EFFICIENT FAULT TOLERANT ROUTINGS IN NETWORKS

Andrei Broder
IBM Research
San Jose, California 95193

Danny Dolev
Hebrew University, Givat Ram
91904 Jerusalem, Israel

Michael Fischer
Yale University
New Haven, Connecticut 06520

Barbara Simons
IBM Research Laboratory
San Jose, California 95193

ABSTRACT: We analyze the problem of constructing a network which will have a fixed routing and which will be highly fault tolerant. A construction is presented which forms a "product route graph" from two or more constituent "route graphs." The analysis involves the *surviving route graph*, which consists of all non-faulty nodes in the network with two nodes being connected by a directed edge iff the route from the first to the second is still intact after a set of component failures. The diameter of the surviving route graph, that is, the maximum distance between any pair of nodes, is a measure of the worst-case performance degradation caused by the faults. The number of faults tolerated, the diameter, and the degree of the product graph are related in a simple way to the corresponding parameters of the constituent graphs. In addition, there is a "padding theorem" which allows one to add nodes to a graph and to extend a previous routing.

1. Introduction.

We consider the problem of constructing a "fault-tolerant" routing in a network with an arbitrary number of nodes. This work is motivated by a practical problem of message routing in a communications network. A *route* is a path from one node to another. The message delivery system must find a route along which to send each message to its destination. If the route is known beforehand, then it can be attached to the message, allowing intermediate nodes to send the message on using only information contained in the message itself. Such a simple forwarding function can be built into fast special-purpose hardware, yielding the desired high overall network performance.

The problem is greatly simplified if one chooses a route in advance for each source/destination pair and uses that route for all messages from the one node to the other. Such a choice of routes is called a *routing table*. If the routing table is computed only once for a given network configuration, considerable effort can be put into its computation. Even this effort, however, must be kept within reasonable bounds since the routing table must be recomputed when the network configuration changes. All routes in a routing table are customarily simple paths and in addition might have other desirable properties such as being minimal length and approximately evenly distributed throughout the network.

In this paper, we are particularly concerned with the fault-tolerant properties of fixed routings. Namely, when a node or link fails, all of the routes which go through the failed component become unusable, leaving certain pairs of nodes unable to communicate in the normal way. However, assuming the network remains connected, communication is still possible by sending a message along a sequence of surviving routes. We analyze the *surviving route graph*, which consists of all non-faulty nodes in the network with two nodes being connected by a directed edge iff the route from the first to the second is still intact after a set of component failures. Then the diameter of the surviving route graph, that is, the maximum distance between any pair of nodes, is a measure of the worst-case performance degradation caused by the faults.

There are several reasons for continuing to use old routing tables even after a fault has occurred. One significant reason is that nodes must communicate in order to compute the new routing table, so some kind of interim communication mechanism is essential. Moreover, one node can broadcast to all others without knowing which routes are still intact -- it simply "floods" the network by sending its message together with a "route counter" along all of its routes; any node receiving the message increments the route counter and rebroadcasts it along all of its routes if the route counter does not exceed the bound on the diameter of the graph. This type of flooding done on route graphs frequently results in fewer

This work was supported in part by the Office of Naval Research under Contract N00014-82-K-0154, and by the National Science Foundation under Grant MCS-8116678.

messages than when done along the edges. In addition, for certain types of fault tolerant protocols, such as those used in Byzantine Agreement, a node at the endpoint of a route must do considerably more processing of messages than one which is an interior point of a route. Consequently, the time it takes for a message to reach all other nodes is proportional to the diameter of the surviving route graph.

A further application for this model is the case of a network that reconfigures itself according to some shortest path strategy at certain (relatively rare) intervals. If one wishes to run a protocol on such a network in which it is assumed that messages between two nodes are always delivered so long as neither of the nodes is either down or disconnected, then the message can be sent over the routes of the surviving route graph. As mentioned above, if one assumes more extensive processing at nodes which are the endpoints of routes, then the maximum delivery time for a message is proportional to the diameter of the surviving route graph. The length of the diameter of the surviving route graph is utilized in a clock synchronization algorithm [HSS] which has been developed for an arbitrary network that might contain faults. The algorithm has already been implemented in the Highly Available Systems Project currently under development at IBM.

