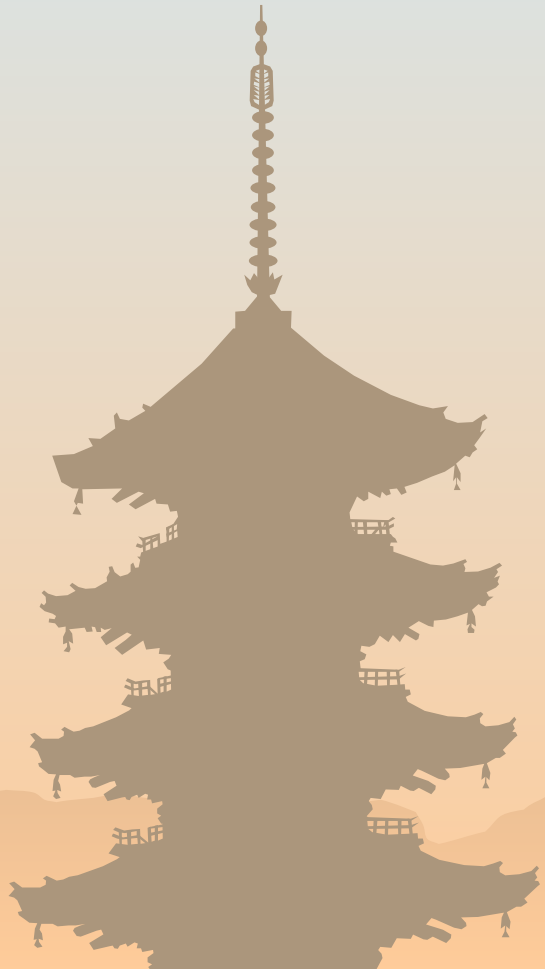
Design of Interconnection Networks for MPP of Next Generation

By S. Peng

Faculty of Computer and Information Sciences

Hosei University

Tokyo, Japan



Introduction

- Interconnection networks are currently being used for many different applications, ranging from internal buses in VLSI circuits to wide area computer networks. Among others, these applications include backplane buses and system area networks, telephone switches,ATM switches, processor/memory interconnects for vector supercomputers, interconnection networks for parallel computers, NOW, COW, LAN, MAN, WAN and networks for industrial applications.

Introduction (Continue)

- Designing parallel computers from commodity components is a popular trend for cost-effectiveness. In these parallel computers, many processors cooperate to solve a large problem. Memory bandwidth can be scaled by distributed shared-memory. Also, RAID allows the implementation of high-capacity, reliable file systems.

Introduction (Continue)

- However, a parallel computer requires some kind of communication subsystems to interconnect processors, memories, disks using system buses and I/O buses, and then, interconnect processors using the interfaces to local area networks. Unfortunately, the communication subsystem always becomes the bottleneck in most applications for these parallel computers.

Introduction (Continue)

- As the interconnection network is the only subsystem that cannot be efficiently implemented by using commodity components, its design becomes very critical for the performance of parallel computers. The discussion of the issues arising from this approach is not covered in this talk.
- In this talk, we will focus on the interconnection networks for MPP of very large scale.

Reference Books

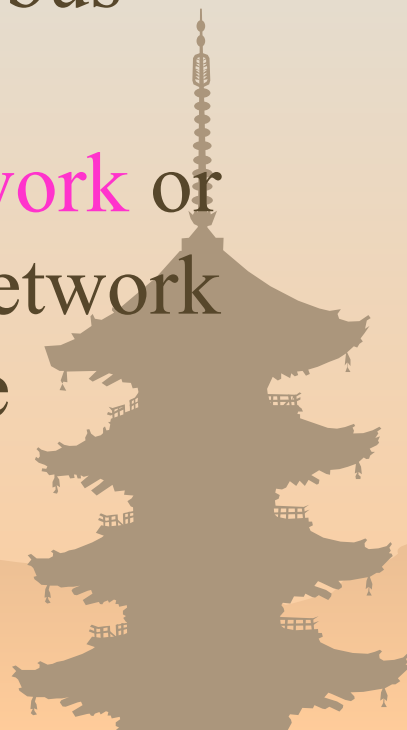
- "Interconnection Networks" by J. Duato, S. Yalamanchili and L. Ni, 1997
- "Scalable Parallel Computing" by K. Hwang and Z. Xu, 1998
- "Parallel Computer Architecture" by D.E. Culler and J.P. Singh, 1999
- "Introduction to Parallel Processing: Algorithms and Architectures", by B. Parhami, 1999

Classification of Interconnection Networks

- A classification scheme based on network topology categorizes the known interconnection networks into four major classes: shared-medium networks, direct networks, indirect networks and hybrid networks.

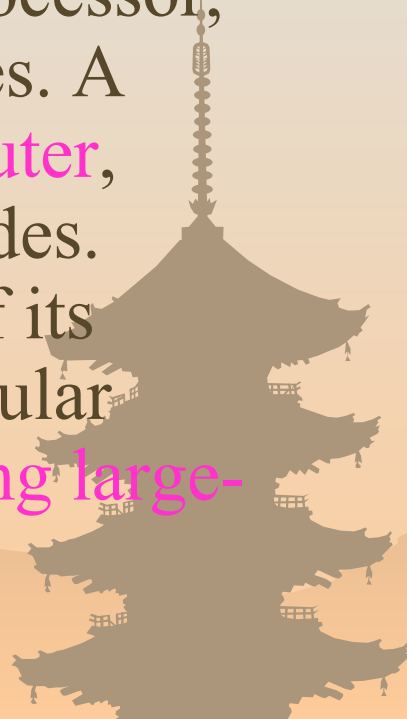
Direct Networks

- Scalability is an important issue in designing parallel computer systems. Bus-based systems are not scalable as the bus becomes the bottleneck when more processors are added. The direct network or point-to-point network is a popular network architecture that scales well to a large number of processors.



