniy Vitkovskiy<sup>†</sup>  
Vassos Soteriou<sup>†</sup>  
Chrysostomos Nicopoulos<sup>‡</sup>

<sup>†</sup>Cyprus University of Technology  
<sup>‡</sup>University of Cyprus



## Agenda

- Introduction
- Motivation and related work
- Partially faulty link recovery mechanism
- Probabilistic link fault model
- Hardware implementation
- Routing algorithms
- Experimental setup and simulation results
- Conclusions



## The need for Link-Level Fault-Tolerance (FT) in NoCs

- Modern CMPs (e.g. Intel Teraflops, Tiler TILE64) rely on Network-on-Chips (NoC) for inter-tile communication
- Parallel links in NoCs; several metal layers, high metal link density
- Technology scaling leads to higher wear-out of HW components
  - *electro-migration (EM)*, *stress induced voiding (SIV)*, *negative bias temperature instability*, *oxide breakdown*
  - up to 10% of on-chip components will be defective [1]
  - up to 20% of on-chip faults will be permanent [2]
  - **EM** identified as the main problem for links: current density is high leading to thermal stressing, metal deformation behavior, mass transport
- Fault-tolerant design becomes mandatory
- NoCs present excellent opportunities for FT design (modular, path diversity, regularity)

[1] E. Karl et al. *Reliability Modeling and Management in Dynamic Microprocessor Systems*. In DAC, pp. 1057–1060, July 2006.

[2] T. Lehtonen et al. *Self-Adaptive System for Addressing Permanent Errors in On-Chip Interconnects*. In IEEE TVLSI, Vol. 18, No. 4, pp. 527–540, April 2010.

20.10.2010

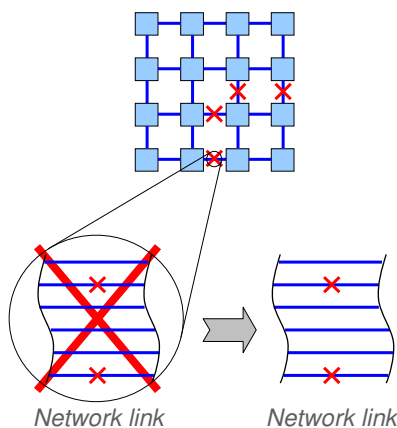
HiPEAC, TF Reliability

3



## Partially faulty links: Motivation

Network on chip topology  
(connectivity graph)



Faulty link(s) lead to:

- network disconnection, segmentation
- communication deadlocks
- possible livelocks

Faulty wires:

- entire link marked as faulty
- a faulty link can still operate as a partially faulty link (PFL) is used
  - maintain connectivity
  - deadlock and livelock avoidance
  - Intuition: would take longer to bypass a fault in some cases

20.10.2010

HiPEAC, TF Reliability

4



## Related Work

- Error Correction Codes (ECC):
  - limited correction capability
  - non-trivial algorithms and associated hardware
  - large HW overhead (timing and area)
- Fault-tolerant routing algorithms:
  - not effective in disconnected networks
  - elaborate, can be complex
  - relatively large HW overhead
  - limited routing capability around faulty links
- Exploit architectural capabilities:
  - make use of NoC inherent modularity and regularity

20.10.2010

HiPEAC, TF Reliability

5



## PFLRM Mechanism Overview

- Consists of four main phases
  1. Error **occurrence** at the link
  2. Error **detection** (basic ECC)
    - at the receiver router, just detect a bit fault (single)
  3. Fault vector generation → **diagnosis**
    - sender router sends test vectors to the receiver router over the faulty link
    - receiver creates fault vectors, as many as equal to the max fault cluster plus 1
  4. Fault **recovery** using **reconfiguration**
    - mask vector generation, derived from fault vectors, at the receiver and sender flit retransmission
    - appropriate flit bit rotation at the sender, multiple rotated flit re-transmission, and flit de-rotation at the receiver

