Large Scale Linear and Mesh Network of PCs Connected by SCSI

G. Young
Computer Science Department
California State Polytechnic University
Pomona, CA, USA

B Cong
Computer Science Department
California State University, Fullerton
Fullerton, CA, USA

P. Ng
Computer Science Department
The Chinese University of Hong Kong
Hong Kong

Abstract - *High-performance computers have been in great demand for applications in different areas. The increase in the processing power of processors cannot solely satisfy our demand. Parallel computers are made to overcome this technology limitation. In the last decade, research topics on parallel computer using network-connected multicomputer have been studied extensively. A cost-efficient high-speed multicomputer system was built using the SCSI bus for the network connection, and it has been shown that it can reduce the communication overheads and hence increase the overall performance [5]. In order to build highly scalable multiple computers based on this design, we have to take into consideration of different network topologies. Since SCSI bus [2,3] possesses some unique properties, it induces some interesting properties on the design of the network topology. In this paper, we evaluate the performance of the large scale SCSI networks with linear and mesh structures.*

Keywords: message passing, SCSI, diameter, parallel computing, topology.

1 Introduction

Parallel computer system using network-connected multicomputers have been a popular alternative solution for the high performance computing world in recent years[4]. An example of such a low cost system, which makes use of off-the- shelf hardware products, was built by several Pentium based PCs running Unix and SCSI buses. It has been shown to have better performance when compared with the PVM system connected by Ethernet [1,6]. So, it is necessary to consider building large scale SCSI networks. Several network topologies of SCSI network have been proposed before [6]. In this paper, we study the linear and mesh SCSI networks

In Section 2, we describe message passing in SCSI network, and we identify the unique properties of SCSI network in the next section. In Sections 4 and 5, we give our results of the linear and mesh SCSI networks. Finally, we give some concluding remarks in the last section.

2 SCSI Network

In a SCSI network computers are connected by SCSI buses, and we refer to such computers as hosts. Every host can be connected by more than one SCSI bus. When one host wants to pass a message to another host on the same SCSI bus, message can be passed directly through that bus. However, when a host wants to send a message to another host on a different bus but these two buses share a common host, the message can first be passed to the common host, then the common host in turn passes the message to the destination host through the other SCSI bus. In this case, this common host is called a router. In the description above so far, it is similar to the conventional network. However, we will soon see the difference in the next section.

3 Unique Properties of SCSI Network

Computers in the SCSI network are connected by the SCSI buses and this off-the-shelf SCSI bus possesses some unique properties:

1. Each SCSI bus can only connect up to sixteen hosts.
2. Each router in SCSI network can only connect up to four SCSI buses.
3. No host can access the others within a SCSI link if the link failed.

One may immediately notice that a conventional hypercube of dimension greater than four cannot be constructed because such structure has degree greater than four for every node that contradicts the second property mentioned above. On the other hand, message passing between hosts connected by a single SCSI bus is much faster than that between hosts on different SCSI buses through routers. This unique feature gives SCSI network a totally different performance consideration even for some familiar network topologies such as linear and mesh structures.

In the following sections, when considering some large scale SCSI network topologies, we do not restrict our numbers 16 and 4 in properties 1 and 2 above. Instead, we generalize the properties 1 and 2 as follows:

1. Each SCSI bus can only connect up to $\sigma+1$ hosts.
2. Each router in SCSI network can only connect up to θ SCSI buses.

where σ and θ are some fixed integers.

4 Large Scale Linear SCSI Networks

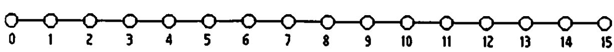


Figure 1 Linear SCSI Network

As in figure 1, the simplest linear SCSI network consists of simply one SCSI link connecting 16 hosts. This simple network has size 16, cost 1, degree 1 and arc connectivity 1. Since every node pairs can access each other directly, every shortest path has length 1. And hence both its diameter and average distance are 1. For other conventional linear network of the same size, for example, node 0 need to pass through 14 routers in order to reach node 15, hence this path has length 15 which is much greater than that of SCSI one. One can immediately see that it has much lower cost, diameter and average distance than that of conventional linear non-SCSI network of the same size.

