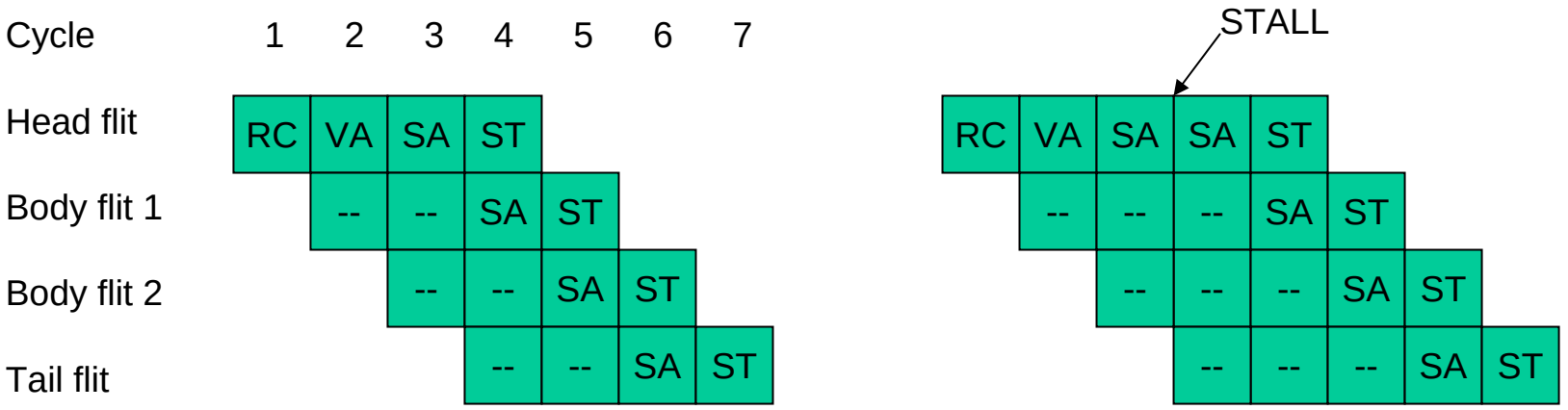
Lecture 22: Router Design

Papers:

- Power-Driven Design of Router Microarchitectures in On-Chip Networks, MICRO'03, Princeton
- A Gracefully Degrading and Energy-Efficient Modular Router Architecture for On-Chip Networks, ISCA'06, Penn-State

Router Pipeline

- Four typical stages:
  - RC routing computation: compute the output channel
  - VA virtual-channel allocation: allocate VC for the head flit
  - SA switch allocation: compete for output physical channel
  - ST switch traversal: transfer data on output physical channel



Data Points

- On-chip network's power contribution
  - in RAW (tiled) processor: 36%
  - in network of compute-bound elements (Intel): 20%
  - in network of storage elements (Intel): 36%
  - bus-based coherence (Kumar et al. '05): ~12%

- Contributors:
  - RAW: links 39%; buffers 31%; crossbar 30%
  - TRIPS: links 31%; buffers 35%; crossbar 33%
  - Intel: links 18%; buffers 38%; crossbar 29%; clock 13%

Unlike traditional off-chip networks, link power is not dominant

Network Power

- Energy for a flit = $E_R \cdot H + E_{wire} \cdot D$
  $= (E_{buf} + E_{xbar} + E_{arb}) \cdot H + E_{wire} \cdot D$

