PROM
Path-Based, Randomized, Oblivious, Minimal Routing

Myong Hyon "Brandon" Cho, Mieszko Lis, Keun Sup Shim, Michel Kinsy, and Srinivas Devadas

MIT Computer Science and Artificial Intelligence Laboratory, Cambridge, MA, USA

NoCArc 2009

Outline

- **Motivation and Strategies**
  Robust performance through better load-balancing

- **PROM:**
  Path-Based, Randomized, Oblivious, Minimal Routing

- **Performance Evaluation**

- **Conclusions**

Outline

- **Motivation and Strategies**
  Robust performance through better load-balancing

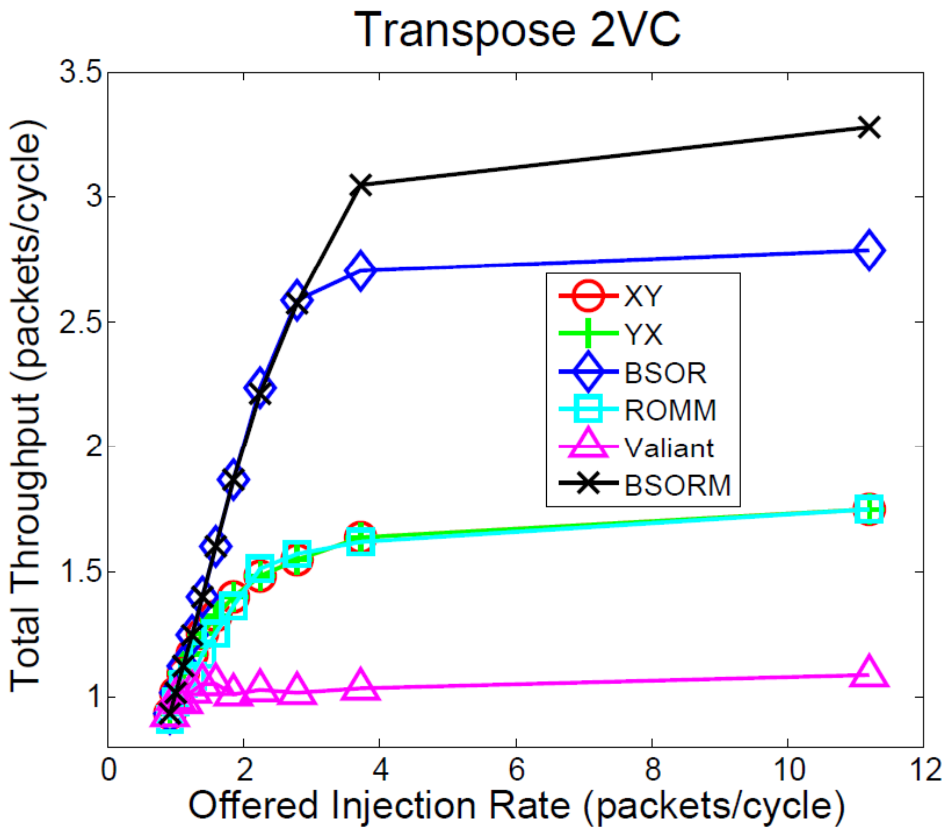
- **PROM:**
  Path-Based, Randomized, Oblivious, Minimal Routing

- **Performance Evaluation**

- **Conclusions**

Routing & Network Throughput

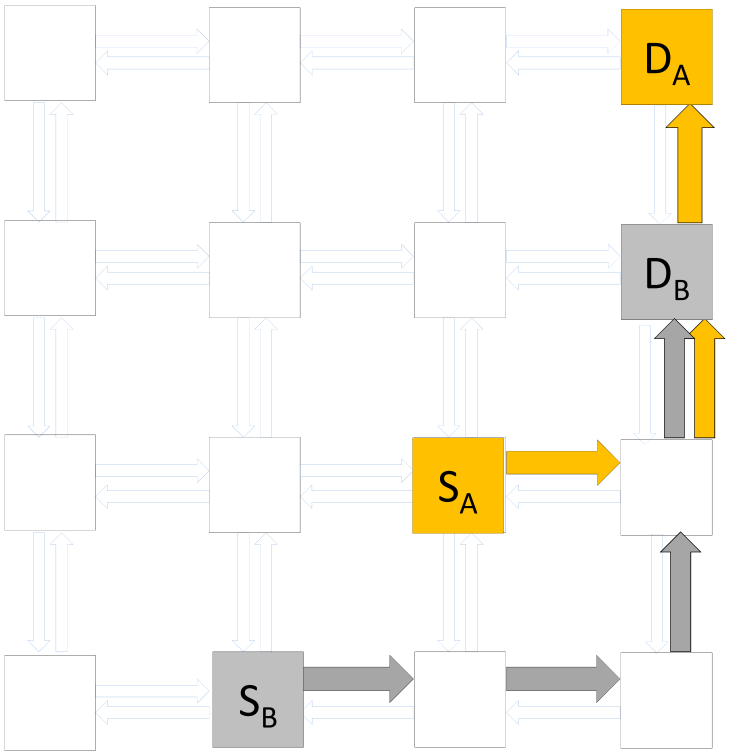
- Routing plays a critical role in network throughput



From “Application-Aware Deadlock-Free Oblivious Routing” [Kinsy et al./ISCA’09]

Dimension-Order Routing (DOR)

- Approaches in one dimension first, then in the other

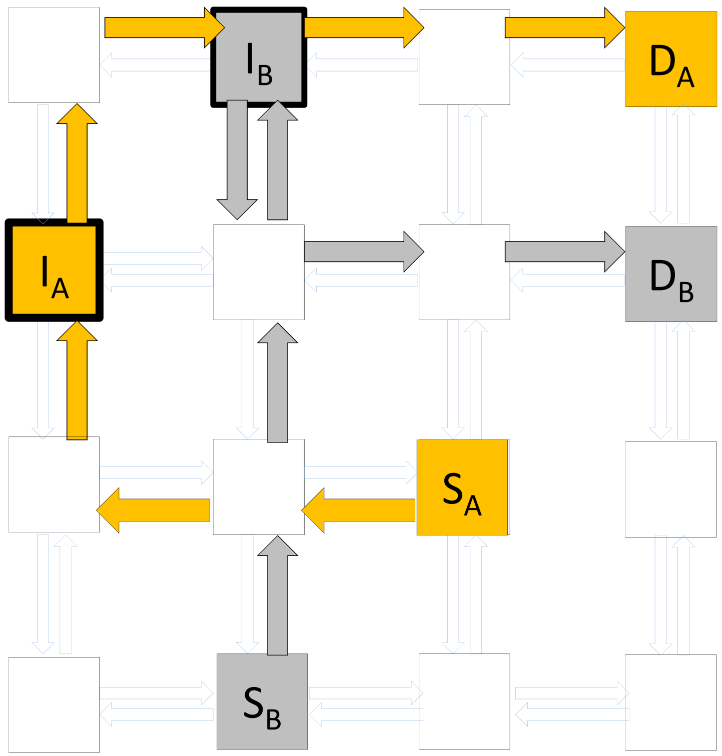


DOR (XY)

- Bandwidth
  No path diversity
- Latency
  Minimal routing
- Deadlock Prevention
  Deadlock-free with 1 VC

