

講義紹介：並列計算プログラミング，先端計算機演習（地球惑星科学専攻）

中島研吾

東京大学情報基盤センター

1. はじめに

本稿では、2008年度夏季集中講義として実施した、「並列計算プログラミング」，「先端計算機演習」について紹介する。

本講義・演習は、21世紀COEプログラム「多圏地球システムの進化と変動の予測可能性（観測地球科学と計算地球科学の融合拠点の形成）」（2003年度～2007年度）において2004年度より開講されたもので〔1〕，2008年度からは「理学系研究科大学院教育高度化プログラム」に認定され、夏季集中講義として開講された。

また、本講義・演習は情報基盤センターが関連部局との協力のもと2009年度から始動する「学際計算科学・工学人材育成プログラム¹」の一環として実施されたものである。

表1に講義日程と内容を示す。

表1：講義日程，内容

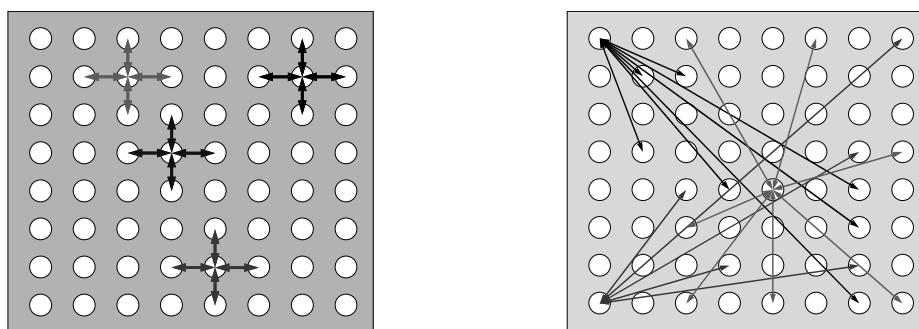
日付	時間	番号	内容
7月23日（水）	1300-1615	CS-01	はじめに，数値解析の基礎，線形ソルバー（I）
7月24日（木）	1300-1615	CS-02	共役勾配法による一次元熱伝導解析プログラム
7月25日（金）	1300-1615	CS-03	MPIによるプログラミング概要（I）
7月30日（水）	1300-1615	CS-04	MPIによるプログラミング概要（II）
7月31日（木）	1300-1615	CS-05	MPIによるプログラミング概要（III）
8月25日（月）	1300-1615	CS-06	課題S1解説，T2K オープンスパコン見学
8月26日（火）	1300-1615	CS-07	課題S2解説，線形ソルバー（II）
8月27日（水）	1300-1615	CS-08	チューニング入門（I）
8月28日（木）	1300-1615	CS-09	並列アプリケーション開発法（I）有限体積法
8月29日（金）	1300-1615	CS-10	並列アプリケーション開発法（II）有限体積法：局所分散データ構造，領域分割
9月08日（月）	1300-1615	CS-11	並列アプリケーション開発法（III）有限体積法並列化
9月09日（火）	1300-1615	CS-12	並列アプリケーション開発法（IV）粒子間熱伝導
9月10日（水）	1300-1615	CS-13	演習（自習）
9月25日（木）	1300-1615	CS-14	チューニング入門（II）
9月26日（金）	1300-1615	CS-15	演習（自習）

¹ <http://nkl.cc.u-tokyo.ac.jp/CSEedu/>

2. 講義内容

本講義・演習は、科学技術計算プログラミングに必須の項目である「SMASH (Science-Modeling-Algorithm-Software-Hardware)」を、できるだけ幅広くカバーし、広い視野を持った人材を育成することを最終的な目標としている。特に、MPI (Message Passing Interface) を使って、差分法、有限要素法、境界要素法等によるアプリケーションを並列化する能力を身につける、ことを重視している。