The above simple linear structure can be easily extended to linear structure of scale n. A linear SCSI network of scale 3 is given in figure 2.

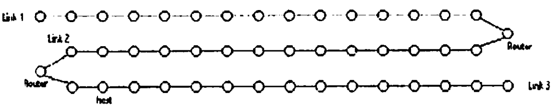


Figure 2 Linear SCSI Network of scale 3

In the above linear structure, adjacent SCSI links are connected by common hosts which serve as routers. For a linear network of scale n, the size is σn + 1 and the cost is n. Every non-router host has degree 1 while every router has degree 2. This large scale network also has arc connectivity as there exist degree 1 hosts in the network.

Labeling: We label the hosts by ascending integers running from zero, i.e. 0,1,2,3, … , that is, every router has label σk where k = 0,1, … ,n.

Now, we state and prove several simple but useful lemmas.

Lemma 1: For a non-router host a ∈ (σ(i-1), σi) and a host b ∈ (σ(j-1), σj] where a < b and $1 \leq i \leq j \leq n$, i.e. they are in the ith and jth SCSI links respectively, the length of the shortest path between them, denoted by d(a, b), is j – i + 1.

Proof: There is an obvious shortest path from a to b
a → σi → σ(i+1) → … → σ(j-1) → b
which has path length j – i + 1.

Corollary 2: For i = j in the above lemma, i.e. a and b are connected by the same SCSI link, d(a, b) is always 1.

Lemma 3: For a router $a = \sigma i$ and a host b in the interval $(\sigma(j-1), \sigma j]$ where $0 \le i < j \le n$ (hence $a < b$), we have $d(a,b) = j - i$.

Proof: There is an obvious shortest path from a to b
$a = \sigma i \rightarrow \sigma(i+1) \rightarrow \ldots \rightarrow \sigma(j-1) \rightarrow b$
which has path length $j - i$.

Now, we state the routing algorithm that can get the above shortest path for the linear SCSI network.

Routing Algorithm for linear SCSI network: Find the shortest path P from hosts x_1 to x_2.

```
begin
   P = φ;
   if (x1 ≠ x2) then
      k = ⌈(x1 / σ)⌉
      if (x2 ≤ kσ) then
         P = {(x1, x2)};
      else
         P = {(x1, kσ)};
         k = k + 1;
         do while (kσ < x2)
            P = P ∪ {((k-1)σ, kσ)};
            k = k + 1;
         enddo
         P = P ∪ {((k-1)σ, x2)};
      endif
   endif
end
```

Next, we give results for network diameter and average distance of the network.

Theorem 4: Network diameter of linear SCSI network of scale n is n.

Proof: By Lemmas 1 and 2, we see that one of the longest shortest path between all host pairs is $0 \to \sigma \to 2\sigma \to \ldots \to n\sigma$, that has path length n − 0 = n.

Theorem 5: Network average distance of linear SCSI network of scale n is

$$\frac{(\sigma-1)(n^2+3n-1)+(n+1)(n+2)}{3(\sigma n+1)}$$

Proof: First of all, we need to find out the total distance of all host pairs. We divide all the host pairs (a, b) into three cases:

Case 1: $a, b \in ((i-1)\sigma, i\sigma]$ where $1 \le i \le n$
There are totally ${}_{\sigma}C_2 = \sigma(\sigma - 1) / 2$ such host pairs. By corollary 2, d(a, b) = 1. Since there are n connected SCSI links, the total distance between these host pairs is $n\ {}_{\sigma}C_2$.

Case 2: $a \in ((i-1)\sigma, i\sigma)$ and $b \in ((j-1)\sigma, j\sigma]$ where $1 \le i < j \le n$
For each fixed $i, j \in [1, n]$, there are $\sigma(\sigma - 1)$ such host pairs. By lemma 1, $d(a, b) = j - i + 1$ so the total distance is

$$\sigma(\sigma-1)\sum_{1\le i<j\le n} d(a,b)$$

$$=\sigma(\sigma-1)\sum_{1\le i<j\le n}(j-i+1)$$

$$=\frac{\sigma(\sigma-1)n(n-1)(n+4)}{6}$$

Case 3: $a = i\sigma$ and $b \in ((j-1)\sigma, j\sigma]$ where $0 \le i < j \le n$
For each fixed $i \in [0, n]$, $j \in [1, n]$, there are σ such host pairs. By lemma 3, $d(a, b) = j - i$ so the total distance is

$$\sigma \sum_{0 \le i < j \le n} d(a,b)$$

$$= \sigma \sum_{0 \le i < j \le n} (j-i)$$

$$= \frac{\sigma n(n+1)(n+2)}{6}$$

Since the above three cases cover all the situations, the total distance between all host pairs is the sum of distances obtained from the 3 cases, i.e.

