

プログラム解析情報可視化システム-PREPS-における

3L-2 データ分散配置表示方式と静的シミュレーション方法について*

草野和寛 †藤田淳 †渡辺幸光 妹尾義樹†

NEC C&C 研究所 † †NEC 情報システムズ‡

1 はじめに

分散メモリシステムで並列実行するプログラムでは、データ分散配置が非常に重要である。データの分散配置が上手く行なわれていないと大量のデータ転送が必要になり、プログラムの実効性能が低下する。これまでプログラムの静的な解析情報を表示するツール [1] やプロファイル実行の結果を表示するツールは開発されてきた。しかし、データの分散配置の決定をサポートするものは少ない。そこで我々は、データの分散配置を決定する情報の提供を目的としたプログラム解析情報可視化システム PREPS [2] を開発した。PREPS では並列計算機内におけるデータの論理プロセッサへの分散方法を記述できるように Fortran を拡張した HPF [3] をターゲット言語としている。本稿では、PREPS におけるデータ分散配置の表示および、データアクセス情報の表示について述べる。

転送量、アクセス方向を色の違いで表示する。この楕円領域をクリックすることにより、データ分散配置およびデータアクセス情報の表示機能をもつ変数表示ウィンドウが立ち上がる。

3 データ分散配置の表示

3.1 データ分散配置表示機能

PREPS では、HPF のデータ分散指示行により指定された各プロセッサへのマッピング情報を視覚的に表示する機能を提供している。データの分散配置情報表示は 2 次元配列で行ない、3 次元以上の配列の場合には任意に選択した 2 つの次元に関する分散配置情報を表示する。この配列要素を、マッピングされたプロセッサ単位に異なる色にすることで分散情報の表示を行なう。データ分散配置の表示例を図 2 に示す。

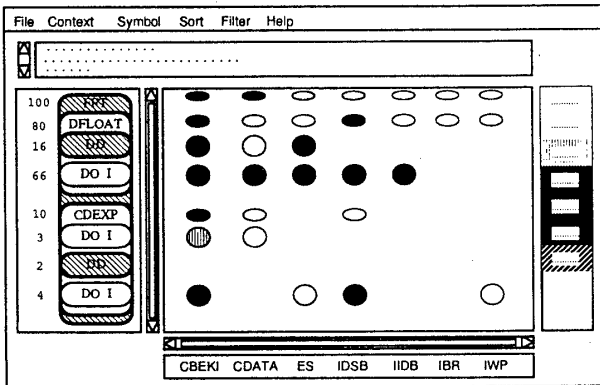
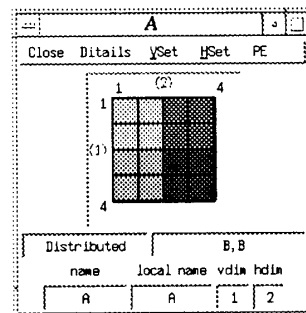


図 1: 表示画面イメージ

2 PREPS のメイン表示

PREPS のメイン表示 (図 1) は、縦軸にプログラムのコントロールフローを横軸に変数名を表示し、交点にあたる楕円領域でアクセス属性、アクセス回数、データ



DIMENSION A (4, 4)
!HPF\$ PROCESSORS PE (2, 2)
!HPF\$ DISTRIBUTE A (BLOCK, BLOCK) ONTO PE

図 2: 配列の分散配置表示例

この表示には、変数の名前とサイズなどに加えて、分散指示情報も表示している。

3.2 配列表示のズーム機能

変数表示ウィンドウはその大きさが固定なため表示要素数には上限がある。そのため表示要素が複数の配列要素に対応する場合があります。配列一要素の状態を見ることができない。そこで、ズーム機能を用いて配列一要素単位から全体まで見ることができるようにした。

ズーム機能は、ズームインとズームアウトの機能を持つ。ズームイン機能は、複数の配列要素に対応している

*The data distribution display and the data access simulation of the Program information visualization system prototype -PREPS-

†Kazuhiro KUSANO, Jun FUJITA, Yukimitsu WATANABE, Yoshiki SEO

‡C&C Research Laboratories, NEC Corporation

§†NEC Informatec Systems, Ltd.

