Maximum-Shortest-Path (MSP): An Optimal Routing Policy for Mesh-Connected Multicomputers

Jie Wu, *Senior Member IEEE*
Florida Atlantic University, Boca Raton

Key Words — Hypercube, Mesh, Mesh-connected topology, Multicomputer, Shortest-path routing, Torus.

Summary & Conclusions — This paper presents a new routing policy, maximum-shortest-path (MSP), within the class of shortest-path routing policies for mesh-connected topologies which include popular 2-D and 3-D meshes, 2-D and 3-D tori, and n-dimensional binary hypercubes (n-cubes). In MSP, the routing message is always forwarded to a neighbor from which there exists a maximum number of shortest paths to the destination. The optimal routing (defined in this paper) maximizes the probability of reaching the destination from a given source without delays at intermediate nodes, assuming that each link in the system has a given failure probability. The results show that:

- the optimal e-cube routing in n-cubes is a special implementation of MSP,
- MSP is equivalent to the Badr & Podar zig-zag (Z^2) routing policy in 2-D meshes which is also optimal.
- the Z^2 routing policy is not optimal in any $N \times N$ torus, where $N > 4$ is an even number.

A new routing algorithm implements MSP in 2-D tori and is at least suboptimal. Two examples are used in 6×6 and 8×8 tori to demonstrate that MSP is optimal for some 2-D tori. This is the first attempt to address optimal routing in the torus network, which still is an open problem.

INTRODUCTION

Acronyms

aka	also known as
MSP	maximum shortest path (routing)
n-D	n-dimensional
Z^2	zig-zag

Notation[1]

n	number of dimensions
m	number of eligible neighbors for a node
k	number of nodes in each dimension
v	a source or intermediate node
u	a destination node in a given network topology
V	$(v_1, v_2, ...v_m)$: 'eligible neighbor vector' of v with respect to destination u
p	Pr{message is successfully forwarded to a neighbor along a given link}
$\bar{p}$	$1-p$
$S(v,u)$	maximum Pr{delivery of a message from v to u}
$P(v,u)$	number of shortest paths from v to u

[1] Other, standard notation is given in "Information for Readers & Authors" at the rear of each issue.

Recent microprocessor developments have sparked interest in large scale multiprocessors composed of 100s or 1000s of processors and in data communication networks allowing for delivery of advanced digital services. In a multicomputer system, a collection of computing nodes (computers) work in concert to solve large application problems. These nodes communicate data and coordinate their efforts by sending & receiving messages through the underlying data communication network. Thus, the performance of such a multicomputer system depends on the end-to-end cost of communication mechanisms. Routing time of messages is one of the key factors that are critical to the performance of multicomputers. Basically, routing is the process of transmitting data from one node (source) to another node (destination) in a given system.

Direct-networks have become a popular means for interconnecting components of multicomputers. In a direct network, nodes are connected to only a few nodes, its neighbors, according to the topology of the network.

Mesh-connected topology is one of the most thoroughly investigated network topologies for parallel processing. It is important due to its simple structure and its good performance in practice. Mesh-connected topologies, also called k-ary n-cube based networks, have an n-dimensional grid structure with k nodes in each dimension such that every node is connected to two other nodes in each dimension by a direct link. Mesh-connected topologies include:

- n-dimensional mesh,
- torus (a mesh with wrap-around links),
- hypercube.

These topologies have desirable properties of regularity, balanced behavior, and many alternative paths. Examples of commercial products based on binary n-dimensional hypercubes (n-cubes) include: the Ncube's nCUBE and the Thinking Machine's Connection Machine, which is a hypercube interconnected bit-serial SIMD machine. Multicomputers that use 2-D meshes include the MIT J-machine [3], the Symult 2010 [4], and the Intel Touchstone [6]. Several commercial multicomputers have been using the toroidal network topology; *eg*, the Tera system (3-ary n-cubes) [1], the CRAY T3D (3-D torus) [5], and the CRAY T3E (3-D torus) [8].

As more components are placed in a multicomputer, the system becomes more complex, and this increases the chances of having one or more component failures. On the other hand, many applications (such as real-time applications) demand timeliness of communication service. To meet such a demand, routing messages can be routed through a shortest path between the source and destination. In a shortest-path routing, only shortest paths (to the destination) are acceptable. This paper:

· focuses on a shortest-path routing policy that can deliver routing messages through a shortest path (many might exist) in the presence of faulty components;

· considers only the case of link faults.

In a shortest-path routing, only shortest paths are acceptable. If a routing message can not be forwarded to the destination through a shortest path, then it is simply discarded. A shortest-path routing policy is optimal [2] if it maximizes the probability of reaching the destination from a given source without delays at intermediate nodes. It is assumed that some of the outgoing links at a node may be unavailable due to competing traffic or physical link failure.

This paper presents MSP, a new routing policy, within the class of shortest-path routing policies for mesh-connected topologies. In MSP, the routing message is always forwarded to a neighbor from which there exists a maximum number of shortest paths to the destination. The optimal e-cube routing in n-cubes is a special implementation of MSP, and MSP is equivalent to the Badr & Podar Z^2 routing policy in 2-D meshes which is also optimal. The Z^2 routing algorithm is proved to be not optimal in any $N \times N$ torus, where $N > 4$ is an even number. This extends a result [9], where only one counter example is given for a 6×6 torus. A routing algorithm that implements MSP in 2-D tori is presented, and proved to be at least suboptimal (optimal for the cases that we check). MSP is the first attempt to achieve optimal routing in the torus network which is still an open problem.