$$\frac{\sigma(\sigma-1)}{2} + \frac{\sigma(\sigma-1)n(n-1)(n+4)}{6} + \frac{\sigma n(n+1)(n+2)}{6}$$

$$= \frac{\sigma n}{6}\left[(\sigma-1)(n^2+3n-1)+(n+1)(n+2)\right]$$

Now, there are totally ${}_{\sigma n+1}C_2$ host pairs, therefore the network average distance of linear SCSI network of scale n is

$$\frac{\frac{\sigma n}{6}\left[(\sigma-1)(n^2+3n-1)+(n+1)(n+2)\right]}{\frac{\sigma n(\sigma n+1)}{2}}$$

$$= \frac{(\sigma-1)(n^2+3n-1)+(n+1)(n+2)}{3(\sigma n+1)}$$

This completes our proof of the theorem.

5 Large Scale Mesh SCSI Networks

Mesh is a very popular topology since it possesses rather good network parameters. So, we build our SCSI network using this topology and hence we have the simple mesh SCSI network.

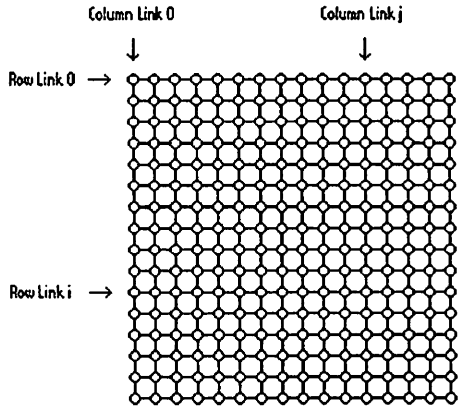


Figure 3 Simple Mesh Network

As in figure 3, this simple network has size 256 (i.e. 16^2), cost 32 (16 row links, 16 column links) and degree 2. It is obvious that failure of any one link cannot disconnect the network. Since there are hosts of degree 2, the arc connectivity is 2 for this simple mesh network. This topology has some interesting properties, that is, all hosts are also routers.

Labeling: For the sake of clear description, we label the hosts as coordinates (i, j) where $i, j \in [0, \ldots, 15]$.

In order to determine the diameter and average distance, we first give some lemmas.

Lemma 6: For two hosts a, b having the same x- or y-coordinate, d(a, b) = 1.

Proof: Suppose $a = (x, y_1)$ and $b = (x, y_2)$ having the same x-coordinate. Then in fact a and b are on the same row SCSI link x, hence a can directly access b without passing through any router. Therefore d(a, b) = 1. The case of hosts having the same y-coordinate is similar.

Lemma 7: For two hosts a, b having different x- and y-coordinates, d(a, b) = 2.

Proof: Suppose $a = (x_1, y_1)$ and $b = (x_2, y_2)$ where $x_1 \neq x_2$ and $y_1 \neq y_2$. Obviously, they are connected by different links. In order to reach each other, they must pass through at least one router, so $d(a, b) > 1$. Now consider the following path

$a = (x_1, y_1) \rightarrow (x_1, y_2) \rightarrow b = (x_2, y_2)$

which has path length 2 and hence the result follows.

Theorem 8: Network diameter of simple mesh SCSI network is 2.

Proof: Result follows from lemma 6 and 7.

Theorem 9: Network average distance of simple mesh SCSI network is 1.882.

Proof: Suppose $a = (x_1, y_1)$ and $b = (x_2, y_2)$ where $x_1, x_2, y_1, y_2 \in [0,\dots 15]$. We divide the situations into two cases.

