5th International Workshop on Network on Chip Architectures
(NoCArc 2012)

Position-Based Weighted Round-Robin Arbitration for Equality of Service in Many-Core Network-on-Chips

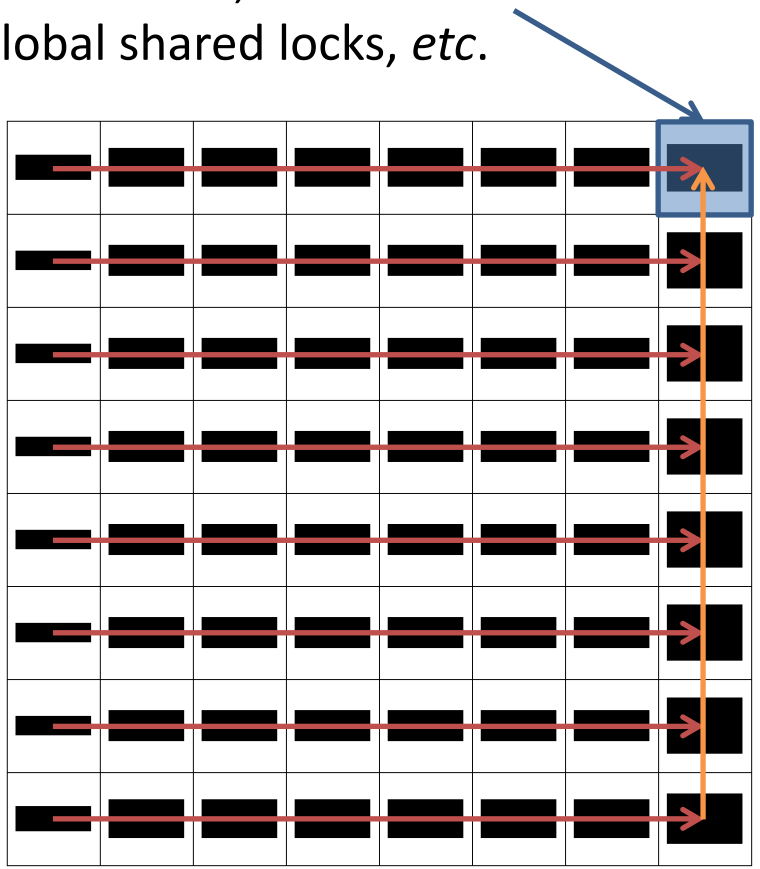
Hanmin Park and Kiyoung Choi

Seoul National University

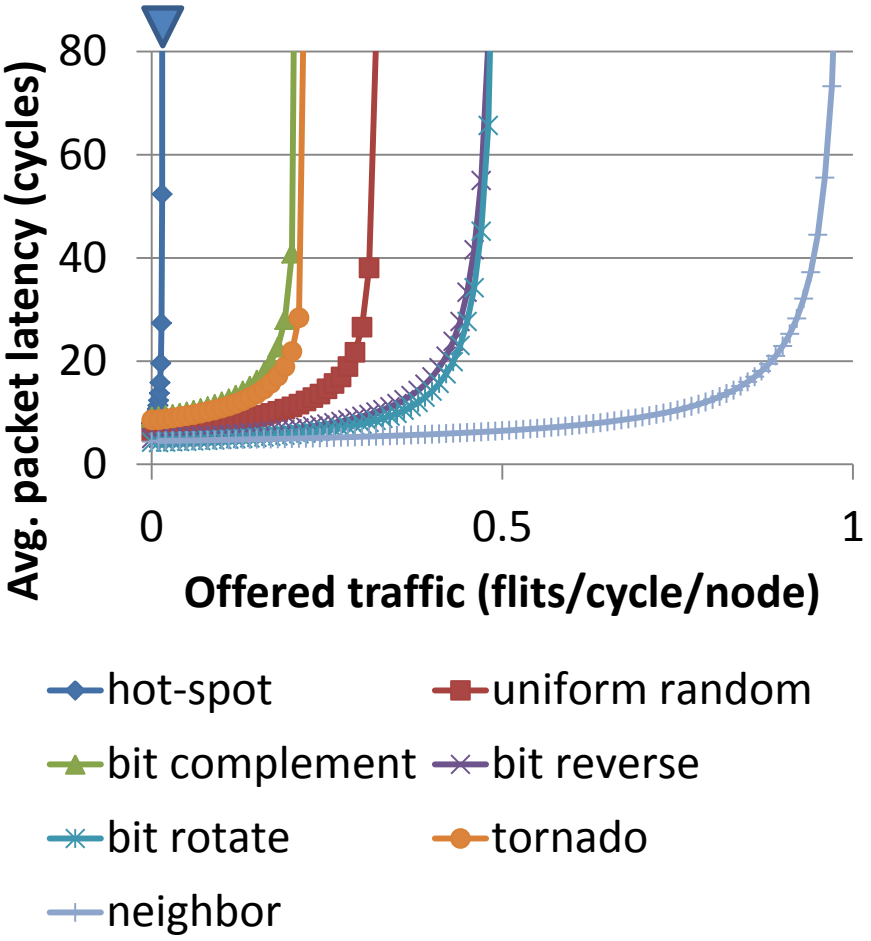
Hot-spot traffic

Input buffer utilization

Shared data,
global shared locks, *etc*.