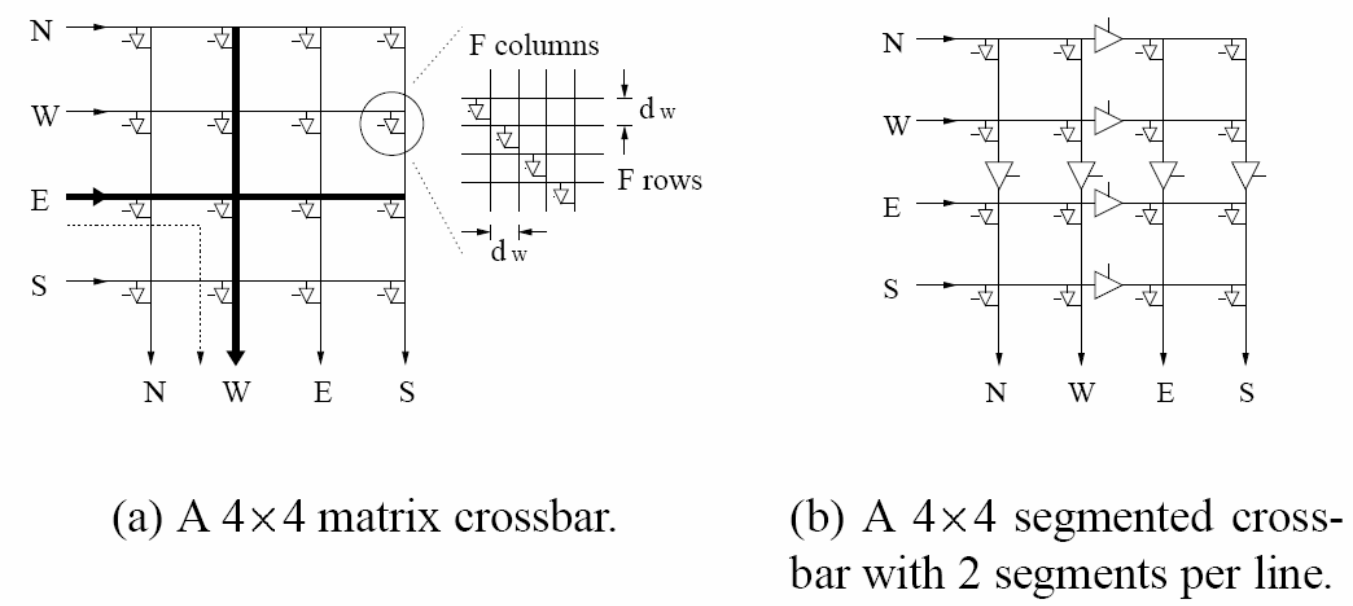
E_R = router energy
E_{wire} = wire transmission energy
E_{buf} = router buffer energy
E_{arb} = router arbiter energy

H = number of hops
D = physical Manhattan distance
E_{xbar} = router crossbar energy

- This paper assumes that $E_{wire} \cdot D$ is ideal network energy (assuming no change to the application and how it is mapped to physical nodes)

- Optimizations are attempted to E_R and H

Segmented Crossbar



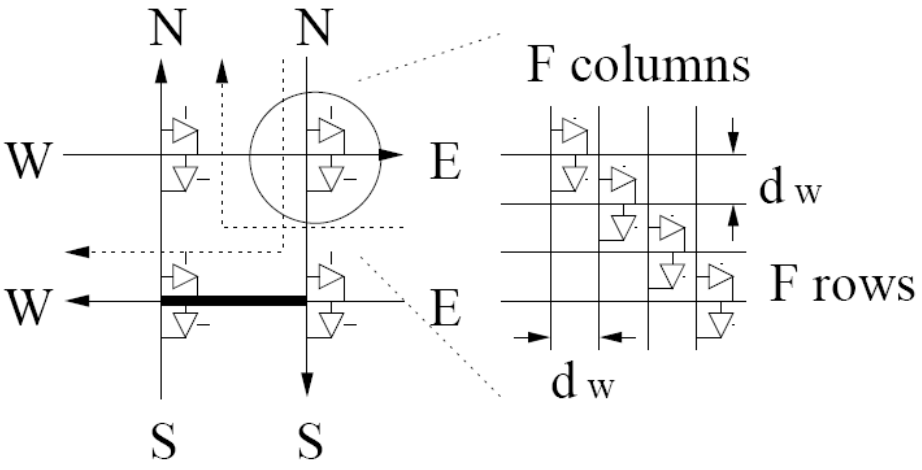
(a) A 4×4 matrix crossbar.

(b) A 4×4 segmented cross-bar with 2 segments per line.

- By segmenting the row and column lines, parts of these lines need not switch → less switching capacitance (especially if your output and input ports are close to the bottom-left in the figure above)
- Need a few additional control signals to activate the tri-state buffers (~2 control signals, ~64 data signals)
- Overall crossbar power savings: ~15-30%

Cut-Through Crossbar

- Attempts to optimize the common case: in dimension-order routing, flits make up to one turn and usually travel straight
- 2/3rd the number of tristate buffers and 1/2 the number of data wires
- "Straight" traffic does not go thru tristate buffers