表示要素を配列一要素に対応させていく機能である。ズームアウト機能は、ズームイン機能によって変えられた表示要素の対応している配列要素数を立ち上げ時の配列要素数に戻す機能である。但し、ウィンドウ立ち上げ時の表示要素が既に配列一要素に対応している場合には使用出来ない。

ズーム機能により、配列一要素単位の分散を正確に表示することができるようになった。

4 データアクセス情報の表示

プログラムのループ中における配列へのアクセスパターンが複雑でわかりにくい場合がある。そこで、PREPSではプログラム中でアクセスされる配列要素を調べて視覚的に表示する機能を提供している。この機能では、アクセスされる配列要素を正確に調べるために静的シミュレーションを行なっている。

4.1 静的シミュレーション機能

PREPSでは、アクセスされる配列の要素位置を求める為にスカラー変数に対してのみシミュレーションを行なう。プログラムファイルから生成された内部情報を元にシミュレーションを行なう為実行プロファイルや実行オブジェクトを必要としない。

シミュレーションはプログラム中の手続き又はループ単位に実行でき、任意のループのみを対象とする事も可能である。シミュレーション中に値が未確定もしくは値が不明である情報(仮引数、配列の間接参照、分岐等)を検出した場合、利用者に情報を求める。これにより正確にアクセスされる配列の要素位置を決定する。

4.2 データアクセス情報表示方式

静的シミュレーションにより求められた配列のアクセス情報は、図2で示した画面を利用して表示する。

この情報は、アクセスされた要素を含む表示ブロックに対してアクセスの種類により異なる色をアクセス順序に従って表示することで表している。アクセスの種類による色分けは以下の通り。

1. 参照 [R]
2. 書き込み [W]
3. 参照の後書き込まれた [R→W]
4. 書き込み後参照された [W→R]
5. 3/4の後さらに参照/書き込みがあった

図3に示したプログラムの変数Aがどのようにアクセスされるかすぐにわからない。しかし、データアクセス情報の表示機能を使用することで図3に示すようなアクセス状態を容易に得ることが出来る。

5 PREPS を利用したデータ転送の最適化

PREPSのデータ分散配置の表示機能とデータアクセス情報の表示機能を利用してデータ転送の最適化を行なう手順を以下に示す。

```
DIMENSION A(4,4),B(4,4)
!HPF$ PROCESSORS PE(2)
```

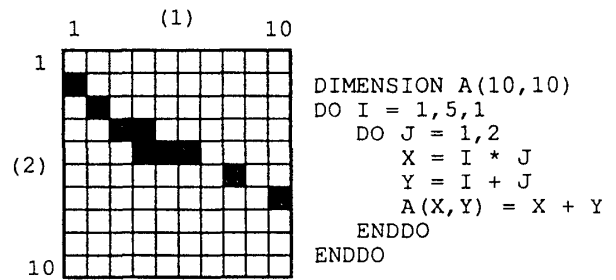


図3: データアクセス情報の表示例

```
!HPF$ DISTRIBUTE (*,CYCLIC) ONTO PE :: A,B
DO I = 1,4,1
DO J = 1,4,1
A(I,J) = B(J,I)
ENDDO
ENDDO
```

前述のプログラムにおいて、ループ中に現われている変数AとBについて、PREPSを利用してアクセス情報の表示を行なう。添字Iのループに対して変数Aは1次元方向に書き込みアクセスが、変数Bは2次元方向に参照アクセスされていくことがわかる。このアクセスの状態から、ループ中で行なわれる演算に必要なデータの半分が異なるプロセッサ上にマッピングされている。つまり、並列実行時にプロセッサ間でデータ転送が必要であることが判る。

そこで、変数Aのデータ分散指示行を以下のように変更する。

```
!HPF$ DISTRIBUTE A(CYCLIC,*) ONTO PE
```

ここで再びアクセス情報の表示を行なうと、演算に必要なデータは常に同一のプロセッサ上にあり、並列実行時にデータ転送は必要ないことがわかる。

以上述べたように、利用者はPREPSの機能を用いてデータ転送の最適化を行なうことができる。

6 まとめ

本稿ではデータ分散配置の表示機能とデータアクセス情報の表示機能について述べた。本機能を利用することにより、データ転送の最適化における利用者の負担を軽減することができる。

なお本システムの表示方式の研究は科学技術振興調整費総合研究「知的生産活動における創造性支援に関する基盤的研究」の一環として行なった。

参考文献

- [1] M.W.Hall, et al., *Experiences Using the ParaScope Editor: an Interactive Parallel Programming Tool*, ACM Sigplan Symp. on PPOPP, pp.33-43, 1993.
- [2] 草野 他, プログラム解析情報可視化システム PREPS, 情報第49回全大予稿集 4U-05, 1994.
- [3] High Performance Fortran Forum, *High Performance Fortran Language Specification Version 1.1*, November 1994