Exact Formulas for the Average Internode Distance in Mesh and Binary Tree Networks

Behrooz Parhami

Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106-9560, USA
*Corresponding Author: parhami@ece.ucsb.edu

Abstract   The average internode distancein an interconnection network (or its average distance for short) is an indicator of expected message latency in that network under light and moderate network traffic. Unfortunately, it is not always easy to find an exact value for the average internode distance, particularly for networks that are not node-symmetric, because the computation must be repeated for many classes of nodes. In this short paper, we derive exact formulas for the average internode distance in mesh and complete binary tree networks.

Keywords   Complete Binary Tree; Diameter; Expected Message Latency; Graph Theory; Interconnection Network; Linear Array; Mesh Network; Parallel Processing; Ring Network; Routing Algorithm; Torus Network.

1. Introduction

A variety of interconnection networks have been studied for linking the nodes in a parallel or distributed system [1-4]. It is often difficult to compare such networks with respect to their suitability for a particular application domain[5], given the multitude of static attributes (diameter, average distance, bisection width, VLSI layout area) and dynamic properties (routing algorithms, deadlock prevention, traffic balance, fault tolerance) to be taken into consideration. New interconnection networks that continue to be introduced, while enriching the repertoire of parallel computer designers, also add to the selection difficulty.

In this paper, we focus on a particular static attribute of a network, its average internode distance Δ, and derive exact formulas for it in the case of two highly popular, but node-asymmetric, networks. The derivation for meshes takes advantage of the fact that average distance in a q-dimensional mesh is the sum of average distances along the q dimensions. The derivation in the case of the complete binary tree utilizes the network's recursive structure.

While the average internode distance Δ is widely acknowledged as an important static parameter of an interconnection network, exact formulas for it have not been published for many useful networks, including the two widely used ones: meshes and binary trees. By contrast, we almost always have exact formulas for network diameter D, in part because it is easier to compute. We take small steps in correcting this deficiency.

We define the average internode distance of an interconnection network to be the mean distance from a randomly chosen node to every node in the network, including the node itself. This inclusion of null paths, that is counting paths from each node to itself in computing the average distance, leads to cleaner results in most cases and has negligible effect on our results. Furthermore, given that inclusion of length-0 paths does not affect the various sums of distances that we compute, we can resort to the more conventional definition by simply dividing at the end the overall sum of all n^2 path lengths in an n-node network by $n(n-1)$ rather than n^2, leading to a slightly larger average.

2. Average Internode Distance in Meshes

Nodes in a q-dimensional $n_1 \times n_2 \times \ldots \times n_q$ mesh are labeled by q-tuples $x_1 x_2 \ldots x_q$, where x_i ($1 \le x_i \le n_i$) is the dimension-i address of the node. Each node $x = x_1 x_2 \ldots x_q$ is connected to up to $2q$ nodes whose labels differ from that of x by 1 in exactly one of the q dimensions. A q-dimensional torus is similarly defined, except that every node has exactly $2q$ neighbors due to the inclusion of wraparound links that connect the last node along each dimension to the first node. For example, node 1234, which is connected to nodes 2234, 1134, 1334, 2244, 2224, and 2233 in a $4 \times 4 \times 4 \times 4$ mesh, is also connected to nodes 4234 and 1231 in the corresponding $4 \times 4 \times 4 \times 4$ torus. Mesh and torus networks have been quite popular in modern supercomputers [1, 3].

The shortest path length in a q-dimensional mesh can be found by adding the distances of the destination node from the source node along each of the q dimensions. Thus, all that is required for finding an exact formula for Δ is to have an exact formula for the average distance in a p-node linear array, also known as a p-path (Fig. 1).
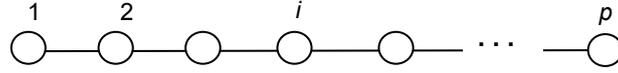
Figure 1. Linear array with p nodes (a p-path).