A future research direction is to extend MSP to systems with node faults.

· Section 2 overviews basic definitions, including shortest-path routing, Z^2 routing, and optimal routing; and proposes the MSP.

· Section 3 demonstrates the use of MSP in various mesh-connection networks, including n-cubes, 2-D meshes, and 2-D tori; and shows that MSP is optimal in both n-cubes and 2-D meshes.

· Section 4 discusses the relationship between MSP and optimal routing in 2-D tori. MSP is proved to be an approximation of the optimal routing policy, and shown to be optimal in 6×6 and 8×8 tori.

2. PRELIMINARIES

Definitions

· Wraparound link: A link that connects two neighbors whose addresses differ by $k-1$ in a dimension.

· (k,n)-torus: k-ary n-dimensional torus

· (k,n)-mesh: A (k,n)-torus without wraparound links.

· Eligible neighbor of v with respect to destination: A neighbor closer to u than from v to u in the network. ◁

Assumptions

1. p is uniform across the whole network.
2. $0 < p < 1$; since $p = 0$ and $p = 1$ are trivial cases. ◁

2.1 Mesh-Connected Topologies

Mesh-connected topologies include (k,n)-tori and (k,n)-meshes. A (k,n)-torus has k^n nodes, each of which is uniquely indexed by an n-tuple:

$(x_1, \ldots, x_i, \ldots, x_n)$, $0 \le x_i \le k-1$.

Each node connects to 2 neighbors in each dimension:

$(x_1, \ldots, x_{i-1}, x_i + 1, x_{i+1}, \ldots, x_n)$,

$(x_1, \ldots, x_{i-1}, x_i - 1, x_{i+1}, \ldots, x_n)$;

addition & subtraction are modulo k.

Two commonly used tori are:

· $(k,2)$-tori aka 2-D tori,

· $(2,n)$-tori aka n-dimensional binary hypercubes (n-cubes).

Two commonly used meshes are:

· $(k,2)$-meshes aka 2-D meshes,

· $(k,3)$-meshes aka 3-D meshes.

This paper focuses on:

· 2-D tori, figure 1a,

· 2-D meshes, figure 1b,

· n-cubes, figure 1c.

2.2 Shortest-Path Routing

Use assumptions 1 – 2.

The m varies from node to node depending on the distance between the current node and the destination node (see figure 2). If a priority order is defined among the v's eligible neighbors, V, to which the routing message is forwarded, this order can be represented by

$(v_1^{(k)}, v_2^{(k)}, \ldots, v_m^{(k)})$

which is a permutation of V; k $(1 \le k \le m!)$ represents permutation k based on a permutation generator.

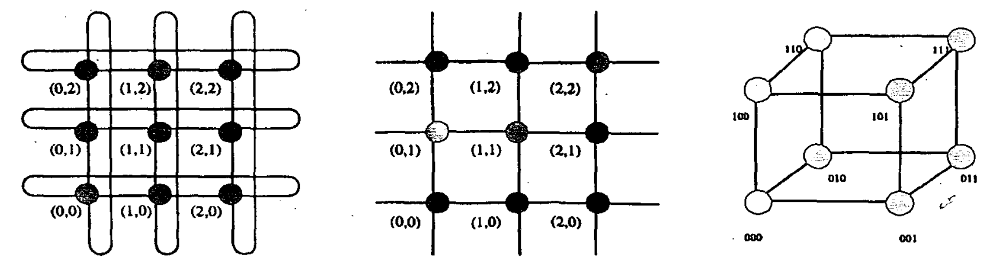
The probability vector associated with the eligible neighbor vector is:

$$(p, p \cdot \bar{p}, p \cdot \bar{p}^2, \ldots, p \cdot \bar{p}^{m-1})$$

$p \cdot \bar{p}^{i-1}$ is element i of this probability vector, where

$p \equiv$ probability that the routing message is successfully forwarded to neighbor i in the given order of eligible neighbors,

$\bar{p}^{i-1} \equiv$ probability that the first $i-1$ tries fail.



a. $(3,2)$ torus (3×3 2-D torus); b. $(3,2)$ mesh (3×3 2-D mesh); c. $(2,3)$ torus (3-cube)

Figure 1: Sample mesh-connected-networks

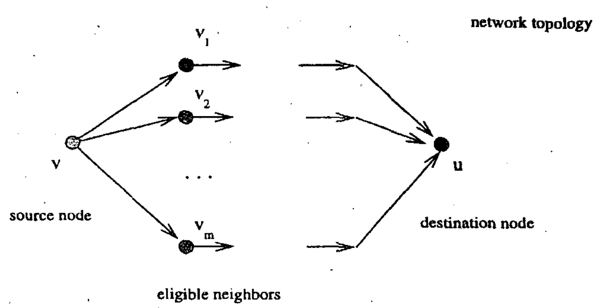


Figure 2: Eligible Neighbor Vector for a Given v & u

Again, $(v_1^{(k)}, v_2^{(k)}, \ldots, v_m^{(k)})$ is a priority order (out of $m!$) of m eligible neighbors of v with respect to destination u. $S(v,u)$ satisfies the recursive equations:

$$S(v,u) = \max_k \left[\sum_{i=1}^{m} p \cdot \bar{p}^{i-1} \cdot S(v_i^{(k)}, u) | 1 \le k \le m! \right], \quad u \ne v,$$
$$S(v,v) = 1 .$$

2.3 Z^2 Routing

The Z^2 routing policy [2] is a shortest-path routing algorithm and is optimal for 2-D meshes. Informally, the Z^2 policy states that the routing message should be sent towards the diagonal which denotes the set of nodes that have an equal number of rows and columns away from the destination node. Z^2 is optimal in 2-D meshes.