Case 1: a, b having the same x- or y-coordinate

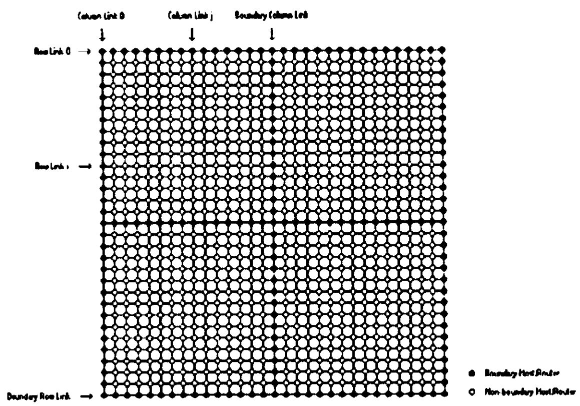
For each fixed a, there are 15 hosts on the same column link and 15 hosts on the same row link. Since there are 256 hosts in the network, the total number of host pair is $30 \times 256 / 2 = 3840$. By lemma 3.4.1, d(a, b) = 1, the total distance between these host pairs is 3840.

Case 2: a, b having different x- or y-coordinate

In this case, the total number of host pair is $_{256}C_2$ - 3840 = 28800. By lemma 3.4.2, d(a, b) = 2, the total distance between these host pairs is 28800 × 2 = 57600.

By the result of the two cases, d(a, b) = (3840 + 57600) / $_{256}C_2$ = 1.882.

In view of the above results, the diameter and average distance of this mesh SCSI network has a much better performance than that of conventional mesh network. It is really an interesting result and mainly due to the unique property of SCSI link that any of the 16 hosts connected by a single SCSI link can access each other without passing through any router. So, the message transmission time will be largely saved and hence we have no reason to limit this mesh topology to the

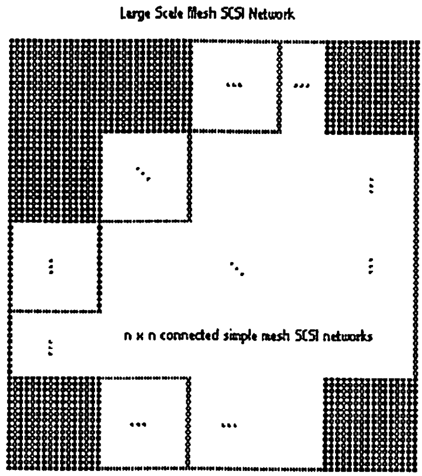


size of 256. Now, we will demonstrate the scalability of this mesh topology.

Figure 4 Mesh SCSI Network of scale 2

Similar to large scale linear SCSI network, we can connected four simple mesh SCSI networks, in the obvious way, into a larger mesh SCSI network as in figure 4.

Following the same ideas, we can repeatedly compose n^2 simple mesh SCSI network into a large scale mesh SCSI network or a mesh SCSI



network of scale n.

Figure 5 Mesh SCSI Network of scale n

We define the mesh SCSI network of scale n as the interconnection network of n^2 simple mesh SCSI networks as shown in figure 5. This giant network topology has size $(\sigma n + 1)^2$ and cost $2(\sigma n+1)$. With the same reason as in the case of simple mesh network, the arc connectivity of the network is also 2. Moreover, all the hosts function as routers and we also label each host according to its x-, y-coordinates. Those hosts connecting two or more simple mesh networks have degree 3 or 4. Note that degree 4 is the upper limit of SCSI network. So, we cannot build 3D mesh SCSI network by the conventional way because it will produce a degree 5 host which is not possible by the off-the-shelf SCSI adapter. Although we may do so by some tricky way, that is, connecting the third dimension simple SCSI network by some degree 2 routers instead of degree 3 ones, we are not going to discuss this in detail in this report. Now we call these hosts/routers together with those hosts in the network physical boundary the boundary hosts/routers. Then all the non-boundary hosts/routers still have degree 2. Now we are going to determine the network diameter and average distance which is non-trivial in this case. Besides, we will see that the calculation of average distance is in fact similar to that of large scale linear SCSI network. As in the previous cases, we first give some useful lemmas.

