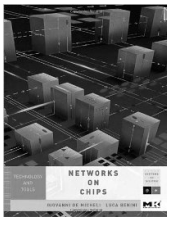
Topology in Multi-core systems

Midia Reshadi, Ph.D
CE Department of Science and Research Branch of Islamic Azad University (SRBIAU)
reshadi@ieee.org
http://www.irprofs.com/reshadi

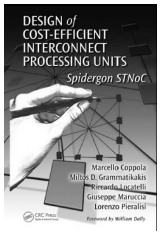
Back ground figure: :On-chip communication architectures, system on chips interconnect

Outline

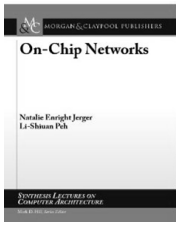
- Main concept
- Examples
- References:



2006



2009



2009

2006

Perspective

- *Definition :* The topology can be described by a directed graph *T(N,L)* where:
- Each $n_i \varepsilon N$ denotes a Processing/ IP core node
- Nodes can incorporate switches, often called routers.
- The directed edge $l_{i,j}(n_i,n_j)\varepsilon L$ denotes a direct physical link between the Processing/ IP core nodes n_i and n_j.
- For every $l_{i,j}(n_i,n_j)\varepsilon L$, $bw_{i,j}$ represents the bandwidth available across the edge $l_{i,j}$.
- Physical Links are wires which can be interrupted by repeaters that have the task to amplify the signal

The types of the topologies

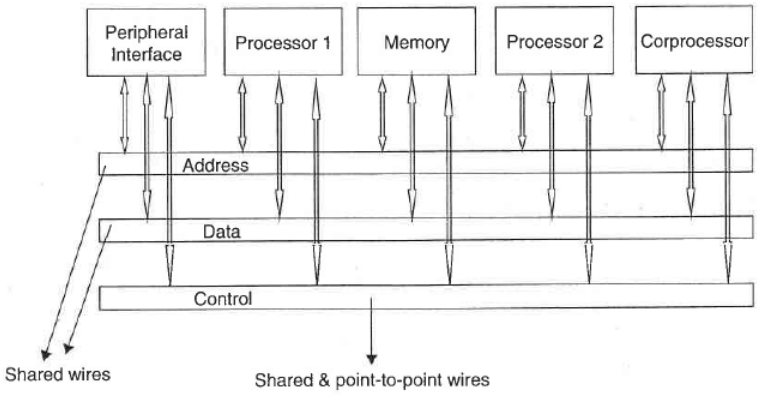
- Shared-medium networks
  - The transmission medium (link) is shared by all nodes
  - only one node at the time can send the information.
- Direct networks
  - Each node has a router and point to point links to other nodes.
- Indirect networks
  - Each node is connected to a switch and switches have point-to-point links to other switches.
- Hybrid networks
  - A mixture of the previous approaches.

Shared-Medium Networks

- Most systems on chip (SoCs) use this class of interconnection
- Role:
  - Initiator (master)
  - Target (slaves)
- Distinctive characteristic: supports broadcast
- Critical issue: arbitration strategy

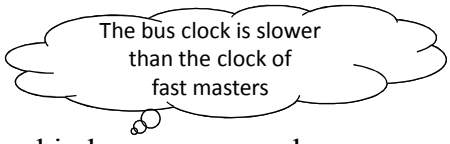
Example of on-chip shared medium networks

- Backplane bus



Note;

- Design issues of on-chip busses:
  - Synchronous
    - All bus interfaces are synchronized with a common clock
  - Asynchronous
    - All devices operate with their own clock and use a handshaking protocol
- Trade-offs:
  - Testability
  - Complexity
  - Debugging
  - simulation
- Currently, all commercial on-chip busses are synchronous



Other issues

- Bus arbitrator and performance
- Various solution has been published
- Industrial examples:
  - Large semiconductor firms:
    - coreConnect by IBM
    - STBUS by ST Microelectronics
  - Core vendors:
    - advanced microcontroller bus architecture (AMBA) by ARM)
  - Interconnect intellectual property vendors:
    - CoreFrame by PalmChip
    - WishBone by Silicore
    - SiliconBackPlane by Sonics

Disadvantages

- Scalability:
  - Busses are used to integrate less than 5 cores
  - Rarely more than 10 cores
- Energy inefficiency
  - Congestion

Direct Networks

- Terminologies
  - Router
  - Input/ output or bidirectional channels
  - node degree:
    - The number of channels connecting a node with its neighbors
  - Network diameter:
    - The maximum distance between two nodes in a network.