Without loss of generality, let the:

- source node be (i,j),
- destination node be $(0,0)$.

Z^2 routing works as follows:

1. Build a rectangle that contains v and u as two opposite corners.
2. Derive a line L that crosses the u and equally divides the angle between two boundary lines (of the rectangle) that both cross u.
3. Route the message toward line L: each intermediate node is selected based on its distance to L; the closest one is selected from the eligible ones. ◁

Figure 3 is an example of Z^2 routing, with $i,j \ge 0$. Line L crosses u and (j,j). The Z^2 routing goes along the line connecting $s = (i,j)$ and (j,j) until it reaches node (j,j); then follows the line connecting (j,j) to $u = (0,0)$.

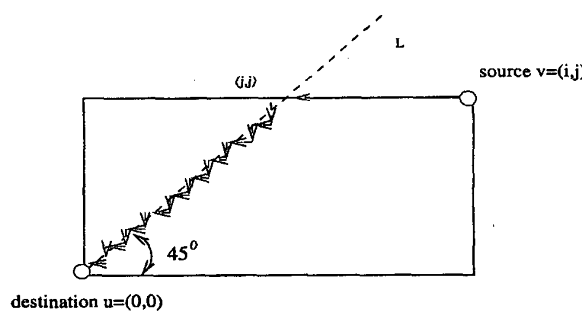


Figure 3: Zig-Zag Routing

2.4 Maximum Shortest-Path Routing

MSP is another shortest-path routing policy. The goal is to show that MSP is optimal for many mesh-connected networks. In MSP, the routing message is always forwarded to an eligible neighbor from which there exists a maximum number of shortest paths to the destination. These shortest paths can overlap by sharing some links. $P(v,u)$ satisfies:

$$P(v,u) = \sum_{i=1}^{m} P(v_i, u), \quad v \ne u$$
$$P(v,v) = 1$$

As an example, consider routing in a 2-D mesh. Let

- source $v = (i,j), i,j \ge 0$,
- destination $u = (0,0)$,
- $P(v)$ replace $P(v,u)$.

The equations become:

$$P(i,j) = P(i-1,j) + P(i,j-1), \quad (i,j) \neq (0,0)$$
$$P(0,0) = 1$$

3. OPTIMAL SHORTEST-PATH ROUTING BASED ON MSP

3.1 n-Cube

An n-cube [7] consists of $N = 2^n$ nodes and $n \cdot (N/2)$ links. $\{0,1\}^n$ is the set of nodes, where u and v are connected iff u and v differ in exactly one bit. The conventional e-cube routing is an example of MSP. In an n-cube, eligible neighbors of each node have the same number of shortest paths to the destination. Any selection policy of an eligible neighbor is MSP, because all eligible neighbors are indistinguishable; any shortest-path routing (including e-cube) is optimal: selection of a neighboring node to which the routing message is forwarded can be random.

3.2 2-D Mesh

Prove that MSP is an optimal shortest-path routing in a 2-D mesh. Let

- $v = (i,j)$, $i,j \geq 0$, be the source node
- $u = (0,0)$ be the destination node

It suffices to prove that

$$P(i-1,j) \leq P(i,j-1) \ \text{ iff } \ S(i-1,j) \leq S(i,j-1),$$

because for $v = (i,j)$ there are at most 2 eligible neighbors: $(i,j-1)$ and $(i-1,j)$. Use dummy values for $S(-1,j) = 0$ and $S(i,-1) = 0$.

Lemma 1: In a 2-D mesh, $P(i,j) = \binom{i+j}{j}$.

Proof: See appendix A.1

Based on lemma 1, plot each $P(i,j)$ in a given 2-D mesh with source node $(0,0)$; these values form a Pascal triangle (see figure 4).

In general, the expression for $S(i,j)$ is rather complex; however, the difference between $S(i-1,j)$ and $S(i,j-1)$ can be represented by the simple expression in theorem 1.

Theorem 1: In a 2-D mesh with source $v = (i,j)$ and $i,j \geq 1$,

$$S(i-1,j) - S(i,j-1) = \left[\binom{i+j-1}{j} - \binom{i+j-1}{j-1}\right] \cdot p^{i+j-1} \cdot \bar{p}^{j}.$$

Proof: See appendix A.2.

Corollary 1: In a 2-D mesh, MSP is an optimal shortest-path routing.

Proof: See appendix A.3

Corollary 2: In a 2-D mesh, the Z^2 routing policy is an optimal shortest-path routing.

Proof: See appendix A.4

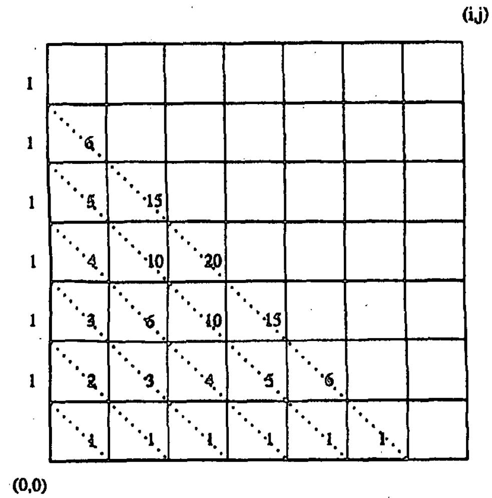


Figure 4: The Pascal Triangle of $P(i,j)$

This is a simpler proof for corollary 2 than the proof in [2]. Because Z^2 is an MSP, it can be used to implement the MSP routing policy in a 2-D mesh.

