

手続き間並列化コンパイラ WPP の試作

1 E - 1

—現状と今後の課題—

飯塚孝好† 佐藤茂久† 蓮見勝久§ 菊池純男‡

†(株)日立製作所

§日立ソフトウェアエンジニアリング(株)

‡新情報処理開発機構

1. はじめに

共有メモリマルチプロセッサ向けの手続き間並列化コンパイラ WPP(Whole Program Parallelizer)を開発中である。既存の並列化支援システム[1]での手続き間スカラ・配列データフロー解析と別途開発中の手続き内並列化コンパイラをベースに、クローニングを伴った定数伝播、手続き間 DO ALL 並列化、手続き間リダクション並列化、手続き間配列プライベート化を組み込んだ。今後、実用レベルのプログラムの並列化を可能とする機能拡張、分散共有メモリマルチプロセッサへの適用などを行い、自動並列化技術の適用性を高めていく予定である。

2. 構成

図1のように、既存のプログラムを既存のユーザーインターフェースで手続き間並列化できる構成とした。即ち、各ソースは従来通りファイル毎にコンパイルし、オブジェクトには各ソースに対する中間語を埋め込み、プログラムの構成手続きが全て揃うリンク時に手続き間コンパイラを起動する。

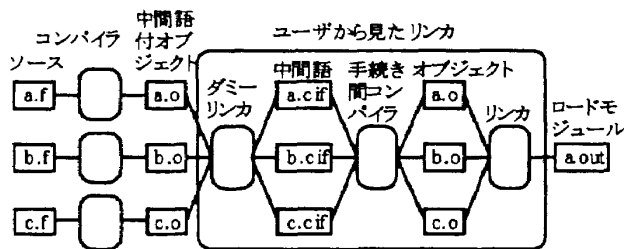


図1 ユーザーインターフェース及び内部構成

手続き間コンパイラは、手続き内コンパイラの処理を流用し、スカラデータフロー解析、配列データフロー解析、並列化のそれぞれのフェーズを手続き

間処理に拡張した。各手続き間処理では、手続き毎のローカルサマリ収集(PPA)、ローカルサマリの手続き間伝播によるグローバルサマリ生成(IPP)を行う。また、各ローカルサマリ収集では、実行済みの手続き間処理で生成したグローバルサマリを利用する。これによりローカルサマリ収集が複数回実施され解析時間が延びるが、解析精度向上および最適化の適用性拡大に役立つ。例えば、ループ長が定数になり、ループ長が大きな並列化可能ループを外側ループにするループ交換が適用されるなど。

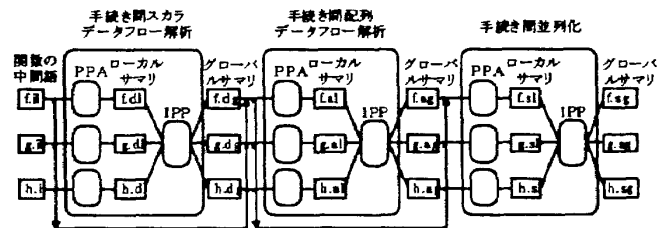


図2 手続き間コンパイラの構成

なお、各手続き間伝播では制御フロー依存な解析情報をコールグラフに対してボトムアップ/トップダウンに伝播する。これを容易にするために、各手続き内の制御フローを階層的に要約した CFS([2]の RSG の制御フロー部)を用いる。

3. 機能

現状では、基本的な手続き間解析・並列化と手続き間最適化のフレームワークを実現している。

(1) 手続き間最適化フレームワーク

コールグラフ管理、クローニング、解析情報メモリ管理、CFS 管理、同一宣言 COMMON の辞書マージ、中間語ファイルの関数単位分割。

(2) 手続き間スカラデータフロー解析[3]

手続き間別名解析、手続き間参照変数解析(定義、使用、確定定義、露出使用、ライブ)、クローニングを伴った手続き間定数伝播。

(3) 手続き間配列データフロー解析

手続き間変数値シンボリック解析、手続き間配列

Prototyping of Interprocedural Parallelizing Compiler "WPP" -Overview-

Takayoshi Iitsuka †, Shigehisa Satoh †, Katsuhisa Hasumi §, Sumio Kikuchi ‡

†Hitachi Ltd., §Hitachi Software Engineering Co., Ltd., ‡RWCP

参照領域解析(定義、使用、確定定義、露出使用、ライブ)。

(4)手続き間並列化[4]

