Low Latency Explicit Communication and Synchronization in Scalable Multi-core Clusters

C. Kachris, G. Nikiforos, V. Papaefstathiou, X. Yang, S. Kavadias, M. Katevenis

Features:

- Direct Access Address Space Extension:
  - SW Allocates part of private cache as scratchpads
- Supports both:
  - Implicit Communication using caches,
  - Explicit Communication via directly accessible local memories (scratchpad)
- Merged Cache controller (CC) and Network Interface for reduced latency

Explicit Communication Mechanisms:

- Remote Stores
  - Fast non-blocking direct transfers
  - Can use write combining to improve data comm. efficiency
  - Store fences may be required in SW
- R-DMA
  - Bulk asynchronous transfers
  - Completion notifications: combination of Acks and Counter

NI Communication & Synchronization primitives:

- Command Buffers
  - Counters: Atomic add-on-store
    - Event response, send notification when counter=0
- Queues: Atomic enqueue and dequeue
  - Single Reader: Many-to-One communication
  - Multiple Reader: Many-to-Many communication
- Locks (using queues)

Cluster Nodes:

- Four MicroBlaze processors (32-bit, in-order, 5-stage pipeline)
- Configurable L2 cache/scratchpad
- Merged NI and CC
- Communication using a 64-bit, 5-port bufferless crossbar switch

FPGA Prototype:

- Multi-FPGA Board (16 nodes based on Xilinx Virtex5 FPGA, Crossbar switch based on Xilinx Virtex II)
- Each node consists of 4 MicroBlaze processors and one DRAM (256 MBytes)

Performance Evaluation:

- Performance evaluation using 4 Multi-core clusters
- Configurations:
  - Each cluster connected to the L2 switch using 2 SATA cables at 3.125Gbps
  - The L2 switch is configured as 4x4 store-and-forward switch with 2 planes
  - Execution time to process (divide and process) 64 MBytes of data:
    - Distribution of the data using Remote DMA (RDMA) with explicit acknowledgments (event responses)

Computer Architecture and VLSI Systems Laboratory (CARV), Institute of Computer Science, FORTH