This problem was introduced by Dolev, et al. [DHSS83]. In it, they establish properties of routings in general networks. They also give a routing for a specific network (a t-dimensional hypercube) which can tolerate up to $t-1$ faults and still have a surviving graph of diameter at most 2. In terms of N, the number of nodes in the graph, their construction tolerates up to $t = \log_2 N - 1$ faults and can be applied whenever N is a power of two. Also, the degree of each node is $t + 1$.

In this paper, we look at the problem of finding good routings for networks where the number of nodes is not a power of two. We have a general construction which allows one to form a "product route graph" from two or more constituent route graphs. The tolerance, diameter, and degree of the product graph are related in a simple way to the corresponding parameters of the constituent graphs, although the construction of the routing on the product graph is definitely non-trivial. Applying this construction repeatedly to simple 2-node graphs yields the cube result of Dolev et al. However, other cardinality graphs can be obtained by starting with a different basis. In addition, we have a "padding theorem" which allows us to add nodes to a graph and extend a previous routing.

As an example, using the 2-node, 3-node, and 5-node starting graphs of figure 1, one can construct a routed graph of any cardinality N of the form $2^i 3^j 5^k$. The resulting graph will tolerate $i+2j+2k-1$ faults, have degree $i+2j+2k$, and have surviving diameter of 2. Alternatively, if the complete graph on 5 nodes is substituted for the 5-cycle, the resulting graph will tolerate $i+2j+4k-1$ faults, have degree $i+2j+4k$, and have surviving diameter of 2. Note that in both cases the fault tolerance is optimal in that any larger group of faults might disconnect the network.

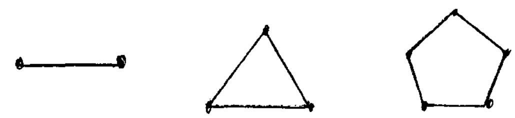

Figure 1: potential building blocks.

In addition to providing a constructive technique for building networks and providing them with fault-tolerant routings, our approach provides the network designer with a powerful tool. As the above example illustrates, s/he can use sparse or dense "basic" graphs in constructing the product graph according as s/he is interested in minimizing the number of links or maximizing the fault-tolerance.

2. Graph Routing

A network is modeled as an undirected graph $G=(V,E)$, with nodes representing processors and edges representing communication links. We do not allow self-loops or parallel edges. A *routing* assigns to any pair of nodes in the network a fixed path between them. All communications between these nodes will travel along this path. When speaking of a path between x and y in G, we use the notation $\pi_G(x,y)$. A *minimal length* routing is one that always gives a path of minimal length.

More formally, define $Path_G(x,y)$ to be the set of simple paths between the nodes x and y in G and Path(G) to be the set of all simple paths in G. A *routing* is a partial function $\rho: V \times V \rightarrow Path(G)$ such that $\rho(x,y) \in Path(x,y)$. (If $Path_G(x,y) = \phi$, then $\rho(x,y)$ is undefined). We call $\rho(x,y)$ the *route from x to y*. A *shortest path routing* is a routing ρ such that for every pair (x,y), $\rho(x,y)$ is a shortest path between x and y. A routing ρ induces the *route graph* $R(G,\rho) = (V, Dom(\rho))$, where $Dom(\rho)$ is the domain of definition of ρ. If ρ is defined for every pair x,y for $x \neq y$, then $R(G,\rho)$ is the complete graph on $|V|$ nodes. We shall abbreviate $R(G,\rho)$ as R and $\pi_G(x,y)$ as $\pi(x,y)$ whenever such an abbreviation is unambiguous.

