

格子状接続並列計算機用マッピング： ローリングマッピングの効率的なPUアレイの実装例

村田 淳, 京都府立大学人間環境学部環境情報学科

murata@kpu.ac.jp

天野英晴, 慶應義塾大学理工学部情報工学科

hunga@aa.cs.keio.ac.jp

要旨

本研究では、過去に提案した、疎密のある物理空間に対する格子状接続並列計算機のためのマッピングであるローリングマッピングの実装方法を考案している。ローリングマッピングでは疎密単位はPUアレイ4倍の大きさである。そこで最初からPUアレイ4倍の大型のPUアレイを考え、これをモジュラマッピングする。この場合必要となる、モジュラマッピング特有のサイクリック構造を大型PUアレイを2回に折り畳むことにより回避することで実現している。

New Structure for Rolling Mapping

Atsushi MURATA, Dept. of Environmental Information, Kyoto Prefectural University

Hideharu Amano, Dept. of Computer Science, Keio University

abstract

In this paper, we propose new architecture that have some PU(Processing Unit) arraies of Nearest Neighbor Mesh(NNM). Using some NNM PU arraies, it is faster than one PU array when calculate the rolling mapping for the mesh with various densities.

1 疎密のある物理空間とそれに対するマッピング

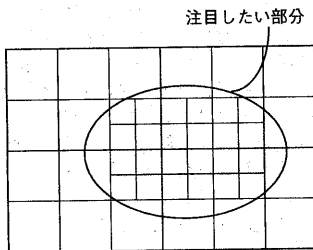


図 1: 物理空間中の注目したい部分

偏微分方程式を差分法(陽解法)で解く場合、一般的には計算する物理空間を均一に離散化し格子点空間を作成し、差分式を計算する [1]。その際には十分な精度を得るために格子点間隔を小さくすることになる [2]。

しかし、広い物理空間の全体にわたって同じように高精度で解こうとすると離散化された格子点空間の格子点数が多くなり、結局は計算量の増大を招くことになる。そこで、広い物理空間全体を均一に高精度を求めめるのではなく注目したい部分は格子点間隔を小さく(高精度に)して、残りのあまり精度を要求されない格子点空間では格子点間隔を大きく(低精度に)取る方法がある [5]。

図 1はその一例である。全体としては直交格子の格好を保ちながら部分的に精度を高くしている。このような離散化された物理空間を本論文では「疎密のある物理空間」と呼ぶ。 [3]

本研究ではこの疎密のある物理空間を格子状接続並列計算機がどのように計算すれば効率良く(つまり計算精度と計算量のバランスを保ち)計算できるかどうかを考えていく。

離散化された物理空間(正方格子点空間)での計算には、複数の計算要素(Processing Unit, 以下PUと呼ぶ)を相互に格子状に接続したPUアレイを持つ格子状接続並列計算機で計算すると、効率よく高速に計算出来ることが知られている [4]。一般的に、対象となる物理空間の格子点数はPUアレイのPU数より遥かに多いため、PUと物理空間の格子点を対応づけるマッピングが必要となる。この場合、最も簡単な方法は直接マッピングである [4]。

直接マッピングは、物理空間中の互いに隣接したいくつかの格子点の集合を1つのPUに対応させる。しかし直接マッピングを図1の様な疎密のある物理空間に用いると、1つのPUが受け持つ格子点数に差が出て、各PUの計算負荷にばらつきが生じる。

この問題を解決するために、モジュラマッピング [4] が提案された。しかし、この方法は計算負荷にはばらつきは生じないが、PUアレイにサイクリックな構造が要求される上、疎密境界部分での遠隔PU間通信が増大する問題点がある。

