

ОГЛАВЛЕНИЕ

Предисловие	3
Введение	5
Глава 1. Основные понятия стандарта MPI	8
1.1. Коммуникаторы	8
1.2. Блокирующие операции Send/Recv	8
1.3. Неблокирующие операции Isend/Irecv	9
1.4. Поиск поступивших сообщений	11
1.4.1. Протокол Eager	11
1.4.2. Протокол Rendezvous	12
1.5. Структура библиотек MPI	12
1.6. Подходы к реализации алгоритмов коллективных операций	16
1.6.1. Алгоритмы без учета топологии коммуникационной сети	16
1.6.2. Алгоритмы для фиксированной сетевой топологии	17
1.6.3. Алгоритмы с аппаратной реализацией	17
1.7. Классификация коллективных операций	17
1.7.1. Блокирующие коллективные операции	17
1.7.2. Неблокирующие коллективные операции	18
1.7.3. Коллективные операции с постоянными значениями аргументов	19
1.7.4. Коллективные операции для виртуальных топологий процессов	20
1.8. Модели параллельных вычислений	21
1.8.1. Модель вычислительной системы	23
1.8.2. Модель Хокни	24
1.8.3. Модель LogP	26
1.8.4. Модель LogGP	29
Глава 2. Широковещательная передача	32
2.1. Операция Vcast	32

2.2. Алгоритмы с линейной сложностью.....	33
2.3. Алгоритм параллельных цепочек	35
2.4. Алгоритм параллельных цепочек переменной длины	37
2.5. Алгоритм бинарного дерева	39
2.6. Алгоритм оптимального k-ичного дерева.....	43
2.7. Алгоритм биномиального дерева	44
2.8. Алгоритм k-номиального дерева.....	48
2.9. Конвейерный алгоритм с сегментацией сообщения...	51
2.10. Алгоритм бинарного дерева с сегментацией сообщения	55
2.11. Алгоритм на основе Scatter и Allgather	57
2.12. Алгоритмы оптимальные по пропускной способности	62
2.13. Операция Scatter.....	62
2.13.1. Линейный алгоритм	64
2.13.2. Алгоритм биномиального дерева.....	65
Глава 3. Широковещательный прием Gather.....	68
3.1. Алгоритм плоского дерева	68
3.2. Алгоритм биномиального дерева	69
Глава 4. Операция Allgather	72
4.1. Алгоритм рекурсивного удваивания.....	72
4.2. Обобщенный алгоритм рекурсивного удваивания....	76
4.3. Алгоритм рекурсивного удваивания для произвольного числа процессов	77
4.4. Алгоритм рекурсивного деления пополам.....	81
4.5. Алгоритм Врука	84
4.6. Обобщение алгоритма Врука	86
4.7. Алгоритм Sparbit	90
4.8. Кольцевой алгоритм.....	91
4.9. Обзор алгоритмов.....	92
Глава 5. Корневая редукция Reduce.....	95
5.1. Линейный алгоритм	96
5.2. Конвейерный алгоритм	103
5.3. Алгоритм параллельных цепочек	105
5.4. Алгоритм биномиального дерева для коммутативной операции	110
5.4.1. Структура дерева обменов.....	110
5.4.2. Анализ эффективности алгоритма	115