3.3 2-D Torus

A 2-D torus is a 2-D mesh with wraparound links at the ends. Therefore, for some destination-source pairs, there are more than 2 eligible neighbors. Specifically, for an $N \times N$ torus where N is even, there is 1 node that has 4 eligible neighbors and $2(N-2)$ nodes ($N/2$ rows or columns away, but not both) for which 3 directions lie along a shortest path. Without loss of generality, consider only source nodes that are $N/2$-column away (see figure 5); nodes that are not on column or row $N/2$ in a 2-D torus, are equivalent to the one in a 2-D mesh. Therefore, any optimal routing in a 2-D mesh is also optimal in a 2-D torus. When either i or j in source node (i,j) is larger than $N/2$, then the shortest path uses wraparound links. To simplify the discussion, let:

- $0 \leq i,j \leq N/2$;
- the destination node always be $(0,0)$;
- the source node be (i,j).

Theorem 2 shows that Z^2 is not optimal for any $N \times N$ torus, $N > 4$ is an even number. It generalizes the result in [9], where only one counter example is given for a 6×6 torus.

Theorem 2: The Z^2 routing algorithm is not optimal in any $N \times N$ torus, $N > 4$ is an even number.

Proof: See appendix A.5.

MSP routing on a 2-D torus works as follows: Find a location k (at the y dimension) along column i (column $N/2$), such that for any source whose j value is larger than k, the routing message should be forwarded along column i until reaching $t = (i,k)$. Then the remaining part is

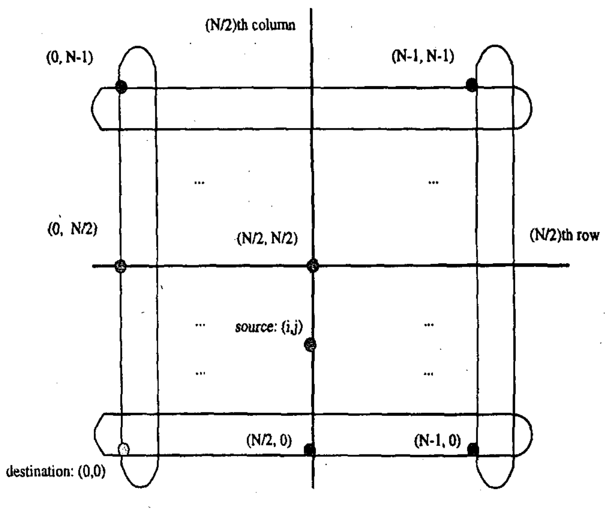


Figure 5: An $N \times N$ 2-D Torus

equivalent to the routing in a 2-D mesh. Therefore, Z^2 routing can be used. For any source node whose 'value at the y dimension' $\leq k$, the optimal routing is equivalent to the one in a 2-D mesh. This special point $t = (i, k)$ is a 'turning point'. Clearly, $k < i = N/2$. The goal is to find the value of k in turning point $t = (i, k)$.

Let $P(i, j)$ be the number of shortest paths from node (i, j) to node $(0, 0)$. Theorem 3 shows the number of shortest paths from (i, j) to $(0, 0)$.

Theorem 3: In a 2-D torus,

· $P(i,j) = \binom{i+j}{j}$, for (i, j) at neither column nor row $N/2$;

· $P(i,j) = 2\binom{i+j}{j}$, for (i, j) at column or row $N/2$ but not both;

· $P(i,j) = 4\binom{i+j}{j}$, for (i, j) at both column and row $N/2$.

Proof: See appendix A.6.

Theorem 4: If the source is at column $(N/2)$ and $N > 2$ is an even number, then the turning point at column $(N/2)$ is $t = (N/2, k)$, where $N/2 = 2k + 1$ or $N/2 = 2k$.

Proof: See appendix A.7.

MSP routing on a 2-D torus works as follows:

A. Let the source be at column $N/2$ (this algorithm also applies when the source is at row $N/2$). Let j be the value at dimension y. Consider two cases.

1. $j > k$: The routing message should be forwarded along the y dimension until it reaches row k; then follow Z^2 routing.

2. $j \leq k$: Follow Z^2 routing directly.

B. If the source is not at column or row $N/2$ then follow Z^2 routing. ◀

Figure 6 shows two cases where the source is on column $N/2$ and it has 3 eligible neighbors:
Case 1: The source is above the turning point.
Case 2: The source is (or below) the turning point.

When the source has 4 eligible neighbors, step #1 can be along either row (x dimension) or column (y dimension), then the remaining steps are the same for situations where the source is on row or column $N/2$.

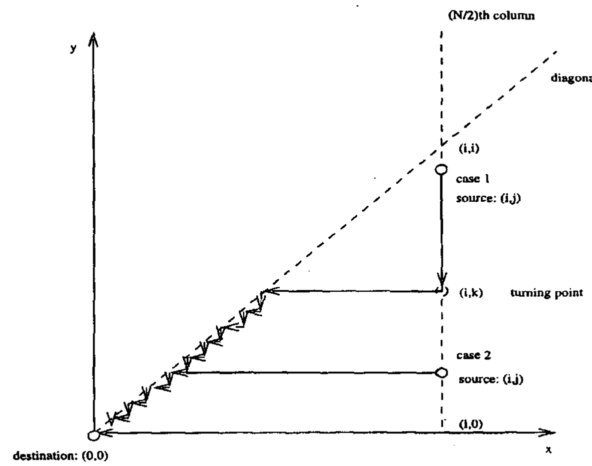


Figure 6: Two Optimal-Routing Examples

4. DISCUSSION

The optimal routing policy and the MSP are shown to be the same in mesh and hypercube routing. The Z^2 algorithm is shown to be not optimal in a general 2-D torus network. An MPS routing algorithm for a general 2-D torus network is presented.