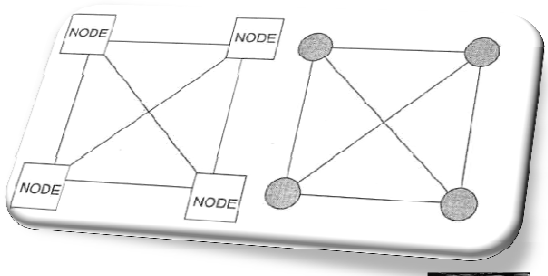
Terminologies

- Regularity
  - A network is regular when all nodes have the same degree
- Symmetric
  - Symmetrical when the network looks alike from every node.
- A path diversity:
  - How many node pairs have multiple (minimal) paths between them
- bisection width:
  - The minimum number of edges that must be cut when the graph is divided into two equal sets of nodes
- Bisection bandwidth:
  - The collective bandwidth of the links associates with these edges

©Midia Reshadi

Example

The diameter=1
Node Degree=3
Type of network: regular & symmetric
Path diversity with non-minimal alternative paths
Bisection width=4



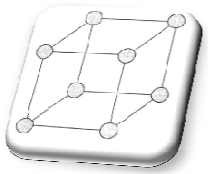
Trade-off in direct networks

- Cost
- Connectivity
  - Higher connectivity → higher performance→ higher area and energy consumption

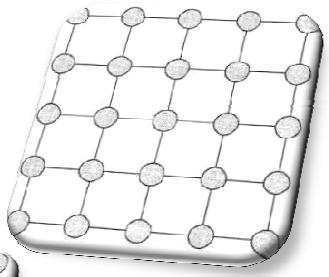
Most practical implementations of direct networks

- An orthogonal topologies
  - Nodes can be arranged in an n-dimensional orthogonal space
  - Every link produces a displacement in a single dimension
- Examples of popular orthogonal networks:
  - n-dimensional mesh
  - Torus
  - hypercube

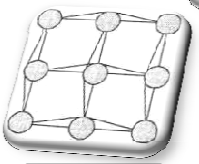
Examples



N=two-dimensional hypercube



Two dimensional mesh



3-ary-2-cube

Hypercube

Binary 1-cube built of two binary 0-cubes labeled 0 and 1

Binary 2-cube built of two binary 1-cubes labeled 0 and 1

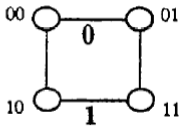
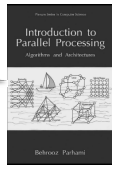
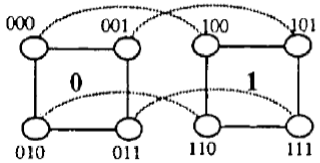
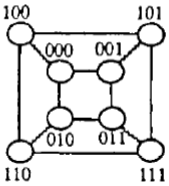


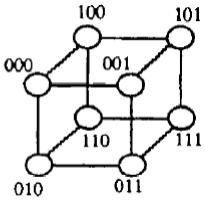
Fig. reference: Based on this book



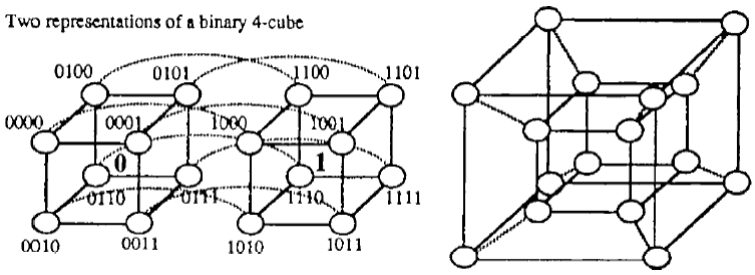
Three representations of a binary 3-cube

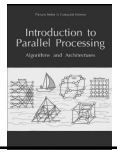


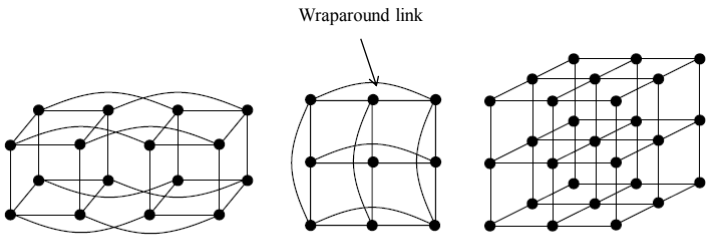




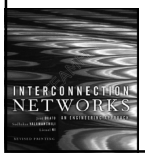
Hypercube





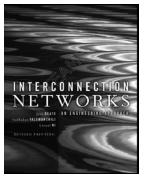


Strictly orthogonal direct network topologies: (a) 2-ary 4-cube (hypercube), (b) 3-ary 2-cube, and (c) 3-ary 3-D mesh.



