

1 $\frac{1}{2}$ メッシュアレーのプロセッサとバスの 同時故障を補償する多項式時間再構成 アルゴリズム

堀田 忠義

高浪 五男

岩手大学工学部 情報工学科, 〒020 盛岡市上田 4-3-5
phone: +81-196-21-6968 or +81-196-21-6464, fax: +81-196-21-6487
E-mail {hori,takanami}@cis.iwate-u.ac.jp

あらまし: VLSI 技術の発展により, 多数の処理要素を用いた並列計算機を, ウェハー上に集積して実現することに関心もたれるようになってきた. その際の大きな問題の一つは, 集積の複雑さが増大するために, 何らかの対策を講じなければ製造中または動作中における故障の発生確率が無視できない程度に増大する, ということである. そこで, 製造時における生存率や動作中における信頼性の向上を図る工夫が必要である. そのための1つの手段として, 本論文では, Roychowdhury 等のアルゴリズムを拡張することにより, 1 $\frac{1}{2}$ トラックスイッチモデルの PE とバスの同時故障に対し, 多項式時間の再構成アルゴリズムを提案する.

A Polynomial Time Algorithm for Reconfiguring the 1 $\frac{1}{2}$ Track-Switch Model with PE and Bus Faults

Tadayoshi HORITA

Itsuo TAKANAMI

Department of Computer and Information Science, Faculty of Engineering, Iwate University,
4-3-5 Ueda, Morioka-shi, 020 Japan
E-mail {hori,takanami}@cis.iwate-u.ac.jp

Abstract: As VLSI technology has developed, the interest in implementing an entire or significant part of a parallel computer system using wafer scale integration is growing. The major problem for the case is the possibility of drastically low yield and/or reliability of the system if there is no strategy for coping with such situations. As a strategy, in this paper, by extending the Roychowdhury's algorithm, we propose a polynomial time algorithm for reconfiguring 1 $\frac{1}{2}$ track-switch model compensating for simultaneous PEs and buses faults.

1 まえがき

現在の情報化社会において、高速で質の高い情報処理技術は、なくてはならないものになっている。今後はそのような事に対して、ますます質の高い技術が求められることは明らかである。その高速で質の高い情報処理を実現するために、古くから、多数の処理要素 (PE) から成る並列計算機に関する研究がなされてきた。

更に最近では、VLSI 技術の発展により、多数の処理要素を用いた並列計算機を、ウェハー上に集積して実現することに関心もたれるようになってきた。その際の大きな問題の一つは、集積の複雑さが増大するために、何らかの対策を講じなければ製造中または動作中における故障の発生確率が無視できない程度に増大する、ということである。そこで、製造時における生存率や動作中における信頼性の向上を図る工夫が必要である。そのための1つの手段として、PEの配列を再構成する方法がいろいろな文献で提案されている。これらの再構成はそれぞれ提案された方式に応じた結合リンクとスイッチからなる結合網によってなされる。このようにして構成されたシステムの信頼性はPEと結合網の信頼性によって定まるが、これらの2つが同時に故障する場合に対する研究はほとんどない。従来多くの研究では、結合網は故障しない(故障フリー)という仮定、またはそれが故障したときはシステムが故障するという仮定のもとでなされている [1][2][3][4][5]。この場合におけるシステムの信頼性は、故障フリーな結合網のもとで再構成に成功する確率と、結合網が故障しない確率の積として与えられる。すなわち、結合網の故障対策はなされていないわけである。PEの複雑度が高く、結合網のそれが相対的に低いときはこのような仮定で充分であるが、そうでない場合には、結合網の故障がシステムの信頼性に無視できない影響を及ぼすことが予想される。

PEと結合網の同時故障に関連する研究には、以下のようなものがある。

- F. Distant 等は、スイッチ付きバス格子からなる結合網をもつ矩形アレーにおいて、結合ユニットと呼ばれるスイッチ対からなるリンクが故障した場合に、それを避けて結合を再構成する網羅的アルゴリズムを与えているが、アレーのサイズに関し指数関数的計算量を必要とするものである [6]。
- 伊藤は、WSI 結合網に3つの異なるスイッチ