そこで、本報告では、PUアレイにサイクリックな構造を要求せず、疎密境界領域上での遠隔PU間通信も少ないマッピング法である、ローリングマッピングについて述べる。

疎密のある物理空間でも各PUに負荷が均一になるように考え出されたモジュラマッピングを元にして、その改良型としてローリングマッピングを考案した。次に、ローリングマッピング用の結合網であるローリングメッシュをPUアレイに付加することにより、物理空間の疎密境界上に位置する格子点間のデータ交換をより高速に行うことができるようになった。最後に、ローリングマッピングに適したPUアレイの新たな実装方法について報告する。

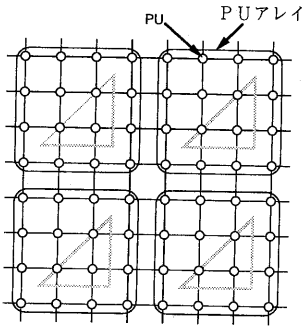


図 2: モジュラマッピングの例

2 モジュラマッピングとローリングマッピング

2.1 モジュラマッピングとその問題点

モジュラマッピング [4] とは、物理空間の格子点座標 (X, Y) と PU アレイ (サイズ $m \times m$) の PU 番号 (p, q) の対応を

$$\begin{aligned} p &= \text{mod}(X, m) \\ q &= \text{mod}(Y, m) \end{aligned} \quad (2)$$

で関係づける方法である。格子点数 8×8 の物理空間に対してサイズ 4×4 の PU アレイをモジュラマッピングした様子を図 2 に示す。

図 2 中の三角形記号は PU アレイの向きを表している。

この方法では、以下の条件の下に、疎密のある物理空間でも各 PU の負荷が均一になるように PU を割り付けることが出来る [5]。

1. 物理空間に疎密をつけるときには、疎の部分も密の部分も正方格子状に離散化する。
2. 疎の部分 (単密) に対して密の部分 (倍密) の格子点密度は 4 倍 (格子点間隔が $1/2$) とする。なお、ここでは簡単のために格子点密度は単密と倍密の 2 種類とする。
3. 疎密の基本単位を PU アレイサイズ ($= m \times m$) と同一とする。

これらの条件の下で、図 2 で示した物理空間に疎密をつけると例えば図 3 の様になる。図 2 の左上 4 分の 1 が倍密になっていて 4 つの PU アレイが担当している。

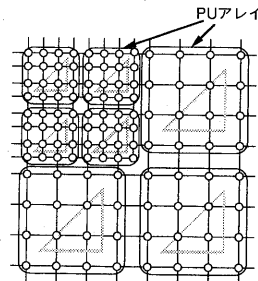


図 3: 一部分を倍密にした結果

モジュラマッピングの問題点は、通信コストの増大を招くことである。まず、均一な物理空間に適用した場合でも、PU アレイの両端を接続しサイクリックにする必要がある。このような構造の格子はトーラスと呼ばれ、かつては PAX [4] をはじめとする多くの格子状接続計算機で用いられていた。しかし、最近のリンク結合型の並列計算機は、リンクの転送速度を上げるため、遠隔通信が必要なトーラス構造を避け、単なる二次元、三次元の格子状接続をとる傾向にある。Paragon [6], J-machine [7] はこの例である。このようなマシン上ではモジュラマッピングは不利である。

さらに、モジュラマッピングを疎密のある物理空間に適用すると遠隔 PU 間でのデータ交換が必要となる。図 3 に示した疎密のある物理空間の疎密境界を部分的に拡大すると図 4 (上) の図になる。この図の中で \leftrightarrow で示された隣接格子点間のデータ交換は、実際の PU アレイでは図 4 の下の図に示す様に、サイクリックリンクを用いた通信と PU アレイの同じ辺上に位置する離れた PU 間通信の合計 2 バスの通信によって行なわれる必要がある。