The next question is how close the MSP is to the optimal routing policy in a 2-D torus network. Ref [9] predicts that the optimal policy for the torus seems unlikely to be of a simple closed form. Theorem 7 shows the relationship between $S(v)$ and $P(v)$.

Theorem 5: Let

$$S(v) = a_0 \cdot p^d + a_1 \cdot p^{d+1} + a_2 \cdot p^{d+2} + \ldots;$$

a_0 is a positive integer; d is the distance between v and the destination.

Then $P(v) = a_0$.

Proof: See appendix A.8. ◁

Basically, MSP can be viewed as an approximation of the optimal policy where high order terms of p are eliminated. The question is whether these two policies are equivalent in a 2-D torus network. That is, it is not known that $S(v) \geq S(u)$ iff $P(v) \geq P(u)$, where $v = (i, j-1)$ and $u = (i-1, j)$.

For at least some cases, MSP routing is shown to be optimal in a 2-D torus. Table 1 shows the P & S for all the nodes in a 6×6 torus, where $(0, 0)$ is the destination

Table 1: P & S for Each Node of a 6 × 6 Torus

v	(0,0)	(1,0)	(2,0)	(1,1)	(3,0)	(2,1)	(3,1)	(2,2)
$P(v)$	1	1	1	2	2	3	8	6
$S(v)$	1	p	p^2	$2p^2-p^3$	$2p^3-p^4$	$3p^3-2p^4$	$8p^4-12p^5+6p^6-p^7$	$6p^4-7p^5+2p^6$

node. $S(v) > S(u)$ iff $P(v) > P(u)$ holds in this case. For example,
$S(3,1) - S(2,2) = p^4 \cdot \bar{p}^2 \cdot (2-p) > 0$,
$P(3,1) - P(2,2) = 8 - 6 > 0$;
$S(2,1) - S(3,0) = p^3 \cdot \bar{p}$,
$P(2,1) - P(3,0) = 3 - 2 > 0$.
Therefore, the MSP in section 3 is an optimal routing in a 6 × 6 torus; the turning point is at node (3,1), where $N = 6$ and $k = 1$. Table 2 shows S values for all the nodes in an 8 × 8 torus, where $(0,0)$ is the destination node.
$S(3,2) - S(4,1) = p^6 \cdot \bar{p}^2 > 0$,
$S(3,1) - S(4,0) = 2p^4 \cdot \bar{p} > 0$;
$S(4,2) - S(3,3) = p^6 \cdot \bar{p}^3 \cdot (2-p) \cdot (p^2 - 5p + 5) > 0$.
Therefore, the turning point is node (4,2), where $N = 8$ and $k = 2$.

It is still an open problem whether MSP is optimal in a general torus network. I conjecture that MSP is optimal in a general 2-D torus.

Table 2: S Values for Each Node of an 8 × 8 Torus

$S(0,0)$	1
$S(1,0)$	p
$S(1,1)$	$2p^2-p^3$
$S(2,0)$	p^2
$S(2,1)$	$3p^3-2p^4$
$S(2,2)$	$6p^4-7p^5+2p^6$
$S(3,0)$	p^3
$S(3,1)$	$4p^3-3p^5$
$S(3,2)$	$10p^5-14p^6+5p^7$
$S(3,3)$	$20p^6-38p^7+24p^8-5p^9$
$S(4,0)$	$2p^4-p^5$
$S(4,1)$	$10p^5-15p^6+7p^7-p^8$
$S(4,2)$	$30p^6-73p^7+71p^8-35p^9+9p^{10}-p^{11}$

APPENDIX

A.1 Proof of Lemma 1

Use induction to prove this lemma. When node (i,j) has 1 eligible neighbor then either $i=0$ or $j=0$; there is only 1 path and this lemma is true in this case. Therefore, only cases when both i and j are not 0, need to be considered. When $i+j=2$, $\binom{i+j}{j} = \binom{2}{1} = 2$, this lemma is true.

Assume that this lemma holds for $i+j=k$. Consider $P(i,j) = P(i-1,j) + P(i,j-1)$ for $i+j=k+1$. Based on the induction, then $P(i-1,j) = \binom{i+j-1}{j}$ and $P(i,j-1) = \binom{i+j-1}{j-1}$. Therefore,

$$P(i,j) = \binom{i+j-1}{j} + \binom{i+j-1}{j-1} = \binom{i+j}{j}.$$

A.2 Proof of Theorem 1

Use induction on $i+j$, where i and j (both are positive) are coordinates of $v=(i,j)$. When $i+j=2$, then $S(i-1,j) - S(i,j-1) = S(0,1) - S(1,0) = p - p = 0$.

Assume that the theorem holds for $i+j=k$. When $i+j=k+1$, consider the 3 cases: $i=j$, $i>j$, $i<j$.