構造を導入し、ラインの開放故障と隣接ライン間の短絡故障を予備のラインで回避する方式を提案し、それをハイパーキューブ網に適用したときの歩留まりについて考察している [7]。

フォールトトレラント超並列計算機アーキテクチャの1つに、Kung 等らが提案した、 $1\frac{1}{2}$ トラックスイッチモデルがある [2]。 $1\frac{1}{2}$ トラックスイッチモデルは、メッシュ(トーラス)結合プロセッサアレーに、冗長な回路としてトラックとスイッチと予備のPEを付加したものである。このアーキテクチャの特長は、付加する切換網が単純であるためハードウェアのオーバーヘッドが少ないこと、再構成後の論理的に隣接したPE間の距離が一定に抑えられること、などである。Roychowdhury 等は、同モデルにおけるPEの故障に対し、多項式時間の再構成アルゴリズムを提案した [8]。また著者らは、同モデルにおけるPEとバスの同時故障に対し、近似解法である、ニューラル再構成アルゴリズムを提案した。しかしながら、 $1\frac{1}{2}$ トラックスイッチモデルのPEとバスの同時故障に対し、多項式時間の再構成アルゴリズムはまだ提案されていない。著者らのニューラル再構成アルゴリズムは、再構成時間が短く、ホストコンピュータによらない自律的再構成が可能である、という特長を持っている。しかし近似解法であるため、PE数が多くなると、再構成率が減少する傾向がある。ここに、メッシュ(トーラス)結合プロセッサアレーの代表的な用途に、高速な画像処理がある。この場合、画像領域を細かく分割し、その各画像領域を多数のPE1つ1つに割り当てる事によって処理を高速にすることが考えられるが、このような場合、PE数が非常に多く、かつ結合網の複雑度がPEのそれに対して無視できない場合が十分に考えられる。従って、 $1\frac{1}{2}$ トラックスイッチモデルのPEとバスの同時故障に対し、多項式時間の再構成アルゴリズムを考察することは、非常に有意義であると、著者らは考える。

そこで本論文では、Roychowdhury 等のアルゴリズムを拡張することにより、 $1\frac{1}{2}$ トラックスイッチモデルのPEとバスの同時故障に対し、多項式時間の再構成アルゴリズムを提案する。2章では、 $1\frac{1}{2}$ トラックスイッチモデルについて述べる。3章で再構成方法について、4章で提案アルゴリズムとその計算量についてそれぞれ述べ、最後に5章で、まとめと今後の課題について述べる。