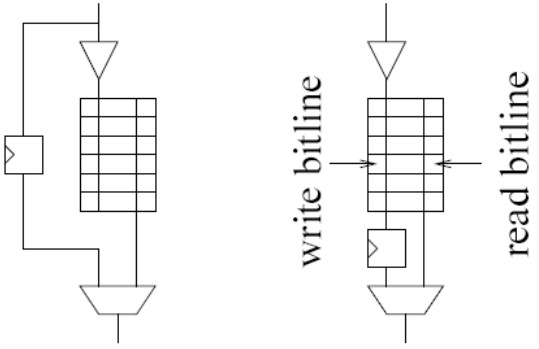


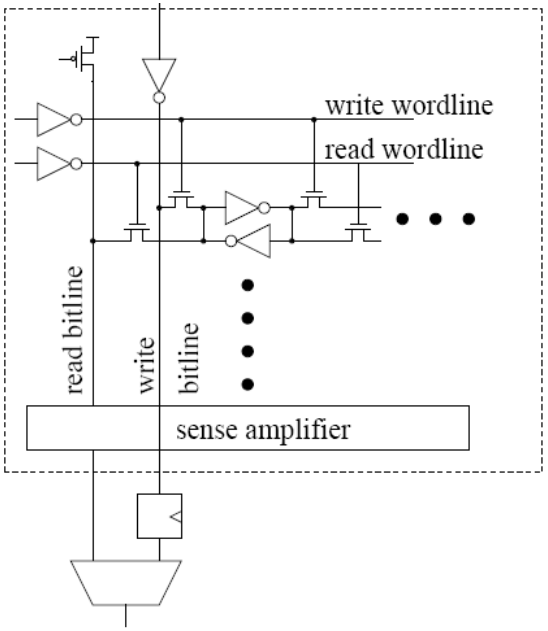
(a) A 4×4 cut-through crossbar.

- Some combinations of turns are not allowed: such as E → N and N → W (note that such a combination cannot happen with dimension-order routing)
- Crossbar energy savings of 39-52%; at full-load, with a worst-case routing algorithm, the probability of a conflict is ~50%

Write-Through Input Buffer

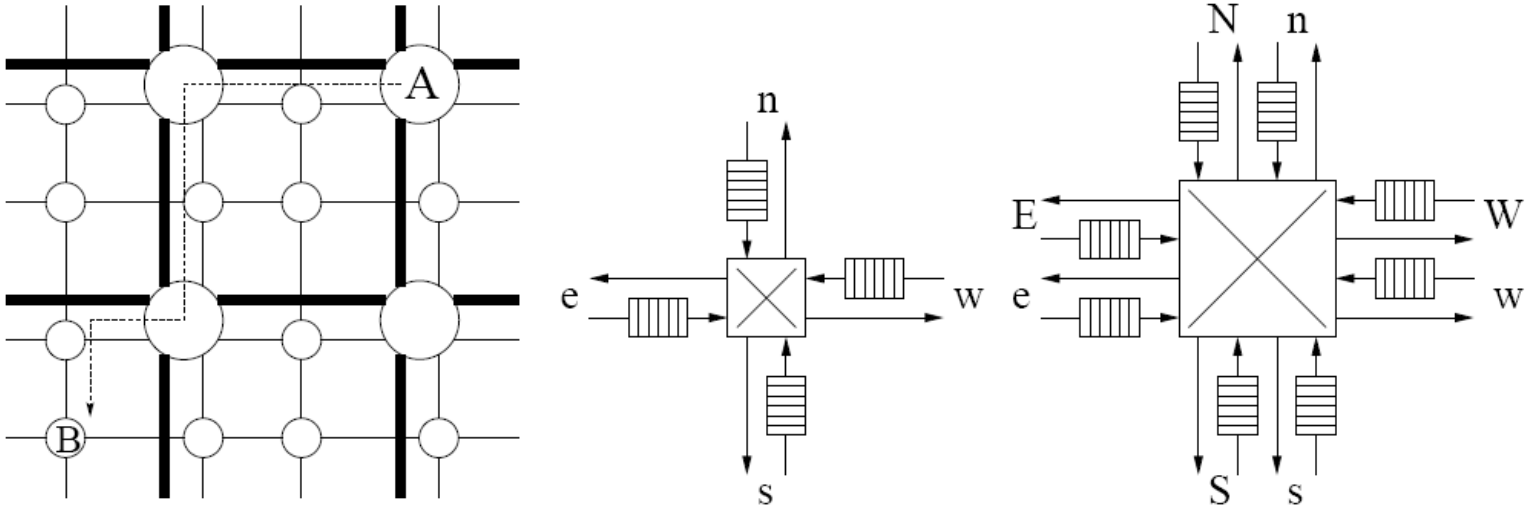
- Input flits must be buffered in case there is a conflict in a later pipeline stage
- If the queue is empty, the input flit can move straight to the next stage: helps avoid the buffer read
- To reduce the datapaths, the write bitlines can serve as the bypass path
- Power savings are a function of rd/wr energy ratios and probability of finding an empty queue