2.2 ローリングマッピング

以上の問題点を解決するため、モジュラマッピングと同様に疎密のある物理空間に適用した場合に負荷の不均一を生ぜず、しかも PU アレイにサイクリック構造を必要としないマッピング方法を考案した。この方法は、物理空間の格子点座標 (X, Y) と PU アレイ (サイズ $m \times m$) の PU 番号 (p, q) の対応を次の様に定義したものである。

$$\begin{aligned} P &= \text{mod}(X, 2m) \\ p &= P \end{aligned} \quad (P < m)$$

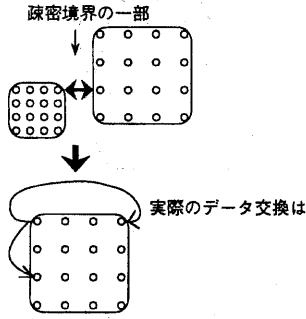


図 4: 疎密境界上実際のデータ交換

$$\begin{aligned}
 p &= -P + 2m - 1 & (P \geq m) \\
 Q &= \text{mod}(Y, 2m) \\
 q &= Q & (Q < m) \\
 q &= -Q + 2m - 1 & (Q \geq m)
 \end{aligned} \quad (3)$$

このマッピング法では、モジュラマッピングと同様、個々のPUが担当する複数の格子点が物理空間中で分散されるため、疎密のある物理空間に適用した場合でも負荷の不均衡を生じない。

格子点数 8×8 の物理空間に対してサイズ 4×4 のPUアレイをマッピングした例を図5に示す。モジュラマッピングが、物理空間中をPUアレイが平行移動して埋めていくイメージであるのに対して、このマッピング法は、図5の様にPUアレイを“ぐるりと”裏返して物理空間を埋めていくイメージである。このイメージから我々は式(3)に定義したマッピング法をローリングマッピングと呼ぶことにする。

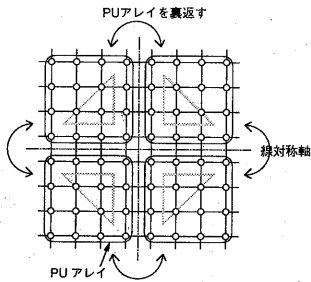


図 5: PUアレイをひっくり返す

モジュラマッピングと比べ、ローリングマッピングはPUアレイをひっくり返す時の線対称軸に相対する格子点とは同じPUが担当することから、PUアレイ構造をサイクリックにする必要がない。さらに、モジュラマッピングでは、疎密の境界上のPU間で、2回の交信が必要であるのと同様に、通信オーバーヘッドもモジュラマッピングに比べ少

くなる。

2.3 疎密のある物理空間に対するローリングマッピング

ローリングマッピングでも、モジュラマッピングと同様、一定の条件の下で疎密のある物理空間にPUを均一の負荷となる様にマップすることが出来るはずである。しかし、この一定の条件は、モジュラマッピングの条件(前節参照)より少し制約が大きい。モジュラマッピングの場合は、式(2)($p = \text{mod}(X, m)$)で示した様に大きさ $m \times m$ の繰り返しパターンで物理空間をマップして、これがモジュラマッピングの疎密の基本単位となる。一方ローリングマッピングでは疎密の基本単位をPUアレイサイズの4倍に定める。なぜならローリングマッピングの場合は、式(3)($P = \text{mod}(X, 2m)$)で示した様に大きさ $2m \times 2m$ の繰り返しパターンで物理空間をマップしているためである。[5]

以上の条件で疎密のある物理空間にローリングマッピングを適用した例を図6に示す。図6ではサイズ 4×4 のPUアレイでマッピングしている。図6の格子点を省略して描くと図7のようになる。さらに図7の疎密境界上のPUの様子を拡大して描くと図8のようになる。図8を見ると、疎密境界に相対する格子点と同じPUが担当するのではないが、PUアレイの同じ辺上に位置する離れたPUが担当することがわかる。従って、ローリングマッピングを疎密のある物理空間に適用する場合には、PUアレイの同じ辺上に位置する離れたPU間で互いにデータ交換する必要がある。この辺は、PUアレイの4辺のうちの2辺だけである(図9参照)。

