Processor Shapes and Topologies for Dense On-Chip Networks

Zhibin Xiao and Bevan M. Baas

Department of Electrical and Computer Engineering, University of California, Davis

1. Overview

- Many-core processor topology: The available inter-processor interconnect affects the total application throughput, energy (power), chip area, and latency
- Key motivation: Commonly-used 2D mesh many-core processor arrays can have local congestion and long latencies for global traffic
- Research target: Topologies that increase application performance, reduce communication energy, avoid global wires, and ease physical implementation

2. Proposed Topologies and Shapes

- Key idea: Novel interconnection topologies and processor tile shapes to increase the number of local inter-processor connections
- Common 2D mesh and the seven proposed topologies
  - Dense On-Chip Networks (NoCs) without long global wires
  - Two 5-neighbour, three 6-neighbour and two 8-neighbour topologies
  - Hexagonal-shaped and “House-shaped” processor tiles

3. Communication Latencies

- Worst-case communication distance for four basic communication patterns
  - One-to-one, one-to-all, all-to-all and all-to-one

4. Application Mapping

- Two complete real-time complex applications
  - H.264/H.265 residual encoder and 802.11a Wi-Fi TM receiver
  - 6-neighbour topologies vs. 4-neighbour 2D mesh
  - 21% lower application area
  - 19% shorter total communication link length

5. Chip Implementation Results

- Non-2D mesh designs have:
  - Larger area per tile (+3.3% to +5.9%) partially caused by higher fan-in and fan-out I/O port logic
  - Higher energy per operation (+3.7% to +8.6%)
  - Non-rectangular tiles have lower clock skew (−40% to −54%)
- Results would improve with optimization of items such as SRAM and tile I/O port placements

Appendix A. Planned Research Project: Further analysis of the communication latencies for a real-time application. Support from the National Science Foundation (NSF) under Grant No. ECCS-1253892 is gratefully acknowledged. Dr. Zhibin Xiao is also supported in part by NSF Career Grant 0190681. The authors gratefully acknowledge the support of the NSF-wide Center for Nanophasics Modeling and Analysis (CNMa) under Grant NO. 0520185.