TOWARDS THE EXPLORATION OF DYNAMIC MULTIPATH ROUTING IN 3D TORUS NETWORKS THROUGH THE CODES/ROSS SIMULATION FRAMEWORK

DANIEL PARKER, SANJIV KAPOOR, IOAN RAICU

Abstract

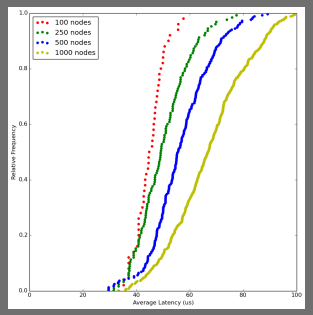
Torus Networks are an important and widely used architecture in modern supercomputers. These supercomputers frequently employ static routing protocols like Dimension Order Routing (DOR). While simple and easy to implement, we theorize that DOR methods incur excessive network latencies due to poor load balancing. In this work, we first characterize such behavior under the current standard protocol, then introduce a simple Randomized Dynamic Routing (RDR) protocol. We demonstrate that even such simplistic dynamic routing protocols offer lower, more scalable latencies over a 3D torus network. As supercomputing moves towards exascale, such scalability will be necessary to cope with unprecedented workloads and network sizes. These results were generated using the CODES highly parallel simulator.

Routing Procedure

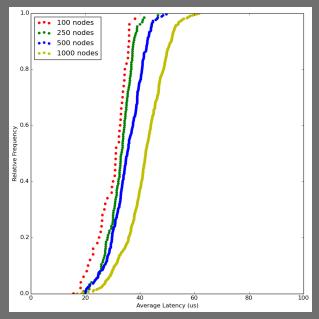
- If at destination: DONE
- Else, for each dimension, determine whether it would be advantageous to travel in the positive direction, negative direction, or not move, using the half length of the network. If travel over that dimension is viable, mark that dimension VIABLE.
- Randomly select a dimension marked VIABLE, travel one node in the appropriate direction, then call RDR again.

Latencies under Different Network Loads

Dimension Order Routing

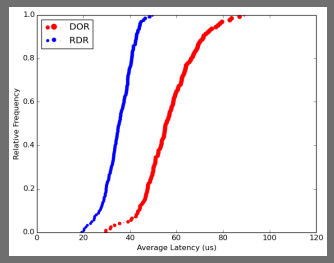


Randomized Dynamic Routing

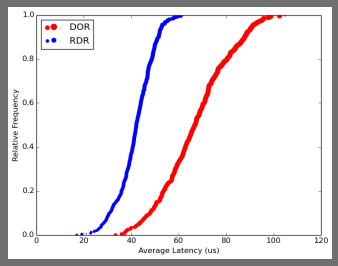


Compared Latencies

500 Active Nodes



1000 Active Nodes



Percent Latency Decrease	Median	95th Percentile
100 Nodes	30.3%	33.6%
250 Nodes	32.5%	41.7%
500 Nodes	36.2%	43.0%
1000 Nodes	36.9%	41.9%

Challenges

- Provide performance increase without excessive overhead
- Routing protocol must scale to large networks and heavy traffic

Proposed Solution

- Use randomization to spread traffic more evenly across the network
- Sufficiently simple to avoid excess overhead and to scale well

Conclusions

- Dimension Order Routing incurs significant latency penalties as network traffic increases
- Our Randomized Dynamic Routing protocol offers significant speedup over Dimension Order Routing and scales better to high traffic
- Dynamic multipath routing has potential to improve supercomputer performance

References

[1] Cope, Jason, Ning Liu, Sam Lang, Phil Carns, Chris Carothers, and Robert Ross. "Codes: Enabling co-design of multilayer exascale storage architectures." In *Proceedings of the Workshop on Emerging Supercomputing Technologies*, pp. 303-312. 2011.

[2] Devetak, F. et al. 2011. Minimizing Path Delay in Multipath Networks. *2011 IEEE International Conference on Communications (ICC)*. (2011).

[3] Junghwan Shin, et al. 2013. Concurrent multipath routing over bounded paths: Minimizing delay variance. *2013 IEEE Global Communications Conference (GLOBECOM)*. (2013).

[4] Myong Hyon Cho, Mieszko Lis, Keun Sup Shim, Michel Kinsy, and Srinivas Devadas. 2009. Path-based, randomized, oblivious, minimal routing. In *Proceedings of the 2nd International Workshop on Network on Chip Architectures* (NoCArc '09). ACM, New York, NY, USA, 23-28.

Daniel Parker: dkparker@uchicago.edu
Computer Science Departments, University of Chicago/Illinois Institute of Technology
Sanjiv Kapoor: Kapoor@iit.edu,
Computer Science Department, Illinois Institute of Technology
Ioan Raicu: iraicu@cs.iit.edu,
Computer Science Department, Illinois Institute of Technology; MCS Division, Argonne National Laboratory

ILLINOIS INSTITUTE OF TECHNOLOGY