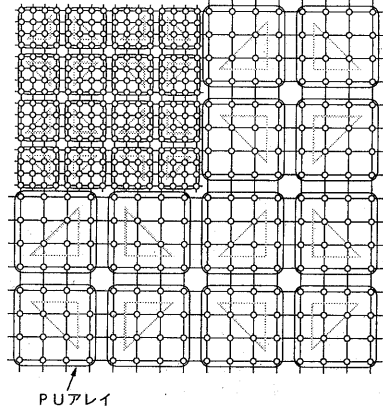


図 6: ローリングマッピングの場合の倍密化

2.4 ローリングメッシュ

今までは、遠隔リンクを持たない平面メッシュ構造を持つ並列計算機に対して、疎密のある物理空間のマッピングについて検討してきた。しかし、PUアレイのサイズが大きくなり、疎密境界が長くなると辺上での交信コスト(つまり疎密境界上のデータ交換)が全体の計算時間を支配し、オー

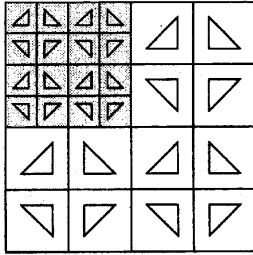


図 7: PU アレイの割り当て

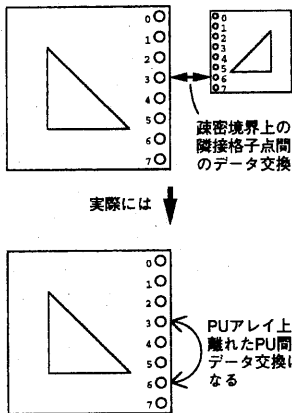


図 8: ローリングマッピングでの疎密境界上のデータ交換

バヘッドになる場合が考えられる。そこで、疎密のある物理空間にローリングマッピングを適用させるには図 10 で示したような、PU アレイの特定の 2 辺で、高速にデータ交換を行うこのローリングマッピング用の遠隔リンクを持つメッシュ結合であるローリングメッシュを提案した。[3]

図 10 にローリングメッシュ(PU アレイサイズ 8×8 の場合) の概念図を示す。PU アレイの縁には、辺上の PU と同数の遠隔交信用ルータを設ける。このことにより、辺上の PU もアレイ内部と同様、4 本のリンクを持つことになる。ルータは 4 本の遠隔リンクを持ち、お互いに接続される。ルータは 4×4 のクロスバから構成され、遠隔プロセッサの番号により、自動的にルーティングを行なう。この構成により、疎密境界上の遠隔プロセッサ間の通信は、全て 1 バスで行なうことができる。

3 PU アレイの 3 次元的構成法

ローリングマッピングは、PU アレイのサイズを基本として、疎密のある物理空間を区切って効率良く数値計算を行なうことができる。しかし、半導体技術の進歩により、巨

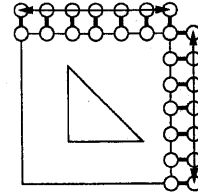


図 9: 辺上に位置する PU 間のデータ交換

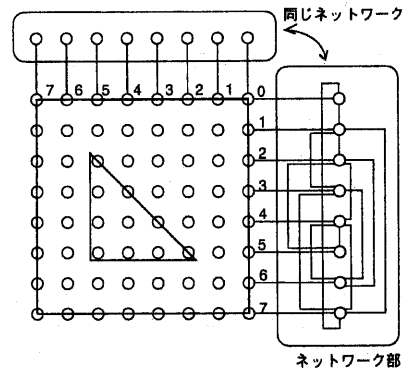


図 10: ローリングメッシュのネットワーク

大な PU アレイから成る並列計算機の構成が可能になった場合、PU アレイ全体のサイズを基本としてマッピングを行なう方法は柔軟性の点で不利になる可能性がある。2 次元の PU アレイを実装する場合でも、PU 数が大きくなると実装は 3 次元的になる。そこで、PU アレイを 3 次元的に実装した場合に、有利になるマッピングの方法と、端に位置する PU 間の接続法を提案する。

