Interconnection Networks

1 Ideal Parallel Computer

A natural extension of the Random Access Machine (RAM) serial architecture is the Parallel Random Access Machine, or PRAM. It consists of p processors and a global memory of unbounded size that is uniformly accessible to all processors. Each processor can simultaneously undertake $O(1)$ work in $O(1)$ time. Each processor can access any memory location in $O(1)$ time.

Depending on how simultaneous memory accesses are handled, PRAMs can be divided into 4 subclasses:

1. **Exclusive-read, exclusive-write (EREW) PRAM**: No two processors can read from the same memory location simultaneously, so each read operation must have exclusive access to the memory cell being read. No two processors can write to the same memory location at the same time. Each write operation must occur without any other processor attempting to write to the same location.

2. **Concurrent-read, exclusive-write (CREW) PRAM**: Multiple processors can read from the same memory location at the same time without any restrictions. No two processors can write to the same memory location at the same time.

3. **Concurrent-read, concurrent-write (CRCW) PRAM**: Multiple processors can read from the same memory location at the same time without any restrictions. Multiple processors can write to the same memory location at the same time.

4. **Exclusive-read, concurrent-write (ERCW) PRAM**: No two processors can read from the same memory location at the same time without any restrictions. Multiple processors can write to the same memory location at the same time.

2 Concurrent Write Model

There are 4 normal ways of concurrent write, including:

1. **Common**: write only if all values are identical

2. **Priority**: follow a predetermined priority order

3. **Sum**: Write the sum of all data items

4. **Arbitrary**: Write the data from a randomly selected processor.

Several classical problems could be solved using PRAM include:

2.1 Sum p numbers on PRAM

If using EREW PRAM, the runtime will be $O(\log p)$, but if using CRCW PRAM, the runtime will be $O(1)$.

2.2 Compute logical OR on PRAM

If using EREW PRAM, the runtime will be $O(\log p)$, but if using CRCW PRAM, the runtime will be $O(1)$.

3 Memory Model

In **Networked Shared Memory Model**, all processors can access all memory as a unified address space. Memory is physically shared among all processors but may be distributed across different nodes, leading to latency variations based on memory location relative to each processor.

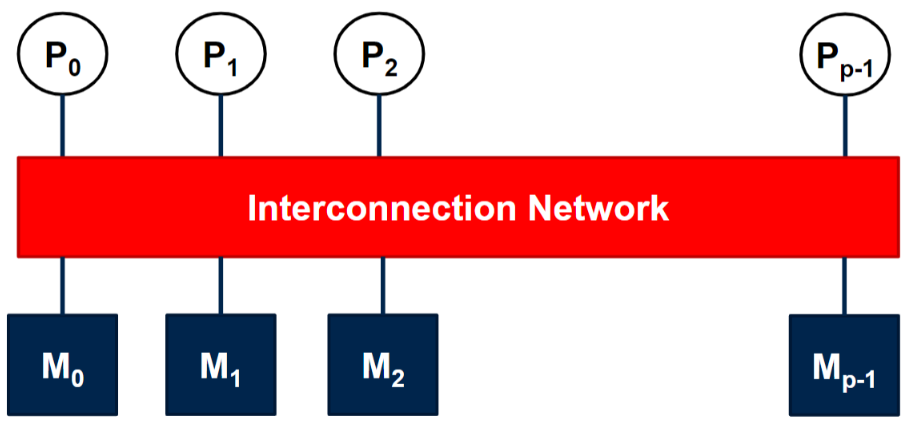

Figure 1: Networked Shared Memory Model

In **Networked Distributed Memory Model**, each processor has its own local memory. The processors are connected by a network, and they do not share a physical memory space.