Labeling: Similar to the simple mesh network, we label the hosts as coordinates (x, y) where $x, y \in [0, \dots, n\sigma]$. *Then a host (x, y) is also a router if and only if* $x = i\sigma$ *and* $y = j\sigma$ *for some integers i, j.*

Lemma 10: Given two hosts $a = (x_1, y_1)$ *and* $b = (x_2, y_2)$. *The path*
$a \to (\sigma i, y_1) \to (\sigma(i+1), y_1) \to \dots \to (x_2, y_1) \to (x_2, \sigma j) \to (x_2, \sigma(j+1)) \to \dots \to b$
is always a shortest path from host a to host b where $(x_2, \sigma j)$ *and* $(\sigma i, y_1)$ *are all the boundary routers between* $a = (x_1, y_1)$ *and* (x_2, y_1) *in the same horizontal level and between* (x_2, y_1) *and* $b = (x_2, y_2)$ *in the same vertical level respectively.*

Proof: The above lemma may seem complicated at the first sight but in fact we just take the most straight forward path as shown in figure 6.

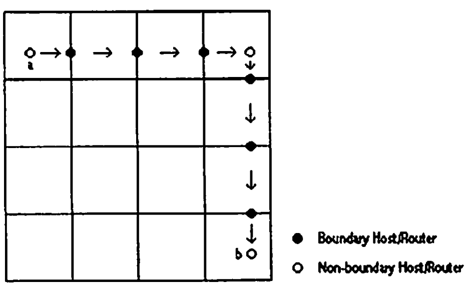


Figure 6 Shortest Path in Mesh Topology

If a and b are in the same simple mesh network, the result follows from the proofs of lemmas 6 and 7. Now suppose that they are in different simple mesh network. No matter which path we choose, a must pass through all boundary column links and boundary row links in between in order to reach b. Then it is obvious that the above path has the minimum necessary length and hence it is a shortest path.

Corollary 11: $d(a, b) = d_y(x_1, x_2) + d_x(y_1, y_2)$ which are the lengths of shortest paths between host pairs x_1, x_2 and y_1, y_2 respectively regarded as hosts in the large scale linear SCSI networks described in the previous section.

Proof: In the shortest path given in lemma 10, when we only consider the hosts having fixed x-coordinate x_1, it is equivalent to the linear SCSI network of scale n. Since the path portion
$a = (x_1, y_1) \rightarrow (\sigma i, y_1) \rightarrow (\sigma(i+1), y_1) \rightarrow \ldots \rightarrow (x_2, y_1)$
falls completely into this sub-network, its path length must be $d_y(x_1, x_2)$. Similarly, the remaining path portion
$(x_2, y_1) \rightarrow (x_2, \sigma j) \rightarrow (x_2, \sigma(j+1)) \rightarrow \ldots \rightarrow b = (x_2, y_2)$
must have path length $d_x(y_1, y_2)$. Hence the result follows.

We know what a shortest path looks like. We now give the routing algorithm that can find the path.

Routing Algorithm for mesh SCSI network: Find the shortest path P from hosts (x_1, y_1) to (x_2, y_2).

begin
  $P_x = \phi$;
  $P_y = \phi$;
  if $(x_1 \neq x_2)$ then
By fixing the y-coordinate y_1, we can apply the Routing Algorithm for linear network to get the shortest path P_x from (x_1, y_1) to (x_2, y_1);
  endif
  if $(y_1 \neq y_2)$ then
By fixing the x-coordinate x_2, we can apply the Routing Algorithm for linear network to get the shortest path P_y from (x_2, y_1) to (x_2, y_2);
  endif
$P = P_x \cup P_y$
end

Here we give the network diameter and average distance of the network.

Theorem 12: Network diameter of mesh SCSI network of scale n is $2\sigma n$.

Proof: It is obvious that the hosts a = (0,0), b = (σn,σn) are one of the most distant pairs in the network. By corollary 11, d(a, b) = $d_y(0, \sigma n)$ + $d_x(0, \sigma n)$ = n + n = 2n.

