Visualization of a Simple Routing Scheme for Meshes

Nelson Ge
Communication Systems
Intergraph Corporation
Huntsville, AL 35894

John K. Antonio
School of Electrical Eng.
Purdue University
West Lafayette, IN 47907

Sharada V. Vitalpur
Electrical & Comput. Eng.
Northeastern University
Boston, MA 02155

Abstract

We propose a simple quasi-static routing scheme for buffered mesh connected network computers. Under the assumption of uniform traffic demands, it is demonstrated numerically that our simple routing scheme achieves near-optimal performance; in the sense of minimizing the total number of outstanding packets in a Jackson-type network. Simulation results are also presented for the case of non-uniform traffic demands. These simulations illustrate that our quasi-static control scheme can dynamically adjust the values of routing probabilities so as to provide acceptable and stable network performance. The results of these simulations are displayed using advanced video graphics.

1 Introduction and Background

1.1 The 2-Dimensional Mesh

Fig. 1 depicts a 9×9 (= 81-node) 2-D mesh network topology. Throughout the paper, the term "n-node mesh" is understood to be a $\sqrt{n} \times \sqrt{n}$ 2-D planar mesh, such as the one shown in Fig. 1.[1] Each node in the graph represents a processor/memory pair. An edge between two nodes, say nodes A and B, actually denotes two separate buffered communication channels; one from A to B and another from B to A.

In this paper we assume more general purpose applications than those studied in references [1–3]. In particular, we consider applications in which there is a significant amount communication demand between nodes which are not nearest neighbors. Thus, the issue of properly routing packets of data through the network becomes an important one.

1.2 Optimal Routing

The overall goal of optimal routing schemes is to route packets of data in such a way that some quantitative measure of network performance is optimized. Obviously, the choice of the performance measure assumed has a direct impact on the resulting "optimal" routing solution, and is by itself an area of interesting research, see for example reference [4].

In this paper we propose a simple *quasi-static* routing scheme in which the packets associated with each origin-destination (OD) pair are sent along one of two possible "L-shaped" shortest hop paths. Each L-shaped path is defined as either a boundary-L (B-L) path or a center-L (C-L) path, depending on whether it uses a majority of links closer to the boundary or the center of the mesh, see Fig. 1. The decision of which L-shaped path to use is made by each origin node in a probabilistic manner.

Under the assumption that packet generation rates for all OD pairs are uniform, we show that a static routing scheme in which B-L/C-L path selection is made uniformly (i.e., 50% B-L and 50% C-L) produces essentially optimal performance, in terms of minimizing the steady-state average number of outstanding packets in the network. An adaptive version is also described in which the probability of routing along B-L paths is changed in an attempt to produce acceptable performance in the case of non-uniform demand patterns.

2 The Network Model

The following formulation uses the same notations and is based on the same approximating assumptions as set forth by Bertsekas and Gallager in reference [5].

2.1 Delay Models

Perhaps the simplest queuing model is the so-called $M/M/1$ queuing system which consists of a single queuing station and a single server. It is assumed that customers (i.e., packets of data) arrive according to a Poisson process with rate F, and the probability distribution of the service rate is exponential with mean C. By applying Little's Theorem, the average delay for a packet to traverse link (i, j) is given by

$$D_{ij} = \frac{1}{C_{ij} - F_{ij}}, \tag{1}$$

where C_{ij} and F_{ij} denote the service rate and arrival rate respectively, associated with link (i, j).

Based on Jackson's Theorem and equation (1), our objective function is defined as a weighted sum of all link delays:

$$D(F) = \sum_{(i,j)\in\mathcal{L}} \frac{F_{ij}}{C_{ij} - F_{ij}}, \tag{2}$$

where links having more traffic flow are given higher relative weightings. Note that each term in the sum represents the average size of the queue associated

[1] It should be noted that the routing scheme presented in this paper can be extended to higher dimensional meshes, however, due to the space limitations we consider only the 2-D case here.

0-8186-2672-0/92 $03.00

with link (i,j). Therefore, $D(F)$ is an estimate of the total number of outstanding packets in the network. For the purposes of this paper, determining routes which minimize $D(F)$, for a given set of OD traffic demands, will constitute the notion of an optimal routing.

2.2 Formulating the Optimal Routing Problem

The following notation is required in order to formally state the optimal routing problem.
Preliminary Notation:

W : The set of OD pairs requesting communication.

w : A generic OD pair in W.

r_w : The arrival rate measured in packets/sec, for the OD pair w.