Advantages

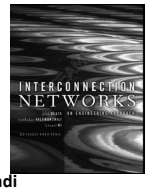
- Simple routing
- Node labeling



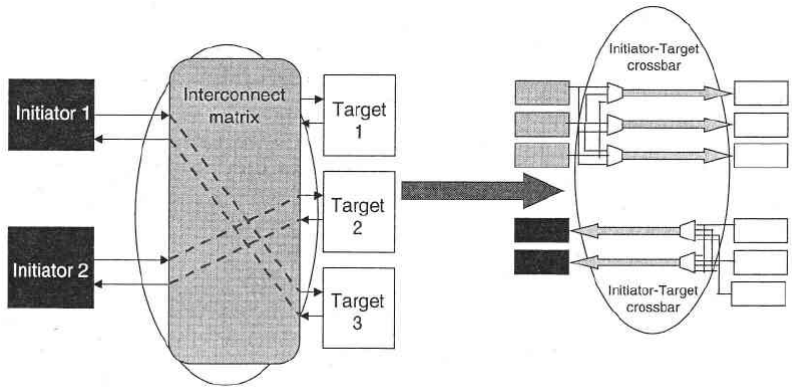
Indirect Networks

- A connection between nodes has to go through a set of switches
- Switches do not perform information processing.
- Their only purpose of switches is to provide a programmable connection between their ports

FPGA



Example

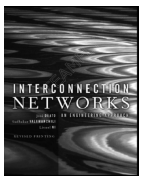


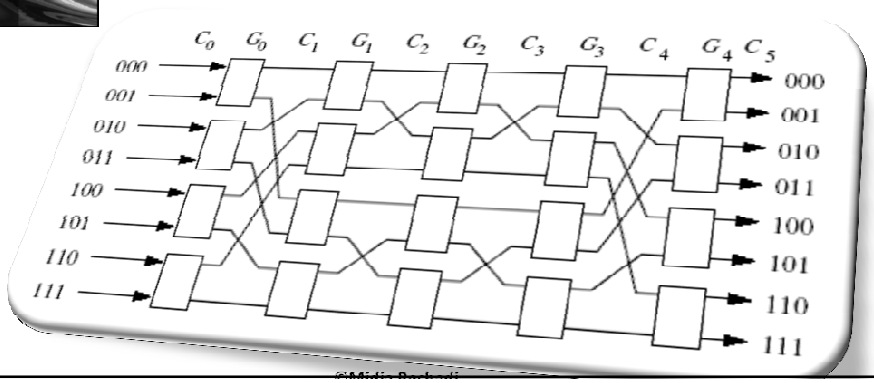
AMBA: multi-layer bus structure

Multistage networks

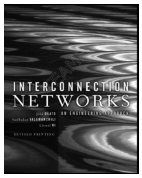
- Blocking networks:
  - Connections between inputs and outputs may not always be possible because of shared resources
  - Information may be temporarily blocked, or dropped,
- Non-blocking networks
- Re-arrangeable networks:
  - Paths may have to be rearranged to provide a connection, and require an appropriate controller.

