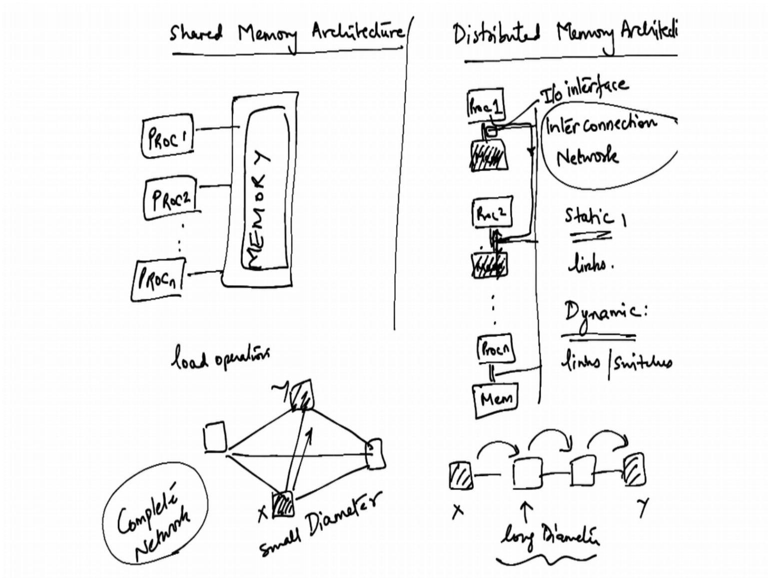
Introduction to Parallel Programming in OpenMP
Dr. Yogish Sabharwal
Department of Computer Science & Engineering
Indian Institute of Technology, Delhi

Lecture – 10
Interconnection networks in Distributed Memory architectures

(Refer Slide Time: 00:01)



So, how do you form these interconnection networks. So, typically there are 2 kinds of interconnection networks one is static. So, in static interconnection networks what happens is that, they are essentially formed by links; these links connect up different nodes together. For instance here is a simple example let us say that you have a four node system, when I refer to a node I mean a processor plus memory unit and it is IO interface.

So, now, I connect these up and I basically put a link between every pair of the nodes. This is one possible interconnection network that I have come up with it. Everything is static over here right if x wants to send data to y, then it essentially uses this link. This particular network is called a complete network. It seems like quite a good network why because anybody can communicate with anybody right let us consider another network. So, suppose you have another network where this is the way the 4 nodes are connected. So, now, if processor x wants to talk to processor y, what does it need to do? It needs to first send the data to this processor, then it is forwarded here and then finally, it is forwarded here, right.

So, there are multiple hops that it has to traverse. So, it takes time for it to do that, at each node certain processing has to be done to forward the data. So, in that sense the complete network seems better why because anybody can talk to anybody right. So, it would not take long it is not going to suffer a lot of latency right if I want to communicate with another processor. So, how do you measure the latency? So, what characteristic of this graph will tell you the latency how good or bad the network is with respect to communicating with somebody far away.

Student: (Refer Time: 01:58) longest part.

Longest part diameter;

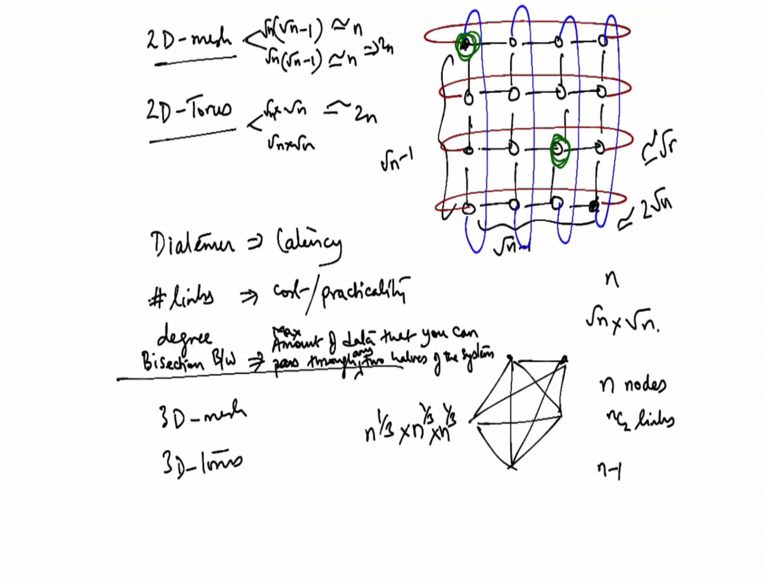
Student: (Refer Time: 02:00).