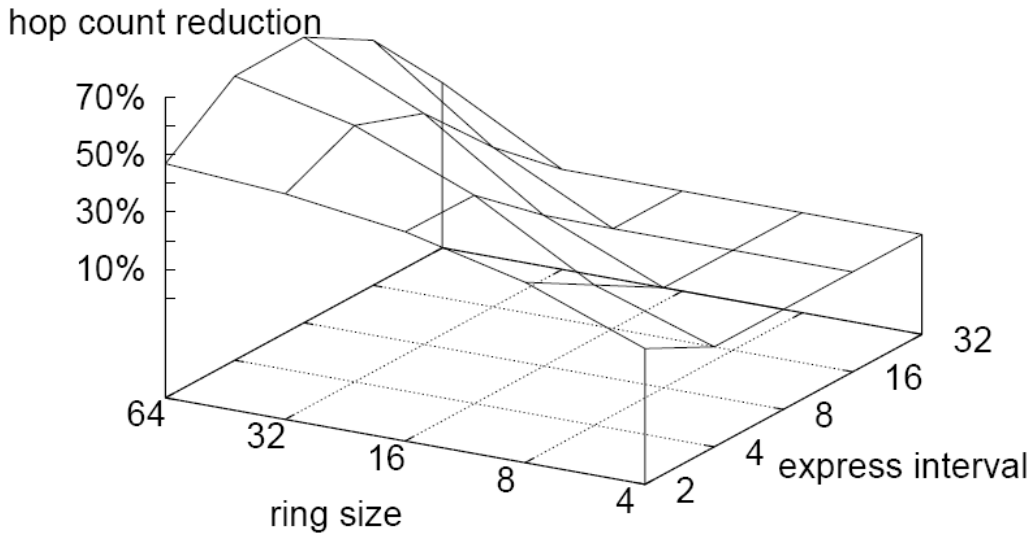
Express Channels

- Express channels connect non-adjacent nodes – flits traveling a long distance can use express channels for most of the way and navigate on local channels near the source/destination (like taking the freeway)
- Helps reduce the number of hops
- The router in each express node is much bigger now



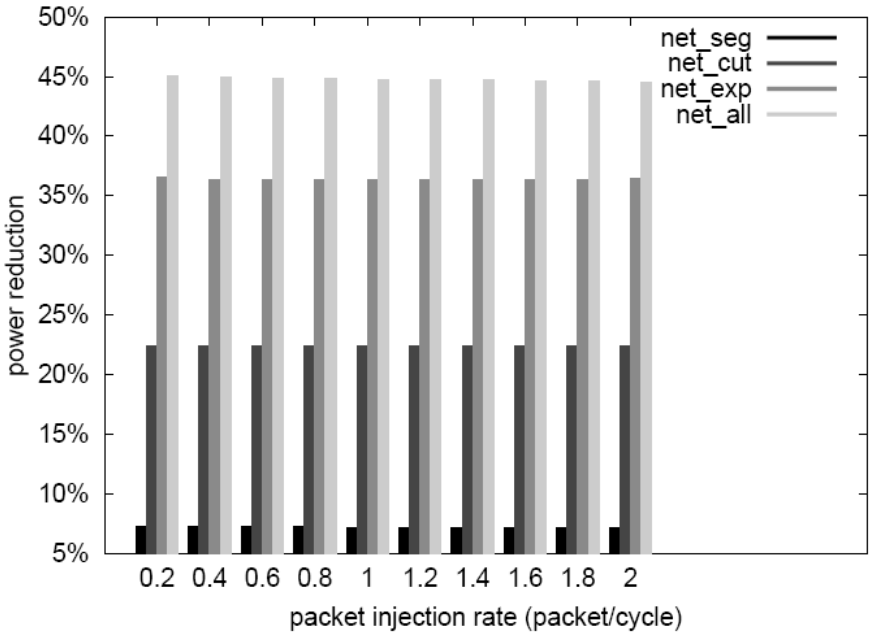
Express Channels

- Routing: in a ring, there are 5 possible routes and the best is chosen; in a torus, there are 17 possible routes
- A large express interval results in fewer savings because fewer messages exercise the express channels