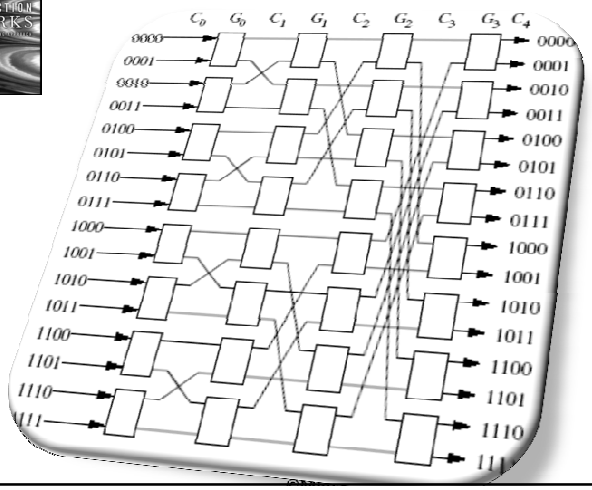
Example: Bene˘s, re-arrangeable network



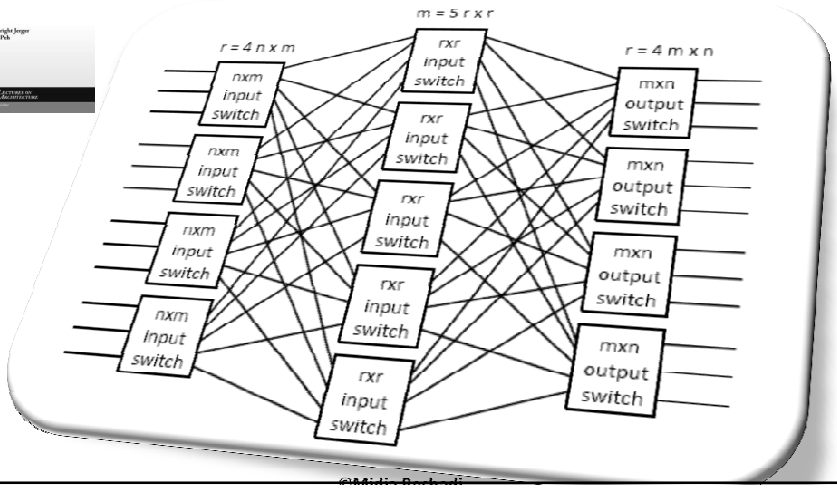


Blocking: Butterfly



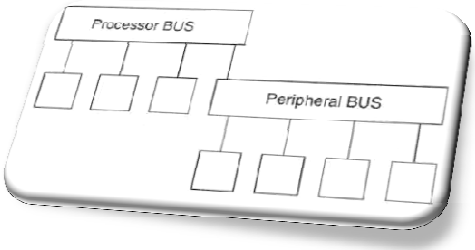


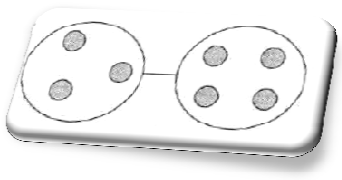
Non-blocking network: the Clos



Hybrid Networks

- Can be abstracted as hypergraphs H(N,C):
  - N is the set of nodes
  - C is the set of channels.



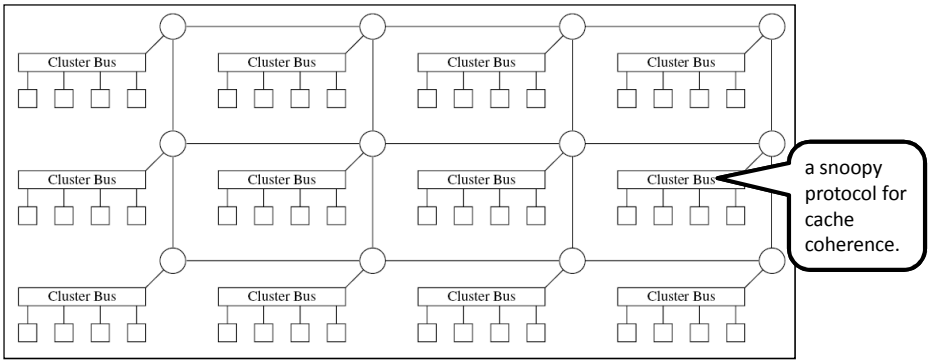


©Midia Reshadi

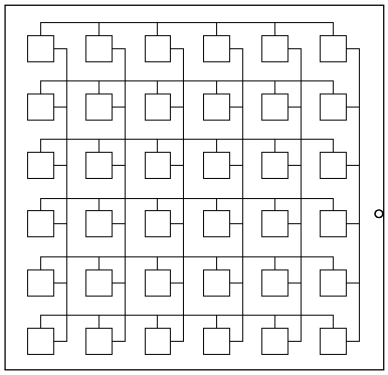
Example

- The AMBA 2.0 standard which includes:
  - AHB protocol
  - advanced system bus (ABS)
  - advanced peripheral bus (APB) protocols
- Multiple bus regions
- Multiple clusters

Examples



Example



Very low diameter
Scalability?

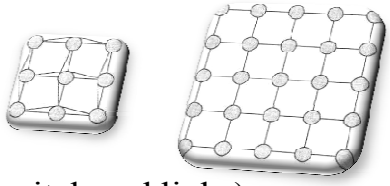
A two-dimensional hypermesh.