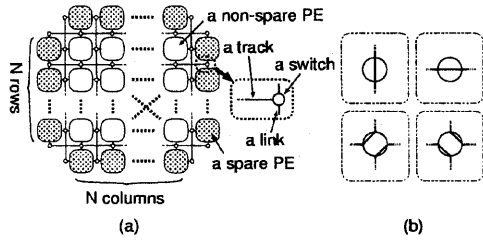


図 1: $1\frac{1}{2}$ トラックスイッチモデル: (a) PE, スイッチ, トラックおよびリンクの配置, (b) スイッチの動作

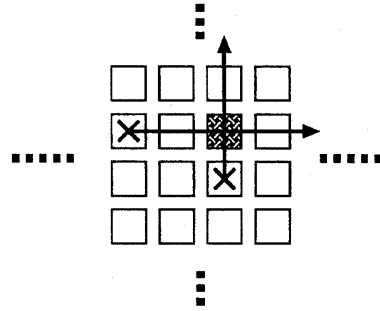


図 3: 交差の関係にある補償バスの例

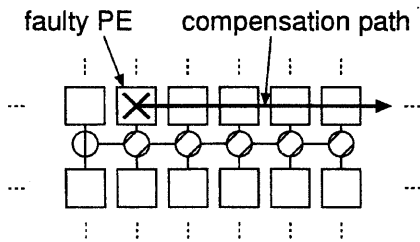


図 2: 補償バスに対する配線の状態

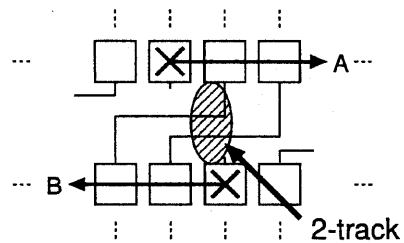


図 4: ニアミスの関係にある補償バスの例

2 アーキテクチャ

図 1に, $1\frac{1}{2}$ トラックスイッチモデルのアーキテクチャを示す. サイズ $N \times N$ のメッシュアレーに, 冗長な回路として, 予備 PE, トラックおよびスイッチが, それぞれ配置されている. リンクまたはトラックのことを, バスと呼ぶ. 本論文では, スイッチは故障しないものとして議論を進める.

3 再構成

以下に述べる「補償バス」による方法によって, 故障 PE と故障バスの使用を回避する.

定義 1 例えば $PE(i, j)$ に故障が起こった場合, まずその同じ行, または列の予備 PE の中から 1 つの非故障なものを選び, その方向に向かって次のように行う. まず故障 $PE(i, j)$ を, その方向の隣の正常な $PE(i', j')$ で置き換える. 同様に, その正常な $PE(i', j')$ を, 同じ方向の隣の正常な $PE(i'', j'')$ で置き換える, といった作業を次々に行い, 最終的にまだ置き換えられていない非故障予備 PE が使用されるまで続け

る. このプロセスにおける PE のインデックスの系列を「補償バス」と呼ぶ. (図 2 参照).

定義 2 任意の異なる方向の 2 つの補償バスが共通のインデックスを含む時, この 2 つの補償バスは「交差」の関係にあるという (図 3 参照). 任意の 2 つの補償バス間に, 1 つでも交差の関係が存在する再構成は失敗である.

定義 3 隣接する行 (または列) の逆方向の 2 つの補償バスが, 2 つ以上の PE でオーバーラップするとき, この 2 つの補償バスは「ニアミス」の関係にあるという (図 4 参照). 任意の 2 つの補償バス間に, 1 つでもニアミスの関係が存在する再構成は失敗である.

ここで, リンクの故障について考える. リンクが故障している PE は他の PE からの接続ができないため, リンクの正常な別の PE に置き換える必要がある. すなわち, そのような PE は故障しているとみなす.

次にトラックの故障について考える.

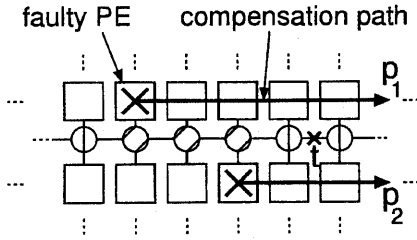


図 5: 同時選択の関係にある補償パスの例

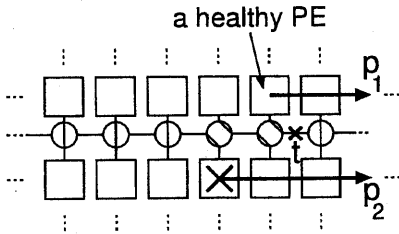


図 6: 同時選択の関係にある健全な PE からの補償パスの発生例

定義 4 ある PE のリンクと、1 つのスイッチを介して接続可能なトラック t があるとき、その PE とトラック t は、「隣接している」という。1 つのトラックに隣接する PE は 4 つある。

隣接する同方向の 2 つの補償パスに対する再構成は、図 5 に示すようにこれらのパスの間のトラックを使用しなくて済む。一方、そのような相手のない独立した補償パスに対する再構成は図 2 に示すように、それに沿って隣接して走るトラックを使用しなくてはならない。以上のことから、次の定義をあてる。

定義 5 隣接する同方向の 2 つの補償パス p_1 と p_2 は、これらの間のあるトラック t が故障しているとき「 t に関して」同時選択」の関係にあるという (図 5, 図 6 参照)。

補題 1 同時選択の関係にある 2 つの補償パスについて、そのどちらか一方のみを用いる再構成は失敗である。なぜなら、一方のみを用いるような再構成では、用いられる補償パスに隣接する故障トラックを用いる必要があるからである。