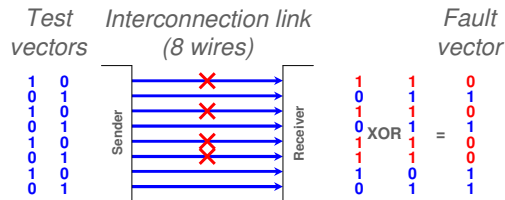
20.10.2010

HiPEAC, TF Reliability

6



## Mechanism Demonstration (1/4)



- Error detection
- Creation of fault vector
- Calculation of max continuous fault segment, size:  $m = 2$ 
  - time required to send a data flit over PFL:  $L = m + 1 = 3$  cycles

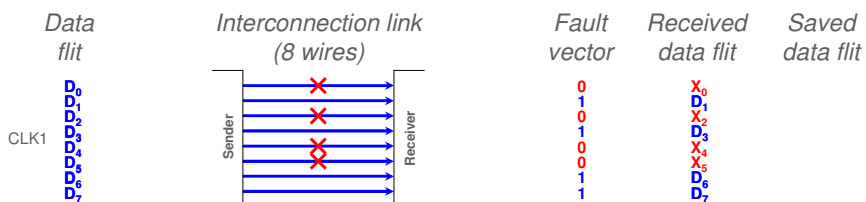
20.10.2010

HiPEAC, TF Reliability

7



## Mechanism Demonstration (2/4)



- Error detection
- Creation of fault vector
- Calculation of max continuous fault segment, size:  $m = 2$ 
  - time required to send a data flit over PFL:  $L = m + 1 = 3$  cycles
- Data transmission and restoration

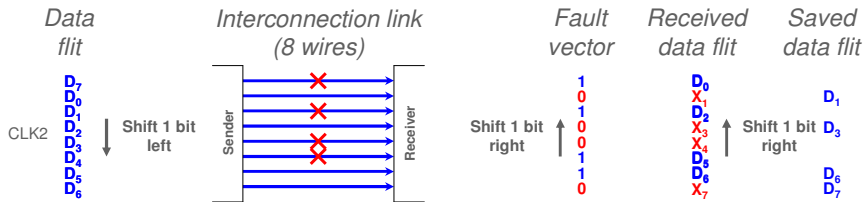
20.10.2010

HiPEAC, TF Reliability

8



## Mechanism Demonstration (3/4)



- Error detection
- Creation of fault vector
- Calculation of max continuous fault segment, size:  $m = 2$ 
  - time required to send a data flit over PFL:  
 $L = m + 1 = 3$  cycles
- Data transmission and restoration

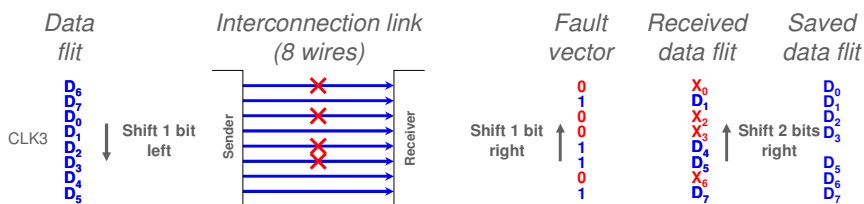
20.10.2010

HiPEAC, TF Reliability

9



## Mechanism Demonstration (4/4)



- Error detection
- Creation of fault vector
- Calculation of max continuous fault segment, size:  $m = 2$ 
  - time required to send a data flit over PFL:  
 $L = m + 1 = 3$  cycles
- Data transmission and restoration

20.10.2010

HiPEAC, TF Reliability

10



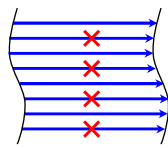
# Attainable Performance (theoretical)

## Best case:

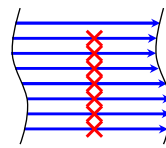
- Half of the wires are faulty
- PFL latency:  $L = 2$  cycles

## Worst case:

- All but one wire is faulty
- PFL latency is equal to the number of wires in the link



