

基于扩展安全级的 Torus 网络容错路由算法研究

梁家荣¹, 徐霜¹, 伍华健²

LIANG Jia-rong¹, XU Shuang¹, WU Hua-jian²

1. 广西大学 计算机与电子信息学院, 南宁 530004

2. 玉林师范学院 数计系, 广西 玉林 537000

1. School of Computer and Electrical Information, Guangxi University, Nanning 530004, China

2. Department of Mathematics and Computer Science, Yulin Normal University, Yulin, Guangxi 537000, China

LIANG Jia-rong, XU Shuang, WU Hua-jian. Research on fault-tolerant routing algorithm in Torus using extended safety levels. *Computer Engineering and Applications*, 2009, 45(29): 99-101.

Abstract: How to design fault-tolerant routing in multiprocessors with fault nodes is an issue of fault-tolerant research. Two-dimensional tori with faulty blocks is studied. A sufficient and necessary condition for minimal routing is proposed. This approach is based on the concept of extended safety level, which is a special form of limited fault information. Based on extended safety level, an approach which is used to establish RMP (Region of Minimal Path) is presented. The research can provide theoretical basis for finding shortest fault-tolerant routing path in Torus.

Key words: fault-tolerant routing; fault block model; extended safety level

摘要: 在存在故障结点的网络中如何设计最小容错路由是网络容错研究中的一个热点问题。以存在矩形故障块的二维 Torus 网络为例, 将扩展安全级运用到 Torus 中, 对于网络中任意一对结点, 给出存在最小路径的充要条件; 并且结合扩展安全级的概念, 给出建立最小通路区的方法, 并用实验验证了方法的可行性。研究为存在故障结点的 Torus 网络寻找最小容错路径提供了理论依据。

关键词: 容错路由; 矩形故障块; 扩展安全级

DOI: 10.3778/j.issn.1002-8331.2009.29.029 文章编号: 1002-8331(2009)29-0099-03 文献标识码: A 中图分类号: TP302.8

1 引言

直连网络是一种常见的网络拓扑形式, 已经广泛应用于多处理器系统, 多计算机系统, 以及集群系统中。随着并行计算机互连网络和 VLSI 技术的迅速发展, 系统中的并行处理机越来越多, 仍采用传统的网络互联结构已不能满足需求。于是人们对并行计算机互连网络拓扑结构进行了大量的研究, 并对其中的一些拓扑结构已研制出了相应的商用和研究用的并行计算机系统^[1-4]。Torus 网络是一种完全对称的直连网络拓扑结构, 它具有很多优秀的网络特性, 如规则对称性, 路径多样性以及良好的扩展性。因此它广泛应用于许多商用系统中, 例如, 2004 年底评出的全球超级计算机 TOP100 中排名首位的 IBM Blue Gene/L 就采用 Torus 网络; 而另一家通信设备制造商, Avici 公司在其推出的世界上第一台太比特路由器中也采用 Torus 网络作为其交换网络拓扑^[5]。

随着多处理器系统中处理器数目的增加, 处理器出现故障的概率也随之增加。因此, 对采用 Torus 结构的多处理器系统进行容错研究是非常必要的。假设 G 是一个网络, 网络 G 的“容错性”是指满足以下条件的最大的整数 k : 在网络 G 中去掉任

意 k 个结点和相应的边后不会破坏网络的连通性^[6]。二维 Torus 网络是二维 mesh 网络的扩展, 它比 mesh 网络具有更好的性能, 近十几年来人们对 mesh 网络的容错进行了大量的研究, 但专门研究 Torus 网容错的文献却相对较少^[7]。在该文中将扩展安全级的概念应用到二维 Torus 中, 给出存在最小路径的充要条件, 最后根据扩展安全级给出寻找最小通路区的方法。主要思想就是在存在故障结点的情况下, 尽量找出最小路径。

2 预备知识

为便于理解, 首先介绍 Torus 中的一些相关概念。

定义 1^[8](n 维 Torus 网络) 它由 $k_0 \times k_1 \times \dots \times k_{n-1}$ 个结点构成, 其中 k_i 表示第 i 维的结点数。网络上的每一个结点都可用一个 n 维向量 $(x_0, x_1, \dots, x_{n-1})$ 表示, 其中 $0 \leq x_i \leq k_{i-1}$ 。结点 $(x_0, x_1, \dots, x_{n-1})$ 和结点 $(y_0, y_1, \dots, y_{n-1})$ 相连接的条件是 *iff* $\exists i$ 使 $x_i = (y_i \pm 1) \bmod k_i$ 而 $\forall j \neq i$, 有 $x_j = y_j$ 。图 1 是一个 5×5 的二维 Torus 网络结构。