Direct Networks (Continue)

- A direct network consists of a set of nodes, each one being directly connected to a (usually small) subset of other nodes in the network. Each node is a programmable computer with its own processor, local memory, and other supporting devices. A common component of these nodes is a router, which handles message passing among nodes. Each router has direct links to the router of its neighbors. Direct networks has been a popular interconnection architecture for constructing large-scale parallel computers.



Direct Networks (Continue)

- Direct networks have been traditionally modeled by a graph. Some basic network properties (node degree, diameter, regularity, and symmetry etc.) can be defined from the graph representation.
- A direct network mainly characterized by three factors: topology, routing, and switching. We discuss briefly these factors.

Topology

- The topology defines how the nodes are interconnected by channels, and is usually modeled by a graph.
- Many topologies have been proposed, trying to balance performance and some cost parameters. In these topologies, messages may have to traverse some intermediate nodes before reaching the destination node. The routing algorithm determines the paths for message transmissions.

Popular Network Topologies

- Most of the implemented networks have an orthogonal topology.
- A network is orthogonal if and only if nodes can be arranged in an orthogonal *n*-dimensional space, and every link can be arranged in such a way that it produces a displacement in a single dimension.
- In a weakly orthogonal topology, some nodes may not have any link in some dimensions.

Popular Network Topologies

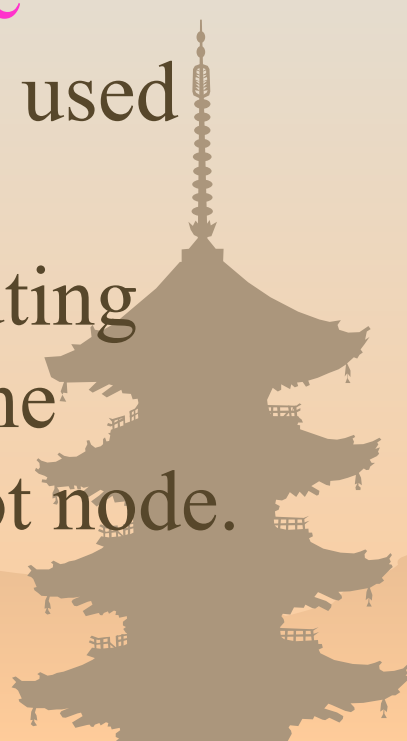
- The most interesting property of orthogonal topologies is that routing is very simple.
- The most popular direct networks are the n-dimensional mesh, the *k*-ary *n*-cube or torus, and the hypercube. All of them are orthogonal (the figures are provided by transparencies).

Trees

- A popular topology is the tree. The drawback of trees as general-purpose interconnection networks is that the root node and the nodes close to it become bottleneck.
- A practical way to implement trees with higher channel bandwidth in the vicinity of the root node is fat-trees.

Trees (Continue)

- One of the most interesting properties of trees is that, for any connected graph, it is possible to define a tree that spans the complete graph. This property can be used to define a routing algorithm for any irregular topology. However, that routing algorithm may be inefficient due to the concentration of traffic across the root node.



Cayley Graphs

- Many topologies have been proposed with the purpose of minimizing the network diameter for a given number of nodes and node degree (e.g., the de Bruijn network and the star graphs). However, routing is complex, and the practical application of these networks is very difficult.
- Star graphs, hypercubes, and cube-connected-cycles are particular cases of Cayley graphs.

Hierarchical Networks

- Some topologies have been proposed with the purpose of reducing node degree while keeping the diameter small. Most of these topologies can be viewed as a hierarchy of topologies (e.g., CCC: the cube-connected cycles).
- The CCC is weakly orthogonal.

Hierarchical Networks (Continue)

- Multiple backplane buses: Different buses are interconnected by routers or bridges to transfer information e.g., bridged LANs).
- Cluster-based networks: Combine the advantages of two or more kinds of networks at different levels in the hierarchy.
- The design of the Stanford DASH parallel computer is cluster-based.

Dual-cube

- The hypercube has been widely used as the interconnection network in a wide variety of parallel systems such as CM-2, Intel nPSC, nCube and SGI Origin 2000.
- We proposed a new topology, called dual-cube to mitigate the problem of increasing number of links in the large-scale hypercube network while keeps most of the topological properties of the hypercube network.

Meta-cube

- Meta-cube is a hierarchical network which can connect a very large number of PEs with just a small amount of links (4 – 6) per node. Therefore, meta-cube has potential to be the interconnection network of high-performance parallel computers of very large scale.
- An MC(3,3) with 6 links per node contains 2^{27} nodes
- Meta-cube includes dual-cube as a special case.

Concluding Remarks

- Design of interconnection networks is still a very popular research area. There are many ongoing projects that are related to this subject. I believe that this trend will continue for decades since the parallel/distributed computer systems are the only solutions for the computational problems challenging human being in the 21st century.
- Topology and communication protocol are among the most important issues for interconnection networks and certainly worth of research effort.