Network link

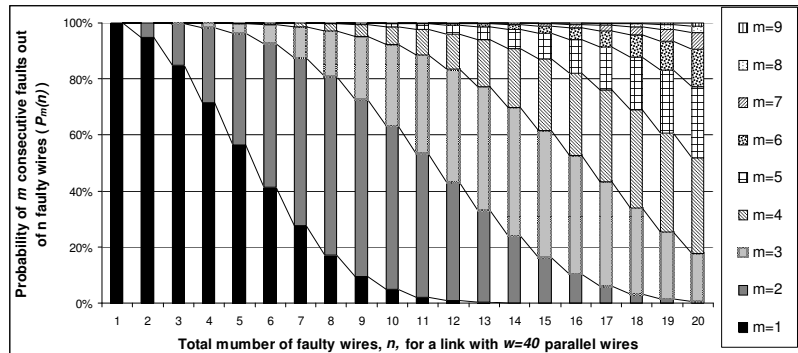


Network link



# PFL Probabilistic Fault Model

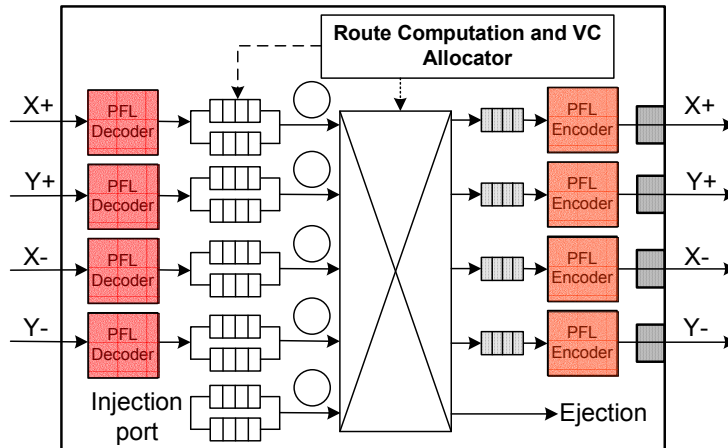
- Random link fault distribution within the link (no low-level model exists)
- Case study:
  - $w = 40$  bits is the link width (number of wires);
  - $m$  is max continuous fault segment size.