定义 2^[8](故障块模型) 在二维 Torus 中, 一个非故障结点的状态开始标记为有效结点, 如果它有两个或者更多的无效或者故障结点, 则该结点的状态被标记为无效。连接无效结点和

基金项目: 国家自然科学基金(the National Natural Science Foundation of China under Grant No.60564001); 教育部优秀人才支持计划专项(No. NCET-06-0756); 广西省自然科学基金(the Natural Science Foundation of Guangxi Province of China under Grant No.0832286)。

作者简介: 梁家荣(1966-), 男, 博士, 教授, 主要研究领域为粗糙集, 数据挖掘; 徐霜(1982-), 女, 硕士, 主要感兴趣的研究方向为: 并行计算, 网络容错; 伍华健(1964-), 男, 副教授, 主要研究领域为计算机理论与研究。

收稿日期: 2008-06-03 **修回日期:** 2008-08-25

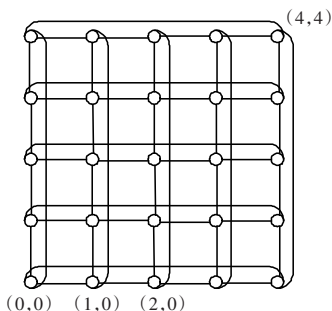


图1 5x5 的二维 Torus 网络

故障结点就构成一个故障块。

如果一个矩形区域的四个顶点的坐标为: $(x, y), (x, y'), (x', y)$ 和 (x', y') , 则可用 $[x: x', y: y']$ 来表示该矩形区域。

定义 3 (扩展安全级) 在一个规模为 $k_1 \times k_0$ 的二维 Torus 中, 结点的扩展安全级用一个四元组 (E, S, W, N) 表示, E 表示在向东方向 $\lfloor k_0/2 \rfloor$ 的距离内该结点与离其最近的故障块的距离, S 表示在向南方向 $\lfloor k_1/2 \rfloor$ 的距离内该结点与离其最近的故障块的距离, W, N 的定义分别与 E, S 的定义类似。如果在相应的方向上在规定的距离内没有故障块则用 $-$ 表示, 如果一个结点的扩展安全级为 $(-, -, -, -)$, 则称该结点是安全的。

3 扩展安全级与最小路径

路由是把消息从源结点发送到目的结点的过程, 在 mesh 网中, 源结点与目的结点之间的最小路径指的是路径长度等于两结点之间的 Hamming 距离的所有路径。但在 Torus 网中由于环的存在, 使得两结点之间的最小路径可能比 Hamming 距离还小, 所以不能像在 mesh 网中那样对结点之间的最小路径进行笼统的分析。下面对 Torus 网中不同的情况分别加以分析。

设源结点的坐标为 (i', j') , 目的结点的坐标为 (i, j) , d 为两结点之间的最小距离, 假设 $i' \leq i, j' \leq j$, 其他情况可类似讨论。当网络中没有故障结点时,

$$i-i'+j-j'; i-i' \leq \frac{k_0}{2}, j-j' \leq \frac{k_1}{2}$$

$$d=i'+k_0-i+j-j'; i-i' > \frac{k_0}{2}$$

$$j'+k_1-j+i-i'; j-j' > \frac{k_1}{2}$$

结合扩展安全级的概念以及以上讨论, 给出以下定义:

定义 4 设在二维 Torus 网中, (i', j') 为源结点, (i, j) 为目的结点, $i' \leq i, j' \leq j$, 源结点 (i', j') 相对目的结点 (i, j) 的扩展安全级为 (E, S, W, N) , 则

- (1) 当 $i-i' \leq k_0/2, j-j' \leq k_1/2$ 时, 若 $i-i' \leq E, j-j' \leq N$;
- (2) 当 $i-i' > k_0/2$ 时, 若 $i+k_0-i' \leq W, j-j' \leq N$;
- (3) 当 $j-j' > k_1/2$ 时, 若 $i-i' \leq E, j'+k_1-j \leq S$, 则

称结点 (i', j') 相对结点 (i, j) 是扩展安全的。

下面给出存在最小路径的一个充分条件:

定理 1 在二维 Torus 网中, 设 (i', j') 为源结点, (i, j) 为目的结点, $i' \leq i, j' \leq j$, 如果源结点 (i', j') 相对目的结点是扩展安全的, 则 (i', j') 和 (i, j) 之间至少存在一条最小路径。