Valiant

- Uses one random intermediate node per each packet

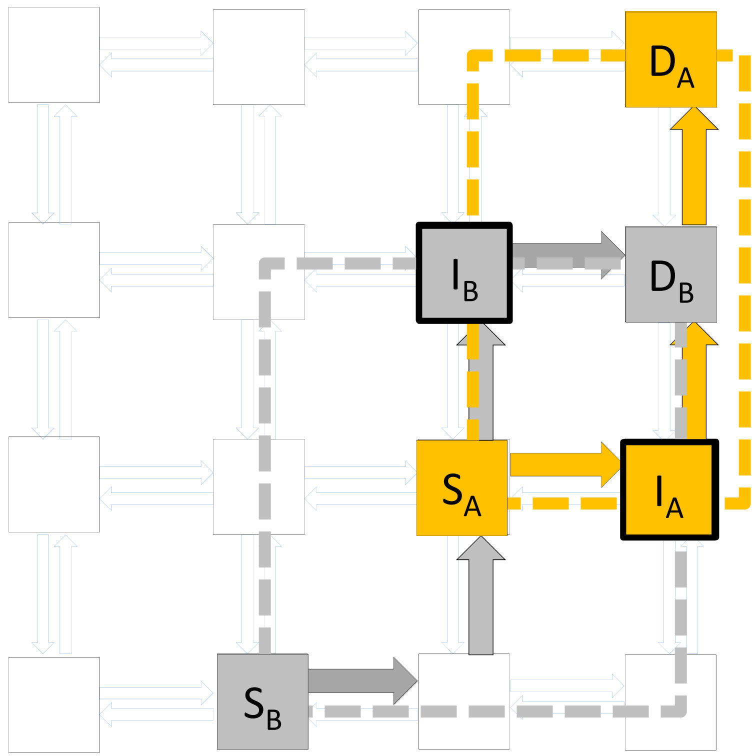


Valiant (XY/YX)

- Bandwidth
  Wide path diversity
  Optimal Worst-case Result
- Latency
  Poor latency
- Deadlock Prevention
  Deadlock-free with >= 2 VCs
  - each phase should use different VCs

n-phase ROMM

- ***n-1*** random intermediate node(s) only in MBR (Minimum Bounding Rectangle)

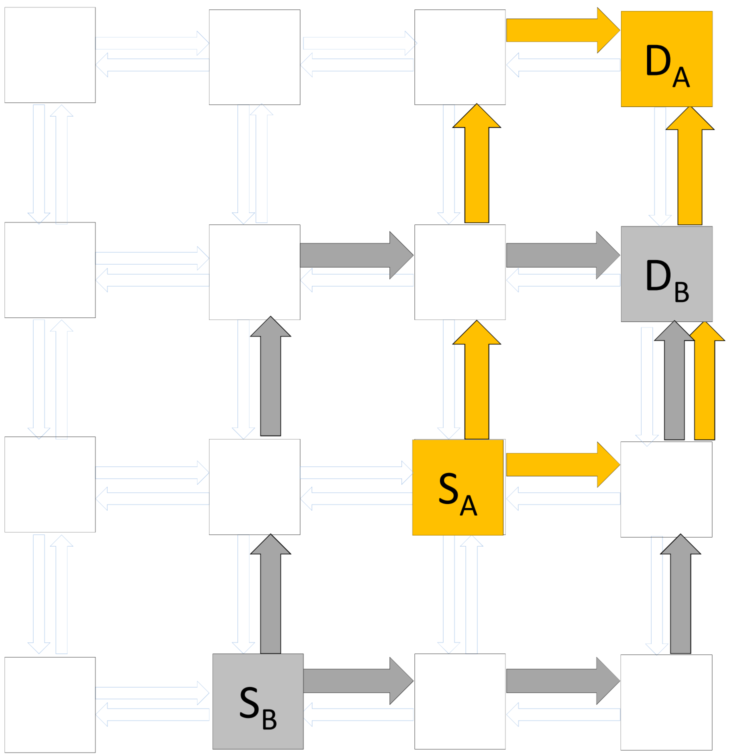


2-phase ROMM (XY/YX)

- Bandwidth
  More path diversity than DOR
  Limited by the value of n
- Latency
  Minimal routing
- Deadlock Prevention
  Deadlock-free with >= n VCs
  - each phase should use different VCs