This is nothing, but the diameter of the graph. So, a complete network has a very small diameter, where as a straight path has a long diameter. So, what is the diameter here? It is 3, what is the diameter here it is 1 right, but which one is more scalable if I want to scale this up to 1000 processors.

Student: (Refer Time: 02:26).

So, the first one will become a pain, if you want to build a complete network over one thou 10,000 nodes 100 1000 nodes it is a nightmare you cannot do it right it is its not practical. Where as in this network a path it is easy to make it, but then it is not a great idea again because just imagine the first node wanting to talk to the last node it is going to take 100 1000 steps for that message to hop away, right.

(Refer Slide Time: 02:57)



So, what are some of the more practical networks? One common network that has been used is the mesh 2 D mesh. So, what does the 2D mesh to? So, this is a 2D mesh of how many nodes 16 nodes all right. So, it is quite simple looks like a matrix right 4 by 4, 4 rows 4 columns, and every node is attached to it is neighbor on the left on the right up and below except for the boundary nodes what is the diameter of this network.

Student: (Refer Time: 03:39).

Let us talk in terms of n, if I had n nodes in a root n by root n configuration, which 2 nodes; should I select to look at the diameter.

Student: (Refer Time: 03:50) diagonally opposite.

Diagonally opposite right this one and this one. So, it takes me root n minus one hops to reach this bottom node and then

Student: Root n minus one hops.

How many hops?

Student: Root n minus 1.

Root n minus 1, another root n minus 1 hops to go from here to here.

Student: (Refer Time: 04:08).

Right. So, let us not bother about the minus 1. So, it is roughly about 2 root n that is the number of hops I have to take, simple trick that is used in practice to reduce the diameter. So this architecture is called a 2D torus. So, what you do is you just add another wraparound link.

Student: (Refer Time: 04:33).

So, the last node gets connected to the first node, both horizontally and vertically.

Student: There is no edge.

Boundary yeah there is no boundary, now all nodes are treated equal right what is the diameter of this curve?

Student: (Refer Time: 04:52) 2.

Two yes suppose to look at the 2 nodes which will take the longest time to communicate with each other. The 2 nodes which took the longest time to communicate with each other in the case of 2D mesh are. Now very close to each other, but that does not mean that these are the farthest nodes now who is the farthest I will I will fix one of the node because no node is special right because it is totally a symmetric network. So, let me fix one of the nodes let me fix this node.

Student: (Refer Time: 05:25).

Which is the node furthest from this node.

Student: (Refer Time: 05:31) 3 comma 3.

3 comma 3.

Student: 3 comma 3.

Halfway down halfway across and what will be the diameter.

Student: 4.

In terms of n.

Student: Root n.

Around root n

Student: (Refer Time: 05:44).

So, you certainly half the diameter just by adding the wraparound that is not a huge cost right yeah what is the number of links in a 2D mesh. So, the number of horizontal links I have is root n minus 1 times root n, there are root n rows and he each row has root n minus one links

Student: Hum

Right and the number of vertical links I have is the same.

Student: (Refer Time: 06:06).

So, roughly this is about n right roughly about n this is also roughly about n. So, I had a total of about 2 n links, and in the 2D torus what do I have? So, if I just look at one row how many links do I have in a row instead of root n minus 1 I have root n. So, it is root n into root n and root n into root n, which is again about 2 n right there is just a difference of 2 root n. As in totecally is the same number of links, but I have got half the diameter let us do some simple tricks that I use to build these networks 2D torus is quite common in practice and in practical networks.