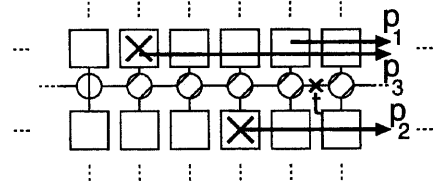


図 7: 重なりのある関係にある補償パスの例

補題 2 補償パス p が、それに隣接するある故障トラック t を持ち、かつ t に関して p と同時選択の関係にある補償パスをもたないならば、補償パス p を用いる再構成は失敗である。なぜならそのような再構成では、故障トラック t を用いる必要があるからである。

定義 6 補題 2 のような補償パス p は、「無効 2」であるという。

定義 7 無効 1 または無効 2 である補償パスは、「無効」であるという。

補題 2 から、図 6 において補償パス p_2 と同時選択の関係にある補償パスが故障トラック t を挟んだ上の行に存在しないならば補償パス p_2 を再構成に用いることができないことが分かる。そこで、故障トラック t に隣接する健全な PE を仮想的に故障とみなし、そこから図 6 に示すような補償パス p_1 が存在するならば、 p_2 と p_1 は同時選択の関係となり、 p_2 を再構成に用いることが可能になる。以上のことから、健全な PE から補償パスを発生させると、再構成率が向上しうることが分かる。

このとき、図 7 の p_1 と p_3 のように、補償パスが重なる場合がありうる。

定義 8 始点の異なる同じ向き の 2 つの補償パスが、PE の共通のインデックスを含むとき、この 2 つの補償パスは、互いに「重なり」の関係にあるという。補償パスの集合の中に重なりのある関係が存在する場合の再構成は不可能である。

以上の考察から、再構成可能な十分条件として以下の定理が成り立つことは、容易に証明できる [2]。

定理 1 (再構成可能定理) 故障アレーは、次の 4 つの条件を満足するように補償パスを選択できるならば、再構成可能である。

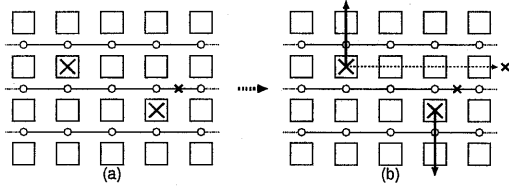


図 8: 故障 PE が故障トラックに隣接している場合の再構成

1. 全ての予備でない故障 PE に対して、連続で直線的な補償バスが必ず 1 つは選択できる。
2. 選択された任意の 2 つの補償バス間に、交差、ニアミスまたは重なり関係が存在しない。
3. 片方のみが用いられるような同時選択の関係にある補償バスの組が存在しない。
4. 無効な補償バスが選択されない。

トラックの故障を考慮した場合の、故障 PE からの補償バスの決定について、注意すべき性質があるので、ここで述べておく。

性質 1 故障 PE が故障トラックに隣接している場合、図 8 の (a) ⇒ (b) のように、その故障トラックに隣接している故障 PE からの補償バスの決定が、故障トラックに関して、他の故障 PE からの補償バスの決定を妨げる場合があるので、このような場合の補償バスの決定には、注意が必要である。

図 9 に、サイズ 4×4 の $1\frac{1}{2}$ メッシュアレーの再構成の手順を示す。プロセッサに故障が発生しても、補償バスを発生させ、再び 4×4 のメッシュアレーを実現している。

再構成可能な補償バスの集合が与えられれば、スイッチの状態は、あらかじめ与えられた表を参照することにより決まることから、以下では補償バスの決定方法についてだけ述べる。

4 再構成アルゴリズム

バスの故障を考慮した再構成アルゴリズムを得る為に、バスの故障を考慮していない場合の多項式時間再構成アルゴリズムである Roychowdry らのアルゴリズム [8] を拡張する。