Theorem 13: Network average distance of mesh SCSI network of scale n is

$$\frac{\frac{\sigma n(\sigma n+1)^2}{3}\left[(\sigma-1)(n^2+3n-1)+(n+1)(n+2)\right]}{\frac{(\sigma n+1)^2\left[(\sigma n+1)^2-1\right]}{2}}$$

$$=\frac{2}{3(\sigma n+2)}\left[(\sigma-1)(n^2+3n-1)+(n+1)(n+2)\right]$$

Proof: We first determine the total distance between all distinct host pairs. Suppose that we consider all possible host pairs including those of the form (a, b), (b, a) and (a, a). For the (a, a), d(a, a) = 0, so it won't affect the total distance. Besides, since d(a, b) = d(b, a) by symmetry, we just double the total distance. Hence the total distance is

$$\frac{1}{2}\sum_{all\ hosts\ a,b} d(a,b)$$

$$=\frac{1}{2}\sum_{x_1,x_2,y_1,y_2}^{\sigma n+1} d((x_1,x_2),(y_1,y_2)) \quad where \quad a=(x_1,x_2),\ b=(y_1,y_2)$$

$$=\frac{1}{2}\sum_{x_1,x_2,y_1,y_2}^{\sigma n+1} d_y(x_1,x_2)+d_x(y_1,y_2)$$

$$=\frac{2}{2}\sum_{x_1,x_2,y_1,y_2}^{\sigma n+1} d_y(x_1,x_2) \quad \text{by symmetry}$$

$$=(\sigma n+1)^2\sum_{x_1,x_2}^{\sigma n+1} d_y(x_1,x_2)$$

$$= 2(\sigma n + 1)^2 \sum_{\substack{x_1, x_2 \\ x_1 < x_2}}^{\sigma n + 1} d_y(x_1, x_2) \quad \text{by symmetry}$$

$$= 2(\sigma n + 1)^2 \times \text{total distance of linear SCSI netowork of scale } n$$

$$= 2(\sigma n + 1)^2 \times \frac{\sigma n}{6} \left[(\sigma - 1)(n^2 + 3n - 1) + (n + 1)(n + 2) \right]$$

$$= \frac{\sigma n(\sigma n + 1)^2}{3} \left[(\sigma - 1)(n^2 + 3n - 1) + (n + 1)(n + 2) \right]$$

Since there are totally $_{(\sigma n+1)}\ _{(\sigma n+1)}C_2$ distinct host pairs, the network average distance is

$$\frac{\frac{\sigma n(\sigma n + 1)^2}{3} \left[(\sigma - 1)(n^2 + 3n - 1) + (n + 1)(n + 2) \right]}{\frac{(\sigma n + 1)^2 \left[(\sigma n + 1)^2 - 1 \right]}{2}}$$

$$= \frac{2}{3(\sigma n + 2)} \left[(\sigma - 1)(n^2 + 3n - 1) + (n + 1)(n + 2) \right]$$

This completes the proof of the theorem.

6 Conclusion

In this paper, we introduce and give results for the large scale linear and mesh SCSI Networks. For future research, other large scale SCSI networks, such as tree, ring, torus and hypercube, could be built and studied. Also we can simulate these giant SCSI networks based on the statistical information taken from the smaller systems. Another interesting research direction is on the embedding issues of SCSI networks.

Reference

[1] Geist G.A., Sunderam V.S., *The evolution of the PVM concurrent computing system*. Proceedings of the 26th IEEE Compcon Symposium, pp. 471-478, 1993.

[2] Schmidt F., *The SCSI bus and IDE interface – protocols, applications and programming*, Addison-Wesley, 1995.

[3] *SCSI-2 Specification*, Document X3.131-1994, ANSI.

[4] Wilkinson B., Allen M., Parallel Programming: Techniques and Applications Using Networked Workstations and parallel Computers, Prentice hall, 1999..

[5] Young G., Yiu, V., *IRI System: Multi-computer Architecture for Parallel and Distributed Processing*, Proceeding of International Conference on Parallel and Distributed Processing Techniques and Applications, pp. 610-618, 1996,.

[6] Young, G., Yiu, V., Wan L, *Parallel Computing on SCSI Network*, Proceedings of IEEE National Aerospace and Electronics, pp. 216-221, 1997.