So, there are different parameters that are used to gauge how good a network is right. One of them is the diameter which basically tells you in some sense the latency how long is it going to take for a request to be serviced if it is send from one node to another node right assuming the furthest nodes right. So, diameter basically tells us the latency. Another is the number of links right this in some sense determines the cost of the network right cost and also maybe how practical it is number of links as well as the degree of a node. So, what is the problem in a complete network right? So, if you have a complete network this is a complete network of 5 nodes.

So, if you have let us say n nodes a complete network of n nodes, what is the total number of links you would have?

Student: NC 2

NC 2 that is quite large is in totecally n square and the other problem is the degree, what is the degree of every node.

Student: n minus 1.

N minus one just imagine every node if you have a 100 1000 nodes every node has 100 1000 links going out of it.

Student: hum

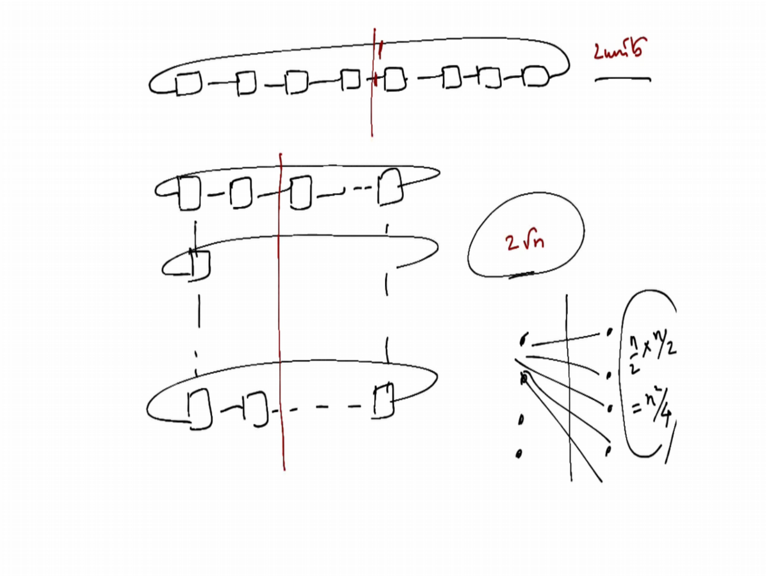
If you cannot even practically design such a system, right. So, although I mean the complete network would be great in terms of designing algorithms, but you do not have that flexibility right we cannot make such networks. People have gone beyond 2D mesh 2D torus or 3D mesh 3D torus right that is quite common ok.

So, you can imagine what the 3D mesh and 3D torus is going to (Refer Time: 08:36) same thing make it a cube, and n to the power 1 by 3 cross n, 3 power 1 by 3 cross n to the power 1 by 3. That could be the configuration and then you can put wraparound links also now wraparound links would be in three directions.

So, each node would have 6 neighbors. So, as you increase the dimension right it starts becoming more and more dense. There are benefits of that density means that the cost goes up because more links practically starts becoming more and more challenging, but the diameter reduces latency reduces.

Another important factor that we consider in a network is how much data can I pump through the network. So, how do you exactly measure that. So, there is something called bisection bandwidth, this is essentially the amount of data that you can pass through 2 halves of the system. Like I should say any 2 halves of the system. So, the maximum amount of data that you can pass through any 2 halves of the system.

(Refer Slide Time: 09:53)



So, let us consider a network which is a straight path, what is the maximum amount of data I can pass through 2 halves of the system. So, let me break it up into these 2 halves 4 here, 4 here how much data I can a pass from any one half to the other half? Let us say that for the time being that the capacity of each link is one unit right. So, how many units can I pass at any given point of time from one half to the other half?

Student: (Refer Time: 10:22) one unit

One unit.

Student: (Refer Time: 10:26).

Two units in this case 2 units there are 2 edges right one is this edge and the other is this edge right. So, I can pass 2 units. If I did not have the wraparound link then it would be one unit right what about a 2D mesh or let me just take a 2D torus right.