Results

- Uniform random traffic (synthetic)
- Write-thru savings are small
- Exp-channel network has half the flit size to maintain the same bisection-bandwidth as other models (express interval of 2)
- Baseline model power breakdown: link 44%, crossbar 33%, buffers 23%
- Express cubes also improve 0-load latency by 23% -- the others have a negligible impact on performance

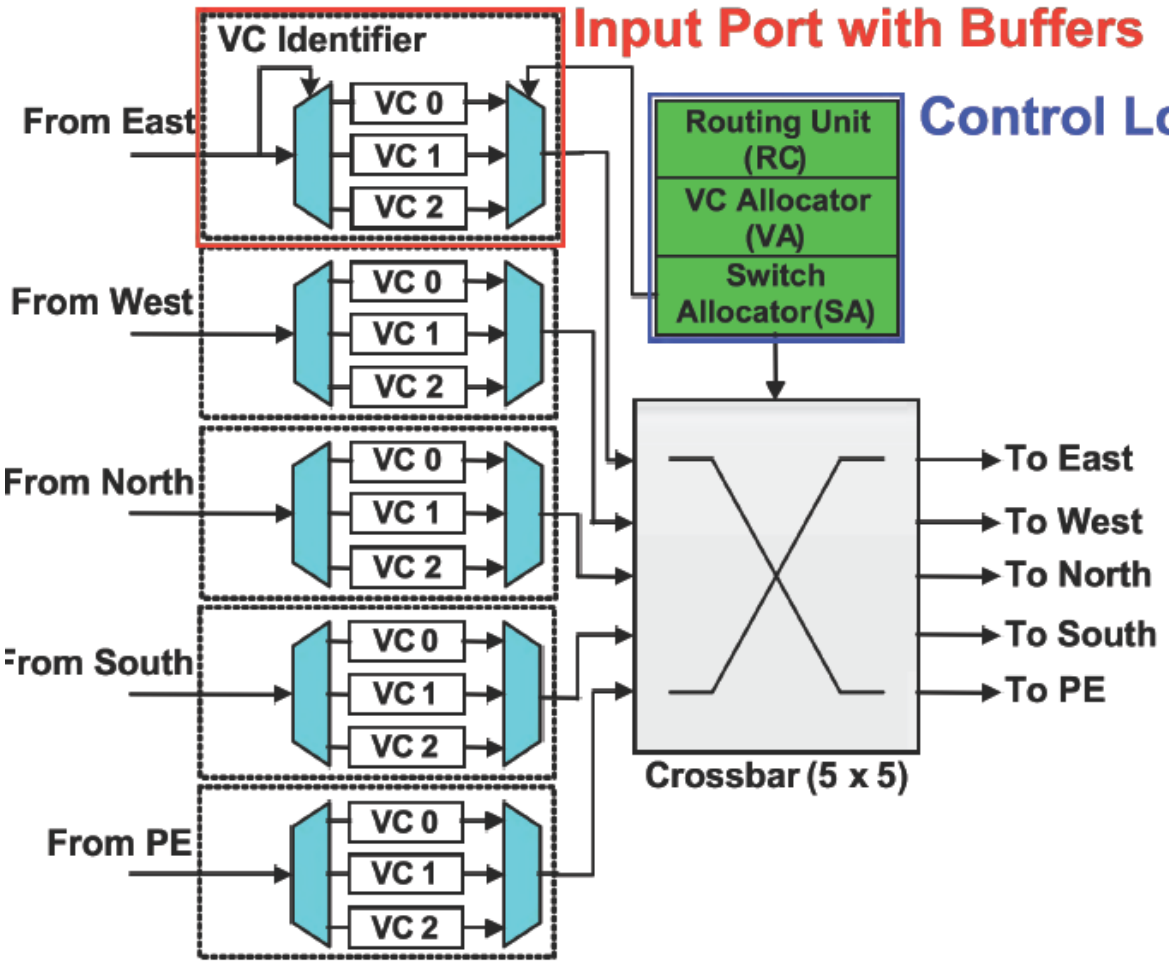


(c) 8×8 network power savings of 4 confi gurations.

Table 3. Average total network power savings (relative to net_base configuration).