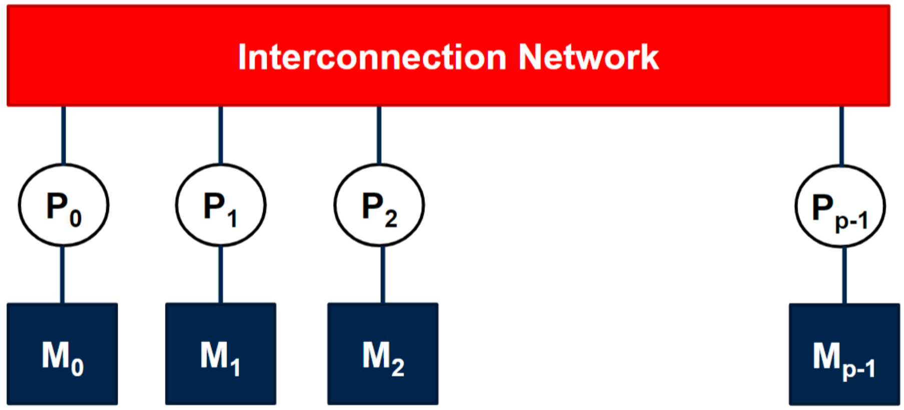

Figure 2: Networked Distributed Memory Model

4 Interconnection Networks

Interconnection networks carry data between processors and to memory. They are made of switches and links (wires, fiber). They could be classified as:

1. **Static networks**: consist of point-to-point communication links among processing nodes and are also referred to as direct networks.
2. **Dynamic networks**: are built using switches and communication links, and are referred to as indirect networks.

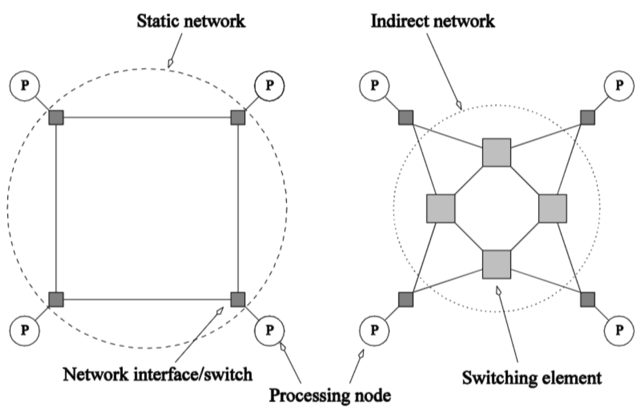

Figure 3: Interconnection Networks

4.1 Static Networks Evaluation

Several parameters are used to evaluate the static networks:

1. **Distance**: between a pair of processors is the number of links on a shortest path connecting the two.

2. **Diameter**: the largest of distances between pairs of processors. The ideal diameter is $\Theta(1)$.

3. **Bisection Width (BW)**: The minimum number of wires you must cut to divide the network into two subnetwork with $\frac{p}{2}$ processors each. The ideal BW is $\Theta(p)$.

4. **Cost**: The number of links or switches (whichever is asymptotically higher) is a meaningful measure of the cost. The ideal cost is $O(p)$.

4.2 Topologies

4.2.1 Completely Connected Network

In this network, each processor is connected to every other processor.

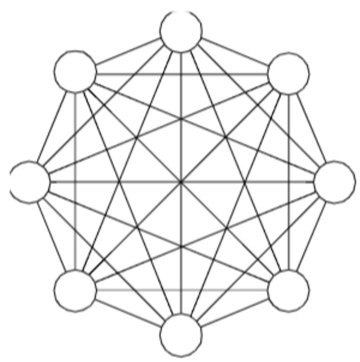

Figure 4: Completely Connected Networks

Important parameters include:

1. Diameter: 1

2. BW:

3. Number of links per node: $p - 1$

4. Cost: the number of links in the network scales as $O(p^2)$

4.2.2 Star Connected Network

In this network, every node is connected only to a common node at the center.

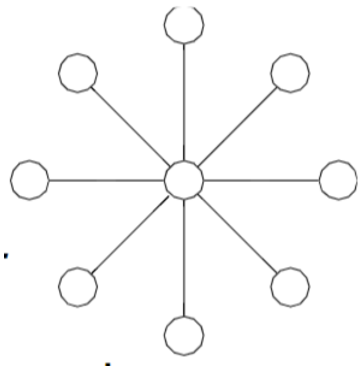

Figure 5: Star Connected Networks

Important parameters include:

1. Diameter: 2
2. Distance between any pair of nodes is $O(1)$, but the central node becomes a bottleneck
3. BW: 1
4. Number of links per node: maximum of $p-1$
5. Cost: $p-1$