O1TURN

- Choose either XY-DOR or YX-DOR per each packet



- Bandwidth

  Two paths are possible

  Optimal Worst-case Result

- Latency

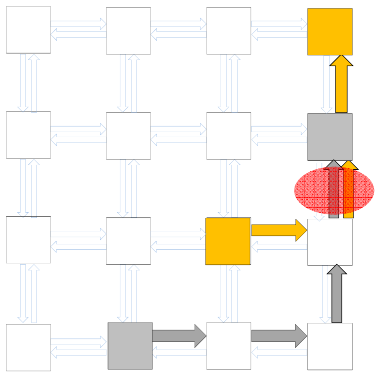
  Minimal routing

- Deadlock Prevention

  Deadlock-free with >= 2 VCs

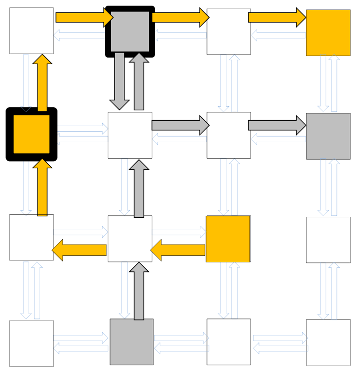
  - each phase should use different VCs

Routing and Performance



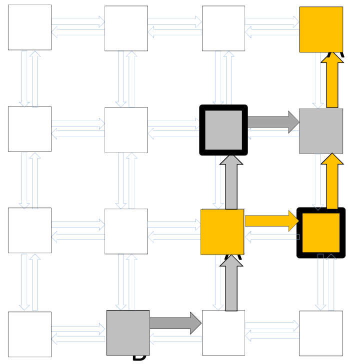
DOR

x0.5



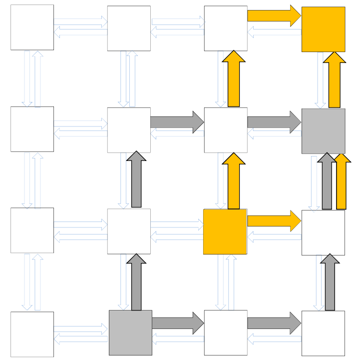
Valiant

x1



ROMM

x1



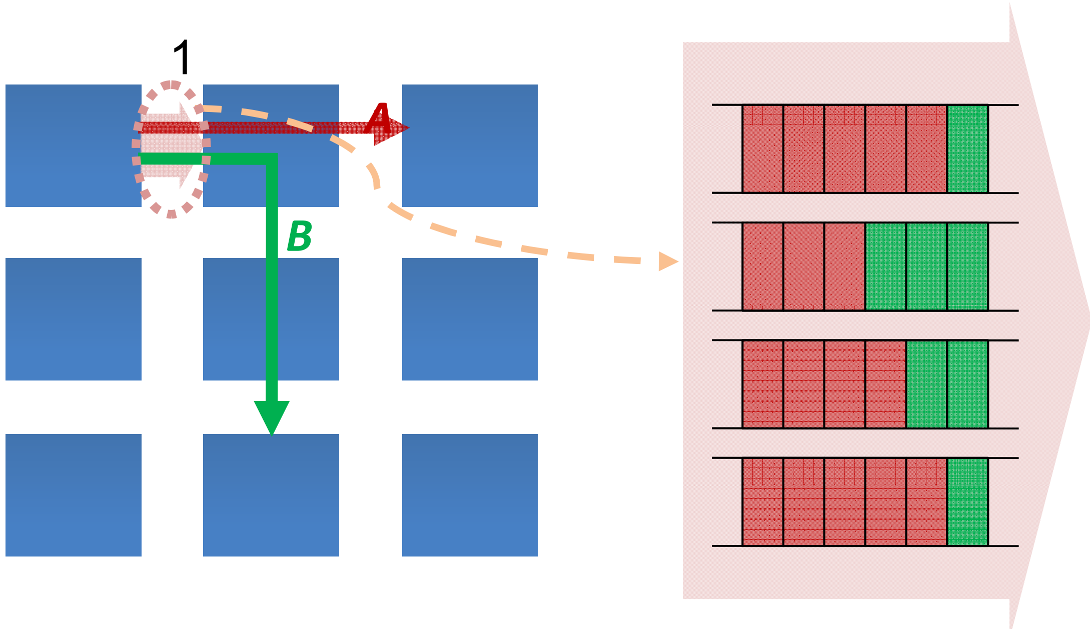
O1TURN

x1

- Depend on traffic patterns
- Optimal “worst-case” result: Valiant, O1TURN
- **In general, path diversity helps lower congestions due to load balancing.**

Congestions & HoL Blocking

- Head-of-Line (HoL) Blocking

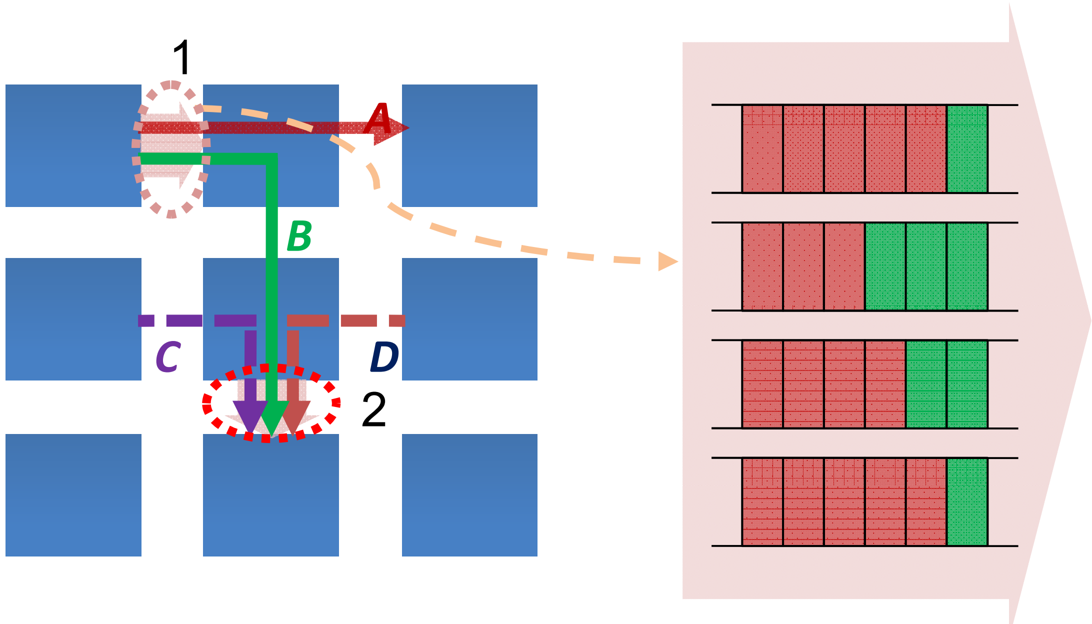


From “Static Virtual Channel Allocation in Oblivious Routing” [Shim et al./NOCS’09]

- Wide, uneven path diversity increases HoL blocking because more flows

Congestions & HoL Blocking

- Head-of-Line (HoL) Blocking



From “Static Virtual Channel Allocation in Oblivious Routing” [Shim et al./NOCS’09]

- Wide, uneven path diversity increases HoL blocking because more flows

Goals & Strategies

- Goal:
  Oblivious routing with robust performance under various traffic patterns

Goals & Strategies

- Goal:
  Oblivious routing with robust performance under various traffic patterns

- Strategies:
  1) A congestion-based approach
  - "Better" traffic distribution with low cost

Goals & Strategies

- Goal:

  Oblivious routing with robust performance under various traffic patterns

- Strategies:

  1) A congestion-based approach

  - “Better” traffic distribution with low cost

  2) For the HoL blocking effect,

  - EVEN traffic load distribution

  - Control path diversity and latency within the MBR

  - Maximize the benefit from less congestions

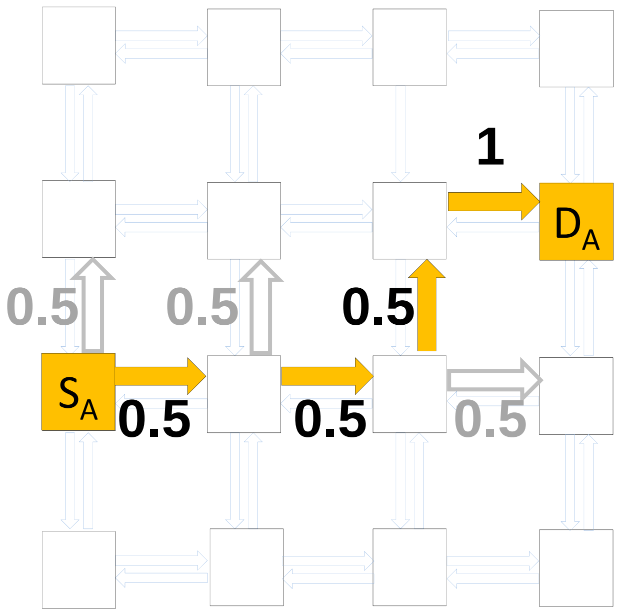
  - Adopt other methods (not related to routing) to reduce HoL

Outline

- Motivation and Strategies
- **PROM:**
  Path-Based, Randomized, Oblivious, Minimal Routing
  - PROM framework
  - PROM variants
  - Deadlock Prevention
- Performance Evaluation
- Conclusions

PROM Framework

- **Minimal Routing**: bounded latency and HoL blocking
- If there are multiple next hops possible,
  Each node chooses one based on a given probability.