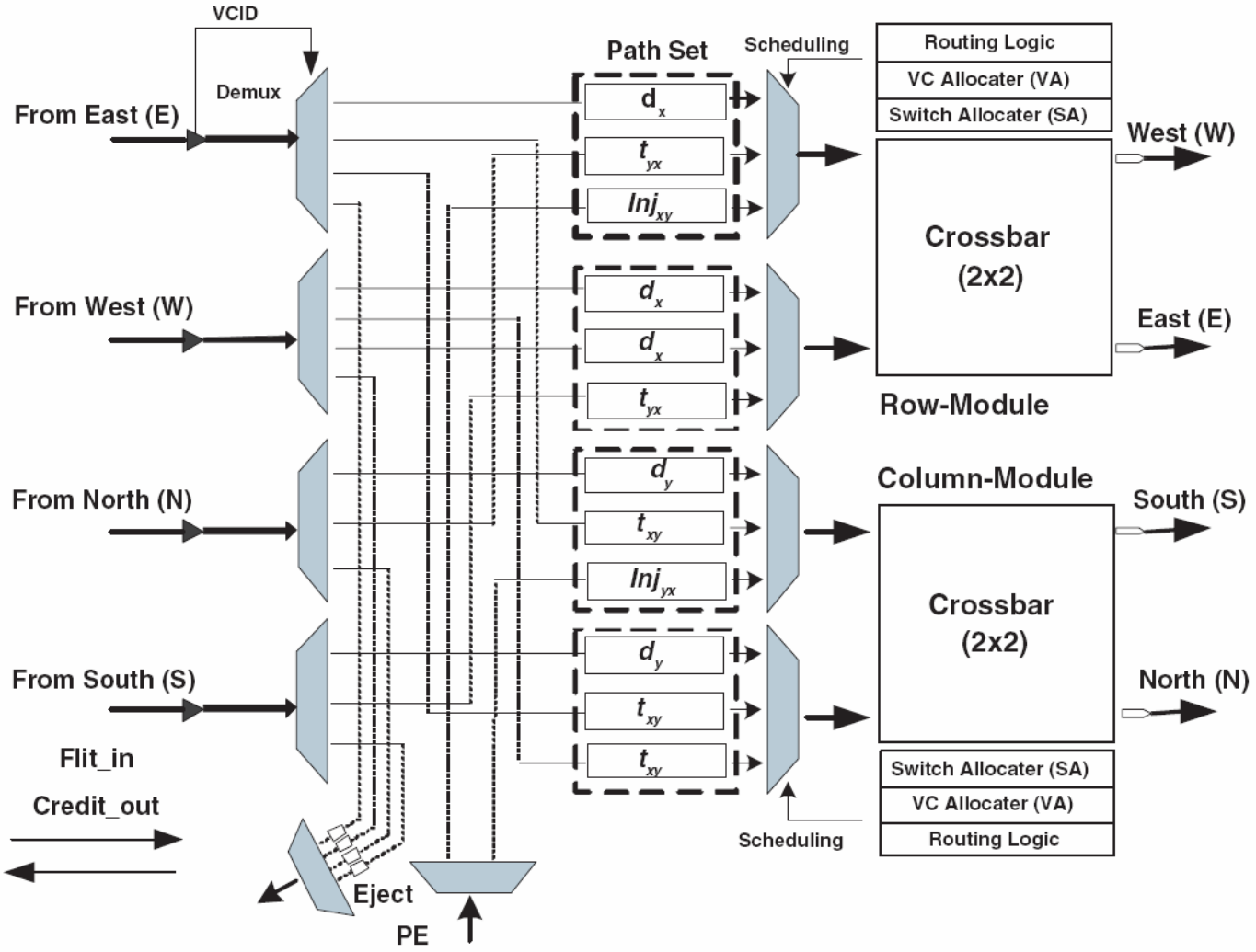
	8×8 torus (random)	4×4 torus (random)	TRIPS traces
net_cut	22.4%	21.6%	20.4%
net_seg	7.2%	6.9%	6.6%
net_wrt	4.9%	4.5%	3.8%
net_exp	36.3%	27.2%	30.9%
net_all	44.9%	36.3%	37.9%