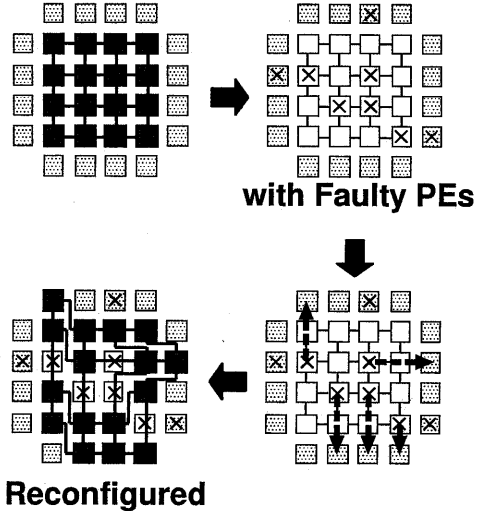


図 9: 再構成の手順

4.1 特殊な場合の再構成

アレーの予備 PE の配置が図 10(a) ~ (d) であるような 4 つの特別な場合の再構成の計算量を求める。

4.1.1 Case 1

これは、補償バスのとりうる方向が、{左, 右} のみの場合 (図 10(a) 参照) である。この場合における再構成の計算量は、 $O(N(N+F+T))$ である (証明略)。ただし、 F は予備でない故障 PE 数、 T は故障トラック数である。

4.1.2 Case 2

これは、補償バスのとりうる方向が、{左, 下} のみの場合 (図 10(b) 参照) である。この場合における再構成の計算量は、 $O(N(N+F+T))$ である (証明略)。

4.1.3 Case 3

これは、補償バスのとりうる方向が、{左, 上, 右} のみの場合 (図 10(c) 参照) である。この場合における再構成の計算量は、 $O(N(N+F+T))$ である (証明略)。

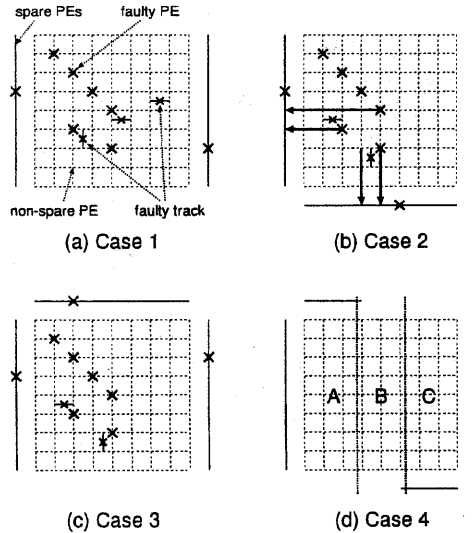


図 10: Case 1~4

4.1.4 Case 4

これは、図 10(d)にあるように、補償バスのとりうる方向が、領域 A では {左, 上, 右} のみ、領域 B では {左, 右} のみ、領域 C では {左, 下, 右} のみ、の場合である。この場合における再構成の計算量は、 $O(N(N+F+T))$ である (証明略)。

4.2 一般の場合の再構成

図 1(a) のように、上下左右の 4 方向に補償バスを選択することが出来る、一般の場合の再構成について考える。この場合のアルゴリズムは、以下の 2 つの部分から成る。ただしこの方法において、特に明記しなくても、前章で述べたように、故障トラックの使用と重なりを避けるために、健全 PE からの補償バスの発生・消滅を行っているものとする。

4.2.1 Part 1

Part 1-1:

1. 予備でない PE アレーの一番下の行を、“調査行” に選ぶ。
2. 調査行中において、そこから補償バスが選択されていない故障 PE の有無を調べる。もしそのような故障 PE が

2.1. 存在し、その故障 PE から下方向への補償バスは

2.1.1. 無効でなく、そのバス上に故障 PE が存在せず、かつ未決定の故障 PE からの補償バスの選択を、

2.1.1.1 妨げないならば、その下方向への補償バスを選択する。

2.1.1.2 妨げる¹ ならば、集合 I を空集合にする。補償バスの選択を妨げられている故障 PE のインデックスを、集合 I に入れる。集合 I のインデックスの故障 PE の下向き補償バスの選択が、更に別の未決定の故障 PE からの補償バスの選択を妨げるならば、集合 I にその補償バスの始点の故障 PE のインデックスを加える。この作業を、選択を妨げられる補償バス E が見付からなくなるまで繰り返す。最終的に得られた集合 I 中のインデックスを 1 つ以上含む行に存在する全ての予備でない故障 PE から、下方向に補償バスを選択できるならば、そのように補償バスを選択し、次のステップへ。選択できない故障 PE が存在するならば、そのうちの一番下の位置にある故障 PE に“y” という印をつけ、Part 1-2: へ。