$$\Delta_{p\text{-path}} = (1/p^2)\Sigma_{0\leq j\leq p-1}[\Sigma_{0\leq i\leq j}(j - i) + \Sigma_{j\leq i\leq p-1}(i - j)] \tag{1}$$

Recall that we include the 0-length path from each node to itself in calculating the average, hence the normalizing term $1/p^2$ in the expression above, in lieu of the often-used $1/[p(p - 1)]$. The two sums within the square brackets are sums of distances from node j to all to nodes to its left and right, respectively. Using the formulas for the sums of consecutive integers and their second powers, we obtain for $p\geq 3$:

$$\Delta_{p\text{-path}} = (1/p^2)\Sigma_{0\leq j\leq p-1}[j(j + 1) - j(j + 1)/2 + (p - j)(p - 1 + j)/2 - j(p - j)] \tag{2}$$
$$= (1/p^2)\Sigma_{0\leq j\leq p-1}[j^2 - (p - 1)j + p(p - 1)/2]$$
$$= (1/p^2)[p(p - 1)(2p - 1)/6 - p(p - 1)^2/2 + p^2(p - 1)/2]$$
$$= (1/3)(p - 1/p)$$

The average internode distance in a q-dimensional $n_1 \times n_2 \times \ldots \times n_q$ mesh is thus:

$$\Delta_{q\text{D-mesh}} = (1/3)[\Sigma_{1\leq i\leq q}(n_i - 1/n_i)] \tag{3}$$

When the dimensions n_i are large, the average internode distance in (3) is roughly one-third of the diameter in (4):

$$D_{q\text{D-mesh}} = \Sigma_{1\leq i\leq q}n_i - q \tag{4}$$

The average internode distance and diameter for a p-ring are similarly derived in (5) and (6), again assuming $p\geq 3$, leading to the average internode distance and diameter for a q-dimensional $n_1 \times n_2 \times \ldots \times n_q$ torus network in (7) and (8):

$$\Delta_{p\text{-ring}} = [p - (p \bmod 2)/p]/4 \tag{5}$$

$$D_{p\text{-ring}} = [p - (p \bmod 2)/p]/2 \tag{6}$$

$$\Delta_{q\text{D-torus}} = (1/4)\Sigma_{1\leq i\leq q}[n_i - (n_i \bmod 2)/n_i] \tag{7}$$

$$D_{q\text{D-torus}} = (1/2)\Sigma_{1\leq i\leq q}[n_i - (n_i \bmod 2)/n_i] \tag{8}$$

Note that the average internode distance of a p-ring is $p/4$ when p is even and it is slightly less when p is odd. Note also that the Δ/D ratio for a ring or torus is always $1/2$.

3. Average Internode Distance in Binary Trees

Deriving an exact formula for the average internode distance of a binary tree is much harder and, to the best of our knowledge, no such formula has been published. We characterize an n-node complete binary tree (Fig. 2) with the parameter $m = n + 1 = 2^l$; we say that the tree has l levels, numbered from 1, for the root, up to l, for the leaves. So, T_m, $m = 2^l\geq 2$, refers to an $(m - 1)$-node complete binary tree. The use of the power-of-2 parameter m in lieu of the network size n leads to some simplifications in mathematical derivations, as we will see shortly.
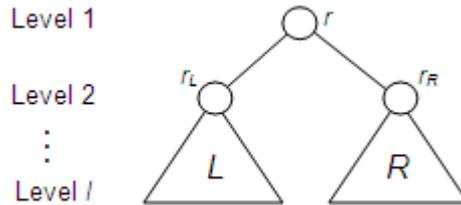

Figure 2. Complete binary tree architecture composed of root r, left subtree L, and right subtree R.

First, some observations. Given that to get from a leaf node in the left subtree to another leaf node in the right subtree the path must go through the root node r, the diameter of T_m is:

$$D_{\text{binary-tree}} = 2l - 2 = 2 \log_2 m - 2 \qquad (9)$$

Intuitively, the average internode distance of T_m should be close to D, because each leaf node in the left (right) subtree is D hops away from each leaf node in the right (left) subtree. So, more than half of the n^2 possible paths in T_m are of length D, with the others ranging in length from 0 to $D - 1$.