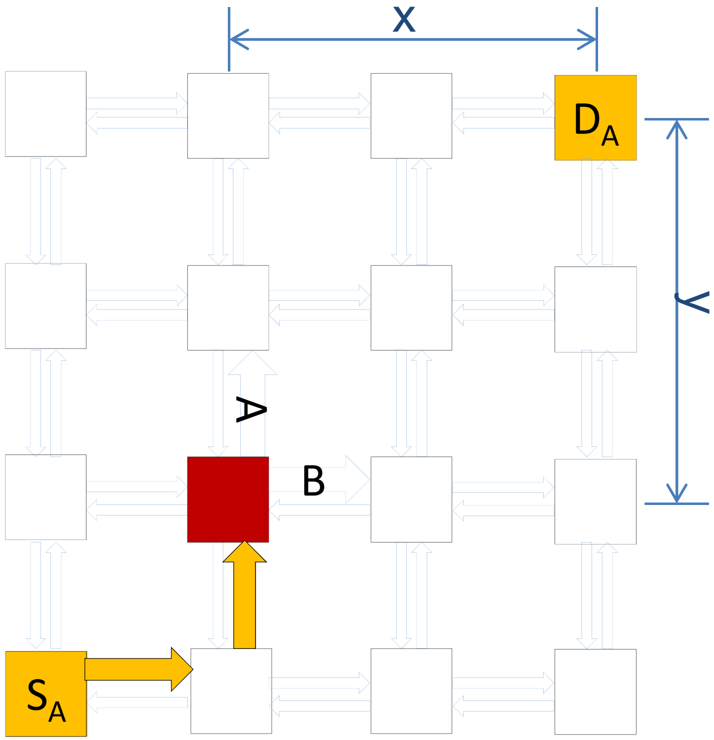


This path has a 12.5% chance of being taken.

- *How these probabilities are set determines the specific instantiation of PROM.*

Uniform PROM

- Each possible minimal route has an **equal** chance of being taken.

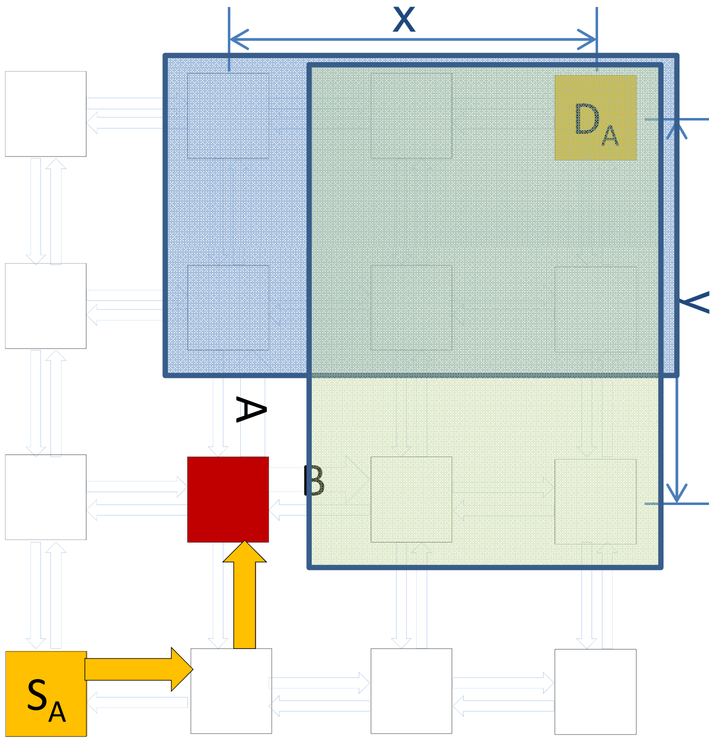


P_A = ?

P_B = ?

Uniform PROM

- Distance to the destination is used to set the probabilities



- A is followed by
[x+(y-1)]!/[x!(y-1)!] minimal routes
- B is followed by
[(x-1)+y]!/[(x-1)!y!] minimal routes

$P_A = y/(x+y)$
$P_B = x/(x+y)$

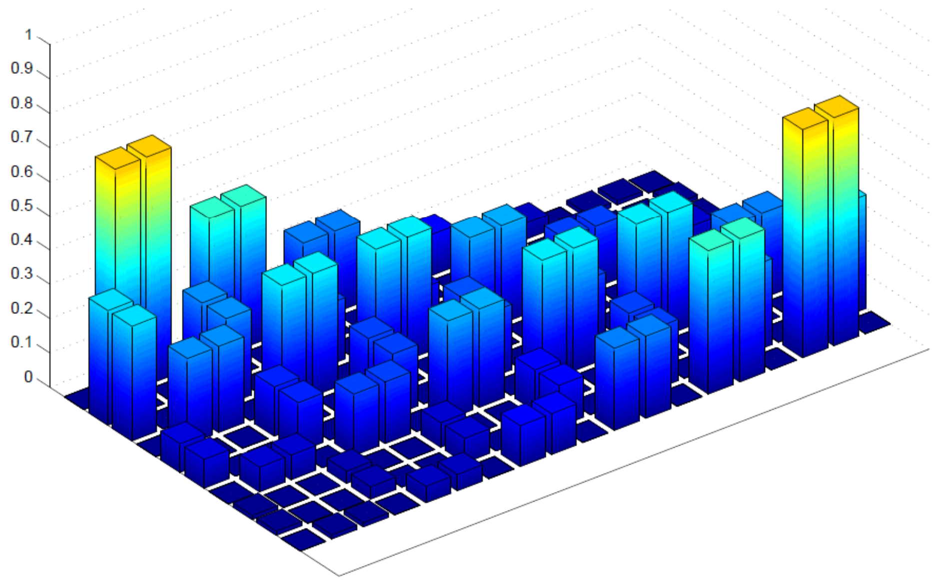
Uniform PROM vs. ROMM for N-by-N mesh

	Uniform PROM	2-phase ROMM	(2N-1)-phase ROMM
Path Diversity	**Max.**	**Limited**	**Max.**
Even * Load Balancing	**Yes**	**No**	**No**
Hardware Cost	**Small** (≥ 2VCs)**	**Small** (≥ 2VCs)	**Large** (≥ 2(N-1) VCs)
Communication Overead	**None**	**Small**	**Large**

* in terms of paths taken
** illustrated in later slides

Parameterized PROM

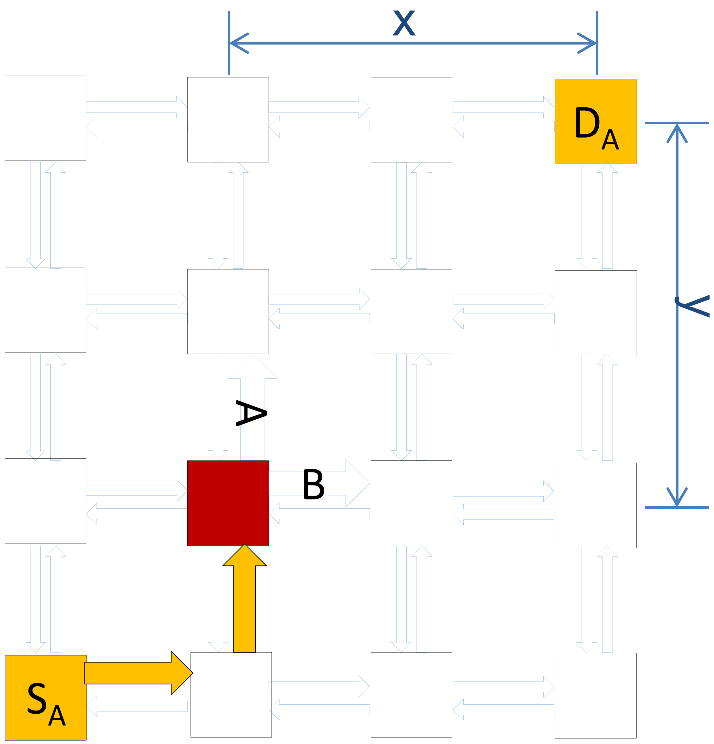
- Uniform PROM:
  Great load-balancing property for **large** mesh networks
- May not fit **small** mesh networks.
  - Links in the middle are more congested.