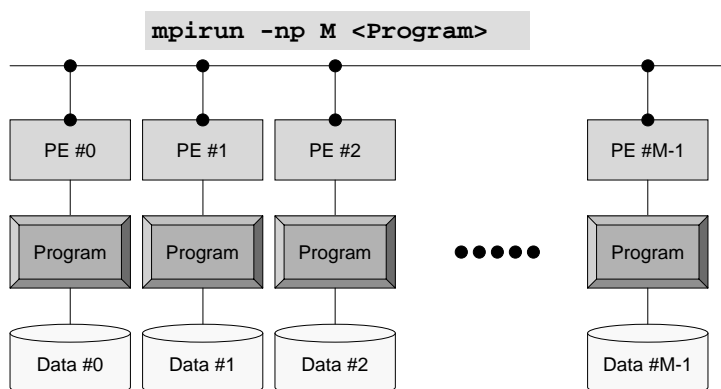
MPI には 400 以上の関数があるが、実際、科学技術シミュレーションで必要になるのは 10 程度である。本講義では、まず、科学技術シミュレーションの手法を図 1 に示す 2 種類 (局所的な手法、大域的な手法) に分類した。これらの手法は連立一次方程式の解法、並列化の手法が図 1 に示すように大きく異なっている。局所的な手法は空間的に近接した領域とのみ相互作用があるため、並列計算時の通信も 1 対 1 通信 (point-to-point communication) が中心である。大域的な手法はグローバルな相互作用があるため、グループ通信 (collective communication) を多用する。



- 局所的な手法 (差分法, 有限要素法等)
 - 疎行列, 1 対 1 通信
- 大域的な手法 (境界要素法, スペクトル法等)
 - 密行列, グループ通信 (collective)

図 1 科学技術シミュレーション手法の分類

講義では、それぞれを「並列化」するために必要な最小限の MPI 関数 (1 対 1 通信, グループ通信) について教え、できるだけ実習によって経験を積んでもらうことにした。MPI の基本的な考え方である SPMD (Single-Program Multiple-Data) (図 2) を実際のアプリケーションの並列化を通して学ぶことに主眼を置いた。SPMD の考え方は計算機科学専攻の学生でも中々理解するのが難しく、教育にあたって苦勞するところなのだそうであるが、アプリケーションという具体的な対象があると逆に理解が進むようである。



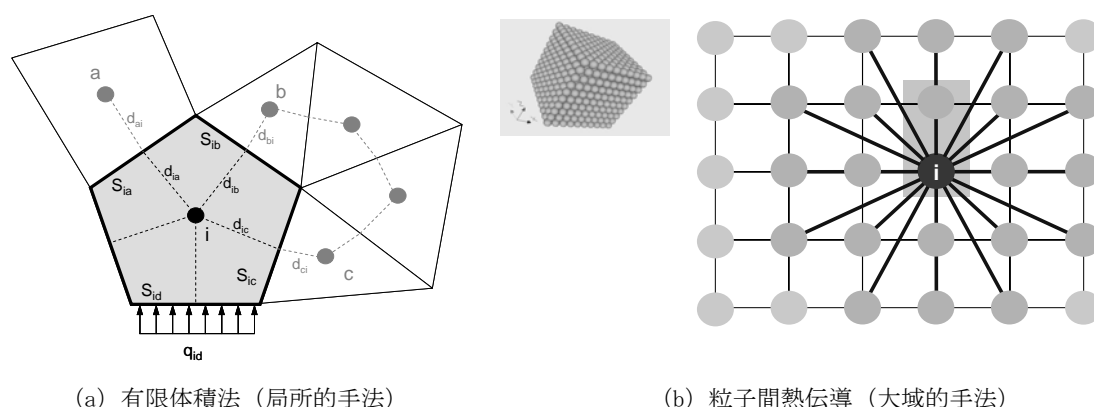
各プロセスは「同じことをやる」が「データが違う」
大規模なデータを分割し、各部分について各プロセス(プロセッサ)が計算する

図 2 SPMD (Single-Program Multiple-Data) の考え方

講義の中では、アプリケーションの中身をよく理解し、並列化するための局所データ構造をどのように考えるか、というような話題が中心であった。表 1 によると、MPI に関する講義は 3 回実施されているが、ほとんどの時間は局所データ構造の考え方の説明に費やした。アプリケーションの中身、アルゴリズムの特性、をよく理解していれば、「並列化」というのはそれほど難しいことでは無い、ということ伝えることには成功したと考えている。実際、講義の中では「並列計算」に直接結びつくものは少なく、むしろ基礎的なモデリングやアルゴリズムに関する基本的な項目について、学習することを重視している。ごく基本的なプログラミングと数値解析に関する知識があれば、これまで、自分でアプリケーションプログラムを開発した経験が無くても充分受講できる内容である。