So, what is the cut that divides it into 2 equal halves? So, let me just take a vertical cut you can try out other kinds of cuts, but you will realize that you know this is the most binding.

So, what do I want? I want to find out what is the maximum amount of data that I can pass through any 2 halves of the system right.

Student: Hum.

So, I want to find that cut which gives the minimum number of links, because that would be the most restrictive let us say consider this vertical cut. So, what is the total amount of data I can pass through this?

Student: (Refer Time: 11:34) 2.

2

Student: 2 root n

Two root n and in a complete network. So, just consider a cut in the complete network right. Every node on the left hand side is connected to every other node on the right hand side. So, how many links going out of every node?

Student: (Refer Time: 11:58).

N by 2.

Student: (Refer Time: 12:01).

Going to the other side right we are very interested in how many of them a crossing the cut.

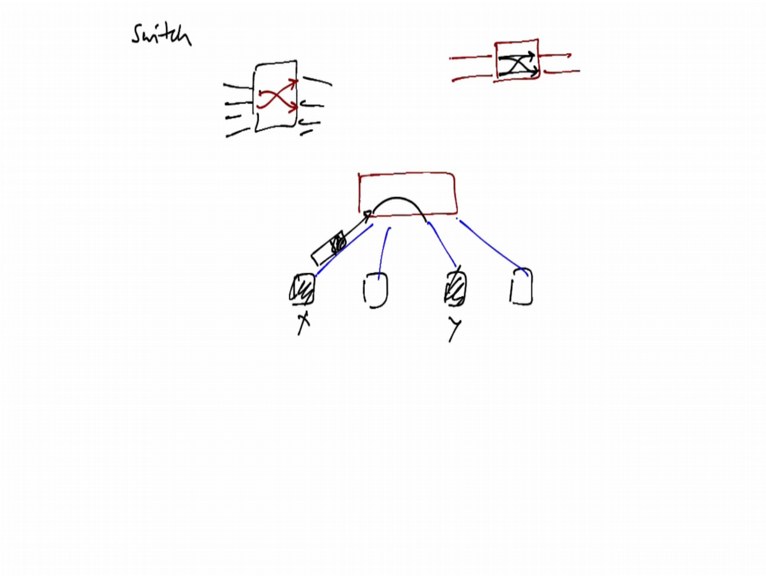
Student: N by 2 (Refer Time: 12:08).

So, n by 2 and how many nodes on the left hand side? N by 2. So, there are n by 2 nodes each of them having n by 2 links going to the other side. So, total number of links is n square by 4.

So, you can see right here you had 2 units you improved it to 2 root n by going to a 2D mesh 2D torus, and then you improved it to order n square right by going to a complete network. So, these are the major factors that determine how great a network is right how good a network is. The latency which is the diameter, the number of links which determines the cost and the degree also which determines how practical it is to build such a network and also the cost and the bisection bandwidth right which determines how much data you can actually push through the network as I said write that 2 kinds of interconnection network static and dynamic.

So, what is the dynamic network? A dynamic network has links and switches. So, what is the switch?

(Refer Slide Time: 13:14)



So, a switch is essentially your device, which has some input ports and some output ports and it has the capability of redirecting data from any of the input ports to output ports.

So, simple 2 cross 2 switch would look something like this, it will have 2 inputs to outputs and then it would have logic inside of either sending them straight or switching them, this the simplest kind of switch you can build. The point here is that this is dynamic in nature because the paths are not established up front, let us to a simple network which is switched right.

So, you have four nodes and you have a switch sitting over here, and all these nodes are connected to the switch and now when I send some data let us say x wants to communicate to y, what will happen? It will send data to the switch and there has to be some header in that data that I am sending, which tells the switch that I want to send this data to y right. So, the switch will read that address from here that this is meant for y. So, I am supposed to send it to y and it will dynamically set up a path and forward this data right. So, we would not get a whole lot deeper into interconnection networks or distributed memory systems, because we are going to focus more on shared memory systems and openMP, right. So, we will stick to shared memory systems.