When $i=j$, $S(i-1,j) = S(i,j-1)$, since S is symmetric. On the other hand,

$$\left[\binom{i+j-1}{j} - \binom{i+j-1}{j-1}\right] \cdot p^{i+j-1} \cdot \bar{p}^j = 0.$$

Therefore,

$$S(i-1,j) - S(i,j-1) = \left[\binom{i+j-1}{j} - \binom{i+j-1}{j-1}\right] \cdot p^{i+j-1} \cdot \bar{p}^j$$

When $i>j$, based on the induction assumption,

$$S(i-1,j-1) - S(i-2,j) = \left[\binom{i+j-2}{j-1} - \binom{i+j-2}{j}\right] \cdot p^{i+j-2} \cdot \bar{p}^{j-1} < 0$$

That is, $S(i-1,j-1) < S(i-2,j)$. Similarly,

$$S(i-1,j-1) - S(i,j-2) = \left[\binom{i+j-2}{j-1} - \binom{i+j-2}{j-2}\right] \cdot p^{i+j-2} \cdot \bar{p}^{j-1} > 0$$

That is, $S(i-1,j-1) > S(i,j-2)$. Therefore,

$$S(i-1,j) = p \cdot S(i-2,j) + p \cdot \bar{p} \cdot S(i-1,j-1),$$
$$S(i,j-1) = p \cdot S(i-1,j-1) + p \cdot \bar{p} \cdot S(i,j-2).$$

Based on the induction assumption,

$$\begin{aligned}
S(i-1,j) - S(i,j-1) &= p \cdot [S(i-2,j) - S(i-1,j-1)] \\
&\quad + p \cdot \bar{p} \cdot [S(i-1,j-1) - S(i,j-2)] \\
&= p \cdot \left[\binom{i+j-2}{j} - \binom{i+j-2}{j-1}\right] \cdot p^{i+j-2} \cdot \bar{p}^j \\
&\quad + p \cdot \bar{p} \cdot \left[\binom{i+j-2}{j-1} - \binom{i+j-2}{j-2}\right] \cdot p^{i+j-2} \cdot \bar{p}^{j-1} \\
&= \left[\left[\binom{i+j-2}{j} - \binom{i+j-2}{j-1}\right] + \left[\binom{i+j-2}{j-1} - \binom{i+j-2}{j-2}\right] \right] \cdot p^{i+j-1} \cdot \bar{p}^j \\
&= \left[\binom{i+j-1}{j} - \binom{i+j-1}{j-1}\right] \cdot p^{i+j-1} \cdot \bar{p}^j.
\end{aligned}$$

When $i < j$, based on the induction assumption,

$$S(i-1,j-1) - S(i-2,j) = \left[\binom{i+j-2}{j-1} - \binom{i+j-2}{j}\right] \cdot p^{i+j-2} \cdot \bar{p}^{j-1} > 0;$$
$$S(i-1,j-1) > S(i-2,j).$$

Similarly,

$$S(i-1,j-1) - S(i,j-2) = \left[\binom{i+j-2}{j-1} - \binom{i+j-2}{j-2}\right] \cdot p^{i+j-2} \cdot \bar{p}^{j-1} < 0;$$
$$S(i-1,j-1) < S(i,j-2).$$

Therefore,

$$S(i-1,j) = p \cdot S(i-1,j-1) + p \cdot \bar{p} \cdot S(i-2,j),$$
$$S(i,j-1) = p \cdot S(i,j-2) + p \cdot \bar{p} \cdot S(i-1,j-1).$$

Based on the induction assumption,

$$\begin{aligned} S(i-1,j) - S(i,j-1) &= p \cdot [S(i,j-1) - S(i,j-2)] \\ &\quad + p \cdot \bar{p} \cdot [S(i-2,j) - S(i-1,j-1)] \\ &= p \cdot \left[\binom{i+j-2}{j-1} - \binom{i+j-1}{j-2}\right] \cdot p^{i+j-2} \cdot \bar{p}^{j} \\ &\quad + p \cdot \bar{p} \cdot \left[\binom{i+j-2}{j} - \binom{i+j-2}{j-1}\right] \cdot p^{i+j-2} \cdot \bar{p}^{j-1} \\ &= \left[\left[\binom{i+j-2}{j-1} - \binom{i+j-2}{j-2}\right] \right. \\ &\quad \left. + \left[\binom{i+j-2}{j} - \binom{i+j-2}{j-1}\right] \right] \cdot p^{i+j-1} \cdot \bar{p}^{j} \\ &= \left[\binom{i+j-1}{j} - \binom{i+j-1}{j-1}\right] \cdot p^{i+j-1} \cdot \bar{p}^{j}. \end{aligned}$$

A.3 Proof of Corollary 1

Based on theorem 1 —

$$S(i-1,j) - S(i,j-1) = \left[\binom{i+j-1}{j} - \binom{i+j-1}{j-1}\right] \cdot p^{i+j-1} \cdot \bar{p}^{j}$$

Based on Lemma 1, this equation is rewritten as:

$$S(i-1,j) - S(i,j-1) = [P(i-1,j) - P(i,j-1)] \cdot p^{i+j-1} \cdot \bar{p}^{j}.$$

That is, $S(i-1,j) \geq S(i,j-1)$ iff $P(i-1,j) \geq P(i,j-1)$. Therefore, MSP is an optimal shortest-path routing for 2-D meshes.

A.4 Proof of Corollary 2

In the Z^2 routing policy, the eligible neighbor closer to line L (see figure 3) has more selections (larger P value) of shortest paths. That is, the Z^2 routing policy is the same as MSP in 2-D meshes. Based on corollary 1, the Z^2 routing policy is an optimal shortest-path routing for 2-D meshes.