いわゆる「並列化」に関係した講義のほか、「線形ソルバー」、「可視化手法」、「チューニング」等についても講義、実習を行った。

「チューニング」については、キャッシュブロッキング、ループアンローリング等スカラープロセッサ向けのテクニックの他、「地球シミュレータ」を念頭において、ベクトルプロセッサ向けのチューニングについても講義した。結果的に講義・演習全体で「SMASH」の「MASH」はほぼカバーするような内容になっている。並列化のためのターゲットアプリケーションとしては、図 3 に示すような有限体積法（局所的手法）、粒子間熱伝導（大域的手法）を扱った。



(a) 有限体積法（局所的手法） (b) 粒子間熱伝導（大域的手法）
 図 3 並列計算プログラミング，先端計算機演習で扱ったアプリケーション

プログラミングの実習のために、21 世紀 COE で導入した AMD Opteron プロセッサ (1.8GHz) 32 台 (16 ノード) から構成される PC クラスタを使用する他、2008 年 6 月より稼動を開始した T2K オープンスパコン (東大) を使用した。実習では、NUMA Control の性能への影響評価等も実施した。講義は情報基盤センター (本郷) 大演習室で実施し、受講者は教育用計算機システム (ECCS2008) から上記 PC クラスタ、T2K オープンスパコンにログインして実習を行なった。可視化ツールとしては ECCS2008 で稼動する「MicroAVS (KGT 社製)」²を使用した。

WEB 環境を有効に活用し、連絡事項の伝達、教材配布のほか、並列計算用ユーティリティプログラムのオンラインマニュアル等にも利用した。

地球惑星科学専攻の学生に FORTRAN ユーザーが多いこともあり、プログラミング言語は FORTRAN 中心であったが、他専攻、他研究科 (物理, 天文, 化学, 工学系, 新領域等) の受講者も居ることから、C 言語で書かれたサンプルコード、模範解答も用意する等配慮した。

評価は、以下のような課題実習レポートによって実施した：

² <http://www.kgt.co.jp/feature/microavs/>

- 課題 S1 : MPI の基本的な機能 (台形法則等, グループ通信)
- 課題 S2 : CG (Conjugate Gradient) 法による一次元熱伝導方程式ソルバーの並列化 (1 対 1 通信)
- 課題 P1 : 有限体積法による三次元熱伝導方程式ソルバーの並列化

講義は表 1 に示すように,基本的に 3 限 (1300~1430), 4 限 (1445~1615) に相当する時間に講義を実施したが, 予習・復習のため講義実施日の 2 限 (1015~1145) は演習室を開放し, 自習時間とした。担当教員が必ず演習室に待機し, 質問への対応を実施した。

詳細については本講義のホームページ³を参照されたい。講義資料等を入手することができる。

5. まとめ

登録者は 34 名 (うち地球惑星科学専攻が 27 名, 天文学専攻 3 名, 物理学専攻 2 名, その他 2 名) であったが, 実際に出席していたのは 20 名程度であった。

本講義も 5 年目となり, 教材も非常に基礎的な部分から細かく説明されていること, また集中講義となったことで, これまでよりも全般的に手応えがあった。また稼動間もない T2K オープンスパコン (しかも正式運用前) を使用できたことは受講者にとっては非常に貴重な体験であったようである。

集中講義と称してはいたが, 実際はかなり日程が分散していたので, 2009 年度は一日 4 コマとし, 内容もよりコンパクトにして, 8 月 17 日~28 日のうち 6 日程度で実施の予定である。

参 考 文 献

- [1] 中島研吾 (2008) 究極の「並列プログラミング教育」を目指して : 地球惑星科学専攻での 4 年間と「学際計算科学・工学 人材育成プログラム」, スーパーコンピューティングニュース (東京大学情報基盤センター) 10-3

<http://www.cc.u-tokyo.ac.jp/publication/news/VOL10/No3/200805nakajima.pdf>

³ <http://nkl.cc.u-tokyo.ac.jp/08s/>