Networks on Chip

ET4351: VLSI System-on-Chip
Computing Fabric for High Performance Applications (COBRA)

Develop a high-performance, dependable many-core array processor architecture

Incorporate a low complexity programming model to reduce programming effort

Leverage the potential of 3D integration to achieve higher performance than planar many-cores

Goal: 1 TOPs by 2013
Applications

Pedestrian Tracking
(Volvo 2010)

Protein Folding
(Justin MacCallum, Stony Brook University)

Data Search and Analysis
(21CT)

Circuit Routing
(Advanced Processor Technologies Group, University of Manchester)

3D CT Scans
(Northstar Imaging)
Naga High-Performance Array Processor

- Clustered array processor with two sub-arrays:
  - Speculative execution based shared memory array for general purpose processing
  - Message passing based wide-issue accelerator for dataflow applications (NagaM)

- Speculative execution using Hardware Transactional Memory for compute applications
  - Parallel tasks execute speculatively, regardless of their data dependencies
  - Data sets are isolated until task execution is determined to be legal
  - Data sets are committed to memory once determined as legitimate
Naga High-Performance Array Processor

- Dataflow applications use a simple but powerful message passing communications protocol
  - Low overhead inter-task communication
  - Address resolution, buffer management performed in hardware
Interconnect

• **Buses**
  + *Simple* and effective for a small number of nodes
  - Scale poorly
  - Requires *access arbitration*, only one node may actively comm.

• **Crossbars**
  + Scale better when compared to buses
  + *Non-blocking*
  - Tend to become huge with increasing number of nodes
  - Timing closure problems arise as well

• **Network-on-Chip**
  + *Highly regularized* interconnect structure
  + Scales very well
  + *Robust* communication
  - Routers occupy some area
  - Communication latencies are larger

Kgil Picoserver, 2006

Theocharides 2005
Network on Chip (NOC)

- Structured network of switches
- Regularized structure means wiring between routers is of fixed length = easier timing closure
- Several topologies – mesh, ring, torus, fat-tree,…
- Data is transferred between network nodes in the form of packets. Packets consist of flits, phits.
- Path diversity between nodes = less congestion and lower packet latency
- Nodes connected to network through Network Interface
Scalability

- NoCs are easy to scale up – just instantiate more routers

- **BUT!** – Is performance adversely affected by scale up?

- With increase in network size, average number of hops increases for a 2D network = higher communication latency!

- As number of nodes increases, the system grows in size = larger area = larger silicon!

- How do you scale without losing performance and without increasing area?
Scalability – 3D

• **Solution:** Go vertical!

• Die stacking + Through Silicon Vias (TSV) = 3D Network-on-Chip

• Area footprint = area of a single die

• **40% performance improvement** over 2D network of same size (Pavlidis and Friedman 2006)

• Each die contains a regular NOC. But, routers have two additional ports, *Up* and *Down*

• TSVs are shorter than lateral wires, have better electrical properties = may be clocked higher
Naga’s Interconnect Architecture

- Wormhole routed, Input Buffered

- Round Robin arbitration at each output port
  - 3 cycle fall through latency
  - Each of the $N$ arbiters poll $N-1$ input ports

- Static Z-X-Y Dimension ordered routing
  - Routes flits to destination tier first, i.e. tier routing resources are conserved for flits in their destination layer
  - In-order traversal and delivery of flits

- On/Off Flow control
Sizing Input Buffers

[Graph showing the relationship between buffer depth and average end-to-end packet latency, as well as average throughput.]
Network Interface

- Two layered:
  - Packetizing/Depacketizing Layer
  - Transfer Layer

- Address Translation converts between tile IDs/Memory addresses and network addresses

- Packets are injected into the network only once the tail is framed
  - Tail flits are marked when packet payload reaches 64B/burst ends
  - Remaining data forms part of the next packet

- Input/Output FIFOs depth = 17 flits (1 complete packet)
Perspectives

• Performance
  – What is the average network latency?
  – What are the effects of using a best-effort network? What kind of latency variations can we expect?
  – How are the vertical links utilized?

• Design
  – How do you reduce design time for 3D architectures?
  – Is there more than one way to place TSVs?

• Thermal
  – What are the implications of TSV topologies on thermal performance of the stack?
Performance – Latency and Jitter

<table>
<thead>
<tr>
<th>Proposal</th>
<th>Latency (cycles/word)</th>
<th>Burst size (words)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ocMPI</td>
<td>32.9</td>
<td>256</td>
</tr>
<tr>
<td>Francesco 2005 – Sh. Mem</td>
<td>25</td>
<td>64</td>
</tr>
<tr>
<td>Francesco 2005 – DMA</td>
<td>9.37</td>
<td>64</td>
</tr>
<tr>
<td><strong>Naga</strong></td>
<td><strong>6.58</strong></td>
<td><strong>64</strong></td>
</tr>
</tbody>
</table>

27 May 2013
Performance – Vertical Link Utilization
Design

• Through Silicon Vias (TSV) complicate the physical design process.
  – Are large (4-8 \( \mu \)m diameter), and require a keep-out zone of upto 10 \( \mu \)m due to mechanical stress margins
  – Keep-out zones vary according to topology of placement

• Optimal architecture from early simulations is therefore NOT always feasible in hardware

• ASIC design flows are time-consuming = iterations cost time
Design

- Naga’s routers use two unidirectional links per port, each of width 37-bits

- Number of different ways to place the 74 signal TSVs constituting the UP port
  - Border, Bundle, Isolated, Shielded

- Optimal topology has:
  - Minimal delay
  - Minimal noise
  - Highest operating frequency with noise
  - Lowest area overhead
  - Smallest impact on existing place & route

- At 45nm, Isolated is optimal. At 32nm, Shielded.
Thermal

- With great stacking, come great thermal gradients

- Power in MPSoCs managed through
  - Dynamic Voltage Frequency Scaling,
  - Clock/Power Gating,
  - Task Scheduling/Migration

- Temperature in 3D stacks is also managed with:
  - Thermal Vias
  - Inter-tier liquid cooling

- Naga uses a novel 3D power management scheme

(THERMAL-AWARE DESIGN FOR 3D MULTI-PROCESSOR, Atienza)
Thermal

• What is the impact of the interconnect?

• TSV topology affects thermal conductance in the vertical direction
  – Vertical thermal conductance = 16x lateral thermal conductance

• Important to understand the thermal implications of the 3D interconnect

• And its constituent effect on system performance
Future

Optical 4x4 hitless silicon router for optical Networks on chip, Droz et al 2008

IBM (2010)

1. WiNOC, Partha Pande, Washington State Univ.
2. Small World-Marculescu, Pande et al.
3. Comparative Performance evaluation of wireless and optical NOC architectures, S Deb and P Pande, SOCC2010
At Circuits and Systems

Many-core Architecture Exploration
- Execution/Programming Models
- Interconnect
- Memory hierarchy
- Dependability management

Many-core software and tooling

Simulation Setup
- Execution driven simulators for many-core
- Dependability simulation

Hardware Prototyping
Networks on Chip

ET4351: VLSI System-on-Chip