Let $\rho(x,y)\rho(y,z)$ be the route from x to y followed by the route from y to z. The function ρ can be extended to a function on V^* in an obvious way: $\rho(x_1,x_2,x_3,...) = \rho(x_1,x_2)\rho(x_2,x_3)....$ In particular, given a path $\pi(x_1,x_k) = x_1x_2...x_k$, then $\rho(\pi(x_1,x_k)) = \rho(x_1,x_2)\rho(x_2,x_3)...\rho(x_{k-1},x_k)$. Let $V_{\rho(x,y)}$ be the set of nodes in $\rho(x,y)$. A routing is *consistent* if for all x,y such that $\rho(x,y)$ is defined and for all z such that $z \in V_{\rho(x,y)}$, $\rho(x,y) = \rho(x,z)\rho(z,y)$.

A *fault* in G is either a node or an edge in G. A route is *affected* by a fault if the fault is contained in it. Note that one fault may affect several routes. Given a set F of faults in G, we define the fault free routing ρ/F to be the reduction of ρ to fault free routes. As above, the fault free routing ρ/F induces the *surviving route graph* $R(G,\rho)/F = (V/F, \mathrm{dom}(\rho/F))$, where V/F consists of all non-faulty nodes in G. We use the notation R/F for $R(G,\rho)/F$ where it is unambiguous.

A (shortest path) routing ρ is called *(d,f)-tolerant* if for every set F of f faults in G, $R(G,\rho)/F$ has diameter at most d. A graph G is called (d,f)-tolerant if there exists a shortest path routing ρ on G that is (d,f)-tolerant. Note that if G is (d,f)-tolerant, then the degree of any node in G is at least f+1.

Lemma 1. If ρ is consistent, then for every set F of faults in G, ρ/F is consistent.

Proof. Immediate. □

Lemma 2. If G is (d,f)-tolerant and f>0, then d>1.

Proof. Let F=(x,y) for some edge (x,y) in G. Since $\rho(x,y)=(x,y)$ for any shortest path routing ρ, $d \geq 2$. □

Lemma 3. Let ρ be a consistent routing of G and let x,y be any pair of nodes in G. Let F be a set of faults such that $\rho(x,y)$ contains a fault but there is a path $\pi_{R/F}(x,y)$ from x to y in R/F which does not contain any faults. Then there exists a node on $\pi_{R/F}(x,y)$ which is not on $\rho(x,y)$.

Proof. Let $V_{\pi(x,y)}$ be the set of nodes in $\pi_{R/F}(x,y)$ and assume to the contrary that $V_{\pi(x,y)} \subseteq V_{\rho(x,y)}$. Let $\rho(x,y) = x_0x_1...x_k$, where $x_0=x$ and $x_k=y$, and let $\pi_{R/F}(x,y)=x'_0x'_1...x'_m$, where $x'_0=x$ and $x'_m=y$. Let I be the largest number less than k such that the edge $(x_I,x_{I+1}) \in F$, and let J be the largest number less than m such that $x'_J = x_i$ for some $i \leq I$. Then $x'_Jx_{i+1}x_{i+2}...x'_{J+1}$ is a route by the consistency assumption with respect to $\rho(x,y)$, and by construction it contains a fault. This contradicts the assumption that $\rho(\pi_{R/F}(x,y))$ is fault-free. □

Lemma 4. Let ρ be a (d,f)-tolerant consistent routing with f>0, and x,y a pair of distinct nodes in G. For every set F of faults with $|F| < f$, there exists a path $\pi_{R/F}(x,y)$ of length at most d such that $\rho(\pi_{R/F}(x,y))$ is fault free and $\pi_{R/F}(x,y)$ contains a node that is not on $\rho(x,y)$.