We begin our derivation of $\Delta_{\text{binary-tree}}$ by calculating the sum $\sigma(T_m)$ of the lengths of the paths from the root r to every node in T_m.

$$\sigma(T_m) = 1 \times 2^1 + 2 \times 2^2 + \ldots + (l - 1) \times 2^{l-1} = (l - 2)2^l + 2 \qquad (10)$$
$$= m \log_2 m - 2m + 2$$

Now T_m can be viewed as consisting of three parts: The root node r, the left subtree L, which is a $T_{m/2}$ with root node r_L, and the right subtree R, also a $T_{m/2}$ with root node r_R (see Fig. 2). To find the sum $S(T_m)$ of the lengths of all paths in T_m, we note that each such path must begin and end in one of the 3 parts, creating a total of 8 cases (ignoring the ninth case of a null path from r to r).

$$S(L, L) = S(R, R) = S(T_{m/2}) \qquad (11)$$
$$S(r, L) = S(r, R) = S(L, r) = S(R, r) = m/2 - 1 + \sigma(m/2) \qquad (12)$$
$$S(L, R) = S(R, L) = (m/2 - 1)^2[2 + 2\sigma(m/2)/(m/2 - 1)] \qquad (13)$$
$$= (m-2)\sigma(m/2) + (m-2)^2/2 \qquad (13)$$

In (12), each of the $m/2 - 1$ paths is one hop longer than a corresponding path beginning at a subtree root. In (13), each of the $(m/2 - 1)^2$ paths is two hops longer than the sum of two paths, one beginning at each of the two subtree roots. Substituting $\sigma(m/2) = (m \log_2 m)/2 - 3m/2 + 2$ in $S(T_m) = 2S(L, L) + 4S(r, L) + 2S(L, R)$ and simplifying, we arrive at the recurrence:

$$S(T_m) = 2S(T_{m/2}) + m^2 \log_2 m - 2m^2 + 2m \qquad (14)$$

The recurrence in (14) has a solution of the form:

$$S(T_m) = am^2 \log_2 m + bm^2 + cm \log_2 m + dm + e \qquad (15)$$

Substituting in (14) and equating the coefficient of the terms of various orders on both sides, we arrive at $a = 2$, $b = -6$, $c = 2$, and $e = 0$, leading to:

$$S(T_m) = 2m^2 \log_2 m - 6m^2 + 2m \log_2 m + dm \qquad (16)$$

Finally, from (16), $d = 6$ follows based on the initial condition $S(T_2) = 0$. Thus, we arrive at the final solution (17) and average internode distance (18):

$$S(T_m) = 2m^2 \log_2 m - 6m^2 + 2m \log_2 m + 6m \qquad (17)$$
$$\Delta(T_m) = (2m^2 \log_2 m - 6m^2 + 2m \log_2 m + 6m)/(m - 1)^2 \qquad (18)$$
$$= 2 \log_2 m - 6 + 2(3m \log_2 m - 3m - \log_2 m + 3)/(m - 1)^2$$

Note that for very large m, the average internode distance of the $(m - 1)$-node complete binary tree T_m asymptotically approaches $2 \log_2 m - 6$, that is, 4 hops less than the diameter in (9); a rather counterintuitive result. As numerical examples, the average internode distance of 7-node, 15-node, 31-node, 63-node, … , and 1023-node complete binary trees are 1.959, 3.271, 4.795, 6.482, … , and 14.026, respectively. The last figure is 3.974 less than the diameter 18 of the network. The Δ/D ratios for the trees just listed are 0.490, 0.545, 0.599, 0.648, … , and 0.779. As expected, the difference $D - \Delta$ increases uniformly, approaching 4 rather quickly. The Δ/D ratio increases and asymptotically tends to 1 as the tree size grows. Given the result just obtained for a 1023-node tree, in a highly parallel system having the binary tree architecture, $D - 4$ is a good estimate for the average internode distance Δ.