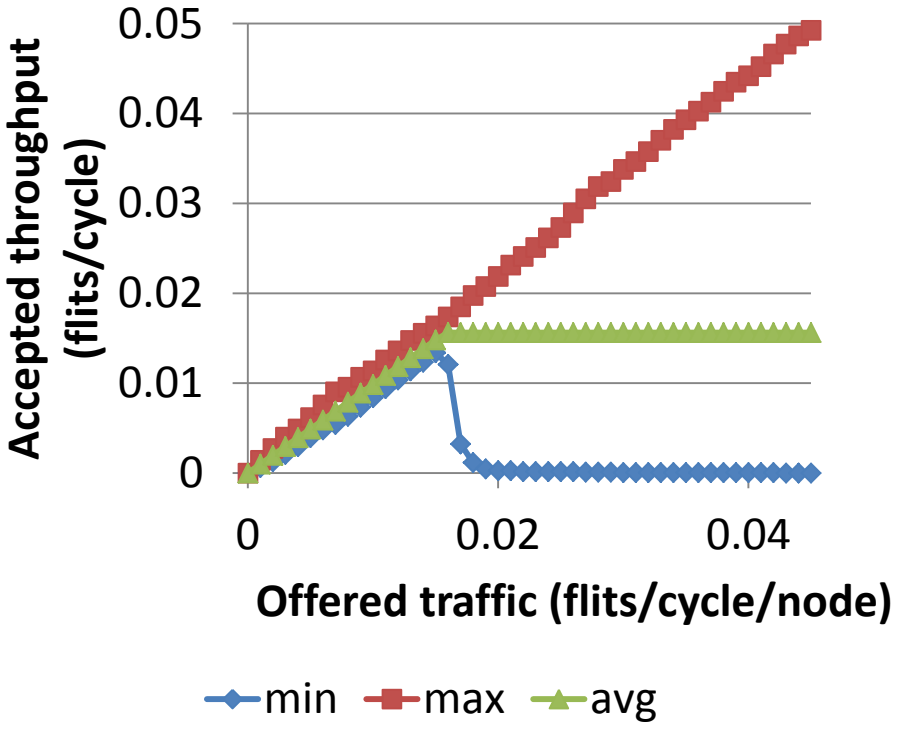
Avg. latency vs. offered traffic



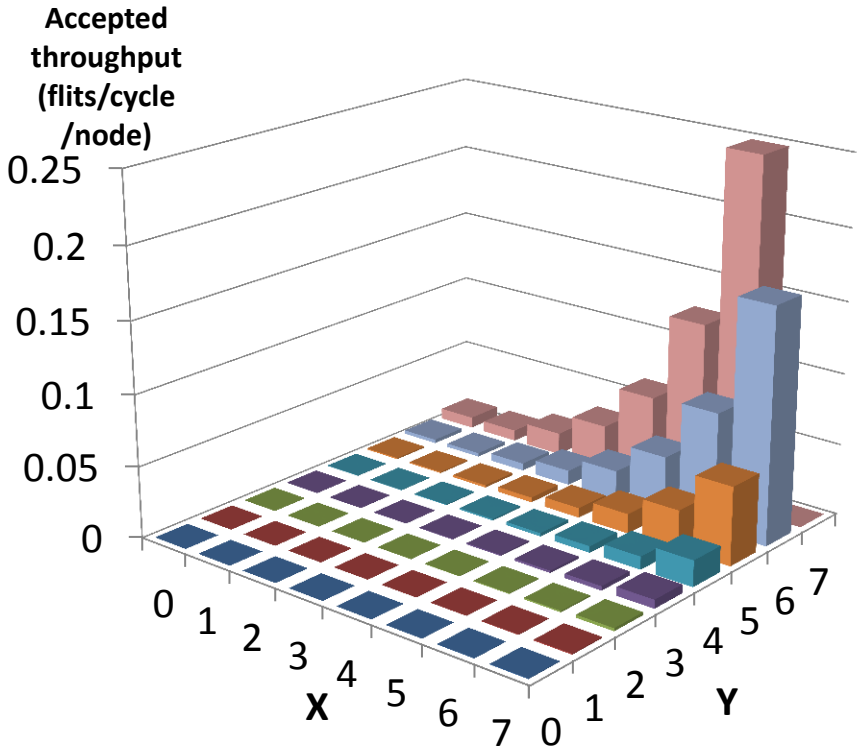
(xy routing and round-robin arbitration are used.)

EoS with hot-spot traffic

Accepted Throughput vs. offered traffic



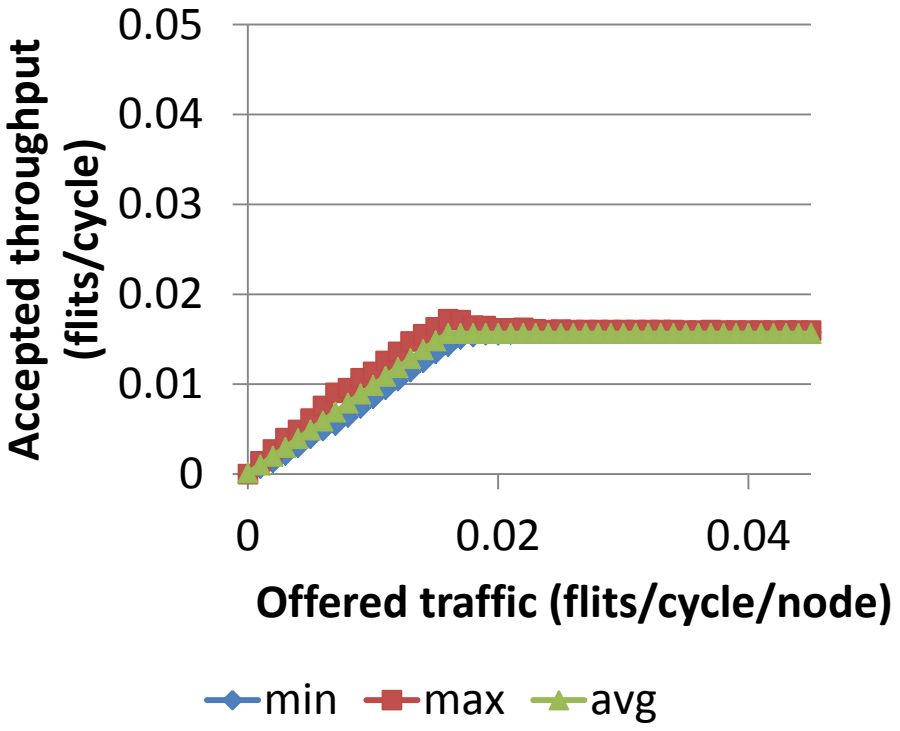
Accepted throughput distributions



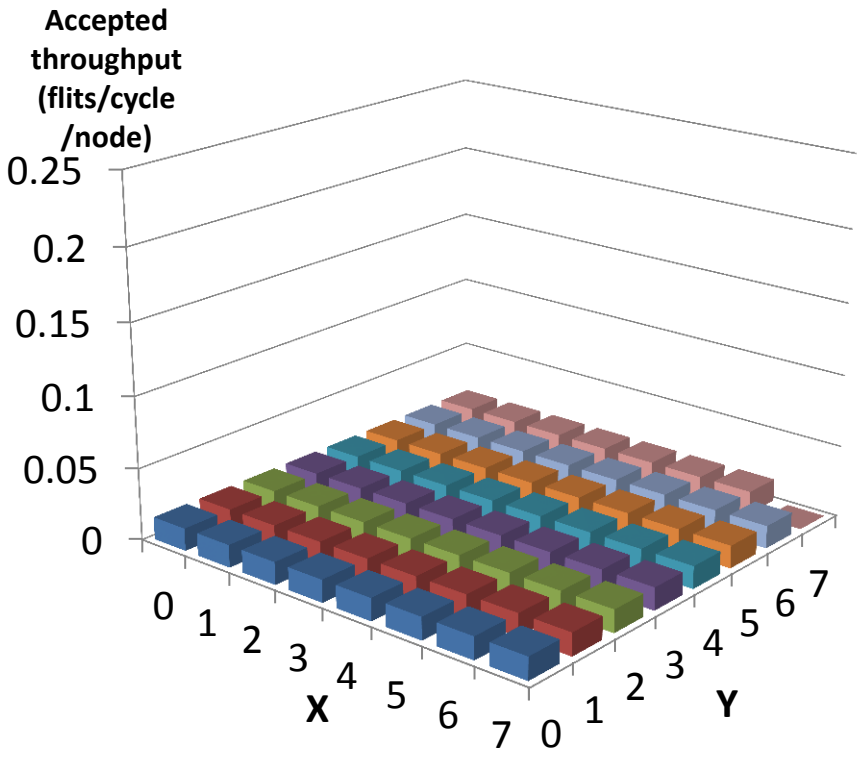
(When the offered traffic is 1.00.)

EoS with hot-spot traffic

Accepted Throughput vs. offered traffic



Accepted throughput distributions



(When the offered traffic is 1.00.)

Previous work

- Large-scale computer networks — Generalized processor sharing, fair queuing, *etc*.
- D. Abts and D. Weisser, "Age-Based Packet Arbitration in Large-Radix k-ary n-cubes," in *Supercomputing*, 2007.
- M. M. Lee, J. Kim, D. Abts, M. Marty, and J. W. Lee, "Probabilistic Distance-based Arbitration: Providing EoS for Many-core CMPs," in *MICRO*, 2010.