Proof. Let F′ be the set of faults F together with an edge from $\rho(x,y)$. F′ contains at most f faults, so by definition there exists a path $\pi_{R/F'}(x,y)$ from x to y such that $\rho(\pi_{R/F'}(x,y))$ does not contain any faults in F′. By Lemma 3, $\pi_{R/F'}(x,y)$ contains a node that is not on $\rho(x,y)$. □

3. Product of routings.

Given two graphs $G = (V_G, E_G)$ and $H = (V_H, E_H)$, their *cartesian product* $G \times H$ is a graph (V, E), where $V = V_G \times V_H$ and $((i,j),(k,l)) \in E$ iff both (i,j) and (k,l) are nodes in V and either i=k and $(j,l) \in E_H$ or j=l and $(i,k) \in E_G$. The *H plane defined by i (G plane defined by j)* in $G \times H$ is the subgraph of $G \times H$ determined by all nodes having the first (second) coordinate equal to i (j). We use the notation H_i and G_j for the H plane defined by i and the G plane defined by j. Isomorphic graphs being considered equal, it can be shown that the cartesian product of graphs is commutative, and that any graph can be uniquely decomposed into a cartesian product of indecomposible graphs. For details see [Sa].

Let ρ_G and ρ_H be given routings for G and H, and let x=(i,j) and y=(k,l). We define the product routing $\rho_G \times \rho_H$ as follows. $\rho_G \times \rho_H(x,y) = \rho_H(x,z)\rho_G(z,y)$, where z=(i,l). In other words, the route is obtained by concatenating the route $\rho_H(x,z)$ of H_i with the route $\rho_G(z,y)$ of G_l. Clearly, if i=k or j=l, then one of these routes is the null route. In this case, we say that x and y are *coplanar*. The routing $\rho_G \times \rho_H$ is a consistent routing iff both ρ_G and ρ_H are consistent. From now on we shall denote $\rho_G \times \rho_H$ by ρ_{GXH}, although of course other routings are possible.

Let x = (i,j) and y = (i′,j′) be two nodes that are not coplanar in $G \times H$, and let F be the set of faults in $G \times H$. We associate to x and y a copy of G, called *G(x,y)*, with the set of faults $F_G(x,y)$. $F_G(x,y)$ is defined as follows:

a) if the edge $(k,l) \in E_G$, then $(k,l) \in F_G(x,y)$ if either the edge between (k,j) and (l,j) or the edge between (k,j′) and (l,j′) is faulty (in $G \times H$).

b) if $k \in V_G$ and $k \notin \rho_G(i,i')$, then $k \in F_G(x,y)$ if $\rho_H((k,j),(k,j'))$ is faulty.

c) if $k \in V_G$, $k \in \rho_G(i,i')$, and $k \neq i,i'$, then $k \in F_G(x,y)$ if either of the nodes (k,j) or (k,j′) is faulty.

$H(x,y)$ and $F_H(x,y)$ are similarly defined. Note that the nodes j, j′ in H(x,y) and i, i′ in G(x,y) are always nonfaulty.

Lemma 5. Any fault in F determines a fault in at most one of $F_G(x,y)$ and $F_H(x,y)$.

Proof. Let x=(i,j) and y=(i′,j′). Suppose that there is an edge fault $f_1 = ((l,k),(l',k)) \in F$. If f_1 determines an edge fault $(l,l') \in E_G$, then it must satisfy condition a) and, therefore, either k=j or k=j′. Suppose by contradiction that f_1 also determines a node fault. Since $k \in V_H$, f_1 can determine a node fault only in $F_H(x,y)$. For condition b) to hold, $k \notin \rho_H(j,j')$, which is clearly impossible, since either k=j or k=j′. If condition c) holds, then $k \neq j,j'$, also impossible.

Suppose that f_1 does not determine an edge fault in E_G, i.e. condition a) does not hold. Then $k \neq j,j'$. Note that at most one of conditions b) and c) can hold, and therefore f_1 can determine at most one fault. (The proof for edge faults of the form ((l,k),(l,k′) is similar).

