Network-on-Chip Architecture

Performance aspect and Firefly network architecture
By Siva Shankar Chandrasekaran and Sree Gowri Shankar
Agenda (Enhancing performance)

- Single to multicore transition
- Why NoC and what is NoC?
- Performance and scalability of NoC
- Traditional NoC Drawback
- Improvements to NoC
- Dynamic heterogeneous Photonic NoC Architecture
- Conclusion
Single core to Multi core transition!

- Initially processors were designed with a single core operating at a particular frequency.
- To increase the performance of system, frequency was increased.
- Beyond a point the frequency could not be increased, because the power consumption and dissipation increased.
- Power Consumption = C*V^2*f
Graph of Clock speed vs Power Consumption
Need for Network on chip

- The performance of systems couldn’t be increased by increasing clock frequency any more
- Therefore multi core chips with many processing elements were introduced
- This introduces the need for inter processor communication
- Inter processor communication leads to communication overheads
- Speedup not always same as the number of processors as there are communication overheads
What is Network-on-Chip?

- NoC is a communication subsystem on an Integrated Circuit (chip), for communication between cores in a System on Chip (SoC).
- NoC are a GALS solution, means Globally Asynchronous Locally Synchronous solution.
- It applies networking theory for on-chip communication
- NoC improves the scalability of SoC’s and is power efficient, which is the need of the hour.
Basic ingredients of a NoC:

- N Computational Resources
  - Processing Elements (PE)
- 1 Connection Topology
- 1 Routing technique
- M ≤ N Switches
- N Network Interfaces
- 1 Addressing system
- 1 Communication Protocol
- 1 Programming model
  - Message passing
  - Shared Memory
Traditionally, dedicated point-to-point connections were used for communication within a chip. This caused large area utilization for the wires. In wired NoC, the links are shared by many signals and parallelism is achieved as all links can operate on different data packets simultaneously. As complexity of the chips increases (more number of cores and computing elements) NoC scales well and provides enhanced performance.
Traditional communication Systems

- They had dedicated bus systems to communicate with the different processing elements
NoC Architecture

- The different processing elements can be considered as nodes and the appropriate topology can be chosen to be implemented on the SoC architecture.
- The switches however must be small, fast and reliable
NoC architecture example
Network on Chip (NoC)

- Adoption of network-based packet communication paradigm.
- Use abstraction and layering to decouple the communication issue from computation.
- Distribute the responsibility of reliable transmission evenly over higher and lower layers of abstraction.

Software
- Application
- Systems

Architecture and control
- Transport
- Network
- Data link

Physical wiring

Protocol stack abstraction
Benini & De Micheli, Computer 2002
Traditional NoC drawback

- More number of computing elements on chips
- Typically multicore chips will have hundreds of components like processing elements, custom logic, GPU units and distributed memory
- Different communicating cores require different bandwidths
- Therefore need for heterogeneous NoC architecture
Expected changes and improvements to NoC

- High bandwidth for communication between processors needed in an energy efficient manner (10 GB/sec approx.)
- Photonic interconnects can help in achieving high bandwidth in an energy efficient manner.
- Performance of GPU increases as bandwidth increases, providing better speedup
- Dense Waveguide Division Multiplexing is adopted. BW depends on number of DWDM assigned to it
- Not all channels require the same amount of bandwidth
- Therefore a Dynamic Bandwidth Allocation scheme is used
Dynamic Heterogeneous Photonic NoC Architecture
Experimental Results
Agenda (Firefly Architecture)

- Why Nano photonic device and firefly architecture?
- Different topology's
- Dragonfly topology and architecture and its disadvantages
- Firefly architecture
- Benefits of Firefly architecture
- Corona vs firefly architecture comparison
Why Nano photonic device and firefly architecture?

- Recent development in silicon nano photonics create new opportunities for on-chip networks
- To exploit the benefits of nano photonics, Firefly - a hybrid network architecture is used
- Firefly uses conventional electrical signaling for communication between cluster of nodes, while inter cluster communication happens nano photonics
- Local communication - Electrical signaling
- Global Communication - Nano Photonics
Nano Photonic Devices
Topologies

- Heritage of networks with new constraints
  - Need to accommodate interconnects in a 2D layout
  - Cannot route long wires (clock frequency bound)

a) SPIN,
b) CLICHE'
c) Torus
d) Folded torus
e) Octagon
f) BFT.
Dragonfly topology and architecture
Disadvantage

- Relies on indirect adaptive routing and multiple traversals for load balancing
- This results in additional complexity
- Packets route within both source and destination which increases the hop count
Firefly architecture
Benefits of Firefly architecture

- Simplifying Routing with Extra Bandwidth
- Efficiently partitioned Optical Crossbars
- Hierarchical Architecture
Corona vs firefly architecture comparison

- Corona Architecture is implemented using a multi-write-single-read optical buses and a 64 x 64 optical crossbar
- Instead of using Multi write optical buses, the firefly uses multi read optical buses assisted with reservation broadcasting
- In Firefly architecture it avoids global arbitration by using localized, electrical arbitration done among small number of ports
Conclusion of Heterogeneous photonic NoC

- High bandwidth in GPU increases its performance.
- Not all communication channels require the same bandwidth, therefore dynamic bandwidth allocation
- Achieves higher performance when compared to homogeneous photonic NoC.
- Suitable for architectures consisting of heterogeneous cores, custom logic, GPGPU’s and memory
Conclusion of Firefly Architecture

- In Firefly architecture which is a hybrid, hierarchical on-chip network architecture that utilizes optical signaling and electrical signaling to achieve energy efficient on-chip network.
- It improves the performance by 54% and efficiency by 38% on traffic patterns with locality.


Questions???