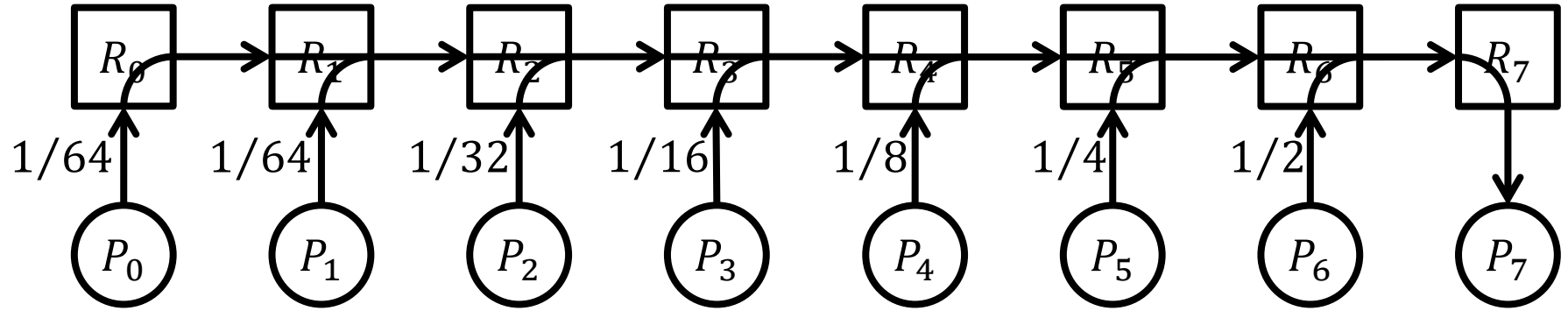
The main section:

Position-Based Weighted Round-Robin Arbitration

Motivation

Probabilistic Distance-Based Arbitration (MICRO-43)

Linear weight (hop count)?



↑ 8-ary 1-mesh, P_7 is the common destination.

Geometric weight

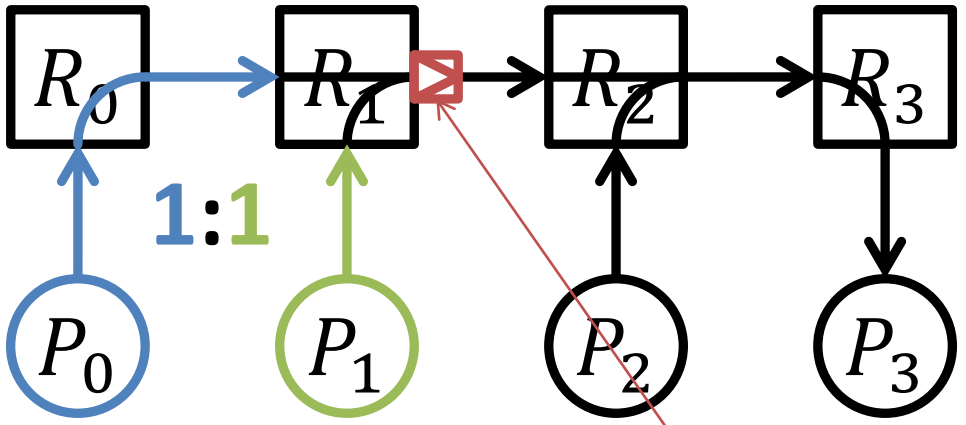
Current coordinate (c_x, c_y)
Source coordinate (s_x, s_y)

$$w = C_x^{h_x} \times C_y^{h_y}$$

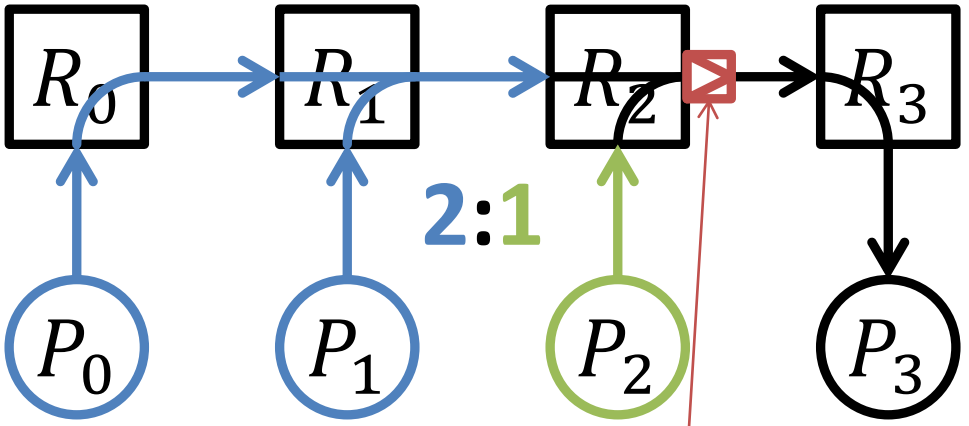
where:

- Hop count $(h_x, h_y) = (|c_x - s_x|, |c_y - s_y|)$
- Contention degree (C_x, C_y)

Motivation



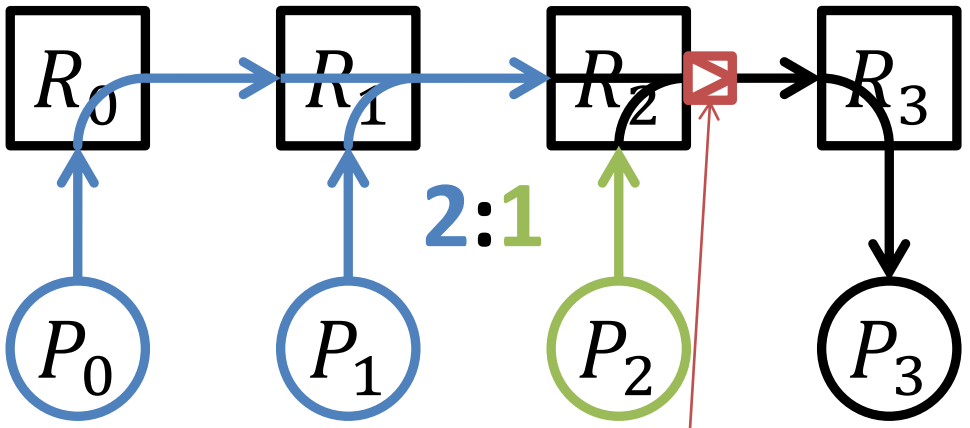
(a) The case of the east output port of R_1.



(b) The case of the east output port of R_2.

Arbitration algorithm design

Position-Based Weighted Round-Robin Arbitration

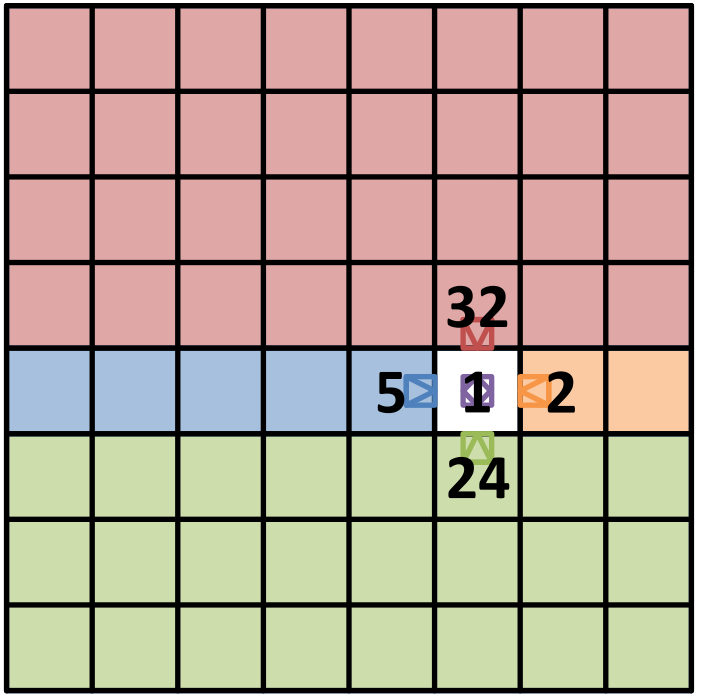


(b) The case of the east output port of R_2.