Now suppose that there is a node fault $f_2 = (k,l) \in F$. By definition, a node fault in G×H cannot determine an edge fault in $F_G(x,y)$ or $F_H(x,y)$. Suppose that f_2 determines a node fault in $F_G(x,y)$. If condition b) holds, then $k \notin \rho_G(i,i')$ and $\rho_H((k,j),(k,j'))$ is faulty. In particular, $(k,l) \in \rho_H((k,j),(k,j'))$. If f_2 also determines a node fault in $F_H(x,y)$, then it must do so by condition c) (since $l \in \rho_H(j,j')$). But for condition c) to hold, $l \neq j,j'$ and one of the nodes (i,l) or (i,l′) is faulty. This implies that k=i or k=i′, which contradicts the assumption that $k \notin \rho_G(i,i')$ (i.e. the assumption of condition b) by which f_2 determines a fault in $F_G(x,y)$).

Finally, suppose that f_2 determines a node fault in $F_G(x,y)$ under condition c). Therefore, $k \in \rho_G(x,y)$, $k \neq i,i'$, and one of the nodes (k,j) or (k,j′) is faulty. This implies that l=j or l=j′. If f_2 determines a node fault in $F_H(x,y)$, then since l=j or l=j′, condition c) cannot hold. For condition b) to hold, $l \notin \rho_H(j,j')$. Since l=j or l=j′, this is clearly impossible. □

Corollary. $|F_G(x,y)| + |F_H(x,y)| \leq |F|$.

Lemma 6. Assume ρ_G is (d_G,f_G)-tolerant, ρ_H is (d_H,f_H)-tolerant, at least one of f_G and f_H is greater than 0, and both are consistent routings. Let x,y be two nodes in G×H that are not coplanar. Then for every set F of faults such that $F_G(x,y)$ (resp. $F_H(x,y)$) contains fewer than f_G (resp. f_H) faults, the distance between x and y in $R(G \times H,\rho_{G \times H})/F$ is at most d_G (resp. d_H).

Proof. Let x=(i,j) and y=(i′,j′). Without loss of generality, assume that $f_G>0$ and that $F_G(x,y)$ contains $f<f_G$ faults. By lemma 2, $d_G \geq 2$. By lemma 4 there exists a path π of length $\leq d_G$ in $R(G,\rho_G)/F_G(x,y)$ from i to i′ such that $\rho_G(\pi)$ is fault-free and π contains a node which is not on $\rho_G(i,i')$.

We first show a fault-free path in G×H from x to y and then prove that its length in $R(G \times H,\rho_{G \times H})/F$ is bounded by d_G. Let k be a node on π that is not on $\rho_G(i,i')$ and let l be the node on π immediately after k (i.e. (k,l) is an edge in $R(G,\rho_G)/F_G(x,y)$). Denote $\pi=\pi_1(k,l)\pi_2$. Note that $k \neq i'$ but that l might equal i′, in which case $\pi_2=\phi$. By the definition of $F_G(x,y)$, since $\rho_G(\pi)$ had no faults in $F_G(x,y)$, $\rho_G(\pi_1)$ is fault free in the G_j. Similarly, both $\rho_G(k,l)$ and $\rho_G(\pi_2)$ are fault-free in the $G_{j'}$. By condition b) of the definition of $F_G(x,y)$, $\rho_H((k,j),(k,j'))$ is fault-free (i.e. $\rho_H((k,j),(k,j'))$ contains no fault from F). Therefore, the path in G×H composed of the corresponding $\rho_G(\pi_1)\rho_H((k,j),(k,j'))\rho_G((k,j'),(l,j'))\rho_G(\pi_2)$ is fault-free. But from the definition of the routing in G×H, it follows that $\rho_H((k,j),(k,j'))\rho_G((k,j'),(l,j'))$ form just one route. Hence, this path is of length at most d_G in $R(G \times H,\rho_{G \times H})/F$.

The proof for $F_H(x,y)$ is similar. The only difference is that we have to take l to be the node immediately preceding k in π to get a path of length d_H in $R(G \times H,\rho_{G \times H})/F$. □

Theorem 1. Let G be (d_G,f_G)-tolerant and H be (d_H,f_H)-tolerant with consistent (d_G,f_G)- and (d_H,f_H)-tolerant routings ρ_G and ρ_H respectively. Then the graph G×H is $(\max\{d_G,d_H,2\}, f_G+f_H+1)$-tolerant.