P_w : For the OD pair w, this is the set of all logical paths connecting the origin node to the destination node.

p : A generic path in P_w.

x_p : The flow rate on the logical path p.

The following constraint equations arise naturally due to conservation of flow.
Constraint Equations:

$$x_p \geq 0 \qquad \text{for all } p \in P_w,\ w \in W, \tag{3}$$

$$\sum_{p \in P_w} x_p = r_w \qquad \text{for all } w \in W \tag{4}$$

and

$$F_{ij} = \sum_{\substack{\text{all paths } p \\ \text{containing } (i,j)}} x_p. \tag{5}$$

Equation (3) is an obvious nonnegativity constraint for the flow rates on all logical paths. Equation (4) ensures that the total flow rate on all paths from O_ω to D_ω is equal to the requested demand, r_ω. Equation (5) states that the total flow on link (i,j), denoted F_{ij}, equals the sum total of all path flows x_p for which link (i,j) belongs.

By combining the above constraints with the objective function of equation (2), we state the optimal routing problem as follows.
The Optimal Routing Problem:

$$\text{Given } r_w, \qquad \text{for each } w \in W,$$

$$\text{minimize } \{D(F)\},$$

such that equations (3), (4) and (5) are satisfied.

2.3 Exact Solution of the Optimal Routing Problem

The optimal routing problem stated above can be solved numerically by using well established techniques from nonlinear programming such as the gradient projection (GP) method, see for example reference [6]. Also, due to the convexity of the objective function and the constraint equations, the solution is a global optimal. The main idea of the GP method is that after shifting path flow in the direction of the negative gradient, the result is orthogonally projected onto the positive orthant.

A sophisticated routing scheme such as the distributed GP routing algorithm *could* be applied to a mesh connected network computer. However, the question of interest is whether the required communication and computational overhead would be outweighed by the amount of gain achieved in network performance, see [7] for more details on the time complexity of the GP algorithm. As an alternative, in the next section we propose a simple routing scheme which requires much less overhead than a true optimal routing scheme (such as the GP algorithm), but in practice is able to achieve acceptable performance.

3 The Simple Routing Scheme

3.1 Some Analytical Insight

Proposition: Consider the optimal routing problem on an n-node mesh in which $C_{ij} = C$, for all links (i,j). Also, assume that the set of logical paths associated with each OD pair are restricted to the shortest hop domain.

If there exists a constant $F < C$ such that $F_{ij} = F$ satisfies constraint equations (3) through (5), then $F_{ij} = F$ for all links (i,j) is an optimal set of link flows.

Proof: See [8].

The above proposition states that having uniform link flows is a *sufficient* condition for optimality (under the assumptions that all link capacities are the same and all active paths reside in the shortest hop domain). From numerical studies we discovered that this condition is not *necessary* whenever the demands are uniform (i.e., all r_w's equal to one, for example). However, we did discover that in practice, optimal link flows are often *nearly* uniform.

The proposition does offer us the following insight: If a routing can be found in which the link flows are *close* to uniform, then, intuitively, the routing should be *close* to optimal. This idea is the basis of our simple routing scheme described below.

3.2 Description of the Simple Routing Scheme

Our routing scheme can be viewed as a "B-L/C-L path constrained" version of the standard optimal routing problem. That is, we constrain the set of active paths associated with each OD pair to be one of the two possible L-shaped paths, namely, either the B-L or C-L path as mentioned in the introductory section.[2] The precise definitions of B-L and C-L paths are included in [8]. It should be understood that proper definitions of B-L and C-L paths are at the heart of the success of our algorithm. It is shown in [8] that the logic required to determine B-L and C-L paths is quite simple.

Our routing scheme is summarized as follows. For each active OD pair, one of the two possible "L-shaped" shortest hop paths is selected by the origin node. The decision of which L-shaped path to use (i.e., B-L or C-L) is based on the current value of a locally stored parameter p_B; which represents the probability that the B-L path is selected (and thus the probability of selecting the C-L path is $p_C = 1 - p_B$). Under "normal operating conditions," meaning the length of each queue in the network is below a given threshold (say T), then the value of the parameter p_B remains unchanged at all nodes. However, if one or more queue lengths exceed the prescribed threshold, then the node(s) associated with these congested queue(s) broadcast an appropriate binary control message which instructs all other nodes to either increase or decrease their current value of p_B by a previously decided amount, say Δp_B. For instance, if a link near the boundary of the network becomes congested, then a control message of "decrease p_B" is broadcast to all other network nodes. The main premise of the algorithm is that whenever the number of packets in each network queue is below the prescribed threshold, then no control action is needed, and thus none is applied. On the other hand, if the size of one or more of the queues grows too large, then the simple control action can potentially shift flow away from the congested areas.