# Augmentations to basic Wormhole Router to Support PFLs

## Enhanced Wormhole NoC Router



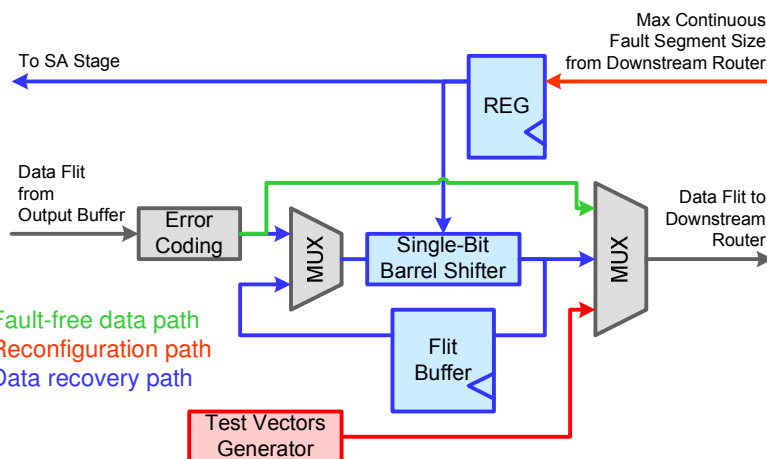
20.10.2010

HiPEAC, TF Reliability

13



# PFL Encoder Hardware



20.10.2010

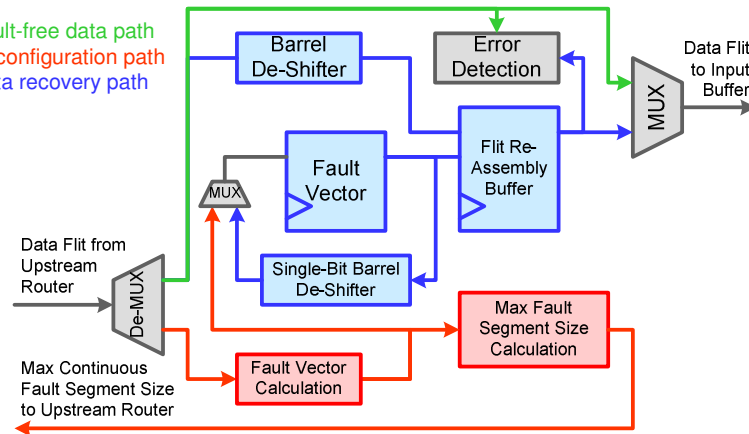
HiPEAC, TF Reliability

14



## PFL Decoder Hardware

- Fault-free data path
- Reconfiguration path
- Data recovery path



20.10.2010

HiPEAC, TF Reliability

15



## Routing Algorithms

Goal of routing algorithms: offer load-balancing to attain as high performance as possible, while bypassing PFLs accordingly

1. Adaptive routing algorithms:
  - a) Routing function:
    - OPT-Y maximally adaptive deadlock free
  - b) Selection function:
    - min congestion, or max credits (MCRD)
    - min effective link utilization (MELU)
    - straight lines (SL)
2. Reference routing algorithm:
  - XY dimension order routing (DOR-XY)

20.10.2010

HiPEAC, TF Reliability

16



## Experimental Setup (1/1)

1. Fault model:
  - a) Quantitative parameter:
    - $\alpha$  – average number of faulty wires per NoC link
  - b) Faults distribution across the link
    - probabilistic model demonstrated earlier
  - c) PFL distribution across the network
    - random: every link can equally contain faulty wires
    - hotspot: pick a limited number of faulty links with faulty wires
2. Data traffic:
  - a) uniform random (synthetic) – stress the network
  - b) real application workloads (PARSEC benchmarks)

20.10.2010

HiPEAC, TF Reliability

17



## Experimental Setup (2/2)

3. Simulator environment:
  - a) Full-system simulator (SIMICS)
  - b) Detailed cycle-accurate network model (Wisconsin Multifacet GEMS with GARNET)
4. System parameters:

Processors	16 UltraSparc III+ cores
L1 Caches	Separate 32 KB I&D, 4-way set associative, 2-cycle latency, 64 B line
L2 Caches	1 MB/core, 10-cycle latency, 64 B line
Main Memory	4 GB, latency 200 cycles, 4 memory controllers
Network	4x4 2D Mesh, 40-bit links, 4-cycle router pipeline, 6 VCs, 6 flits per VC
OS	Solaris 10

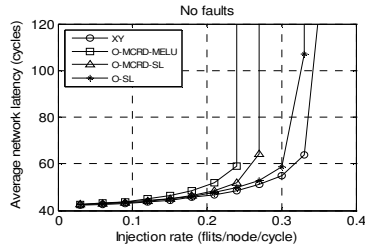
20.10.2010

HiPEAC, TF Reliability

18

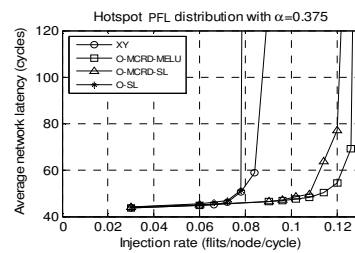
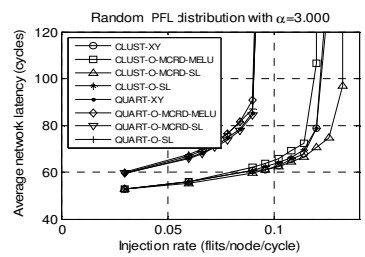


# Simulation Results: Synthetic Traffic



- CLUST – *presented mechanism*
- QUART – *a mechanism from the related work [1]*
- XY – *reference routing algorithm*
- O – *OPT-Y routing function*
- MCRD – *max credit selection function*
- MELU – *max effective link utilization selection function*
- SL – *straight lines selection function*

[1] M. Palesi et al. Leveraging Partially Faulty Links Usage for Enhancing Yield and Performance in Networks-on-Chip. In IEEE TCAD of ICS, Vol. 29, No. 3, pp. 426-440, March 2010.