Figure 3 depicts the parameters Δ and D as functions of the number l of levels in the binary tree, confirming that the trend line for Δ eventually becomes parallel to, and stays about 4 units below, the line for D.
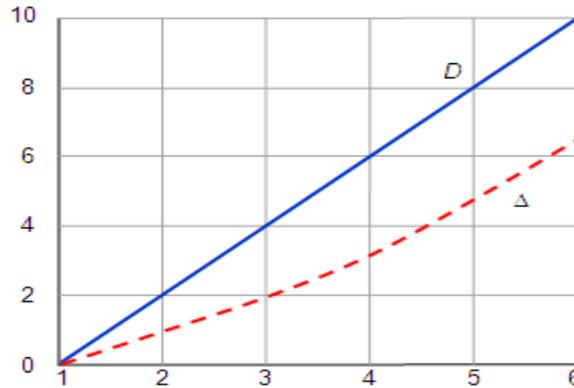
Figure 3. Average internode distance Δ and diameter D of an l-level binary tree as functions of l.

4. Conclusion

We have derived exact formulas for the average internode distance Δ in q-dimensional meshes and complete binary trees. Previously, we have proven [6] that in a node-symmetric network, the two parameters Δ and D are related by:

$$D/2 \leq \Delta \leq D \qquad (19)$$

Of course, neither meshes nor binary trees are node-symmetric, so, the fact that they don't satisfy (19) isn't surprising. Our results here indicate that whereas meshes do not quite satisfy the relationships in (19), they do satisfy the laxer relationship:

$$D/3 \leq \Delta \leq D \qquad (20)$$

Trees violate the relationship (19) only for very small network sizes (see Fig. 3), but their average-internode distance approaches the diameter asymptotically. In other words, given any arbitrarily small value ε, the following relationship holds for all but a finite number $h(\varepsilon)$ of complete binary trees.

$$(1 - \varepsilon)D < \Delta < D \qquad (21)$$

Overall, it is still true that in interconnection networks of practical interest, average internode distance and diameter are related to each other.

We close by reiterating that network diameter, average internode distance, and other topological properties are not as unimportant as some researchers have claimed in the past. The space of possibilities for network architectures and associated routing algorithms is vast, the choice not being limited to low- versus high-dimensional meshes/tori and variations on wormhole switching [7-8], although these choices have been dominant in the recent past. As we continue to discover both new network topologies and novel interconnection technologies, optimality criteria and the optimal balance may change considerably. So, it is quite dangerous to generalize from a small number of high-level purely experimental studies.

REFERENCES

[1] J. Duato, S. Yalamanchili, and L. Ni., *Interconnection Networks: An Engineering Approach*, Morgan Kaufmann, 2002.

[2] H. Haddadi, M. Rio, G. Iannaccone, A. Moore, R. Mortier, "Network Topology: Inference, Modeling, and Generation," *IEEE Communications Surveys & Tutorials*, Vol. 10, No. 2, pp. 48-69, 2008.

[3] B. Parhami, *Introduction to Parallel Processing: Algorithms and Architectures*, Plenum, 1999.

[4] J. Xu, *Topological Structure and Analysis of Interconnection Networks*, Kluwer, 2001.

[5] K. J. Liszka, J. K. Antonio, and H. J. Siegel, "Problems with Comparing Interconnection Networks: Is an Alligator Better than an Armadillo?" *IEEE Concurrency*, Vol. 5, No. 4, pp. 18-28, October-December 1997.

[6] B. Parhami and C.-H. Yeh, "Why Network Diameter Is Still Important," *Proc. International Conf. Communications in Computing*, June 2000, pp. 271-274.

[7] B. M. Maggs, "A Critical Look at Three of Parallel Computing's Maxims," *Proc. Int'l Symp. Parallel Architectures, Algorithms and Networks*, 1996, pp. 1-7.

[8] J. Duato, "Why Commercial Multicomputers Do Not Use Adaptive Routing," *IEEE Computer Architecture Technical Committee Newsletter*, pp. 20-22, Summer/Fall 1994.