To show the potential "controllability" that can be achieved by changing the value of p_B, refer to the plots shown in Fig. 2. The height of the surfaces represent the average amount of flow associated with each underlying node, for various values of p_B. In all cases, it is assumed that $r_w = 1$ for all possible w's (i.e., a uniform communication demand). Note that as the value of $p_B \rightarrow 1$, the average flow on the links near the boundary of the mesh increase, as expected. Likewise, as the value of $p_B \rightarrow 0$, the height of the surface around the center of the network increases.

3.3 Uniform Traffic

Table 1 summarizes the results of a numerical study based on a 5×5 mesh with capacities $C_{ij} = 40$, for all links (i,j), and $r_w = 1$, for all $w \in \{(i,j) : i \neq j\}$ (i.e., uniform traffic). Note that the optimal value of the objective function (obtained by executing the GP algorithm of [6]) is also obtained by using our simple static routing scheme with half the flow routed along B-L paths and half along the C-L paths (i.e., $p_B = 0.5$). Also, from the maximum/minimum link utilization column note that at the optimal solution the link flows are "close" to being uniform.

[2] Actually, for those OD pairs residing on the same row or column, there is only one shortest hop path. Therefore, for such OD pairs both the B-L and C-L paths are defined to be the same path.

Table 1: Comparison between optimal routing and B-L/C-L routing with uniform traffic.

Routing	Max/Min Link Utilizations	Total Delay (Value of $D(F)$)
Optimal	0.750/0.500	160.000
0 % B-L	1.200/0.200	unbounded
10 % B-L	1.110/0.260	unbounded
20 % B-L	1.020/0.320	unbounded
30 % B-L	0.930/0.380	215.663
40 % B-L	0.840/0.440	169.863
50 % B-L	0.750/0.500	160.000
60 % B-L	0.840/0.450	171.307
70 % B-L	0.930/0.400	238.002
80 % B-L	1.020/0.350	unbounded
90 % B-L	1.110/0.300	unbounded
100 % B-L	1.200/0.250	unbounded

4 Simulation and Visualization of the Routing Scheme with Non-Uniform Traffic

A simulation program was developed to test the routing scheme's ability to dynamically adapt to non-uniform traffic patterns. Also, for the sake of comparison, we implemented a standard "random routing scheme" in which packets are routed along randomly selected shortest hop path interconnecting active OD pairs. Three separate traffic patterns were simulated: (1) *uniform* — the probability of packet generation is the same for all origin-destination pairs, (2) *uniform + to-center* — the probability of generating packets having a destination at the center of the mesh is higher than others and (3) *uniform + from-center* — the probability of generating packets from an origin node at the center of the mesh is higher than others.

The results of the simulation studies are summarized in Table 2. For all cases, the queues were initially empty when the simulation began and the value of the parameter p_B was initialized to 0.5 at all nodes. For the uniform traffic case, both the random and BL/CL routing schemes provided stable performance in the sense that the queues lengths remained bounded over time. For the other two (non-uniform) traffic patterns, it can be observed from the visualization graphic (on the video) that the random routing scheme does not provide stable performance, while, on the other hand, the BL/CL scheme does. The tabulated delays were calculated based on approximately 200 simulation time units. If more time had been simulated, then the average delays associated with the random routing scheme would increase (because it was not able to provide stable performance for the non-uniform traffic patterns).

5 Conclusions and Extensions

A simple routing algorithm is proposed for network computers with buffered mesh connected network topologies. The routing scheme operates by shifting flow, in a probabilistic manner, based on a

Table 2: Comparison of Random Routing and BL/CL Routing.

Average Packet Delay			
Traffic Pattern	Routing Scheme		Relative Gain
	Random	BL/CL	
Uniform	20.10	16.35	23.93%
U + From-Center	18.56	14.96	24.06%
U + To-Center	20.39	16.22	25.71%

coarse measure of the state of the network queues.

Theoretical studies are currently underway in an attempt to better understand how the convergence rate and performance of our algorithm are affected by the choice of the threshold value, T, and probability stepsize, Δp.