2.1.2. 無効であるか、またはバス上に故障 PE が存在するならば、その故障 PE に“y” という印をつけ、Part 1-2: へ。

3. もし全ての予備でない故障 PE から補償バスが選択

3.1. されているなら、アレーは再構成可能であり、アルゴリズムは終了。

3.2. されていないならば、現在よりも 1 行上の行を新たに調査行として選択し、ステップ 2 へ。

Part 1-2: “調査列” を一番左の列から 1 列ずつ右にとり、補償バスは左方向にとり、Part 1-1 と同様

¹ 図 8 のように、故障トラックに隣接し、1 つ上の列の故障 PE からの補償バスの選択を妨げる場合

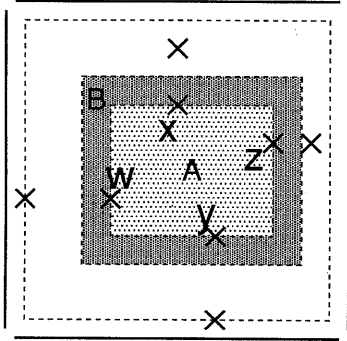


図 11: Part1 終了後

に行く。ただし左方向に補償パスをとれない故障 PE につける印は" w" である。

Part 1-3: “調査行”を最上行から 1 行ずつ下にとり、補償パスは上方向にとり、Part 1-1 と同様に行く。ただし上方向に補償パスをとれない故障 PE につける印は" x" である。

Part 1-4: “調査列”を一番右の列から 1 列ずつ左にとり、補償パスは右方向にとり、Part 1-1 と同様に行く。ただし右方向に補償パスをとれない故障 PE につける印は" z" である。

また Part 1 で再構成が可能かどうか判断できなかった場合の Part 1 終了後の状態は、図 11 に示す。故障 PE "w", "x", "y", "z" の位置は、まだ補償パスの方向が定まっていない領域(図 11 の領域 A と B)の内側にある。このような場合、次の Part 2 に進む。

4.2.2 Part 2

領域 A での補償パスの選択について考える。ここではまず、故障 PE "x" および"y" からの補償パスの全ての場合について考える。それを図 12(a) ~ (g) に示す。

図 12 を見て分かるように、図 12(a) ~ (f) の場合は、補償パスによって分けられた領域が Case 1 ~ 4 のいずれかになっている。従って、Part 2 においてはまず、図 12(a) ~ (f) の場合について、分割された領域に対してそれぞれ対応する Case 1 ~ 4 のアルゴリズムを適用し、再構成が不可能と判断された場合にのみ、図 12(g) の場合を考える。図 12(g) の場合は、故障 PE z と w からの補償パスの方向によって、図 13(a) ~ (d) に示すように更に 4 つの場合がある。

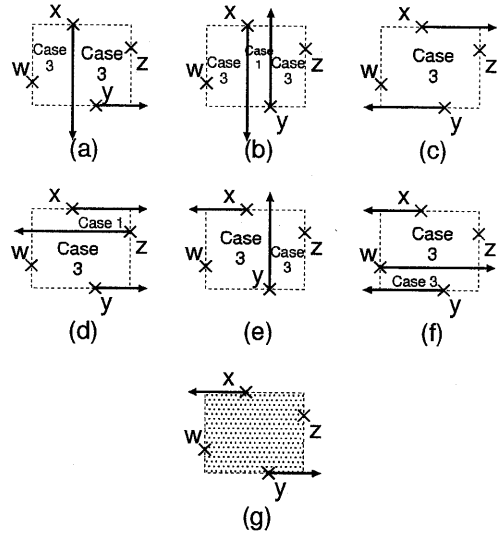


図 12: x,y から選択可能な方向の補償パスの全ての場合

図 13(a) ~ (d) の場合のうち、(a) ~ (c) の場合は、補償パスによって分けられた領域が Case 1 ~ 4 のいずれかになっている。従って、先の場合と同様にして、それぞれ対応する Case 1 ~ 4 のアルゴリズムを適用し、再構成が不可能と判断された場合にのみ、(d) の場合を考える。(d) の場合に対しては、再び Part 1, Part 2 のアルゴリズムを適用する。

