

目次

1章 序論	1
1.1 より大きなコンピュータパワーへの要求	1
1.2 並列コンピューティングの必要性	3
1.3 厄介な問題	5
1.4 MPI	6
1.5 本書の構成	7
1.6 標記に関する約束	9
2章 並列コンピューティング概観	11
2.1 ハードウェア	11
2.1.1 Flynn の分類	11
2.1.2 古典的フォンノイマンマシン	12
2.1.3 パイプラインとベクトルアーキテクチャ	12
2.1.4 SIMD システム	14
2.1.5 一般的な MIMD システム	15
2.1.6 共有メモリ MIMD	16
2.1.7 分散メモリ MIMD	19
2.1.8 通信とルーティング	24
2.2 ソフトウェアの問題	26
2.2.1 共有メモリシステムのプログラミング	26
2.2.2 メッセージパッシング	30
2.2.3 データパラレル型言語	34
2.2.4 RPC とアクティブメッセージ	35
2.2.5 データマッピング	36
2.3 要約	38
2.4 文献	40

viii	目次	
2.5	練習問題	40
3章	こんにちは！	43
3.1	最初のプログラム	43
3.2	実行	44
3.3	MPI	46
3.3.1	一般的なMPIプログラム	46
3.3.2	プロセス世界を知る	47
3.3.3	メッセージ：データとエンベロープ	47
3.3.4	メッセージの送信	50
3.4	要約	53
3.5	文献	55
3.6	練習問題	55
3.7	プログラミングの課題	56
4章	応用：数値積分	57
4.1	台形公式	57
4.2	並列化台形公式	60
4.3	並列システムにおけるI/O	64
4.4	要約	67
4.5	文献	68
4.6	練習問題	68
4.7	プログラミングの課題	68
5章	集団通信	71
5.1	木構造通信	71
5.2	ブロードキャスト	76
5.3	タグ、安全性、バッファ、同期	78
5.4	レデュース	80
5.5	内積	82
5.6	オールレデュース	84
5.7	ギャザとスカタ	85

	目次	ix
5.8	オールギャザ	89
5.9	要約	91
5.10	文献	94
5.11	練習問題	94
5.12	プログラミングの課題	95
6章	通信のためのデータのグルーピング	97
6.1	パラメータ count	97
6.2	派生データ型とMPI_Type_struct	98
6.3	そのほかの派生データ型コンストラクタ	104
6.4	型の一致	107
6.5	パック/アンパック	109
6.6	送信方法の決定	113
6.7	要約	115
6.8	文献	117
6.9	練習問題	118
6.10	プログラミングの課題	119
7章	コミュニケータとトポロジ	121
7.1	マトリクスの積	121
7.2	Foxアルゴリズム	123
7.3	コミュニケータ	126
7.4	グループ、コンテキスト、コミュニケータの働き	128
7.5	MPI_Comm_split	131
7.6	トポロジ	132
7.7	MPI_Cart_sub	135
7.8	Foxアルゴリズムの実装	137
7.9	要約	140
7.10	文献	144
7.11	練習問題	145
7.12	プログラミングの課題	146

8章 I/Oの扱い	149
8.1 stdin, stdout, stderrの扱い	150
8.1.1 属性キャッシング	151
8.1.2 コールバック関数	153
8.1.3 I/O プロセスランクの同定	155
8.1.4 I/O プロセスランクのキャッシング	157
8.1.5 I/O プロセスランクの取得	161
8.1.6 stdinからの読み込み	162
8.1.7 stdout への書き込み	164
8.1.8 stderr への書き込みとエラーチェック	166
8.2 限定されたstdin へのアクセス	168
8.3 ファイルI/O	171
8.4 配列I/O	173
8.4.1 データの配置	174
8.4.2 実 例	176
8.4.3 入力配置	177
8.4.4 派生データ型	178
8.4.5 派生データ型の範囲	179
8.4.6 入力プログラム	182
8.4.7 配列のプリント	184
8.4.8 例 題	186
8.5 要 約	188
8.6 文 献	194
8.7 練習問題	194
8.8 プログラミングの課題	195
9章 プログラムのデバッグ	197
9.1 逐次プログラムのデバッグの復習	197
9.1.1 ソースコードの検査	198
9.1.2 デバッグアウトプットの追加	200
9.1.3 デバッガの使用	203
9.2 さらに逐次プログラムのデバッグ	207
9.3 並列プログラムのデバッグ	207
9.4 非決定論的振る舞い	207

9.5 例 題	211
9.5.1 プログラム?	211
9.5.2 プログラムのデバッグ	216
9.5.3 並列デバッガについて	217
9.5.4 頼みのprintf/fflush	220
9.5.5 並列プログラムの古典的なバグ	221
9.5.6 最初の修正	223
9.5.7 多くの並列プログラムのバグは逐次プログラムのバグ	224
9.5.8 異なるシステムでの異なるエラー	227
9.5.9 マルチプルプロセスへ	227
9.5.10 I/O についての混乱	230
9.5.11 デバッグ終了	232
9.6 MPIのエラーハンドリング	233
9.7 要 約	235
9.8 文 献	238
9.9 練習問題	238
9.10 プログラミングの課題	238
10章 並列プログラムの設計とコーディング	241
10.1 データパラレルプログラム	241
10.2 ヤコビ法	242
10.3 並列ヤコビ法	244
10.4 並列プログラムのコーディング	250
10.5 例題:ソート	251
10.5.1 メインプログラム	252
10.5.2 入力関数	255
10.5.3 全対全スキャタ/ギャザ	257
10.5.4 キーの再配置	259
10.5.5 仕上げまでのもう一息	261
10.5.6 Find_alltoall_send_params	262
10.5.7 完 成	266
10.6 要 約	266
10.7 文 献	268
10.8 練習問題	268