Main issues in choosing the topology

- Static cost –silicon area:
  - Number of ports of the router
  - Buffer size
  - Number of global wire
  - Wire length
- Dynamic cost (power-dissipation):
  - Number of network components that are active
  - Number of routers data has to traverse to arrive at the destination
  - Power dissipation in links

Main issues in choosing the topology

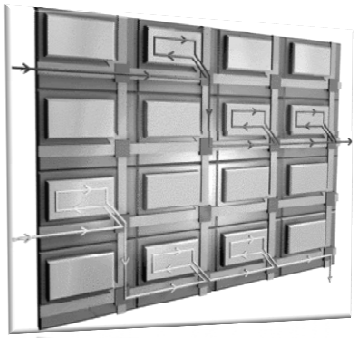
- Performance and cost:
  - Bisection bandwidth
  - Network diameter
  - Higher area cost (number of switch and links)
  - Trade-off
- Application specific & reconfigurable & general purpose



©Midia Reshadi

Purposed topologies

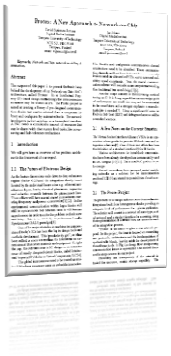
- Mesh:
  - RAW chip multiprocessor
  - Nostrum

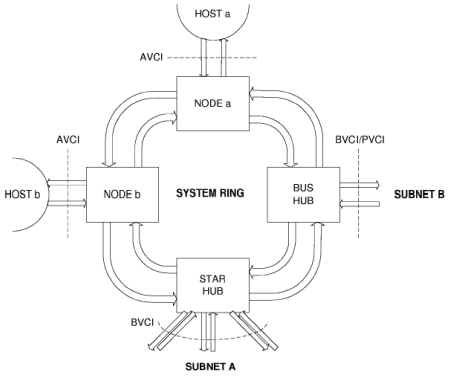


©Midia Reshadi

Purposed topologies

- Torus-based(k-ary l-cube) (one dimensional torus or ring) proteo



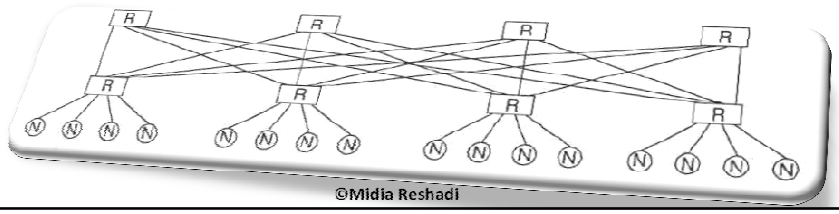


Proteo: A New Approach to Network-on-Chip

©Midia Reshadi

Indirect topology:: fat tree

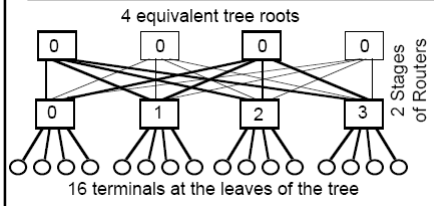
- SPIN network
- Node limitation: 256
- Simple and effective routing
- High area cost for the high number of nodes

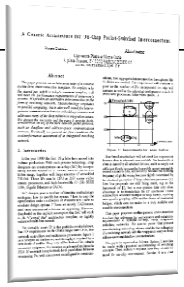


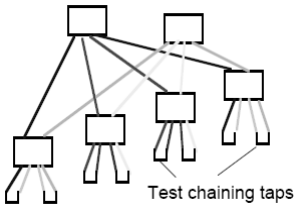
©Midia Reshadi

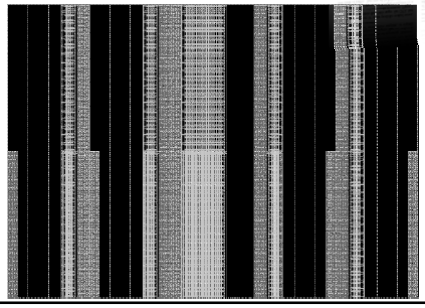
SPIN

A Generic Architecture for On-Chip Packet-Switched Interconnections



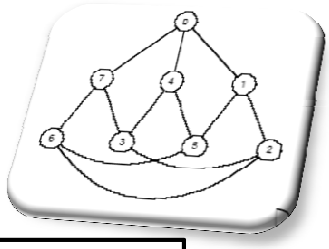


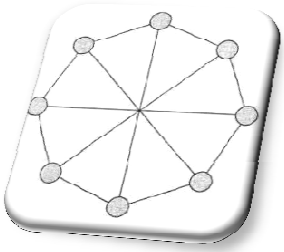




Direct network example: octagon

- Design byST Microelectronics for network processors
- Messages between any two processors require at most two hops,