5.4.3. Левостороннее дерево обменов.....	124
5.5. Алгоритм биномиального дерева для некоммутативной операции.....	124
5.5.1. Структура дерева обменов.....	124
5.5.2. Анализ эффективности алгоритма.....	126
5.6. Алгоритм бинарного дерева.....	127
5.7. Алгоритм Рабенсейфнера.....	132
Глава 6. Глобальная редукция Allreduce.....	138
6.1. Кольцевой алгоритм.....	138
6.2. Алгоритм рекурсивного удваивания.....	140
6.3. Алгоритм ReduceScatter-Allgather с рекурсивным удваиванием.....	145
6.4. Кольцевой алгоритм ReduceScatter-Allgather.....	153
6.5. Сравнение алгоритмов.....	156
6.6. Операция Exscan.....	157
6.6.1. Алгоритм рекурсивного удваивания.....	158
6.7. Операция Scan.....	160
6.7.1. Алгоритм рекурсивного удваивания.....	160
6.8. Операция Reduce-scatter.....	161
6.8.1. Алгоритм деления пополам.....	163
6.8.2. Алгоритм Butterfly.....	166
6.8.3. Алгоритм попарных обменов.....	170
Глава 7. Операция Alltoall.....	172
7.1. Операция Alltoall Scatter/Gather.....	172
7.2. Линейный алгоритм.....	172
7.3. Алгоритм попарных обменов.....	173
7.4. Алгоритм Брука.....	174
Глава 8. Барьерная синхронизация.....	177
8.1. Операция Barrier.....	177
8.2. Линейный алгоритм с центральным счетчиком.....	179
8.3. Алгоритм рекурсивного удваивания.....	181
8.4. Рассеивающий алгоритм.....	183
Глава 9. Алгоритмы для систем с общей памятью ...	185
9.1. Операция Vcast.....	186
9.1.1. Формирование сегмента разделяемой памяти..	187
9.1.2. Шаги алгоритма при вызове Vcast.....	189
9.1.3. Оценка времени выполнения алгоритма.....	192

9.2. Операция Barrier	197
9.2.1. Алгоритм с глобальным счетчиком	199
9.2.2. Плоское дерево	200
9.2.3. Плоское дерево Gather/Release	201
9.2.4. Объединяющее дерево	202
9.2.5. MCS-барьер	204
9.2.6. Турнирный алгоритм	206
9.2.7. Рассеивающий алгоритм	207
9.2.8. Алгоритм выбора корня операции Barrier	207
Глава 10. Иерархические алгоритмы коллективных операций	209
10.1. Иерархические алгоритмы на базе коммутаторов .	211
10.1.1. Иерархический алгоритм Bcast	213
10.1.2. Иерархический алгоритм Barrier	214
10.1.3. Иерархический алгоритм Reduce	215
10.1.4. Иерархический алгоритм Allreduce	216
10.1.5. Иерархический алгоритм Gather	217
10.1.6. Иерархический алгоритм Allgather	218
10.1.7. Иерархический алгоритм ReduceScatter	219
10.1.8. Иерархические алгоритмы Exscan и Scan	219
10.2. Иерархические алгоритмы с формированием групп процессов	220
10.2.1. Построение иерархии групп процессов	221
10.2.2. Алгоритм операции Reduce	224
10.2.3. Алгоритм операции Bcast	226
10.2.4. Алгоритм операции Allreduce	226
10.3. Совмещение операций на разных уровнях иерархии	228
10.4. Динамическая оптимизация алгоритмов	230
10.4.1. Графы алгоритмов Allgather	231
10.4.2. Модель вычислительной системы	234
10.4.3. Метод оптимизации алгоритмов Allgather	235
Глава 11. Неблокирующие коллективные операции ..	239
11.1. Построение расписания коллективной операции	241
11.2. Продвижение выполнения операций расписания	246
11.3. Пути оптимизации расписаний	248
Глава 12. Коллективные операции для гибридных вычислительных систем	251

12.1. Архитектура гибридных вычислительных систем	251
12.2. Алгоритм Allreduce для гибридной системы	254
12.3. Версия Allreduce с параллельным выполнением операций	254
Глава 13. Экспериментальная оценка эффективности коллективных операций	257
13.1. Измерение времени выполнения коллективных операций	257
13.2. Систематические ошибки измерений	259
13.3. Измерения с запуском операций по глобальным часам	262
13.4. Синхронизации показаний часов процессов	264
14. Приложение	266
14.1. Функции округления и операций с битами	266
14.1.1. Функции пол и потолок	266
14.1.2. Операции с битами	266
14.2. Деревья	269
14.2.1. Основные определения	269
14.2.2. Бинарные деревья	270
14.2.3. Обходы бинарных деревьев	271
14.2.4. Биномиальные деревья	272
Литература	274