首先以定义 4 中第(1)种情况(即最小路径不经过环)为例证明定理 1。

证明 设 t 为区域 $[i': i, j': j]$ 间的故障块的数量,

当 $t=1$ 时, 如图 2 所示, 八条直线将区域 $[i': i, j': j]$ 分成八个部分, C, D 分别为源结点和目的结点, P 为直线 $x=i'$ 与直线 $y=j'$

的交点, Q 为直线 $x=i$ 与直线 $y=j'$ 的交点; 由于 $i-i' \leq E, j-j' \leq N$, 所以没有故障块穿过线段 CP 和 CQ , 按照 $X-Y$ (或 $Y-X$) 的路由方式, 从 (i', j') 到 (i, j) 都存在最小路径。

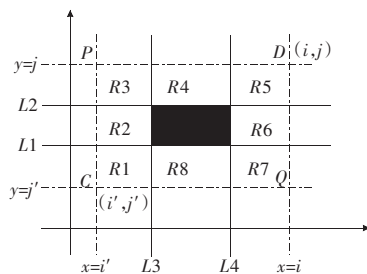


图2 只有一个故障块的 2D Torus 中的最小路由

假设 $t=k-1$ 时, $(i', j'), (i, j)$ 间存在最小路径; 则

当 $t=k$ 时, 如果没有故障块穿过 PD 之间的区域, 则从 $C \rightarrow P \rightarrow D$ 存在一条最小路径。否则, 假设此区间内存在一个故障块 B , 如图 3 所示, 假设 u 为环绕 B 的两条直线 $L1$ 与 $L2$ 的交点, 则从 u 开始沿着直线 $L2$ 向北, 当到达直线 $y=j$ 后向东转, 然后继续向东则可到达目的结点 D , 所以从 u 到 D 存在最短路径。

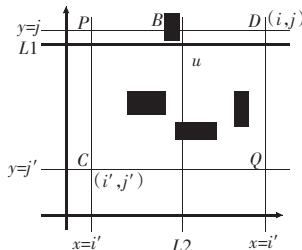


图3 寻找最小路径

现在证明从 (i', j') 到 u 存在最小路径。注意到从 u 到 (i', j') 的任何一条最短路径都不会经过 B 中的任一结点, 也就是说 B 的存在并不影响从 (i', j') 到 u 的最小路径的存在, 所以可以去掉 B , 此时故障块的数量为 $k-1$, 又因为已经假设 $t=k-1$ 时, 从 (i', j') 到 (i, j) 存在最小路径, 所以根据该假设可知, 此时从 (i', j') 到 u 存在最小路径。

综合以上两点可知, 当 $t=k$ 时从 (i', j') 到 (i, j) 存在最小路径。

所以用归纳法可证明定理 1 的正确性。

定义 4 中的第(2)和第(3)种情况是最小路径需要经过环的, 为了便于研究, 把环展开, 如图 4 所示, 假设 C_1 为源结点, D_1 为相应的目的结点, C_1 和 D_1 的坐标满足条件 $j-j' > k_1/2, D_1'$ 为展开环将网络中第 j 行到第 k_1-1 行的结点向下平移后 D_1 的位置; C_1 和 D_2 的坐标满足条件 $i-i' > k_0/2, C_1'$ 为展开环将第 0 列到第 i' 列的结点平移到网络的右边界后 C_1 的位置, 图中的实线框为网络的边界, 此时可以发现对这样的两对结点完全可以在区域 $R1$ 和 $R2$ 中分别进行分析, 而这两种简化后的情况与(1)中的情况是完全一样的, 源结点和目的结点都处在同一矩形区域内, 所以可以按照同样的方法来证明。

定理 1 给出的只是存在最小路径的一个充分条件, 当源结点相对目的结点是扩展安全时, 结点之间也可能存在最小路径。下面在定理 1 的基础上给出存在最小路径的一个充要条件。

定理 2 设在二维 Torus 网中, (i', j') 为源结点, (i, j) 为目的结点, $i' \leq i, j' \leq j$ 且 $i-i' \leq k_0/2, j-j' \leq k_1/2$, 如果存在一组结点 (x_i, y_i) 满足 $i' \leq x_1 \leq x_2 \leq \dots \leq i, y' \leq y_1 \leq y_2 \leq \dots \leq j$, 并且 (i', j') 相对 $(x_1, y_1), (x_i, y_i)$ 相对 $(x_{i+1}, y_{i+1}) (1 \leq i \leq n-1), (x_n, y_n)$ 相对 (i, j) 都