As a final note, we mention that our routing scheme is also applicable to meshes which include the so-called wrap-around links, i.e., symmetric meshes in which the degree of each node is exactly four. An in-depth discussion of this topic can be found in [8].

Acknowledgements

We are thankful to Mr. C. L. Lewis and Prof. A. A. Maciejewski for their help in displaying the simulation data using video graphics.

References

[1] R. E. Cypher, J. L. C. Sanz, and L. Snyder, "Algorithms for Image Component Labeling on SIMD Mesh-Connected Computers," *IEEE Trans. Computers*, vol. C-39, pp. 269–276, Feb. 1990.

[2] R. Miller and Q. F. Stout, "Mesh Computer Algorithms for Computational Geometry," *IEEE Trans. Computers*, vol. C-38, pp. 321–340, March 1989.

[3] L. Lapidus, G. F. Pinder, *Numerical Solution of Partial Differential Equations in Science and Engineering*. New York, New York: John Wiley & Sons, 1982.

[4] D. A. Reed and H. D. Schwetman, "Cost-Performance Bounds for Multimicrocomputer networks," *IEEE Trans. Computers*, vol. C-32, pp. 83–95, Jan. 1983.

[5] D. Bertsekas and R. Gallager, *Data Networks*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[6] D. P. Bertsekas, B. Gendron and and W. K. Tsai, "Implementation of an optimal multicommodity network flow algorithm based on gradient projection and a path flow formulation," Mass. Inst. Technol., Cambridge, MA, Tech. Rep. LIDS-P-1364, Feb., 1984.

[7] J. K. Antonio, G. M. Huang, and W. K. Tsai, "Deriving time complexities for a class of distributed gradient projection-based optimal routing algorithms," *Proccedings of the 29th IEEE Conference on Decision and Control*, Honolulu, Hawaii, Dec. 5–7, 1990, pp. 931–936.

[8] N. Ge, "A Simple Routing Control Scheme for Buffered Mesh Connected Parallel Computer Networks," M.S. Thesis, School of Electrical Engineering, Purdue University, December, 1990.

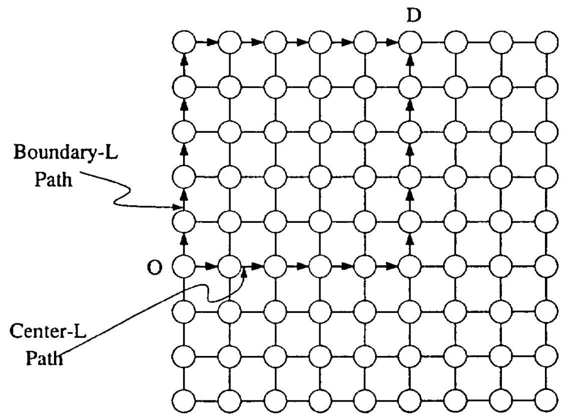

Fig. 1 An 81-node mesh network. The B-L and C-L paths associated with the given OD pair are indicated with arrowed links.

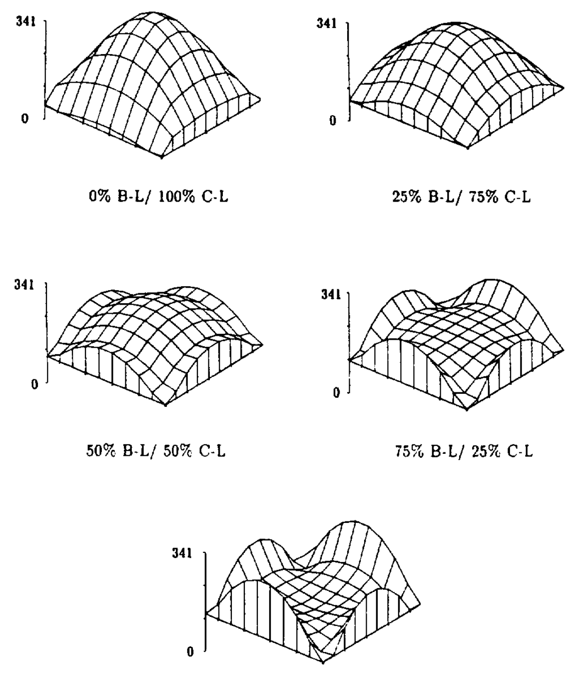

Fig. 2 Link flow distributions for various values of p_B with uniform traffic demands.