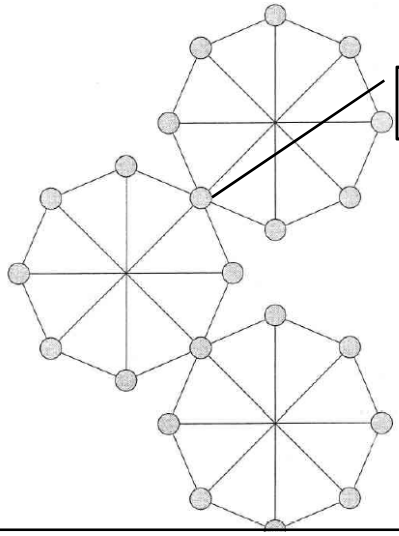




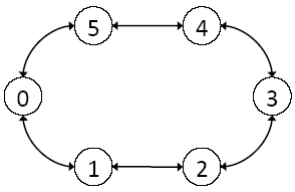
On-chip Communication Architecture for OC-768 Network Processors

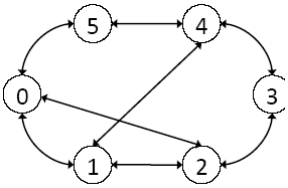
©Midia Reshadi

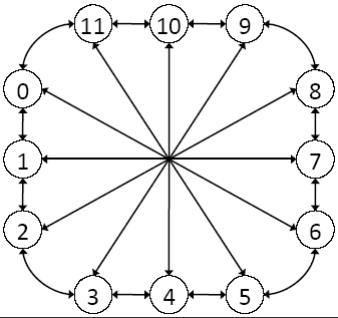
Octagon

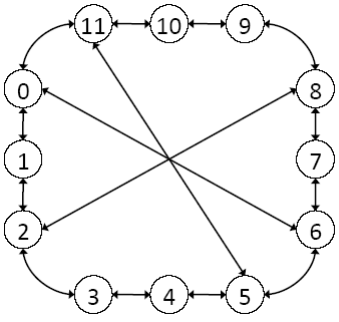


Spidergon

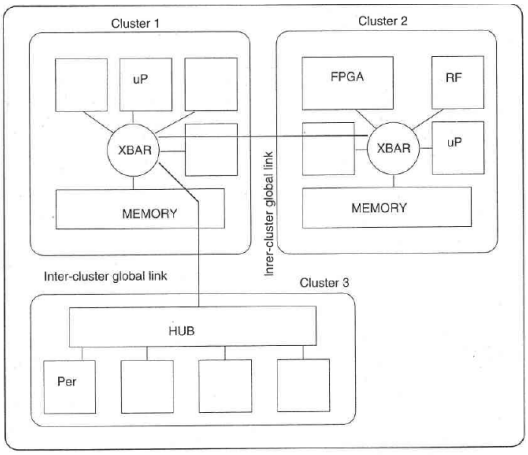








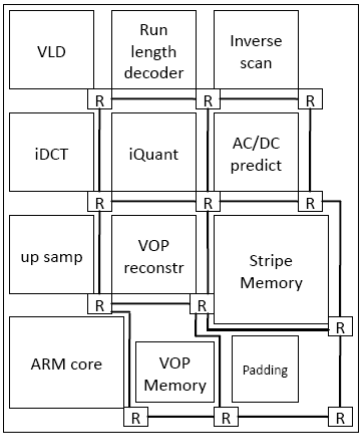
Hierarchical networks



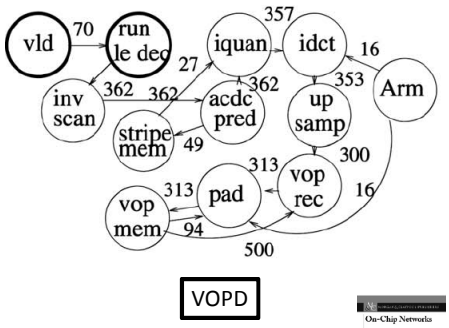
Irregular topology

- Regular topologies → MPSOC
- Heterogeneous IP core → custom topology design

Example

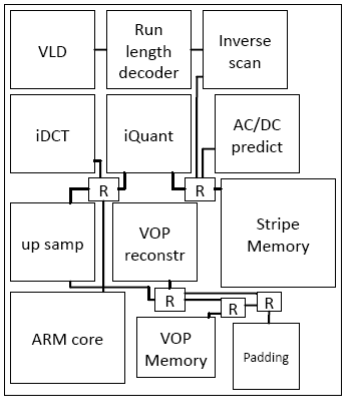


(a) Mesh

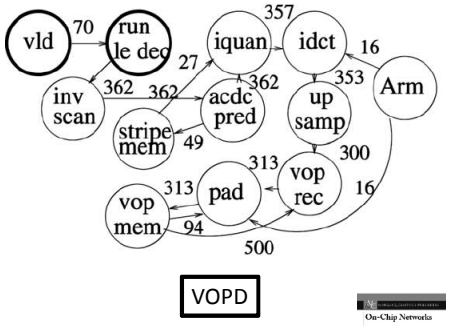


VOPD

Example

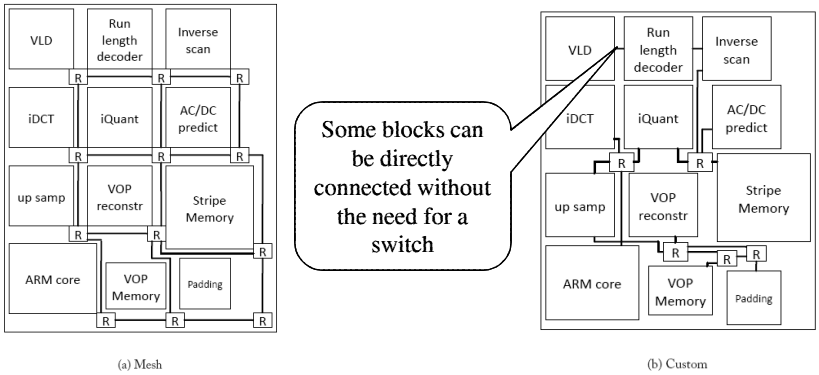


(b) Custom



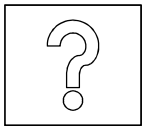
VOPD

Comparison