Proof. Let $\rho_{G \times H} = \rho_G \times \rho_H$. We will show that $\rho_{G \times H}$ is $(\max\{d_G,d_H,2\},f_G+f_H+1)$-tolerant. It suffices to show that for any pair of nodes x=(i,j) and y=(i′,j′) and every f_G+f_H+1 faults in the product graph, there exists a path of length bounded by $\max\{d_G,d_H,2\}$ from x to y in the graph $R(G \times H,\rho_{G \times H})/F$. The proof is by cases.

Case 1: i=i′.

Case 1.1: H_i contains f_H or fewer faults. We are done because H_i itself is (d_H,f_H)-tolerant.

Case 1.2: H_i contains at least f_H+1 faults. Both node x and node y have at least f_G+1 corresponding adjacent nodes in their respective G planes. Each pair of corresponding adjacent nodes has a route joining them in an H plane. These planes are both mutually distinct and also different from H_i. Therefore, these adjacent nodes define at least f_G+1 node disjoint paths, each utilizing a different

H plane, connecting x and y. Since we have at most f_G faults among these node disjoint paths, at least one of them is fault-free. Each of these f_G+1 disjoint paths is composed of exactly two routes in $\rho_{G \times H}$. The first route consists of the edge from x to the H plane and the second consists of the path in that H plane followed by the edge to y. Therefore, each one is of length 2 in $R_{G \times H}/F$.

Case 2: $j=j'$. The proof is similar to case 1.

Case 3: $i \neq i'$ and $j \neq j'$. The remainder of the proof is an analysis of G(x,y) and H(x,y).

Case 3.1: Either $|F_G(x,y)| < f_G$ or $|F_H(x,y)| < f_H$. Then the result follows from lemma 6.

Case 3.2: Both $|F_G(x,y)| = f_G$ and $|F_H(x,y)| = f_H$. If either $\rho_G(i,i')$ or $\rho_H(j,j')$ contains a fault, then the result follows from techniques similar to those of lemma 6. So suppose that both $\rho_G(i,i')$ and $\rho_H(j,j')$ are fault-free in G(x,y) and H(x,y) respectively. By lemma 5 there can be at most one fault in $G \times H$ which has not determined a fault in either $F_G(x,y)$ or $F_H(x,y)$. If this fault is not in either $\rho((i,j),(i,j'))$ or $\rho((i,j'),(i',j'))$, then the route from x to y is fault-free, and the distance from x to y in $G \times H$ is one. If there is a fault in either of the above routes, then by the definition of $F_G(x,y)$ and $F_H(x,y)$, it must be the point (i,j'). Therefore, the point (i',j) must be fault-free and a path of length two from x to y in $R_{G \times H}/F_{G \times H}$ can be obtained by concatenating $\rho_G(i,i')$ in G_j with $\rho_H(j,j')$ in $H_{i'}$.

Case 3.3: Either $|F_G(x,y)| = f_G$ and $|F_H(x,y)| = f_H+1$ or $|F_G(x,y)| = f_G+1$ and $|F_H(x,y)| = f_H$. Assume $|F_G(x,y)| = f_G$ and $|F_H(x,y)| = f_H+1$. If $\rho_G(i,i')$ has a fault, the proof follows using previous techniques. So suppose that $\rho_G(i,i')$ is fault-free. If $\rho_H(j,j')$ is fault free in H_i, then $\rho((i,j),(i,j'))\rho((i,j'),(i',j'))$ is a path of length one. If $\rho_H(j,j')$ is fault free in $H_{i'}$, then $\rho((i,j),(i',j))\rho((i',j),(i',j'))$ is a path of length two. So suppose that $\rho_H(j,j')$ contains a fault in both H_i and $H_{i'}$. Since both H planes contain at least one fault, neither contains more than f_H faults. Therefore, we can travel from (i,j) to (i,j') in H_i along a path of length no greater than d_H. If $\rho((i,j')(i',j'))$ is concatenated to this path, the length of the path is not increased. The proof is similar if instead we have $F_G(x,y)=f_G+1$ and $|F_H(x,y)| = f_H$. □