3.1 PU アレイの折り畳み

前述のように、ローリングマッピングを疎密のある物理空間に適応する際は、物理空間の格子点の密度単位を PU アレイの大きさそのものではなく、その 4 倍 (縦 $2 \times$ 横 2) にする必要がある。すなわち、疎密のある物理空間に対するローリングマッピングは PU アレイサイズが 4 倍の大きさの PU アレイによるモジュラマッピングと考えることが出来る。

ここで、疎密のある物理空間に対するローリングマッピング考えたサイズの 4 倍のサイズを持つ PU アレイ (図 11) を想定する。そしてこの PU アレイを疎密のある物理空間にモジュラマッピングを施すことを考える。モジュラマッピングであるので、各 PU の負荷は均一化される。

しかし、モジュラマッピングでは一般的には PU アレイにサイクリックなリンクを要求する。このため、図 12 のように、PU アレイを $1/4$ に折り畳む。このようにすると折り畳まれた PU アレイは、ローリングマッピングで最初に

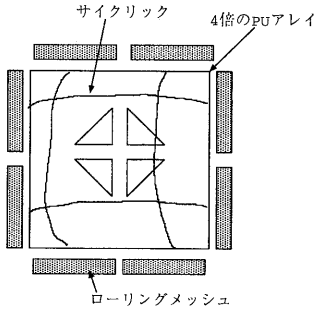


図 11: 4 倍の大きさの PU アレイ

考えた大きさの PU アレイの大きさと等しくなる。この場合、PU アレイの辺上に位置する PU は折り畳まれた反対側の辺の PU と極めて近い位置に配置されることになる。

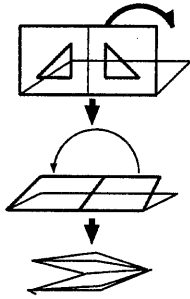


図 12: 2 回折り畳む

このことにより、PU アレイ上に長いサイクリック用の配線をする必要はなくなる。

しかし、この PU アレイでは、折り線になっていない 2 辺上に位置する PU 間でデータ交換を行わなければならない。ここで図 12 を上から見た PU アレイ 4 枚を第 1 層から第 4 層までとする。そうすると折り線になっていない 2 辺では図中の A 辺では、第 1 層と第 4 層が繋がり、第 2 層と第 3 層が繋がる。一方図中の B 辺では第 1 層と第 2 層が繋がり、第 3 層と第 4 層が繋がることになる。この場合 A 辺で 2 個のローリングメッシュネットワークが必要である。また B 辺でも同様に 2 個にローリングメッシュネットワークが必要となる。

3.2 ローリングメッシュを 2 つにする結合方法

前節では PU アレイの辺上の PU 間の結合を 4 つの PU アレイの内の 2 つごとに結合したため、ローリングメッシュは各辺上の 2 層に必要となり、合計で 4 つのローリングメッシュを必要とした。しかし、4 層に分かれている PU アレイ

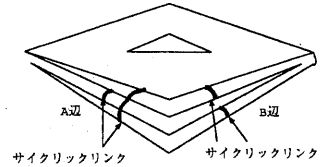


図 13: 4 層の PU アレイ

の辺上の PU は高々 4 個である。そこで、垂直に位置する 4 つの PU 間だけに特殊な配線を施し、相互が同じデータを共有できるようにすれば、ローリングメッシュは 4 層の内のどれか 1 層に作れば良いことになる。

この 4 層間の特殊な配線は、4 つの PU 間だけ高速に通信できればよので、図 14 の様にグローバルバスにするか、或いは 4×4 のクロスバを介するかすればよい。