Conventional Router



Slide taken from presentation at OCIN'06

The RoCo Router

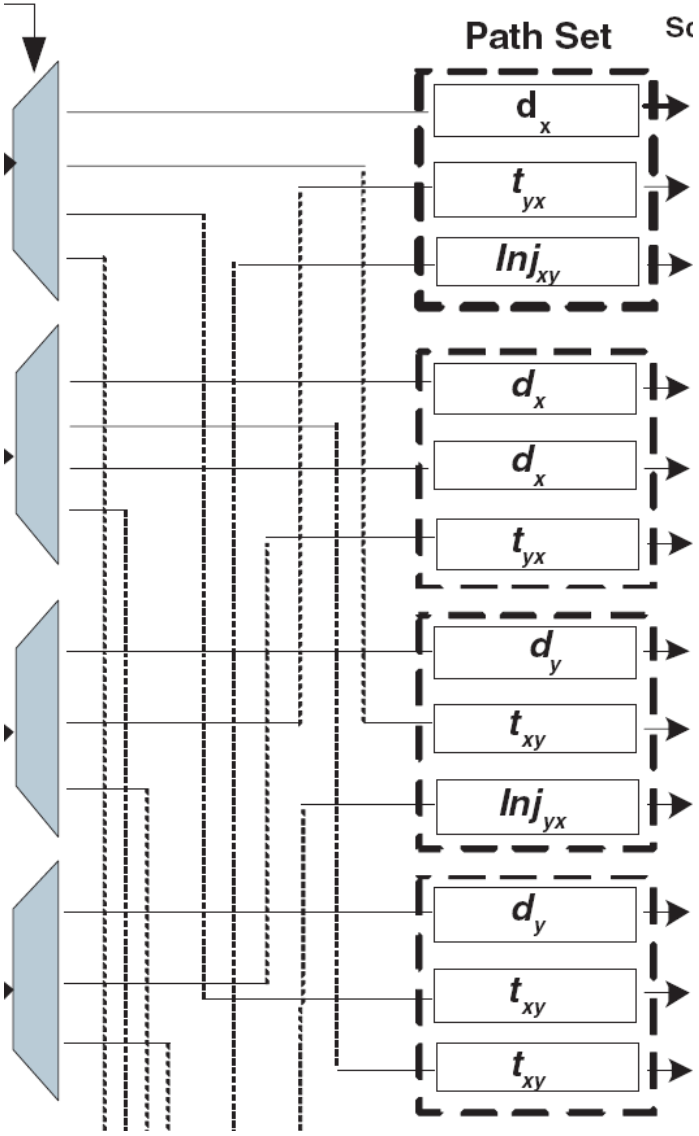


VC Allocation

- XY routing is deadlock-free; need a minimum of 8 VCs to allow every possible flit traversal: 2 d_x, 2 d_y, 2 t_{xy}, 1 Inj_{xy}, 1 Inj_{yx}
- XY-YX routing needs 2 more VCs to enable deadlock freedom
- Adaptive routing needs 12 VCs
- Additional constraints on VCs may lower performance

Input Port	Row-Module		Column-Module	
	Port 1	Port 2	Port 1	Port 2
Adaptive	$d_x\ t_{yx}\ Inj_{xy}$	$d_x\ d_x\ t_{yx}$	$d_y\ t_{xy}\ Inj_{yx}$	$d_y\ t_{xy}\ t_{xy}$
XY-YX	$d_x\ t_{yx}\ Inj_{xy}$	$d_x\ d_x\ t_{yx}$	$d_y\ t_{xy}\ Inj_{yx}$	$d_y\ d_y\ t_{xy}$
XY	$d_x\ d_x\ Inj_{xy}$	$d_x\ d_x\ Inj_{xy}$	$d_y\ t_{xy}\ Inj_{yx}$	$d_y\ d_y\ t_{xy}$

Table 1. VC Buffer Configuration for the Three Routing Algorithms

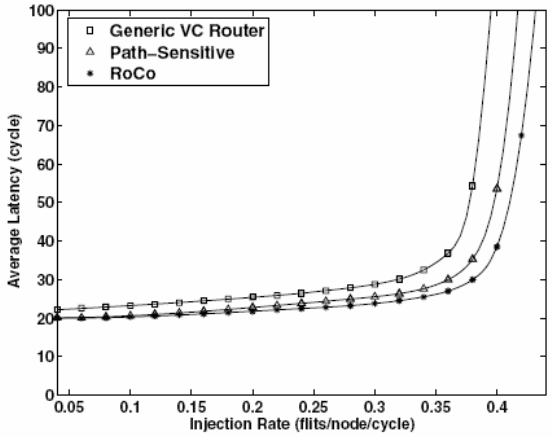


RoCo Router

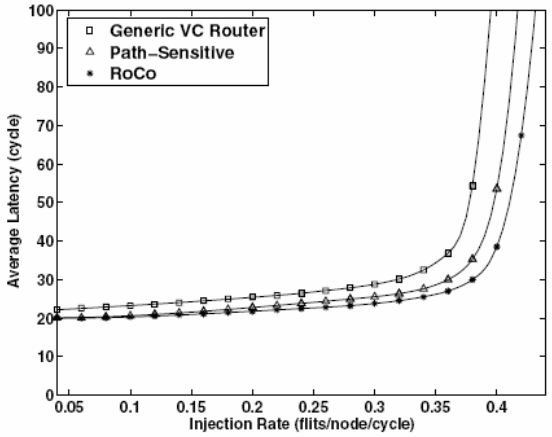
Key features:

- Early ejection mechanism for flits destined for the PE (saves 2 cycles since they don't have to go through SA and xbar stages)
- Flits are steered to the appropriate crossbar thanks to routing info computed in previous stage – enables use of 2 2x2 crossbars instead of 1 5x5 crossbar
- Results show much lower contention probability for RoCo (?!)
- Need fewer and smaller arbiters: 2x2 xbar arbiter algorithm: (mirroring)
  Generic case: for each input port, one arbiter selects the winner
  for each output port, an arbiter selects the winner
  RoCo: for each input port, two arbiters select two winners
  for each 2x2 xbar, one arbiter selects the winner for one port
  and the outcome is mirrored on the 2nd port

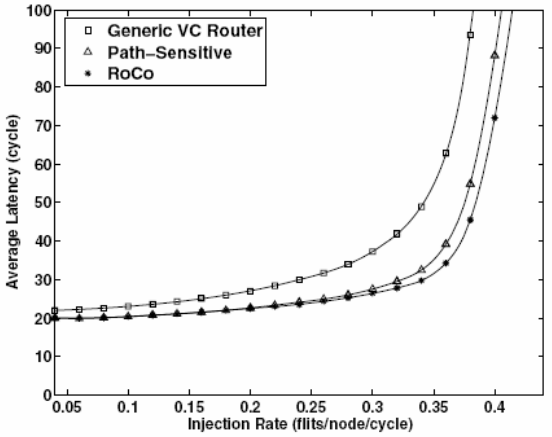
Results



(a) Deterministic Routing



(b) XY-YX Routing



(c) Adaptive Routing

Figure 8. Uniform Random Traffic

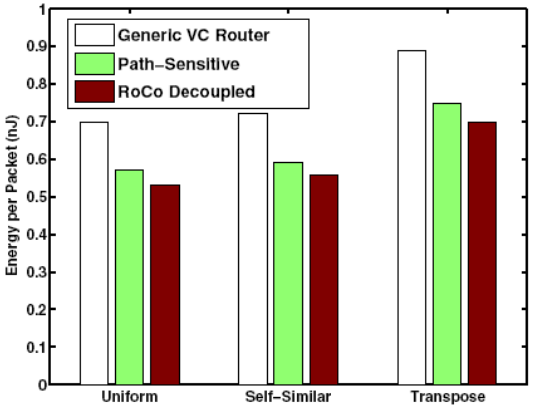
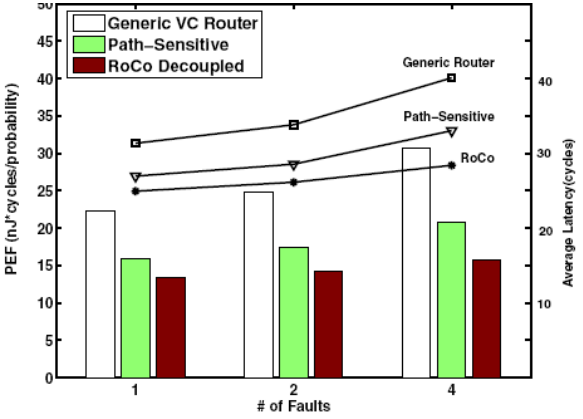
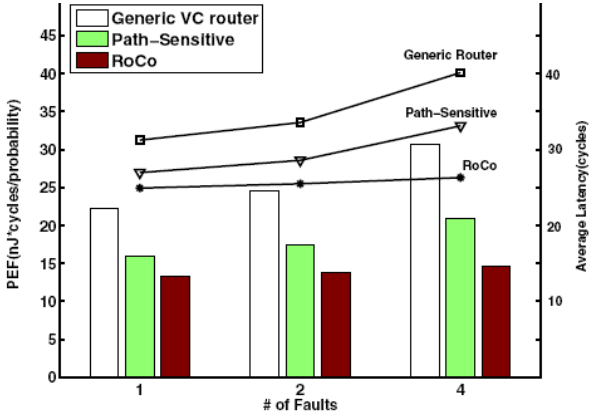


Figure 13. Energy per Packet



(a) Critical Region Fault



(b) Non-Critical Region Fault

Figure 14. Performance-Energy-Fault (PEF) Product

Title

- Bullet