領域 B の故障 PE の補償パスの方向の決定を、以下のように行う。領域 B の故障 PE において、領域 A の一番外側の故障 PE からの補償パスと同時選択の関係があれば、その方向に決定し、そうでなければ外側方向に決定する。この方法において、領域 A の補償パスとニアミスや交差の関係は生じず、かつ故障トラックも使用しない。領域 B に存在する故障 PE の数は高々 F 個であり、補償パスの方向はバクトラックなしに決定可能であるから、領域 B の補償パスの方向の決定の計算量は、 $O(F)$ である。この領域 B の補償パスの選択は、領域 A でのそれが決定されてから行うものとする。

Part 1 から Part 2 の図 13(c) までは、Case 1 ~ 4 の定数回の繰り返しであるから、 $O(N(N+FT))$ である。このプロセスは、高々 F 回繰り返される。また領域 B のアルゴリズムの計算量は、 $O(F)$ である。従って、この場合のアルゴリズム全体の計算量

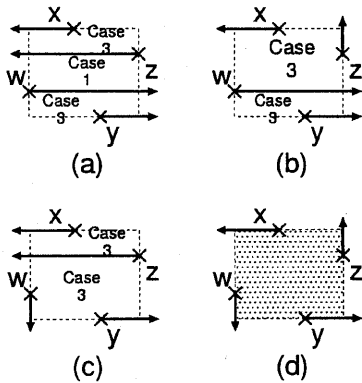


図 13: z, w から選択可能な方向の補償パスの全ての場合

は, $O(N \cdot F \cdot (N+F+T))$ である.

5 むすび

本論文において, $1\frac{1}{2}$ トラックスイッチモデルに対し, PEとバスの同時故障を補償する多項式時間の再構成アルゴリズムを提案した. 本手法のソフトウェア実現や, 本手法と他手法との比較は, 今後の課題である.

6 謝辞

本研究は, 電気通信普及財団の助成金による.

参考文献

- [1] M. Sami, and R. Stefanelli. "Reconfigurable architectures for VLSI processing arrays". *Proc. IEEE*, pp. 712-722, May 1986.
- [2] S. Y. Kung, S. N. Jean, and C. W. Chang. "Fault-tolerant array processors using single-track switches". *IEEE Trans. Comput.*, Vol. 38, No. 4, pp. 501-514, January 1989.
- [3] I. Takanami, K. Kurata, and T. Watanabe. "A neural algorithm for reconstructing mesh-connected processor arrays using single-track switches". *Int'l Conf. on WSI*, pp. 101-110, January 1995.
- [4] J.S.N. Jean, H.C. Fu, and S.Y. Kung. "Yield enhancement for WSI array processors using two-and-half-track switches". *Proc. IEEE Int'l Conf. WSI*, pp. 243-250, 1990.
- [5] T. A. Varvarigou, V. P. Roychowdhury, and T. Kailath. "A polynomial time algorithm for reconfiguring multiple-track models". *IEEE Trans. Comput.*, Vol. 42, No. 4, pp. 385-395, April 1993.
- [6] F. Distante, M. G. Sami, and R. Stefanelli. "Reconfiguration techniques in the presence of faulty interconnections". *Int'l Conference on WSI*, pp. 379-388, 1989.
- [7] H. Ito. "A defect-tolerant design for WSI interconnection networks and its application to hypercube". *IEEE Int'l Workshop on Defect and Fault Tolerance in VLSI Systems*, pp. 80-87, 1993.
- [8] V.P. Roychowdhury, J. Bruck, and T. Kailath. "Efficient algorithms for reconstruction in VLSI /WSI array". *IEEE Trans. Comput.*, Vol. 39, No. 4, pp. 480-489, April 1989.