Some blocks can be directly connected without the need for a switch

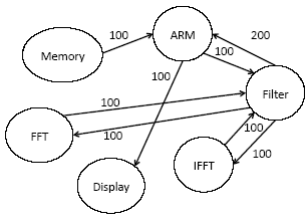
# of switches are reduces from 12 to 5
Significant power and area savings
Max mesh switch size= 5x5
Max custom switch size= 3x3



SPLITTING ANDMERGING

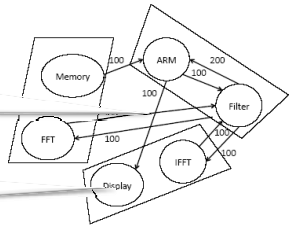
- Splitting
  - Large fully connected switch
  - iteratively be split into smaller switches
  - The bandwidth must satisfy the volume of communication
- Merging
  - It starts with large number of switches
  - Merging them into gather

Topology synthesis algorithm example



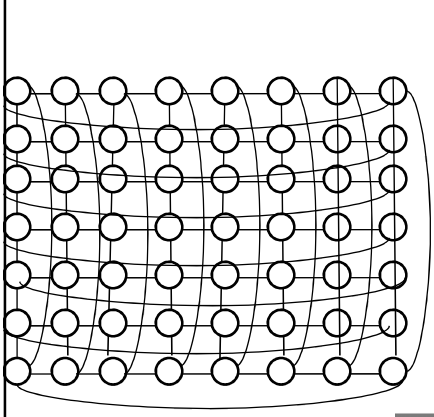
The edges that cross partitions have lower weights than the edges within partitions

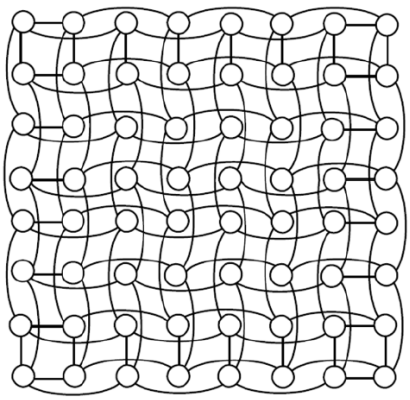
The number of nodes of each partition is same.