- Added a parameter to adjust traffic distribution.

Parameterized PROM

- Previously,

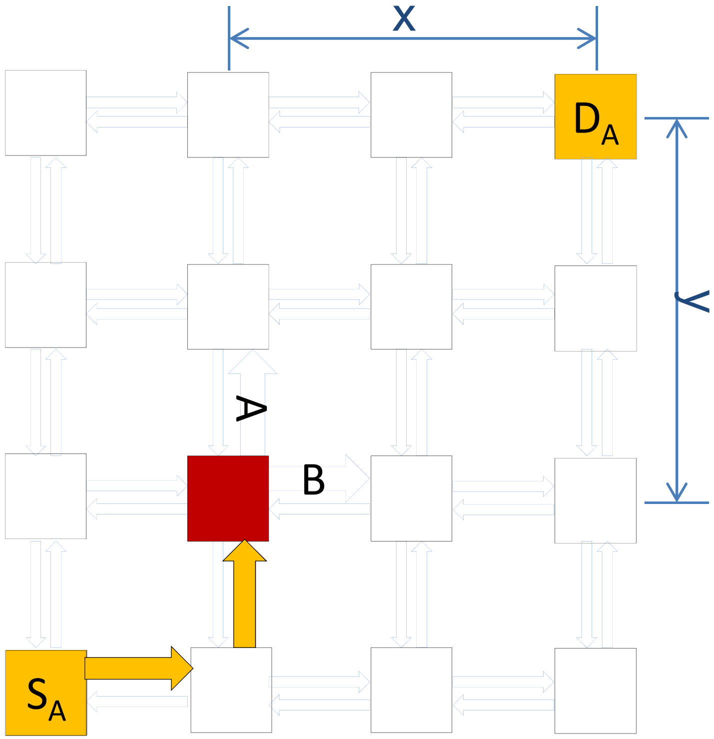


$P_A = y/(x+y)$

$P_B = x/(x+y)$

Parameterized PROM

- An additional parameter *f*:



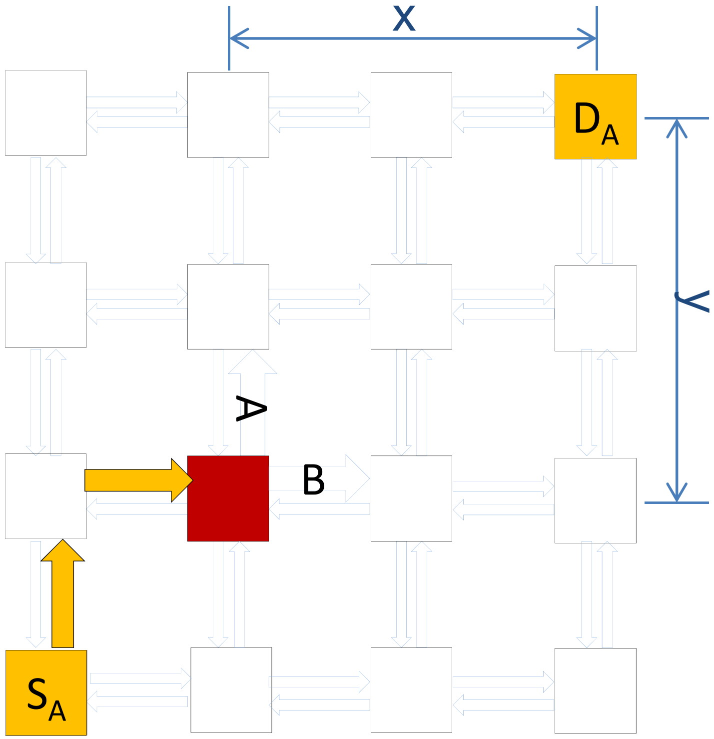
$P_A = (y\mathbf{+f})/(x+y\mathbf{+f})$

$P_B = x/(x+y\mathbf{+f})$

- If *f* is large, A is preferred (not making a turn).

Parameterized PROM

- Probabilities depends on **ingress direction**

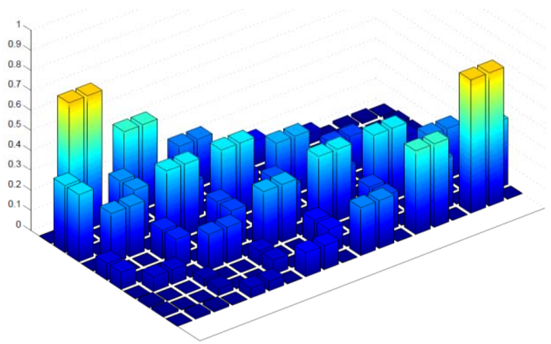


$P_A = y/(x+y+\mathbf{f})$

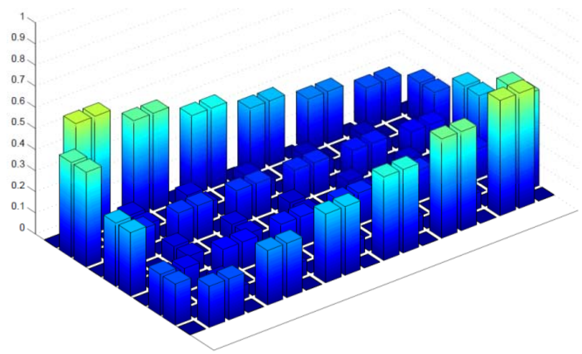
$P_B = (x+\mathbf{f})/(x+y+\mathbf{f})$

- When *f* is large, there is preference “not to make a turn”

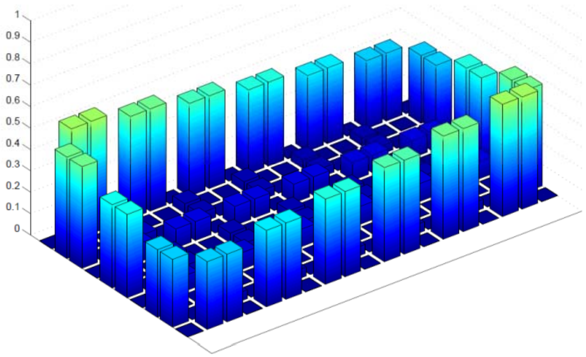
Parameterized PROM



f=0



f=10

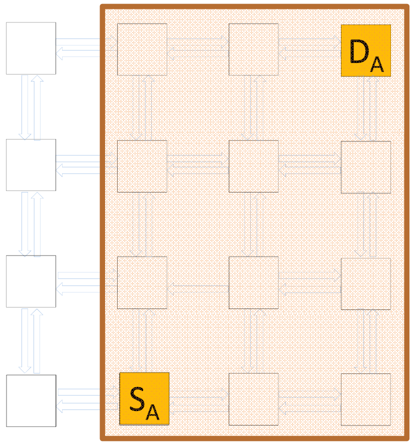


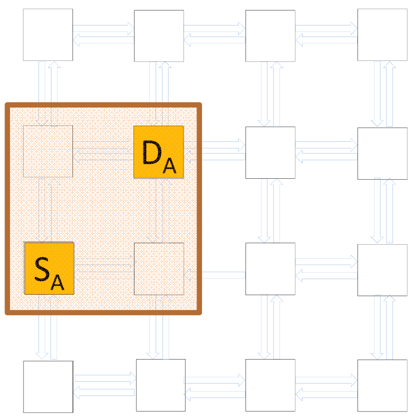
f=25

- Increasing f pushes traffic from the diagonal of the MBR towards the edges

Variable Parameterized PROM (PROMV)

