Technology and Circuits for On-Chip Networks

Dave Albonesi, Keren Bergman, Nathan Binkert,
Shekhar Borkar, Chung-Kuan Cheng,
Danny Cohen, Jo Ebergen, Ron Ho

Technology drivers for networks

Consider two systems as drivers

- Enterprise-class, large-scale CMP-style machine
  - > Willing to spend on power to get performance (150W budget)
  - > Cost is important
- Hand-held personal electronics gizmo
  - > Cost is primary ($25)
  - > Battery life drives a hard power limit (0.2W budget)

Characteristics and constraints

How technology and circuits affect systems

- Bandwidth density ← Establish this first
- Communication latency ← Ensure this is okay
- Power & power density ← Check if this is possit

- Design time and cost
  > Innovations require technology and CAD ecosystems
- Reliability and fault tolerance

Enterprise-class CMP system

Assumptions for year 2015

- 22nm technology, 256 cores on a 400mm^2 die
- Mesh-style routing grid with 2*15*16=480 total links
- 150W total power, ~20% due to network
  - Of this 20%, 1/3 is the channel, 1/3 is the switch, 1/3 is the buffe
  - This turns out to be ~10W for each of the components
- 0.7V power supply, 150ps clock period (7GHz)
- Wires use 0.25mW/Gbps/mm & travel at 100ps/mm

CMP system network channel

Application requirement: 2TBps bisection BW

- Latency?
  - 1.25mm at 100ps/mm = 125ps < 150ps cycle time
  - 1 cycle per link hop: latency is not an immediate problem

- Power?
  - 2TBps = 16Tbps → 1Tbps per individual link
    - In a 16x16 grid, there are 2*15*16=480 individual links, each 1.25mm long
  - 480Tbps at 1.25mm → 150W @ 0.25mW/Gbps/mm
  - And 150 >> 10: Power is a problem!

CMP system network buffers

- Buffers needed at each router
  - Flits are 16B wide, we have 5 bidirectional ports → 160B = 128(
  - Depth set by timing: 4 flits deep (2 cycles FC, 2 cycles CRC)
  - So need 1280*4 = 5120b. 8X: 40Kb/router, or 10Mb/chip
- Latency okay: SRAM access time < 150ps
- Power?
  - 22nm technology: assume a 0.16μm^2 SRAM cell (2-port)
  - Equivalent switching cap of 15% of the area → 2fF/cell
  - Over 10Mb, this is 20nF → CV^2f gives 20*0.7*0.7*7GHz= 70W
  - 70W > 10W
  - Need low power, high performance memories

Research Agenda

- General areas to be addressed
  - Lower power communication
  - Lower power, fast memory
- Areas for future research (for year 2015 systems)
  - Low swing wires
  - Role of 3D integration for on-die interconnects
  - Role of photonics for on-die interconnects
  - Optimized metalization
    - Tradeoff high C lower layers for upper level layers

Personal electronics gizmo

- A heterogenous mix, few cores connected together
- Look at having 5% of power be network power
  - Assume 200mW is the hard power limit, so 10mW in network
  - Assume 5mW of this is for the channel (simple network)
- Assume 50mm^2 die (from Wingard), so links ~ 7mm
  - 0.25 mW/Gbps/mm = 1.75mW/Gbps at 7mm links
  - For 5mW we can afford 2.8Gbps total bandwidth
  - This seems thin…