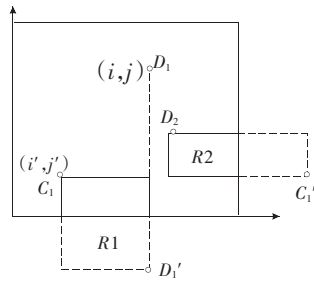


图4 展开环后的网络

是扩展安全的,则 (i', j') 和 (i, j) 之间至少存在一条最小路径。

证明 充分性:当源结点 (i', j') 和目的结点 (i, j) 的坐标满足条件 $i' \leq i, j' \leq j$,且 $i-i' \leq k_0/2, j-j' \leq k_1/2$ 时, (i', j') 到 (i, j) 的最小路径是不需要经过环的。如果 (i', j') 相对 (x_1, y_1) 是扩展安全的,则由定理1可知, (i', j') 和 (x_1, y_1) 之间至少存在一条最小路径,因为 (x_i, y_i) 相对 (x_{i+1}, y_{i+1}) ($1 \leq i \leq n-1$), (x_n, y_n) 相对 (i, j) 也都是扩展安全的,所以同理可知 (x_i, y_i) 和 (x_{i+1}, y_{i+1}) , (x_n, y_n) 和 (i, j) 之间也都存在最小路径;又因为从 (x_i, y_i) 到 (x_{i+1}, y_{i+1}) 都是前进型的,没有回溯,所以由定理1的证明可知,将所有的相邻两点的最小路径连接起来构成的路径即为 (i', j') 和 (i, j) 之间的最小路径。

必要性:因为在从源结点沿着最小路径到达目的结点的过程中,路由的每一步都是向着靠近目的结点的方向的,所以最小路径是前进型的,如果 (i', j') 和 (i, j) 之间存在一条最小路径,由于两个矩形故障块的距离至少是3,则必定可以在这条最小路径上找到一组结点 (x_i, y_i) ,满足 (x_i, y_i) 相对 (x_{i+1}, y_{i+1}) 的扩展安全级中的 $E > 1$ 或 $N > 1$,而由定义4可知此时 (x_i, y_i) 相对 (x_{i+1}, y_{i+1}) 是扩展安全的。

由定理2给出以下推论:

推论1 在2维Torus网中,设 (i', j') 为源结点, (i, j) 为目的的结点, $i' \leq i, j' \leq j$,且 $i-i' < k_0/2$,如果在区域 $[0: i', j': j]$ 和 $[i: k_0-1, j': j]$ 内存在一组结点 (x_i, y_i) ,满足 $0 \leq x_k \leq x_{k-1} \leq \dots \leq x_1 \leq i', i \leq x_n \leq x_{n-1} \leq \dots \leq x_{k+1} \leq k_0-1, j' \leq y_1 \leq y_2 \leq \dots \leq y_n \leq j$,且 (i', j') 相对 (x_1, y_1) , (x_i, y_i) 相对 (x_{i+1}, y_{i+1}) ($1 \leq i \leq n-1$), (x_n, y_n) 相对 (i, j) 都是扩展安全的,则 (i', j') 和 (i, j) 之间至少存在一条最小路径。

推论2 在2维Torus网中,设 (i', j') 为源结点, (i, j) 为目的的结点, $i' \leq i, j' \leq j$,且 $j-j' > k_1/2$,如果在区域 $[i': i, 0: j]$ 和 $[i': i, j: k_1-1]$ 内存在一组结点 (x_i, y_i) ,满足 $i' \leq x_1 \leq x_2 \leq \dots \leq x_n \leq i, j' \geq y_1 \geq y_2 \geq \dots \geq y_k \geq 0, k_1-1 \geq y_k \geq \dots \geq y_n \geq j$,并且 (i', j') 相对 (x_1, y_1) , (x_i, y_i) 相对 (x_{i+1}, y_{i+1}) ($1 \leq i \leq n-1$), (x_n, y_n) 相对 (i, j) 都是扩展安全的,则 (i', j') 和 (i, j) 之间至少存在一条最小路径。

对于推论1和推论2仍可以采用图4中的方法,将环展开进行证明,其证明方法与定理2完全相同。

4 最小通路区

对于一对给定的源结点和目的结点,如果它们之间存在最小通路区,则构成最短路径的所有结点都在最小通路区内,而最小通路区外的结点都不是构成最短路径的中间结点。下面介绍利用扩展安全级建立最小通路区的方法。