- If the MBR of a network flow is large:
  - Outer edges are more likely close to the network edges
  - It's better to push traffic toward the edges (larger *f*)



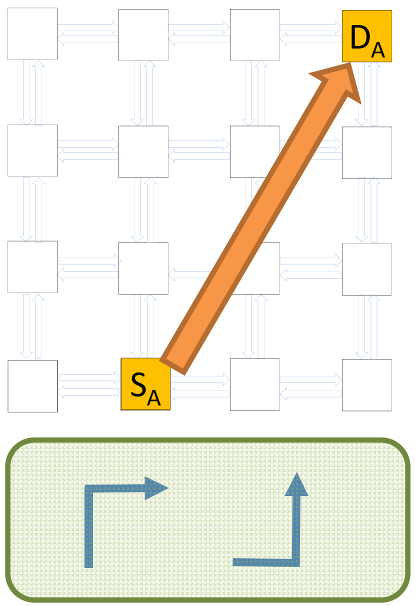


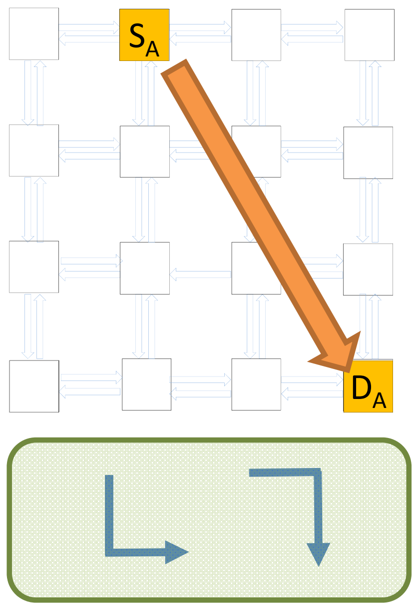
- For each network flow, f is determined by:

$$f = f_{max} \cdot (xy/N^2)$$

Deadlock Prevention

- In PROM, deadlock can be simply avoided by:
  - Network must have **two** disjoint sets of VCs.
  - If the destination is **left** to the source, use the **first** set of VCs.
  - If the destination is **right** to the source, use the **second** set of VCs





…conforms to North-Last turn model.

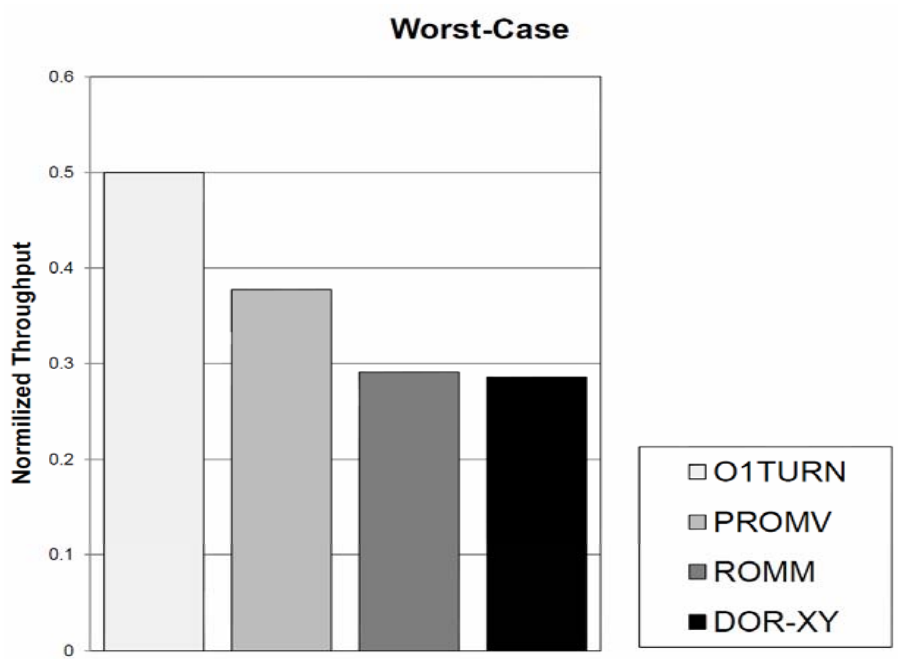
Outline

- Motivation and Strategies
- PROM:
  Path-Based, Randomized, Oblivious, Minimal Routing
- **Performance Evaluation**
- Conclusions

Ideal Throughput

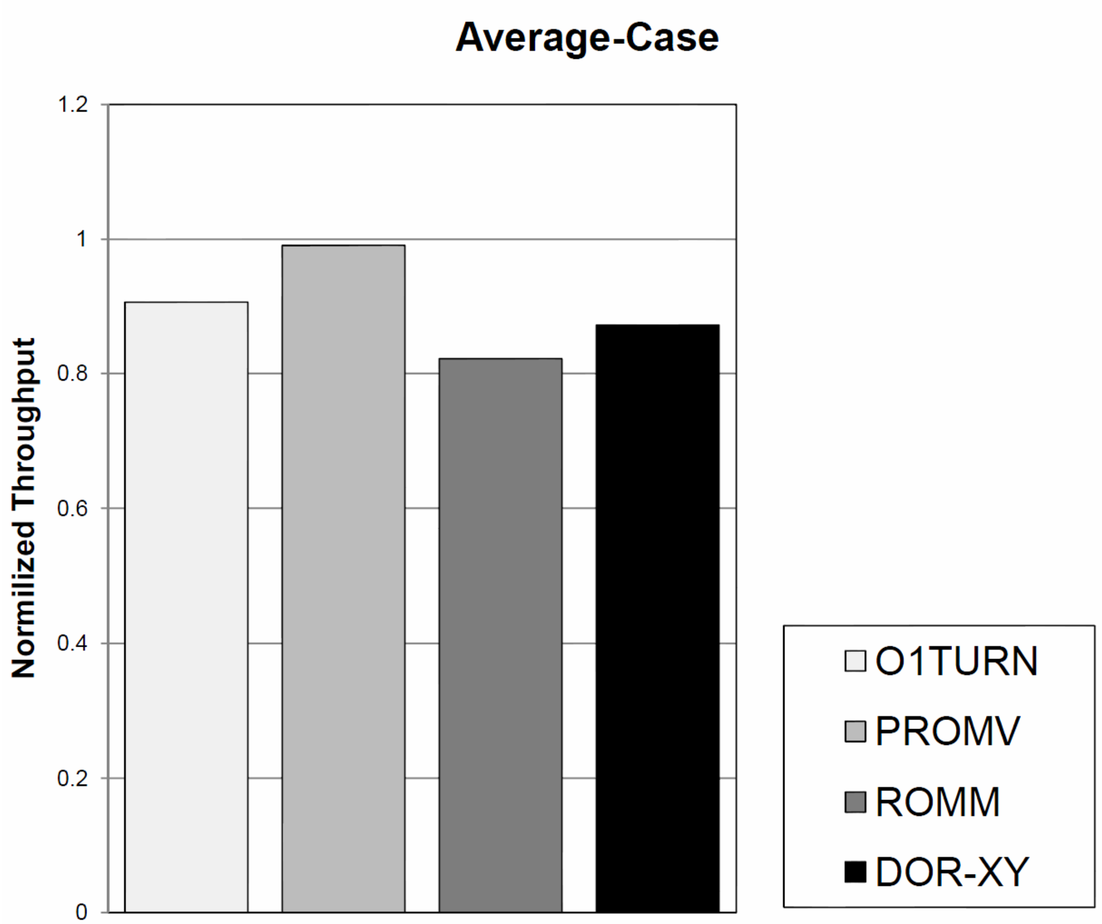
- Ideal Performance assuming:
  - Fair scheduling
  - No HoL blocking issue