A.5 Proof of Theorem 2

It suffices to show one counter example for any given $N \times N$ torus, where $N > 4$ is an even number. Let node $(k, k-1)$ be at column the $(N/2)$ with respect to the destination $(0,0)$, *ie*, $k = N/2$. This node has 3 eligible neighbors: $(k-1,k-1)$, $(k,k-2)$, $(k+1,k-1)$. (A neighbor is eligible if it is along one of the shortest paths from the current node to the destination node.) Now, show that $S(k-1,k-1) < S(k,k-2)$: the diagonal node is not the neighbor that has the largest S value. Because $N \geq 6$, $k = \frac{N}{2} \geq 3$, the node $(k,k-2)$ has 3 eligible neighbors: $(k-1,k-2)$, $(k+1,k-2)$, $(k,k-3)$. Based on the torus topology:

$$\begin{aligned} S(k-1,k-1) = &\max[p \cdot S(k-2,k-1) + p \cdot \bar{p} \cdot S(k-1,k-2), \\ &\quad p \cdot S(k-1,k-2) + p \cdot \bar{p} \cdot S(k-2,k-1)]; \end{aligned}$$
$$\begin{aligned} S(k,k-2) \geq\ &p \cdot S(k-1,k-2) + p \cdot \bar{p} \cdot S(k+1,k-2) \\ &+ p \cdot \bar{p}^2 \cdot S(k,k-3). \end{aligned}$$

Because nodes $(k-2,k-1), (k-1,k-2)$, and $(k+1,k-2)$ are identical with respect to the destination $(0,0)$, then $S(k-2,k-1) = S(k-1,k-2) = S(k+1,k-2)$. Therefore,

$$\begin{aligned} S(k-1,k-1) &= \\ &\quad p \cdot S(k-2,k-1) + p \cdot \bar{p} \cdot S(k-1,k-2) \\ &= p \cdot S(k-1,k-2) + p \cdot \bar{p} \cdot S(k+1,k-2) \\ &< p \cdot S(k-1,k-2) + p \cdot \bar{p} \cdot S(k+1,k-2) \\ &\quad + p \cdot \bar{p}^2 \cdot S(k,k-3) \\ &\leq S(k,k-2). \end{aligned}$$

A.6 Proof of Theorem 3

Use the same approach as in the proof of lemma #?? to show that when node (i,j) is not at column or row $N/2$ and it has 2 eligible neighbors then $P(i,j) = \binom{i+j}{j}$.

When (i,j) is at row or column $N/2$ (but not both), then there are 2 eligible neighbors (where either $i = 0$ or $j = 0$) or 3 eligible neighbors (where $i \neq 0$ & $j \neq 0$). For the '2 eligible neighbors case' there are 2 shortest paths; and $2\binom{i+j}{j}$ is equal to either $2\binom{j}{j}$ or $2\binom{i}{0}$. In either way, the result is 2. Therefore, we only need prove the '3 eligible neighbors case'.

When $i = j = 1$ in a $2 \times N$ or an $N \times 2$ torus, then $P(i,j) = 2\binom{i+j}{j} = 2\binom{2}{1} = 4$. Assume this result holds for $i + j = k > 2$.

Consider the case where (i,j) $(i = N/2)$ has 3 eligible neighbors $(i-1,j)$, $(i,j-1)$, $(i+1,j)$. Then nodes $(i-1,j)$ and $(i+1,j)$ are not at column $N/2$ and node $(i,j-1)$ is still at column $N/2$. Based on the induction assumption and the result for nodes not at column or row $N/2$, then $P(i-1,j) = P(i+1,j) = \binom{i+j-1}{j}$ and

$P(i,j-1) = 2\binom{i+j-1}{j-1}$. Therefore,

$$P(i,j) = P(i-1,j) + P(i+1,j) + P(i,j-1) = 2\binom{i+j-1}{j} + 2\binom{i+j-1}{j-1} = 2\binom{i+j}{j}.$$

Similarly, the same result can be obtained for the case: (i,j) is at row $N/2$; *ie*, $j = N/2$.

For node (i,j) that has 4 eligible neighbors: $(i-1,j)$, $(i+1,j)$, $(i,j-1)$, $(i,j+1)$, each neighbor is at column or row $N/2$; and $P(i+1,j) = P(i-1,j)$ and $P(i,j-1) = P(i,j+1)$. Based on the result for nodes with 3 eligible neighbors:

$$P(i,j) = P(i-1,j) + P(i,j-1) + P(i+1,j) + P(i,j+1) = 2\binom{i+j-1}{j} + 2\binom{i+j-1}{j-1} + 2\binom{i+j-1}{j} + 2\binom{i+j-1}{j-1} = 4\binom{i+j}{j}.$$

A.7 Proof of Theorem 4

It suffices to prove that: at the turning point $t = (i,k)$, where $i = N/2$, then

$$k = \max[l | P(i,l-1) \le P(i-1,l)].$$

Based on the properties of binomial numbers, it is equivalent to show that $P(i,k) > P(i-1,k+1)$ and $P(i,k-1) \le P(i-1,k)$:

$$2\binom{k+i}{k} > \binom{k+i}{k+1}$$
$$2\binom{k+i-1}{k-1} \le \binom{k+i-1}{k}.$$

For $i = 2k$,

$$\binom{k+i}{k+1} = 3\binom{k}{k+1} = \frac{2k}{k+1}\cdot\binom{3k}{k} < 2\binom{3k}{k} = 2\binom{k+i}{k};$$
$$\binom{k+i-1}{k} = \binom{3k-1}{k} = 2\binom{3k-1}{k-1} = 2\binom{k+i-1}{k-1};$$