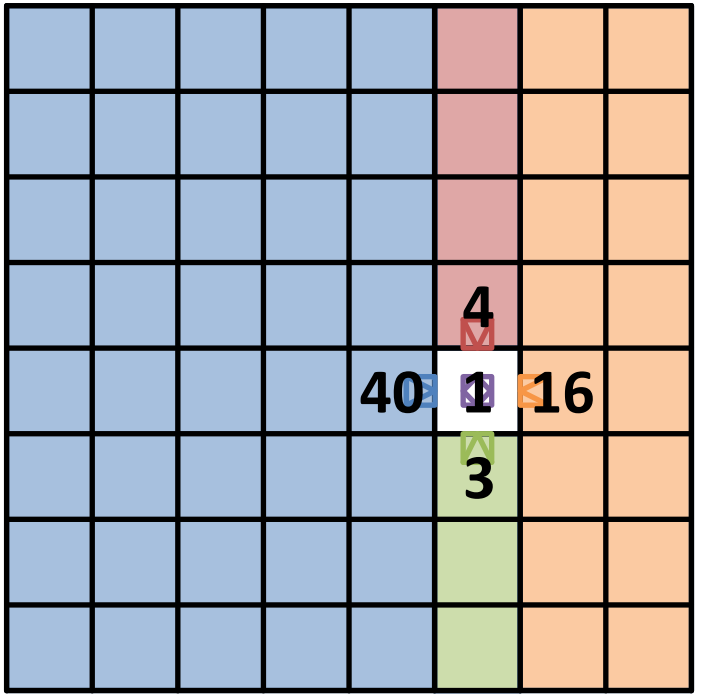
1. Count the number of nodes to be served by each input port.
2. Output port arbitration.

Extension to (8 × 8) 2D meshes

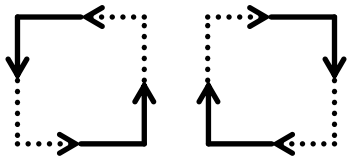
1. Count the number of nodes to be served by each input port.



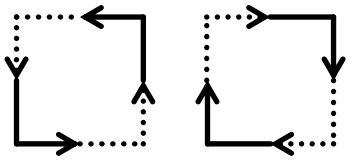
(a) xy routing algorithm



(b) yx routing algorithm

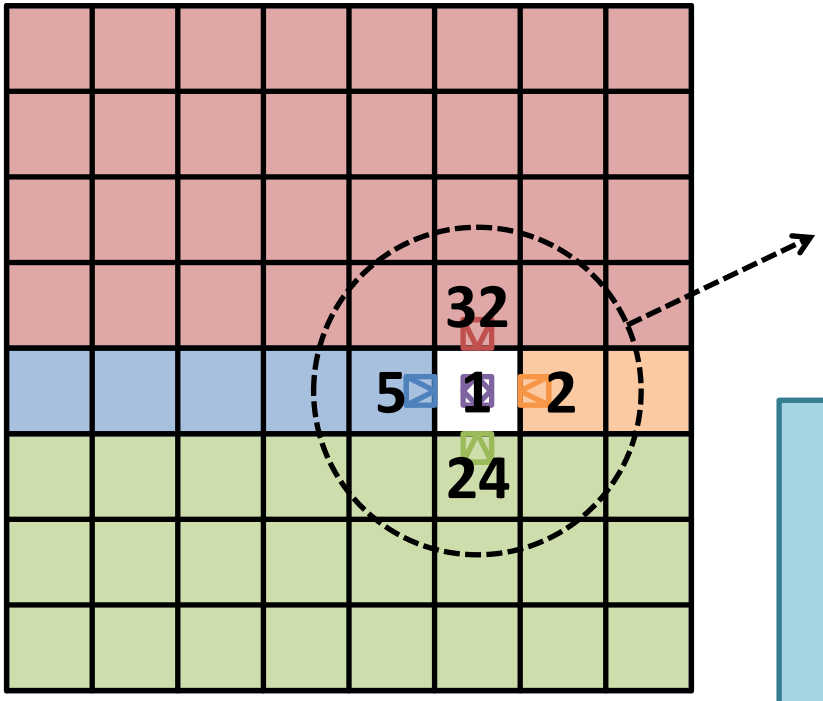


Turn models
[Glass and Ni 1994]