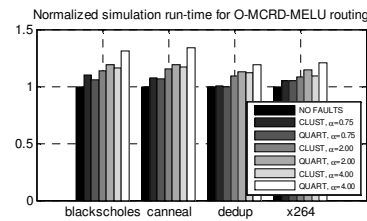
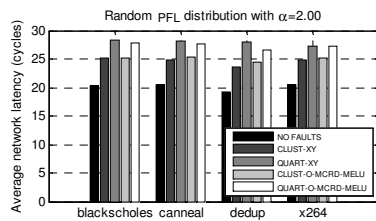
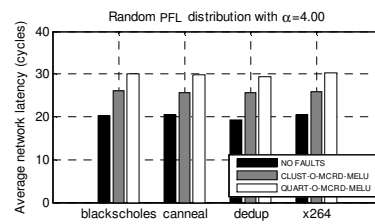
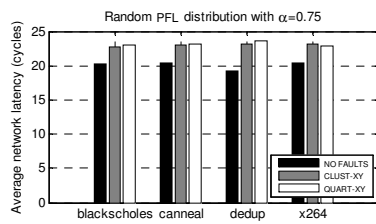
20.10.2010

HiPEAC, TF Reliability

19



# Simulation Results: Full-System Simulation



20.10.2010

HiPEAC, TF Reliability

20



## Synthesis Results – Overheads, Power

Synopsys Synthesis Results (singl conf. path)  
TSMC 65 nm library (tcbn65gplustc)

	Timing		Area		Power	
	ns	GHz	Kgates	Normalized	mW	Normalized
Router (2 VCs)	0,69	1,45	73,24	100,00%	34,77	100,00%
LLFT mechanism	0,69	1,45	16,78	22,92%	10,56	30,38%
Total	0,69	1,45	90,02	122,92%	45,33	130,38%
Router (3 VCs)	0,76	1,32	160,93	100,00%	66,52	100,00%
LLFT mechanism	0,76	1,32	16,48	10,24%	10,33	15,52%
Total	0,76	1,32	177,40	110,24%	76,84	115,52%
Router (4 VCs)	0,83	1,20	282,52	100,00%	119,08	100,00%
LLFT mechanism	0,83	1,20	16,95	6,00%	10,56	8,87%
Total	0,83	1,20	299,47	106,00%	129,64	108,87%
Router (6 VCs)	0,99	1,01	824,10	100,00%	305,40	100,00%
LLFT mechanism	0,99	1,01	17,54	2,13%	10,54	3,45%
Total	0,99	1,01	841,64	102,13%	315,93	103,45%

\* Further hardware optimization work in progress

20.10.2010

HiPEAC, TF Reliability

21



## Ongoing Work

- Hardware optimizations to reduce power and area (gate count) overheads
- Fully fault-tolerant routing algorithm:
  - can handle partially faulty links
  - determine tradeoffs between a PFL and a PFL marked as non-traversable
  - adaptive and load-balancing
- Demonstration of dynamic nature of PFL mechanism

20.10.2010

HiPEAC, TF Reliability

22



## Conclusions

- Dynamic PFL mechanism tolerates permanent and transient faults in NoC communication links
- Dynamic, can handle permanent and transient faults
- Graceful performance degradation
- Increase in link latency correlates to wire fault distribution clustering
- Maintains network connectivity under a large number of faults
- Guaranteed data delivery
- Relatively simple implementation, hw optimizations
- Additional performance to the mechanism presented in the related work under the described fault model



## Questions?

Thank you



## References

- Palesi, M.; Kumar, S.; Catania, V.;  
*"Leveraging Partially Faulty Links Usage for Enhancing Yield and Performance in Networks-on-Chip,"* Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on , vol.29, no.3, pp.426-440, March 2010.
- Agarwal N. et al.  
*"GARNET: A Detailed on-Chip Network Model Inside a Full-System Simulator,"* ISPASS, pp. 33–43, April 2009.
- Bienia C. et al.  
*"The PARSEC Benchmark Suite: Characterization and Architectural Implications,"* PACT, pp. 72–81, Oct. 2008.
- Duato J. et al.  
*"Interconnection Networks: An Engineering Approach".* Morgan Kaufmann Publishers Inc. ISBN: 1558608524, 2002.
- Semiconductor Industry Association.  
*International Technology Roadmap for Semiconductors 2009.*