4. Padding Graphs.

Theorem 2. Let G = (V,E) be (d,f)-tolerant with every node in G having degree no greater than μ_G. Then for $|V| < N < |V| + (|V|/\mu^2)$, G can be extended to a graph $G'=(V',E')$ and a routing ρ' such that G' is (d,f)-tolerant, $|V'| = N$, and the maximum degree in G' is no more than $\mu+1$.

Proof. We extend G to a graph $G'=(V',E')$ with $|V'|=N$ as follows. Let $M=N-|V|$. Match one of the new nodes, say x, to a node in the original network, say x', and connect x to all of x's neighbors in G. Next, choose another new node, say y, and match it to a node in the original network, say y', which has no neighbors in common with x' in G. Connect y to all the neighbors of y'. This procedure can be repeated so long as there exist nodes in G which are neither matched to a new node nor have neighbors in common with an already matched node. Note that each iteration eliminates at most μ^2 nodes from G, since both x' and each of its neighbors have degree at most μ.

Let ρ be a routing in G which is (d,f)-tolerant. We extend ρ to ρ' as follows. For $x,y \in V$, $\rho'(x,y)=\rho(x,y)$. For $x \in V'-V$ and $y \in V$, let x' be the node in V to which x is matched. If y is a neighbor of x', then $\rho'(x,y)=(x,y)$. If y is not a neighbor of x', then let w be the neighbor of x' which lies on $\rho(x',y)$. We define $\rho'(x,y)$ to be the same as $\rho(x',y)$ with the edge (x',w) replaced by the edge (x,w). $\rho'(y,x)$ is similarly defined to be $\rho(y,x')$ with its last edge (w,x') replaced by the edge (w,x). For $x,y \in V'-V$, let x' and y' be the nodes in V to which x and y are matched, and let w and v be the neighbors of x' and y' respectively which lie on $\rho(x',y')$. Note that by construction, $w \neq v$. Then, $\rho'(x,y)$ is the same as $\rho(x',y')$ with the edge (x',w) replaced by (x,w) and (v,y') replaced by (v,y).

The consistency of ρ' follows trivially from the consistency of ρ. Since G tolerates f faults and since all new nodes are connected to at least $f+1$ distinct nodes in G, it is easy to show that G' tolerates f faults.

We now show that $R(G',\rho')/F$ has diameter no greater than d for $|F| \leq f$. Let $x,y \in V'-V$ with x matched to x' and y to y' $(x',y' \in V)$, and assume G' contains at most f faults. Let F' consist of the set F with the following two changes: 1) $x',y' \notin F'$, 2) if $(x,w) \in F$, then $(x,w) \notin F'$ but $w \in F'$. Note that $|F'| \leq f$. Therefore, there exists a path in $R(G,\rho)/F'$ from x' to y'. Replacing x' by x and y' by y gives a path from x to y in $R(G,\rho)/F$. We leave to the reader the verification that the distance between nodes in $R(G',\rho')$, when at least one of the nodes is in V, remains no greater than d.

If the maximum indegree in the original graph is μ, then in the new graph we have degree $\mu+1$. It is straightforward to generalize this construction to handle the case where $|V| < N < 2|V|$. As N increases in size relative to $|V|$, however, the maximum degree of G' increases accordingly. □

A construction similar to the one in the padding lemma can be used to extend a graph with $|V|$ nodes to a graph with up to $2|V|$ nodes at the cost of at most doubling the maximum degree while maintaining the same diameter and fault tolerance.

5. Other bounds.

For a graph G of n_G nodes, denote by η_G the minimum degree of the nodes in G.