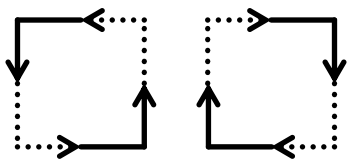


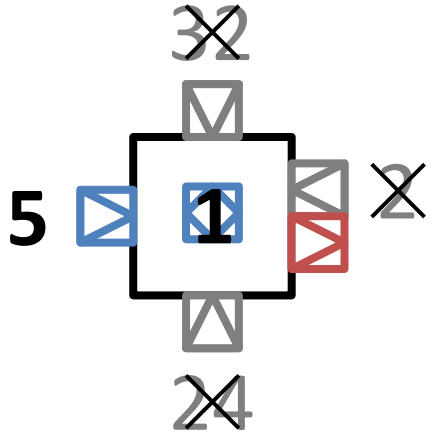
Extension to (8 × 8) 2D meshes

2. Output port arbitration.

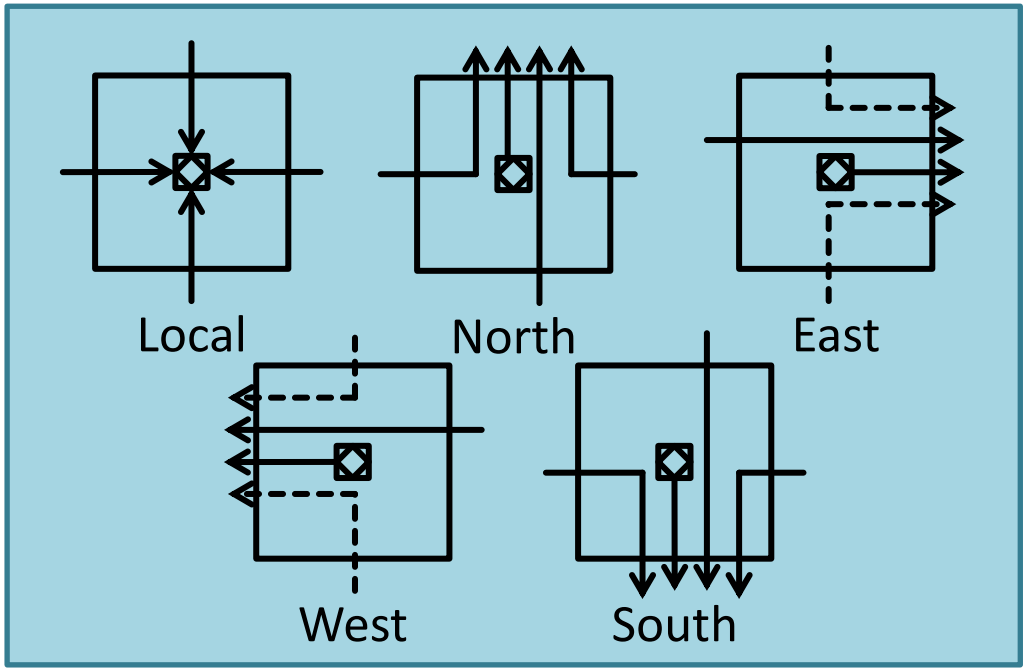


(a) xy routing algorithm

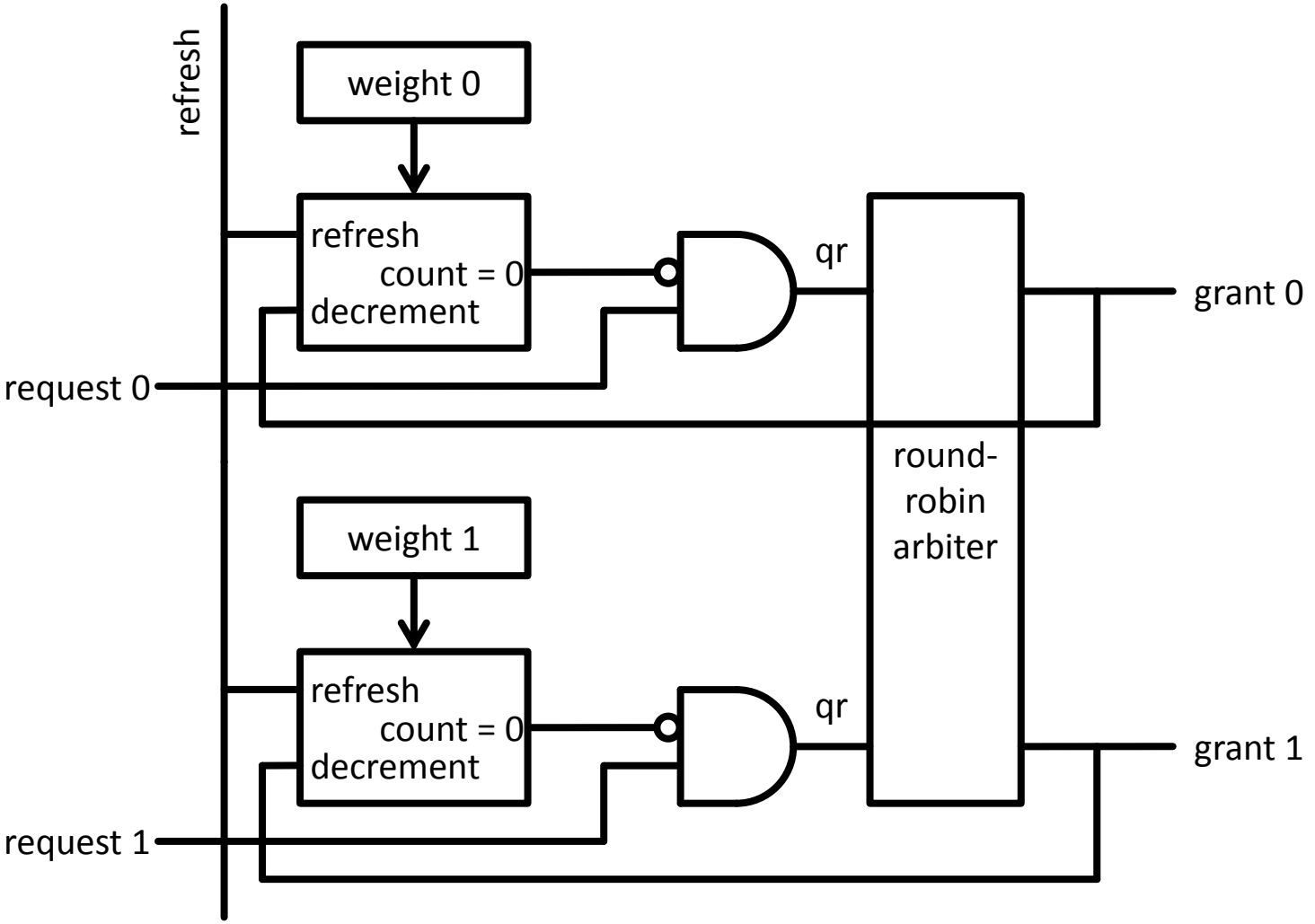




← East output port arbitration



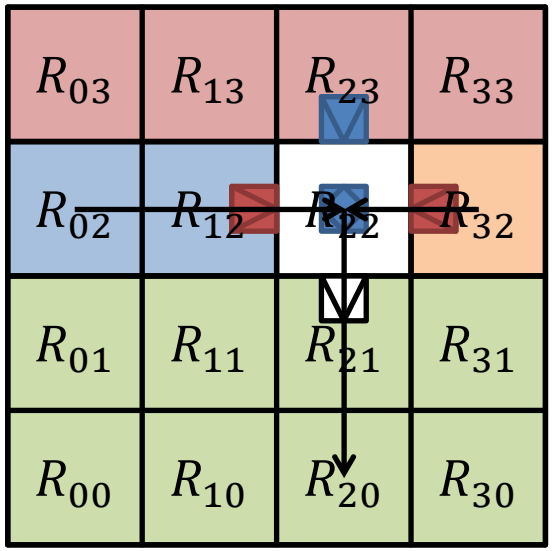
Hardware implementation



William James Dally and Brian Towles, *Principles and practices of interconnection networks*, Morgan Kaufmann, 2004.

Hardware implementation

Second stage arbitration



Active input port

Inactive input port

Weights of input ports

- $LNEWS = 1:4:1:2:8$

Refresh cycle (South output port)

- $1 + 4 + 1 + 2 = 8$

Two-stage arbitration

	Condition	**Action**
First stage	$\exists req: counter(req) \neq 0$	Masked → $counter(req) = 0$
Second stage	$\forall req: counter(req) = 0$	Masked → $counter(req) \neq 0$

Hardware implementation

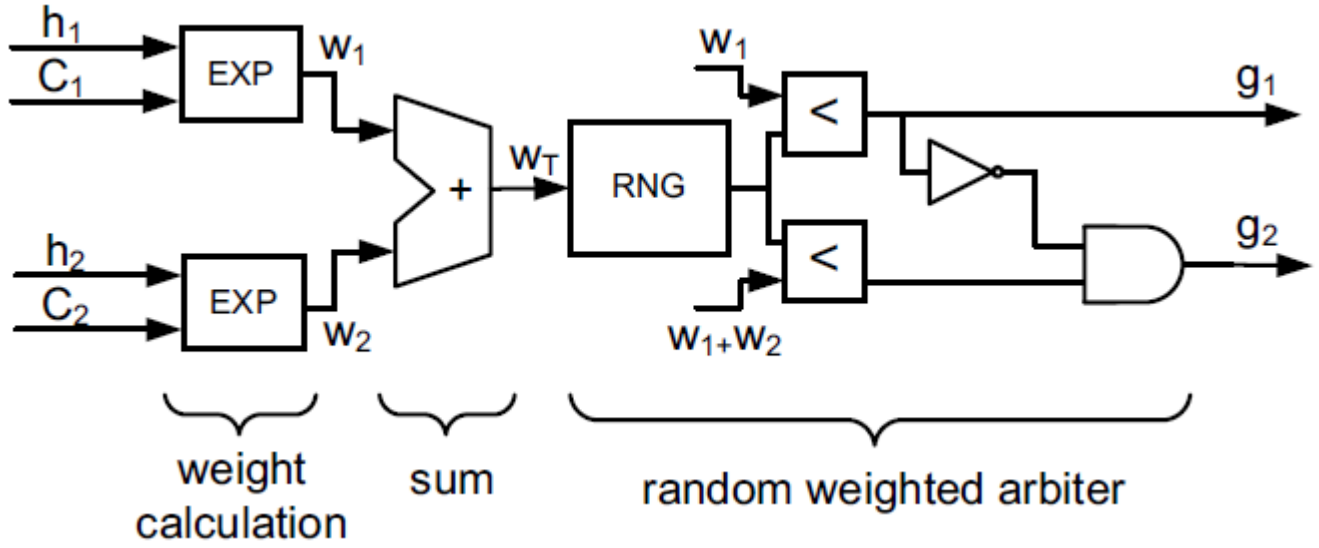
Probabilistic Distance-Based Arbitration (MICRO-43)

Geometric weight

$$w = C_x^{h_x} \times C_y^{h_y},$$

where h_i's are hop counts and
C_i's are contention degrees ($i \in \{x, y\}$).