设 (i', j') 为源结点, (i, j) 为目的结点, $i' \leq i, j' \leq j$,且 $j-j' > k_1/2$,其他情况可类似讨论。如图5所示,同样为了便于分析将环展开, D' 为展开环将网络中第 j 行到第 k_1-1 行的结点向下平移后 D 的位置, P 为直线 $y=j'$ 与 $x=i$ 的交点, Q 为直线 $x=i'$ 和直线 $y=j$ 的交点。

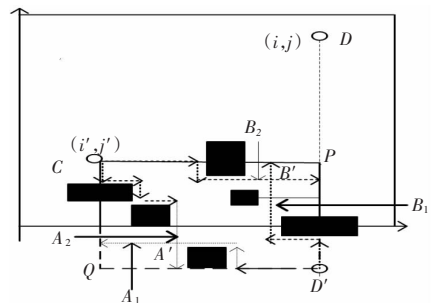


图5 建立最小通路区

首先从目的结点开始发送一个消息,如果目的结点的扩展安全级中的 $N > 1$,则消息向北,当消息所到达的结点的扩展安全级中的 $N=1$ 时,如果此时 $E > 1$,则消息向西转,然后继续向西,当消息所到达的结点的 N 再次满足 $N > 1$ 时,则消息由西向北转,然后继续向北,按照这种方式直到到达直线 $y=j'$,这条路径称为路径 A_1 ;再从源结点出发发送一个消息,如果源结点的扩展安全级中的 $E > 1$,则消息向东,如果消息所到达的结点的 E 总是满足 $E > 1$,则消息一直向东直到到达直线 $y=j'$,如果在此之前,消息所到的结点的扩展安全级中的 $E=1, S > 1$,则消息由东向南转,然后继续向南,当消息所到达的结点的 E 再次满足条件 $E > 1$ 时,则消息由南向东转,然后继续向东,按照这种方式直到到达直线 $x=i$,这条路径称为路径 A_2 ,如果路径 A_1 和 A_2 都存在,设 A_2 第一次与 A_1 相交于点 A' ,则路径 A 如下:从目的结点沿着 A_1 到达 A' ,再从 A' 沿着 A_2 到达源结点。

路径 B 可利用同样的方法建立,具体过程可参照图5。

这样由路径 A 和路径 B 封闭的区域即为 (i', j') 和 (i, j) 之间的最小通路区;这样由路径 A 和路径 B 封闭的区域即为 (i', j') 和 (i, j) 之间的最小通路区。

根据上面建立最小通路区的过程可以看出,当源结点相对目的结点是非扩展安全时,利用该方法仍可建立最小通路区。

引理1^[9] 设结点 (i', j') 为源结点,结点 (i, j) 为目的的结点,其中 $i' \leq i, j' \leq j$,如果路径 A 和路径 B 同时存在,则被路径 A 和路径 B 封闭的区域即为从源结点到目的结点的最小通路区。

下面给出最小容错路由算法,算法只考虑存在最小路径的情况,主要思想就是首先分析两结点之间是否存在最小通路区,如果存在则建立最小通路区,然后让消息在最小通路区里进行路由。

算法 Min-Routing():设源结点为 (i', j') ,目的结点为 $(i, j), i' \leq i, j' \leq j$

(1)计算 $i-i'$ 和 $j-j'$ 的值。

(2)从源结点发送一个消息到目的结点,目的结点收到源结点发送的消息后发送两个回应消息来建立路径 A_1 和 B_1 。根据(1)中计算的结果决定这两个回应消息的路由方向:

①如果 $i-i' \leq k_0/2, j-j' \leq k_1/2$,则一个消息向东,一个消息向北,向东的消息建立路径 B_1 ,向西的消息建立路径 A_1 ;

②如果 $j-j' > k_1/2$,则一个消息向西,一个向北,向西的消息建立路径 A_1 ,向北的消息建立路径 B_1 ;

③如果 $i-i' > k_0/2$,则一个消息向东,一个向南,向东的消息建立路径 A_1 ,向西的消息建立路径 B_1 ;

消息在路由的过程中如果遇到故障块,则向靠近源结点的方向转,绕过故障块后继续按原来的方向路由。

(3)同时从源结点发送两个消息来建立路径 A_2 和 B_2 ,这两个消息的路由方向与目的结点发送的两个消息的方向相反。