- Worst-case Performance

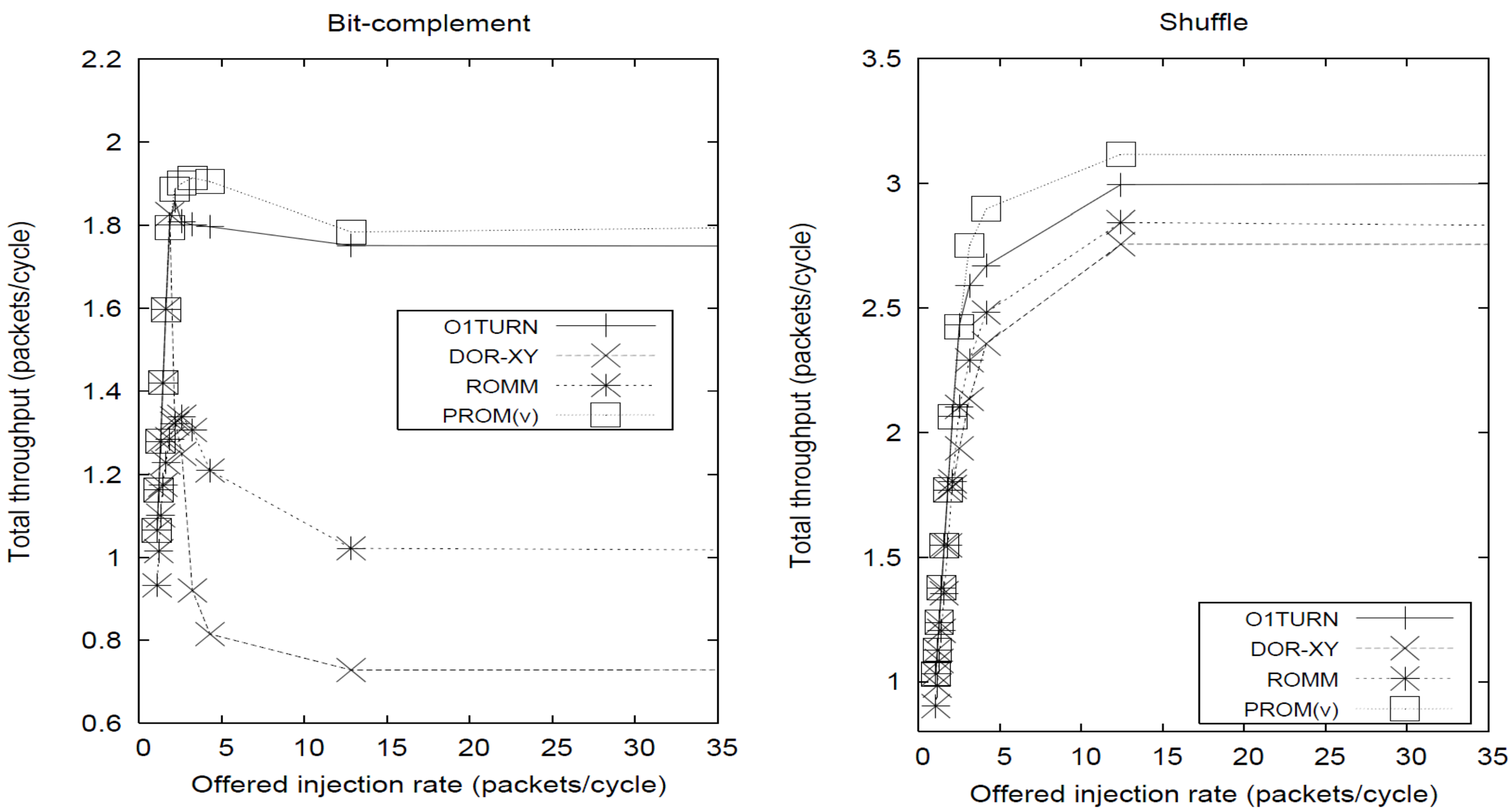


Ideal Throughput

- Average-case Performance (10K random traffic patterns)