Hardware implementation

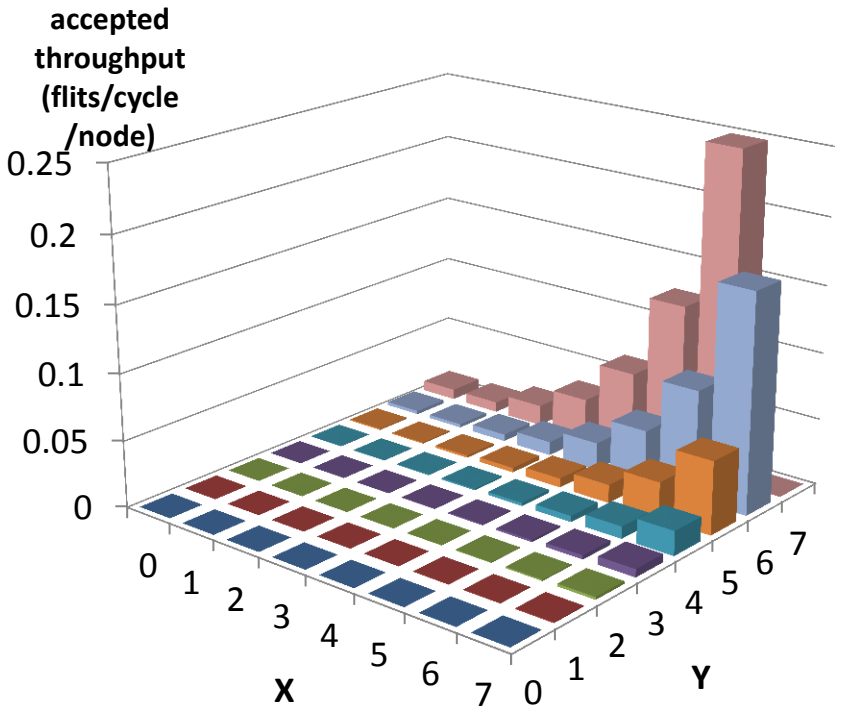


Throughput distribution & Effects on other traffic patterns

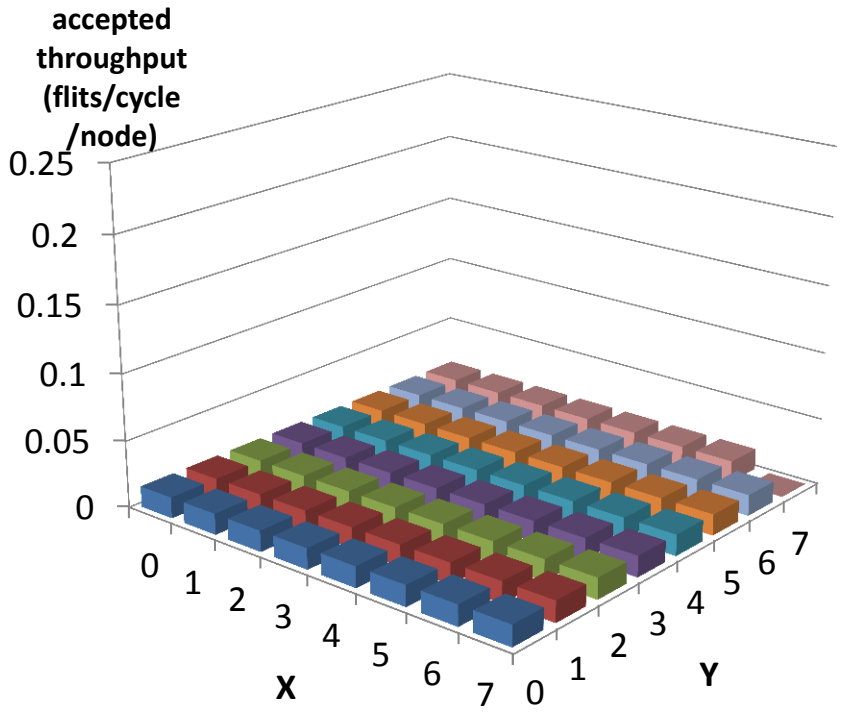
EXPERIMENTAL RESULTS

Experimental results

Accepted throughput – One hot-spot



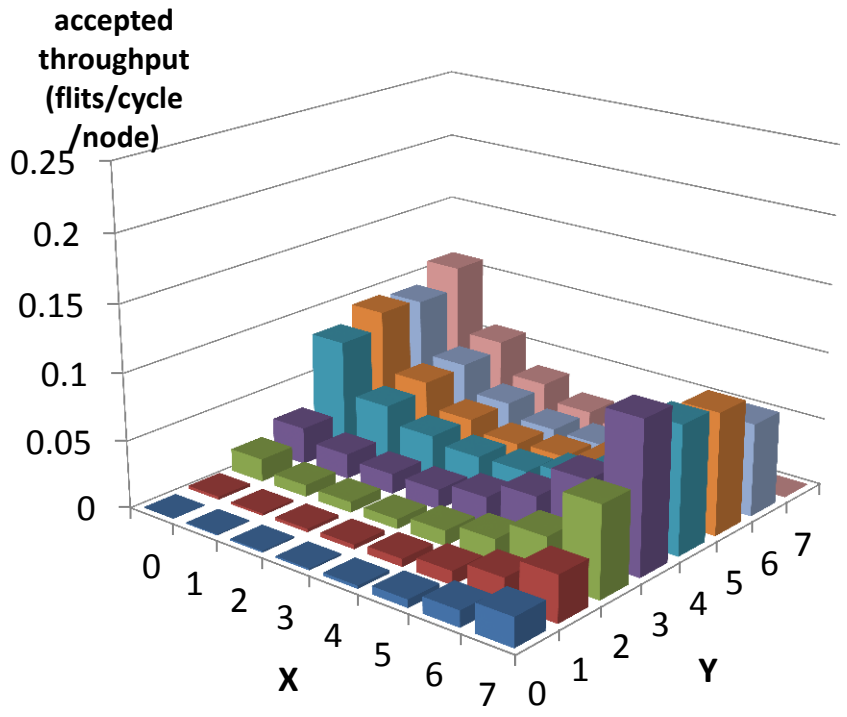
Round-robin



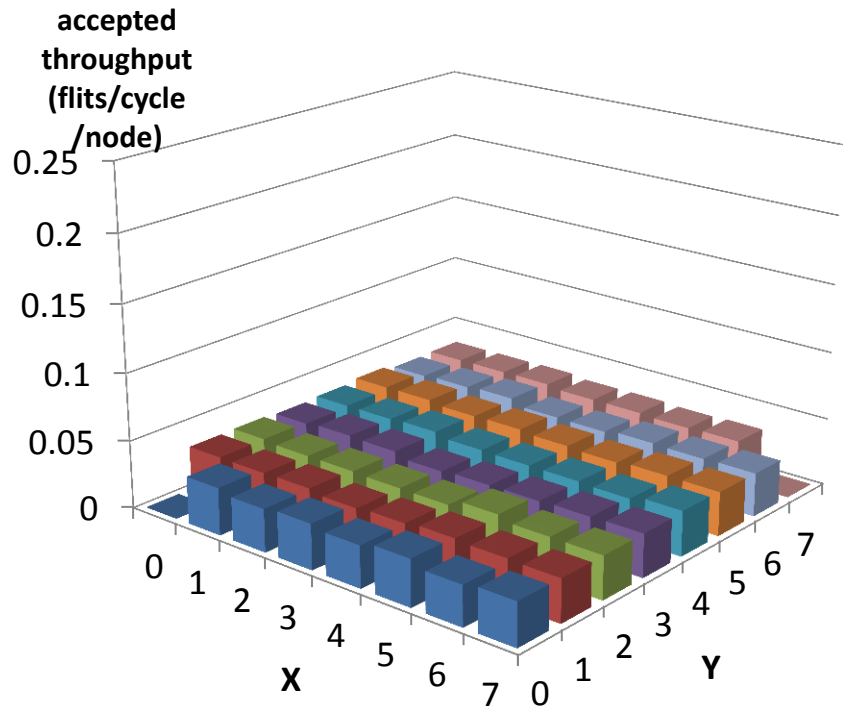
Position-based

Experimental results

Accepted throughput – Two hot-spots