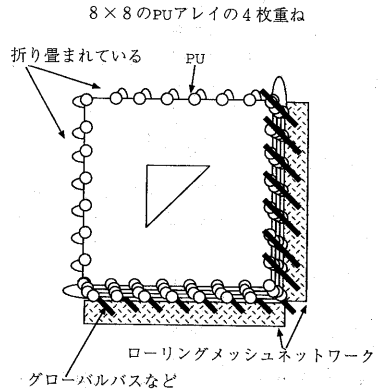


図 14: 折り畳まれた PU アレイ

3.3 実装法の拡張

前述の方法は、疎密の基本単位 (PU アレイの 4 倍) だけでなく、その 2 倍、4 倍…の更に大きな PU アレイを、前節に説明したように折り畳んで拡張することができる。つまり、図 14 の様に各 PU 間をグローバルバスやクロスバ、或はポケット通信などで繋ぐことにすれば PU 間のデータ交換がより柔軟に行うことができる。図 15 に 4 枚一組の PU アレイを 4 段に重ねた様子を示す。この方法を用いれば、PU アレイ 16 枚が同時に計算対象である物理空間を計算できる。疎密に分かれた物理空間も疎の部分と密の部分で同時に PU アレイが計算できる。

4 結論

本論文では、格子状接続並列計算機用の新しいマッピングアルゴリズムであるローリングマッピングについて、過

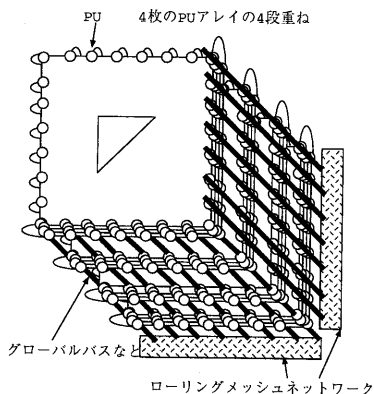


図 15: 4 段 16 層の大きさの PU アレイ

去に提案した方法 (PU アレイの大きさでマップする) を改良し, PU アレイの 4 倍の大きさの大型の PU アレイを折り畳むことによる実装方法を提案した。更に 4 倍を超える PU アレイ数, 例えば 16 枚の PU アレイをローリングマッピングする方法を提案した。提案した方法により, 大型のアレイでもその $1/4, 1/16$ などのサイズを単位としたマッピングが可能になる。

今後の課題としては, 遠隔 PU の通信用付加リンク (グローバルバスなど) の詳細な設計とシミュレーションによる評価が必要である。さらに, 実際のアプリケーション上を用いた評価も必要である。

参考文献

- [1] 矢嶋信男, 野木達夫: “発展方程式の数値解析”, 岩波書店 (1977).
- [2] 川合敏雄, “スーパーコンピュータへの挑戦”, 岩波書店, 1985.
- [3] 村田淳, 天野英晴, “数値計算における疎密のある物理空間に対するマッピング: ローリングマッピング”, 情報処理学会論文誌, 第 35 巻第 2 号 (1994).
- [4] 星野力, “PAX コンピュータ - 高並列処理と科学計算 -”, (1985).
- [5] 村田淳, 大沢暁, 天野英晴, 相磯秀夫: 正方格子状接続並列計算機と疎密のある物理領域とのマッピングアルゴリズム, 第 34 回情報処理学会全国大会論文集, pp.91-92 (1987).
- [6] Paragon XP/S product overview, Intel 1991.
- [7] Dally, W.J. Chien, A. Fiske, S. Horwat, W. Kenn, J. Larivee, M. Lethin, R. Nuth P. and Wills, S.: The J-machine: A fine-grain concurrent computer, In Proc. IFIP 11th Computer Congress, pp.1147-1153, Aug. 1989.
- [8] Osawa, G., Murata, A., Amano, H. and Aiso, H.: Rolling Mapping for Physical Simulations, Proc. of Second Intl. Conf. on Supercomputing, (1987).