(b) Min-cut Partition

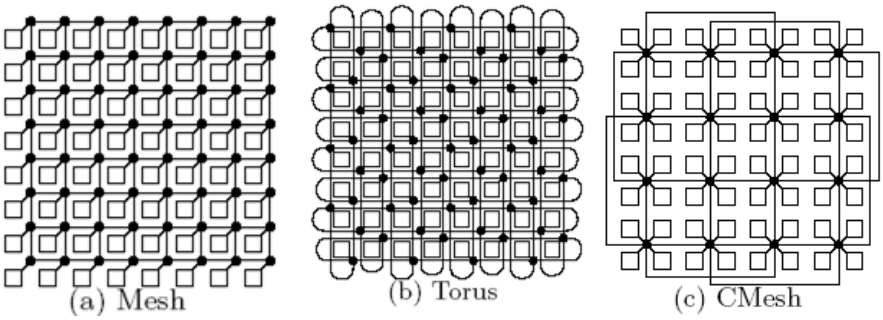
Layout and implementation





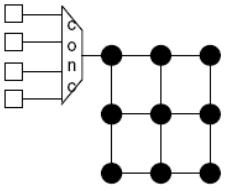
Layout of a 8x8 folded torus:
Reducing link number
Reducing link length
Same link length

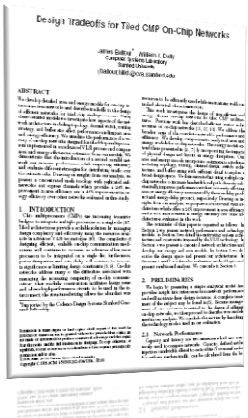
Concentrator



Design Tradeoffs for Tiled CMP On-Chip Networks
James Balfour , William J. Dally

Concentrator



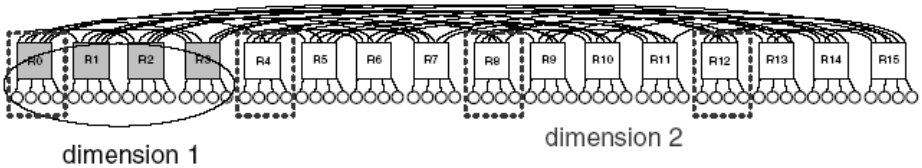

Design Tradeoffs for Tiled CMP On-Chip Networks

Design Tradeoffs for Tiled CMP On-Chip Networks
James Balfour , William J. Dally

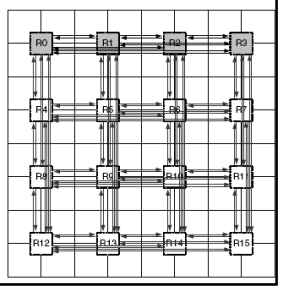
Concentrator

- Nodes (routers) shared by four IP cores
- The use of concentration:
  - Reduction of the number of routers
  - Reduction of the hop count
  - Saving the area
- Implementing a policy for sharing injection bandwidth
  - Static sharing bandwidth: 1/c bandwidth
  - Dynamic sharing bandwidth
- In bursty communication, the injection port bandwidth can become a bottleneck.

Other topologies



Flattened Butterfly Topology for On-Chip Networks
John Kim, James Balfour, and William J. Dally



Implication of abstract metrics on on-chip implementation

- Node degree:
  - useful as a proxy for router complexity
  - higher degree implies greater port count
  - More input buffers
  - More requester to allocator
  - More critical path delay
- Link complexity→ link width → area→ power

Implication of abstract metrics on on-chip implementation

- Hop count:
  - Is used as a abstract value for overall network latency and power
  - But? Router pipeline depth
  - Example page 41 of on-chip networks book

Trade-off between:

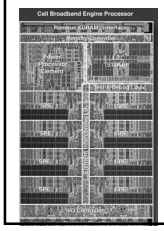
- HOP COUNT VS NODE DEGREE
- ↑ Node degree ≈ ↓ ave hop count

Maximum channel load

- Good abstraction for network saturation throughput and maximum power
- Channel load assignment based on traffic pattern

Bibliographic notes

- **IBM Cell:**
  - Four unidirectional rings, two in each direction
  - Each ring is 16 bytes wide
  - It runs at 1.6GHz
  - Can support 3 concurrent transfers
  - Total network bisection bandwidth is 307.2GB/s



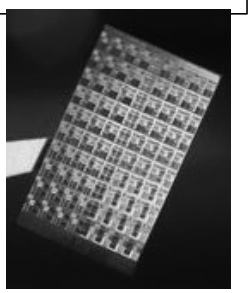
©Midia Reshadi

TileraTILE64

- 8 × 8 meshes,
- Each channel consisting of two 32- bit unidirectional links
- At the chip frequency of 1GHz, a bisection bandwidth = 320GB/s

©Midia Reshadi

Intel TeraFLOPS



- 8 × 10 mesh
- Two 38-bit unidirectional links.
- an aggressive clock=5GHz
- 65nm process tech
- a bisection bandwidth of 380GB/s or 320GB/s

ST Microelectronics STNoC

- Topology: Spidergon, a pseudo-regular topology