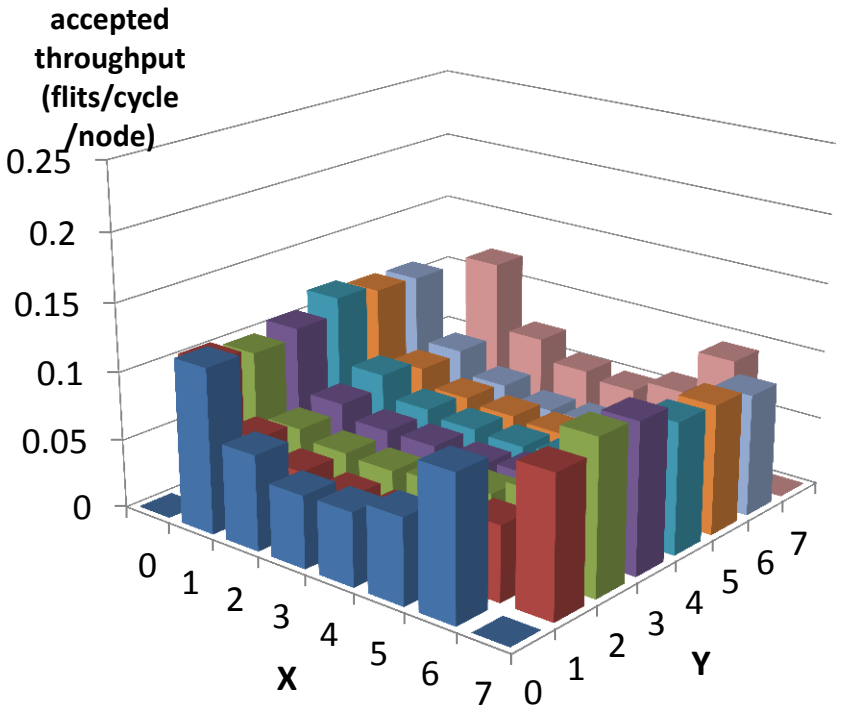
Round-robin



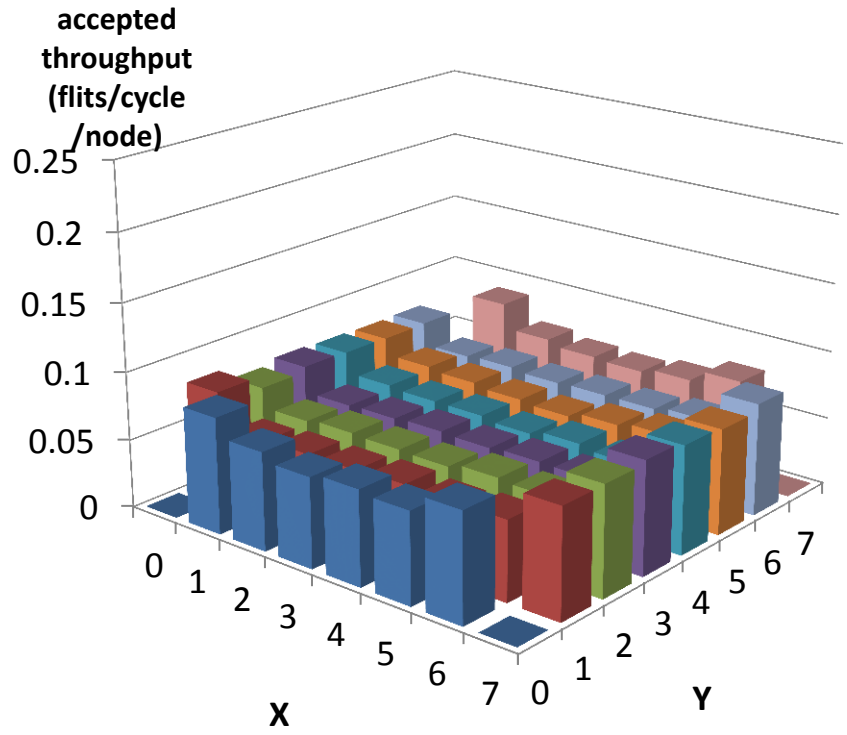
Position-based

Experimental results

Accepted throughput – Four hot-spots



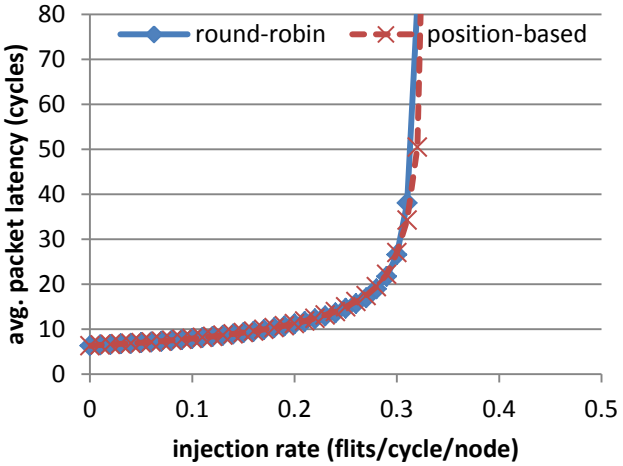
Round-robin



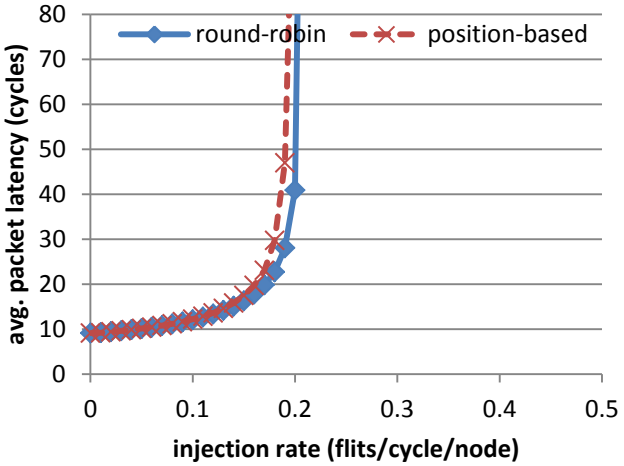
Position-based

Experimental results

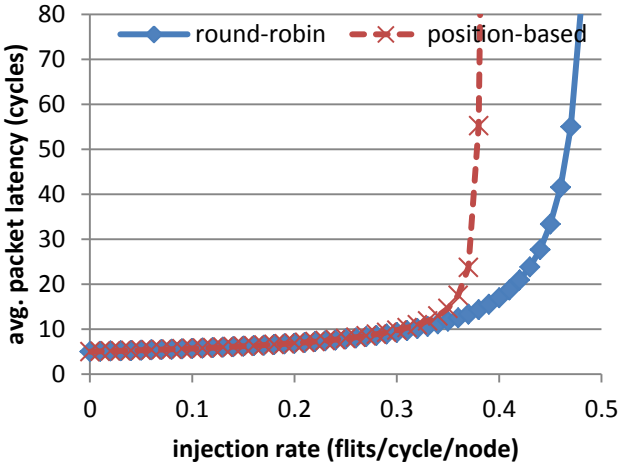
Latency-throughput curve – Synthetic traffics