Theorem 3. $G \times H$ is (3,f)-tolerant, where $f = \max\{\min\{\eta_H, n_G-1\}, \min\{\eta_G, n_H-1\}\}$.

Proof. Let $G \times H$ have fewer than f faults. Without loss of generality assume $f = \min\{\eta_H, n_G-1\}$. Therefore, there is at least one fault-free H plane; denote it by H_k. Define ρ_{GXH} as before, with the difference being that ρ_G and ρ_H are arbitrary (not necessarily shortest path) routings on G and H.

Let x,y be any two nodes in $G \times H$. Assume first that y is not a neighbor of x. To each of the η_H neighbors u of x in its H plane, associate a different neighbor v of y in its H plane, or u itself if u is also a neighbor of y. Let U be the set of pairs constructed in this manner together with the pair (x,y). The set U defines in an obvious way η_H+1 paths from x to y, all of them going through H_k and disjoint outside H_k. Therefore, at least one of them is fault free and has length no more than 3 in the induced graph.

In the case that x and y are neighbors, if the edge (x,y) is not faulty, the distance is 1. Otherwise, the corresponding set U will have η_H pairs with at worst η_H-1 faults on them. □

Using similar observations about faultiness in G and H planes one can obtain other bounds similar to the one in Theorem 3.

6. Remarks.

The proof of the main theorem can be greatly simplified if the following conjecture due to Joe Halpern is true. Let G be a (d,f)-tolerant graph and let ρ be a (d,f)-tolerant consistent routing on G. Then between every pair of nodes in $R(G,\rho)$ there are at least f+1 node disjoint paths $\pi_1, \pi_2, ..., \pi_{f+1}$ of length d or less such that the paths $\rho(\pi_1), \rho(\pi_2), ..., \rho(\pi_f)$ are node disjoint. This property does not hold for inconsistent routings.

7. Open problems.

Our "building blocks" usually will be small graphs with a prime number of nodes, $p_1, p_2, ...$ Starting from these blocks, we can construct (2,f)-tolerant graphs that have $p_1^i p_2^j p_3^k ...$ nodes. If we want to construct a (2,f)-tolerant graph with N nodes and if the gaps in such a sequence are not greater than $O(N/(\log N)^2)$, then we can use a generalization of the padding theorem to construct such graphs where the maximum degree is less than $\log N + c$ for some constant c independent of N. Hence we have the following number theoretic question: what is the minimum number of prime numbers such that, for any N, the gaps in the above sequence are no greater than $O(N/(\log N)^2)$? It seems plausible that the answer is 3 and that the desired bound can be obtained using 2-, 5-, and 7-cycles. (For 2-, 3-, 5-, and 7-cycles the maximum gap up to 10,000 nodes is 199). For known results on this problem, see [Ti1], [Ti2], and references therein.

In general, we would like to know what is the optimum N node graph and what is its optimum routing for any N given a desired (d,f)-tolerance.

Acknowledgements. We would like to thank Joe Halpern both for pointing out an error in an earlier version of Theorem 3 and for simplifying some definitions, Peter van Emde Boas for informing us of the work of Tijdeman, and Jeffrey Lagarias for the discussion about Diophantine approximations.

References

[DHSS] D. Dolev, J. Halpern, B. Simons, and R. Strong, "A new look at fault tolerant network routing," to appear.

[HSS] J. Halpern, B. Simons, and R. Strong, "An efficient fault-tolerant algorithm for clock synchronization," *IBM Research Report RJ4094,* Nov., 1983.

[Sa] G. Sabindussi, "Graph Multiplication", *Math. Zeitschr.,* 72(1960), 446-457.

[Ti1] R. Tijdeman, "On integers with many small prime factors," *Compositio Mathematica,* Noordhoff International Publishing, 26(1973), 319-330.

[Ti2] R. Tijdeman, "On the maximal distance between integers composed of small primes," *Compositio Mathematica,* Noordhoff International Publishing, 28(1974), 159-162.