xii	目次	
10.9	プログラミングの課題	268
11章	パフォーマンス	271
11.1	逐次プログラムのパフォーマンス	271
11.2	例題：逐次台形公式	273
11.3	I/Oについて	274
11.4	並列プログラムのパフォーマンス解析	275
11.5	通信コスト	277
11.6	例題：並列台形公式	279
11.7	時間計測	282
11.8	要約	283
11.9	文献	284
11.10	練習問題	284
11.11	プログラミングの課題	285
12章	さらにパフォーマンス	287
12.1	アムダールの法則	287
12.2	処理とオーバヘッド	289
12.3	オーバヘッドのソース	291
12.4	スケラビリティ	292
12.5	パフォーマンス評価に使える問題	294
12.5.1	ワークステーションクラスタにおけるリソースの競合	294
12.5.2	負荷バランスとアイドルリング	295
12.5.3	通信と計算のオーバーラップ	297
12.5.4	集団通信	298
12.6	パフォーマンス評価ツール	299
12.6.1	MPI プロファイリングインタフェース	300
12.6.2	Upshot	302
12.7	要約	305
12.8	文献	307
12.9	練習問題	308
12.10	プログラミングの課題	308

目次	xiii	
13章	進んだ1対1通信	309
13.1	例題：オールギヤザのコーディング	309
13.1.1	関数のパラメータ	310
13.1.2	リングパスオールギヤザ	310
13.2	ハイパーキューブ	314
13.2.1	ハイパーキューブデータ交換におけるもう一つの問題	316
13.2.2	ハイパーキューブアルゴリズムの詳細	318
13.3	送受信	324
13.4	ヌルプロセス	326
13.5	非ブロッキング通信	327
13.5.1	非ブロッキング通信によるリングオールギヤザ	329
13.5.2	非ブロッキング通信によるハイパーキューブオールギヤザ	331
13.6	持続的通信要求	333
13.7	通信モード	336
13.7.1	同期モード	337
13.7.2	レディモード	338
13.7.3	バッファモード	340
13.8	1対1通信の補足	342
13.9	要約	343
13.10	文献	347
13.11	練習問題	347
13.12	プログラミングの課題	348
14章	並列アルゴリズム	349
14.1	並列アルゴリズムの設計	349
14.2	ソート	350
14.3	逐次バイトニックソート	350
14.4	並列バイトニックソート	355
14.5	木探索と組合せ最適化	359
14.6	逐次木探索	360
14.7	並列木探索	363
14.7.1	Par_dfs	365
14.7.2	Service_requests	368
14.7.3	Work_remains	368

14.7.4	分散停止検出	370
14.8	要 約	371
14.9	文 献	372
14.10	練習問題	373
14.11	プログラミングの課題	374
15章	並列ライブラリ	375
15.1	ライブラリの使用：賛否両論	375
15.2	複数の言語の使用	376
15.3	ScaLAPACK	378
15.4	ScaLAPACK プログラムの例	381
15.5	PETSc	388
15.6	PETSc の例	390
15.7	要 約	396
15.8	文 献	397
15.9	練習問題	397
15.10	プログラミングの課題	398
16章	おわりに	399
16.1	これからどこへ行くべきか	399
16.2	MPI の将来	400
付録 A	MPI コマンドの要約	403
A.1	1対1通信関数	403
A.1.1	ブロッキング送信と受信	403
A.1.2	通信モード	404
A.1.3	バッファアロケーション	405
A.1.4	非ブロッキング通信	405
A.1.5	プローブとキャンセル	409
A.1.6	持続的通信要求	410
A.1.7	送 受 信	412
A.2	派生データ型と MPI_pack/Unpack	413

A.2.1	派生データ型	413
A.2.2	MPI_Pack と MPI_Unpack	416
A.3	集団通信関数	417
A.3.1	バリアとブロードキャスト	417
A.3.2	ギャザとスキヤタ	418
A.3.3	リダクション操作	421
A.4	グループ、コンテキスト、コミュニケータ	422
A.4.1	グループ管理	423
A.4.2	コミュニケータ管理	426
A.4.3	インタコミュニケータ	427
A.4.4	属性キャッシング	429
A.5	プロセストポロジ	430
A.5.1	一般トポロジ関数	430
A.5.2	カーテシアントポロジ管理	430
A.5.3	グラフトポロジ管理	432
A.6	環 境 管 理	434
A.6.1	実装情報	434
A.6.2	エラー処理	435
A.6.3	時間関数	436
A.6.4	起 動	436
A.7	プロファイリング	437
A.8	定 数	437
A.9	型定義	440
付録 B	インターネット上の MPI	443
B.1	MPI の実装	443
B.2	MPI の FAQ	444
B.3	MPI ウェブページ	444
B.4	MPI ニュースグループ	444
B.5	MPI-2 と MPI-IO	444
B.6	MPI 並列プログラミング	444
	参 考 文 献	447
	索 引	451