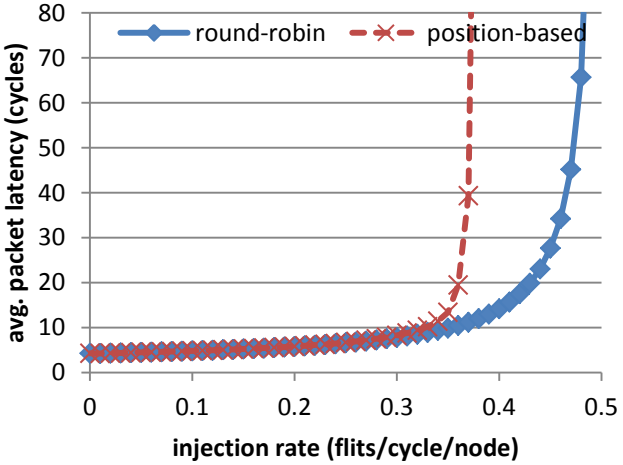
(a) Uniform random



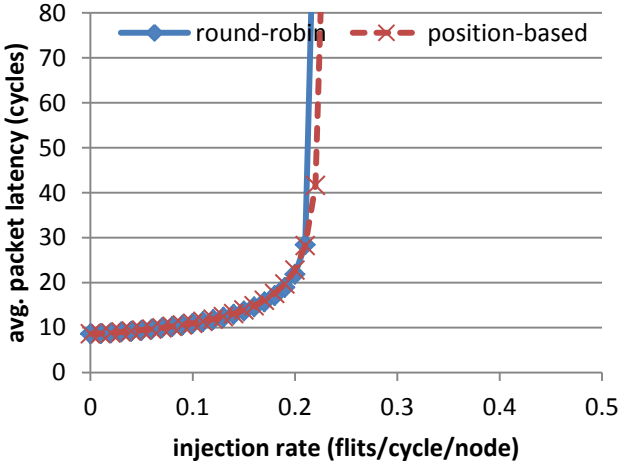
(b) Bit complement



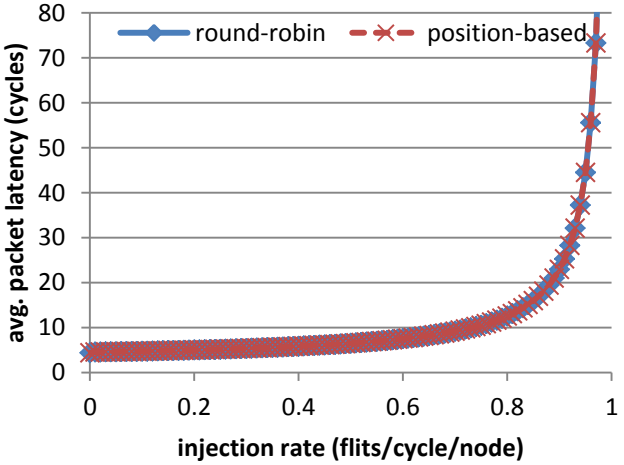
(c) Bit reverse



(d) Bit rotate



(e) Tornado



(f) Neighbor

Conclusion

- Position-based weighted round-robin arbitration is proposed for EoS in many-core NoCs.
- It is shown that the deterministic characteristics of NoC can be exploited to provide EoS with simple weighted round-robin method.
- Optimized to the hot-spot traffic, the proposed arbitration scheme does not degrade much against other traffic patterns.

Q & A

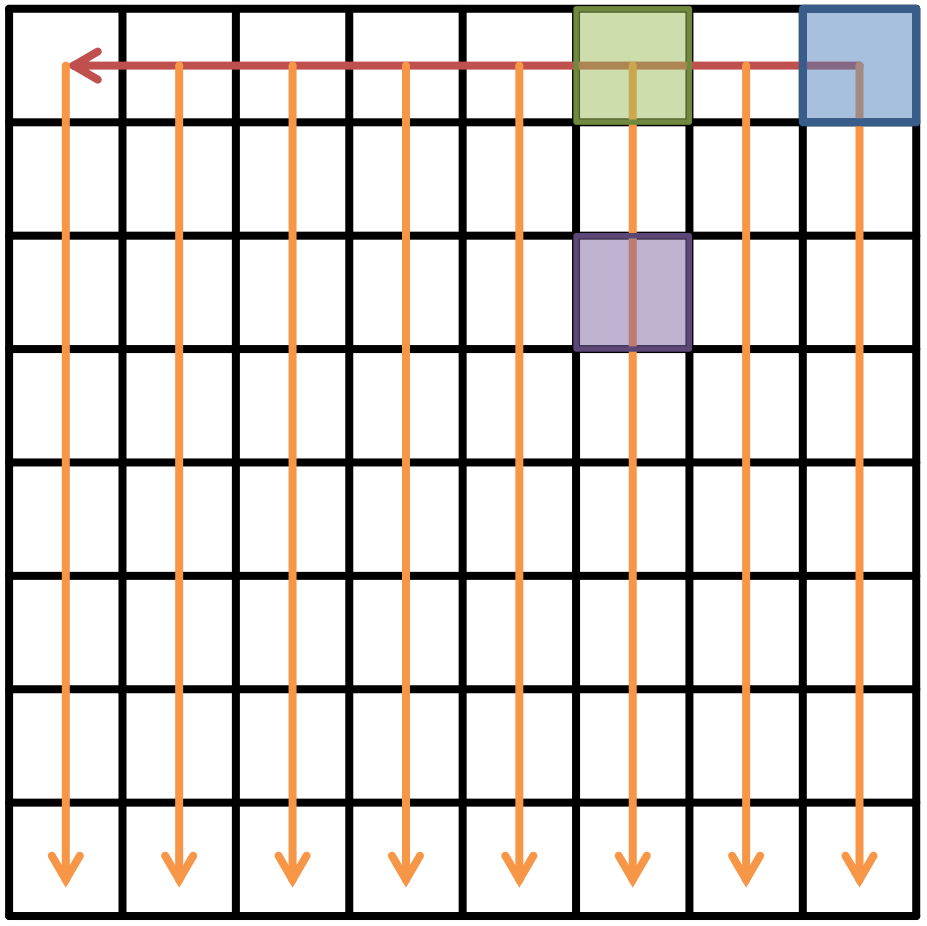
THANK YOU

Appendix

BACKUP SLIDES

Read operation

From $(7,7)$ *to the other nodes*



- No contention
  - *e.g.* $(5, 7)$
    $p_{in}^{east} \rightarrow p_{out}^{west}$ and
    $p_{in}^{east} \rightarrow p_{out}^{south}$
  - *e.g.* $(5, 5)$
    $p_{in}^{north} \rightarrow p_{out}^{south}$
- No saturation
  - Only one node is sending packets.

Experimental results

Latency-throughput curve — Hot-spot traffic