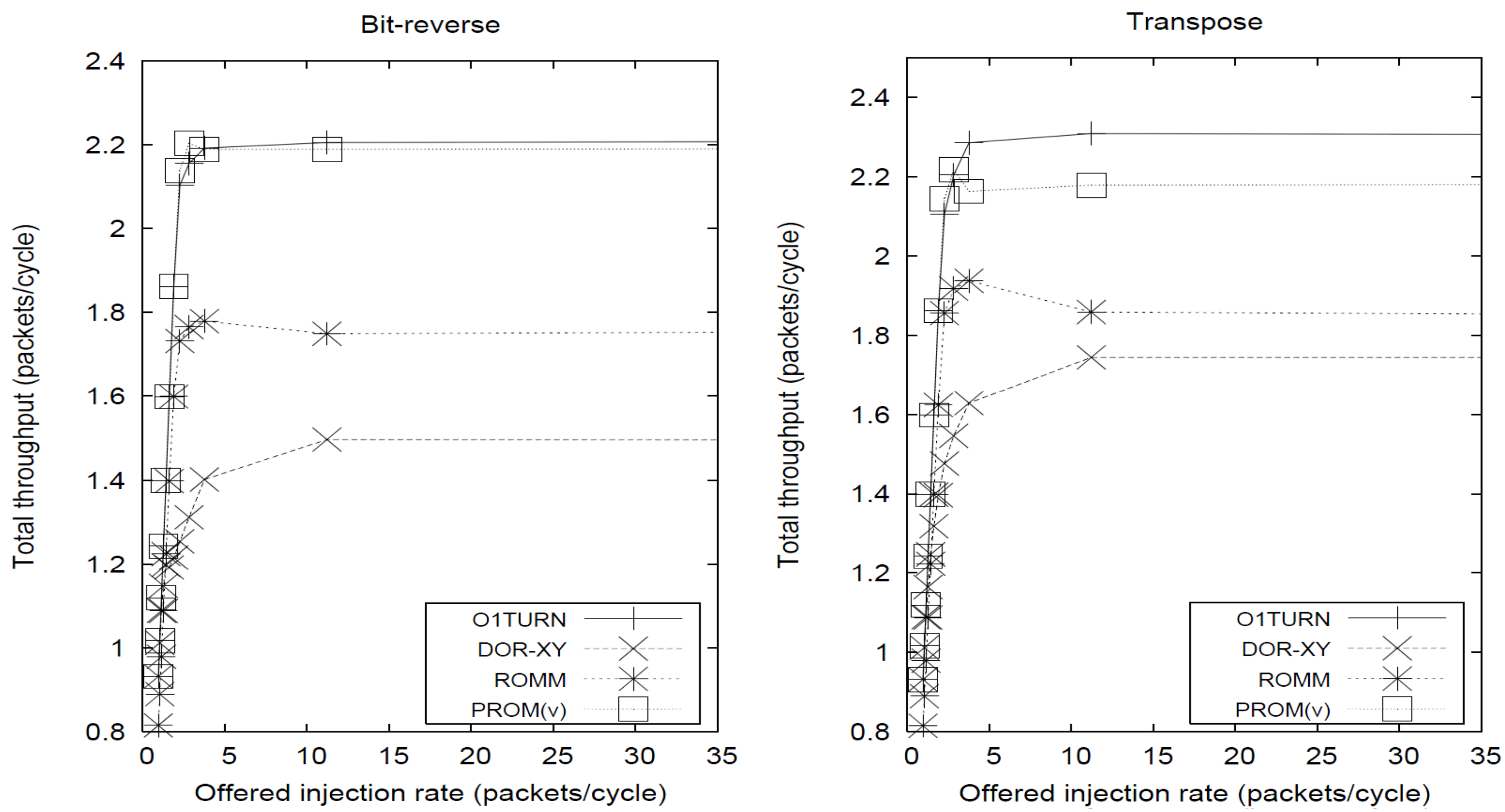


Performance with Dynamic VC Allocation



- Ideal throughput of O1TURN and PROMV are similar.
- In spite of HoL blocking, PROMV and O1TURN are equivalent.

Performance with Dynamic VC Allocation



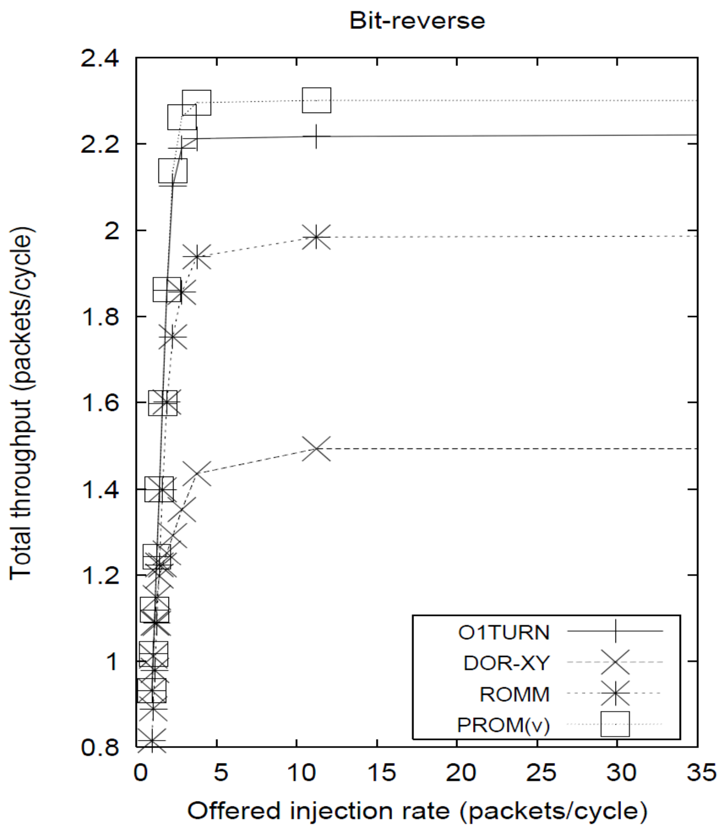
- Ideal throughput of PROMV is better.
- The amount of HoL blocking depends on applications.

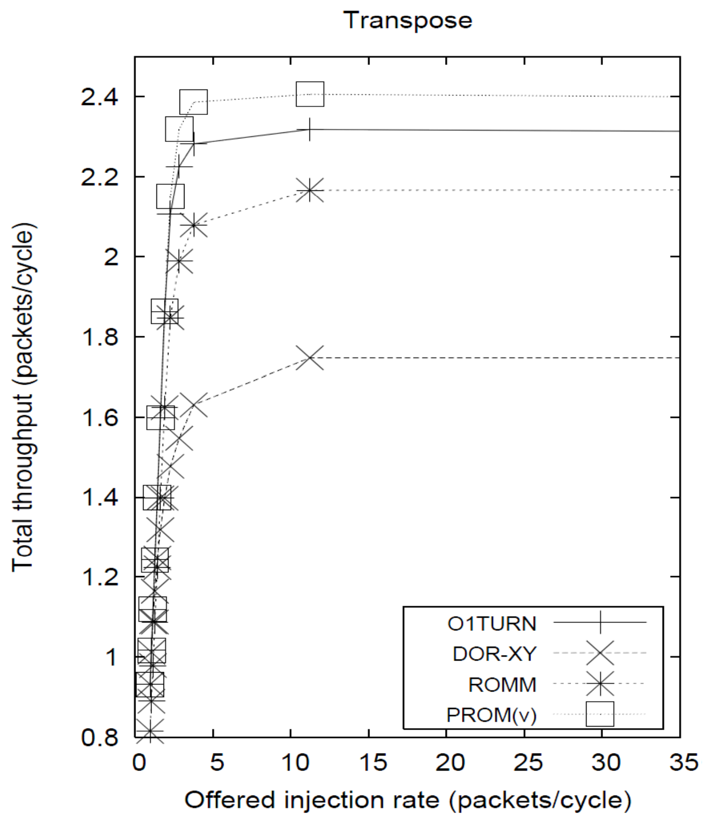
Exclusive Dynamic VC Allocation

- Exclusive Dynamic VC Allocation(EDVCA) mitigates the HoL blocking issue [Lis et al./CSAIL Tech Report'09]:
  - In each cycle, a flow is allowed to hold only one VC at one node.

- HoL blocking is not completely eliminated, but significantly reduced.

Performance with EDVCA





- PROMV's better ideal throughput results in better actual throughput.

Limitations & Future Works

- Need to be evaluated with various ‘real-world’ applications:
  - FPGA implementation of a many-core system

- Need more analytic study on HoL blocking
  - Hard to find a good metric on HoL blocking for given routes
  - Important to design a routing scheme that can *actually* perform better

- PRAM: Path-Based, Randomized, ***Adaptive***, Minimal Routing
  - A good platform for an adaptive routing scheme

Conclusions

- Through better load balancing,
  PROM has robust performance under various loads.

  - Higher ideal throughput with various traffic patterns.

  - It results in better throughput with EDVCA.