For $i = 2k+1$,

$$\binom{k+i}{k+1} = \binom{3k+1}{k+1} = \frac{2k+1}{k+1}\cdot\binom{3k+1}{k} < 2\binom{k+i}{k}$$
$$\binom{k+i-1}{k} = \binom{3k}{k} = \frac{2k+1}{k}\cdot\binom{3k}{k-1} > 2\binom{k+i-1}{k-1}$$

A.8 Sketch of Proof of Theorem 5

This is proved using induction on the distance between the source and destination nodes. Let $(v_1, v_2, \ldots, v_m)$ be the priority order of neighbors of v that corresponds to the maximum probability of delivery of a routing message from v. Based on the induction assumption, assume that

$$S(v_i) = a_0^{(i)} \cdot p^{d-1} + a_1^{(i)} \cdot p^d + a_2^{(i)} \cdot p^{d+1} + \ldots$$
$$P(v_i) = a_0^{(i)}; \quad 1 \le i \le m, \quad a_0^{(i)} \text{ is a positive integer.}$$

Then

$$S(v) = \sum_{i=1}^{m} p \cdot \bar{p}^{i-1} \cdot S(v_i) = (a_0^{(1)} + a_0^{(2)} + \ldots + a_0^{(m)}) \cdot p^d + \ldots .$$
$$P(v) = P(v_1) + P(v_2) + \ldots + P(v_m) = a_0^{(1)} + a_0^{(2)} + \ldots + a_0^{(m)} .$$

Therefore, $a_0 = a_0^{(1)} + a_0^{(2)} + \ldots + a_0^{(m)} = P(v)$ and a_0 is a positive integer.

REFERENCES

[1] R. Alverson, D. Callahan, D. Cummings, *et al*, "The Tera Computer System", *Technical Report*, 1991; Tera Computer Company.

[2] S. Badr, P. Poder, "An optimal shortest routing for network computers with regular mesh-connected topologies", *IEEE Trans. Computers*, vol 38, 1989 Oct, pp 1362 – 1371.

[3] W.J. Dally, "The J-machine: System support for Actors", *Actors: Knowledge-Based Concurrent Computing* (Hewitt & Agha, *Eds*), 1989; MIT Press.

[4] C. L. Seitz, *et al*, "The architecture and programming of the Ametek Series 2010 multicomputer, *Proc. Third Conf. Hypercube Concurrent Computers and Applications*, vol I, 1988 Jan, pp 33 – 36; ACM.

[6] S.L. Lillevik, "The Touchstone 30 gigaflop DELTA prototype", *Proc. 6th Distributed Memory Computing Conf*, 1991, pp 671 – 677.

[7] Y. Saad, M.H. Schultz, "Topological properties of hypercubes", *IEEE Trans. Computers*, vol 37, 1988 (7), Jul, pp 867 – 872.

[8] S. Scott, G. Thorson, "The Cray T3E network: Adaptive routing in high performance 3-torus, *HOT Interconnects IV*, 1991; Stanford Univ.

[9] T. Weller, B. Hajek, "Comments on: An optimal shortest-path routing policy for network computers with regular mesh-connected topologies", *IEEE Trans. Computers*, vol 43, 1994 Jul, pp 862 – 864.

AUTHOR

Dr. Jie Wu; Dep't of Computer Science and Eng'g; Florida Atlantic Univ; Boca Raton, Florida 33431 USA
Internet (e-mail): jie@cse.fau.edu

Jie Wu (S'88, M'89, SM'93) received his BS (1982) in Computer Engineering and MS (1985) in Computer Science from Shanghai University of Science and Technology, and his PhD (1989) in Computer Engineering from Florida Atlantic University. In 1985 – 1987 he taught at Shanghai University of Science and Technology. Since 1989 he has been with the Department of Computer Science and Engineering, Florida Atlantic University where he is a Professor and the Director of CSE graduate programs. Dr. Wu has published over 100 papers in various journal and conference proceedings: *IEEE Trans. Software Eng'g*, *IEEE Trans. Computers*, *IEEE Trans. Parallel and Distributed Systems*, *IEEE Trans. Reliability*, *J. Parallel and Distributed Computing*, *The Computer J*, and *Concurrency: Practice and Experience*. His research interests are parallel/distributed systems, fault-tolerant computing, interconnection networks, Petri net applications, and software engineering. He serves on the editorial board of the *Int'l J. Computers and Applications*. He served on the Program Committee for the 1996, 1998, 2000 IEEE Int'l Conf. Distributed Computing Systems, and the 1999, 2000 Int'l Conf. Parallel Processing. He is a Program co-Chair'n of the 1999 ISCA Int'l Conf. Parallel and Distributed Computing Systems. He is a co-Guest-Editor of the Special Issue on "Challenges in Designing Fault-Tolerant Routing in Networks" for *IEEE Trans. Parallel and Distributed Systems*, and a co-Guest-Editor of the Special Issue on "Routing in Computer Networks" for *J. Parallel and Distributed Computing*. He is the author of *Distributed System Design* from CRC press. Dr. Wu received the 1996 – 1997 Researcher of the Year Award at Florida Atlantic Univ, and the 1998 Outstanding Achievements Award from IASTED. He is an IEEE Computer Society Distinguished Visitor, and a Member of ACM and ISCA, and a Senior Member of IEEE.

Manuscript TR1998-092 received: 1998 June 20; revised: 1999 March 16
Responsible editor: S. Banerjee
Publisher Item Identifier S 0018-9529(99)08986-1