手続き間 DO ALL 並列化、手続き間プライベート化、手続き間リダクション並列化。

4. WPP での新規開発機能

現状の実現機能は、キャッシュなどアーキテクチャ依存な部分を除いては、Stanford 大学の SUIF コンパイラ[5]とほぼ同等である。更に SPECfp95 の apsi の並列化に有効な解析・最適化を新規開発した。

4.1 ループ飛び出し有りループの配列参照領域解析

apsi のサブルーチン BLSOLV では、配列 DKZM と DKZH の値を 3 箇所(ループの前とループ飛び出しを含むループ DO 10 と後続ループ DO 30) で定義している。

```

DKZM(1)=DZ; DKZH(1)=DZ
DO 10 K=2,NZ
  KE=K
C   DKZM(K)とDKZH(K)が確定定義される
C   Cの値は0<=KLEV<=1
  CALL SRFLAY(ZET,DKZM(K),DKZH(K),...,DKSM,...,KLEV)
  IF(DKSM.GT.0) GOTO 20
10 CONTINUE
20 KE=KE+KLEV
  DO 30 K=KE+NZ
C   DKZM(K)とDKZH(K)が確定定義される
  CALL EMKLAY(ZET(K),DKZM(K),DKZH(K),...)
30 CONTINUE

```

KE 及び KLEV の値を詳細に調べると、各配列の確定定義範囲は(1:NZ)であることが分かる。これが解析できると、プライベート化により BLSOLV の呼び出し元 DKZMH のループが並列化可能となり、実行比率 8%の部分で粒度が 2 桁向上する。従来はループ飛び出しがあるループの確定定義は空集合としていたが、WPP ではループ飛び出し時のループ制御変数の値を持った変数 KE を用いて確定定義範囲を計算することにより、上記の解析を可能とした。

4.2 擬寸法配列の終値保証

apsi では、データ依存上は並列化可能で粒度が大きな部分が有り、実行比率約 50%の部分で 1 ミ秒程度の粒度を持っている。しかし配列長が不明となる擬寸法配列が多用されており、プライベート化に伴った終値保証が難しい。論文[5]の粒度データを見る限り、SUIF でも並列化されていない。そこで、定義と確定定義の配列参照領域を用いて安全に終値保

証を行う方法を開発した[4]。

なお、apsi では FFT 処理を呼び出す人手並列化可能なループが多数有り、これらも並列化できれば並列化率は 100%近くなり、粒度も大きい。しかし、この部分の静的な解析は非常に困難である。

5. 課題

実アプリケーションに対して比較的適用性の高い課題としては次のようなものがあり、今後検討が必要と考えている。

(1)手続き間非線型配列参照解析

ループ制御変数の係数が変数の場合や、整合配列の参照領域を手続き間伝播する場合は、非線型解析が必要である。非線型解析の適用対象は広いので、Range Test[6]の手続き間への拡張および数式処理の強化などによって実現していきたい。

(2)共有メモリプロセッサ向け配列配置・構造変換

共有メモリプロセッサでは、プロセッサ毎のキャッシュ参照特性が実行性能に大きく影響する。配列配置・構造変換により性能向上に結び付けたい。

6. おわりに

本発表では、共有メモリマルチプロセッサ向けの手続き間並列化コンパイラ WPP の概要を述べた。今回は SUIF とほぼ同等の機能を実現したが、実用プログラムの自動並列化にはまだ多数の課題があり、今後検討を重ねていく予定である。

参考文献

- [1] 菊池ほか:並列化システム「Parassist」の試作-機能と構成-, 情報処理学会第 44 回大会(1992)
- [2] T.Iitsuka, Flow-sensitive Interprocedural Analysis Method for Parallelization, Proc. of IFIP WG10.3 Working Conference on Architecture And Compilation Techniques for Fine and Medium Grain Parallelism, Elsevier Science Publishers, 1993, pp.65-76
- [3] 佐藤ほか:手続き間並列化コンパイラ WPP の試作-定数伝播とクロウニングの評価-, 情報処理学会第 56 回大会(1998)
- [4] 佐藤ほか:手続き間並列化コンパイラ WPP の試作-変数プライベート化技術-, 情報処理学会第 56 回大会(1998)
- [5] M.W.Hall, et.al., Detecting Coarse-Grain Parallelism Using an Interprocedural Parallelizing Compiler, Supercomputing 95
- [6] W.Blume, et.al., Advanced Program Restructuring for High-Performance Computers with Polaris, CSRD-TR-1473, University of Illinois at Urbana-Champaign, 1996