



## ПАРАЛЛЕЛЬНОЕ ПРОГРАММИРОВАНИЕ В СТАНДАРТЕ MPI

Если у вас есть яблоко и у меня есть яблоко, и мы обмениваемся этими яблоками, то у вас и у меня остается по одному яблоку. А если у вас есть идея и у меня есть идея, и мы обмениваемся идеями, то у каждого из нас будет по две идеи.

*Дж.Б. Шоу*

Суперкомпьютеры и их применение

---

---

---

---

---

---

---

---

---

---

## Содержание

2

- Модель передачи сообщений для параллельного программирования в системах с распределенной памятью
- Модели SPMD и MPMD запуска параллельных программ
- Стандарт Message Passing Interface (MPI)
- Основные понятия и функции MPI

Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

---

---

## Модель передачи сообщений

3

Процесс 0    Процесс 1    Процесс N-1



Сеть

Процессор 0    Процессор 1    ...    Процессор N-1

Память 0    Память 1    ...    Память N-1

Сообщ.

- Параллельное приложение состоит из нескольких процессов, выполняющихся одновременно.
- Каждый процесс имеет частную память.
- Обмены данными между процессами осуществляются посредством явной отправки/получения сообщений.
- Процессы, как правило, выполняются на различных процессорах.

Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

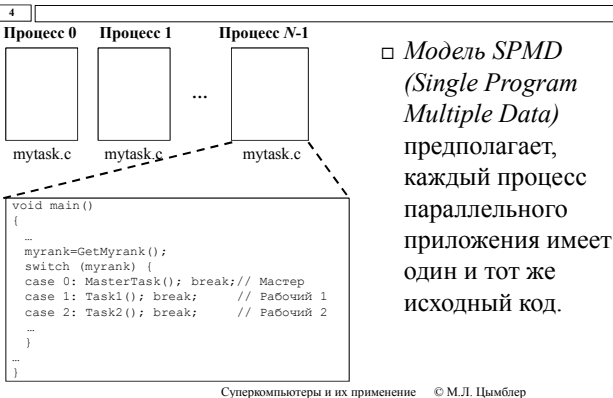
---

---

---

---

## Модель SPMD




---

---

---

---

---

---

---

---

---

---

---

---

## Модель MPMD




---

---

---

---

---

---

---

---

---

---

---

---

## Стандарт MPI

- 6
- *MPI (Message Passing Interface)* – стандарт, реализующий модель обмена сообщениями между параллельными процессами. Поддерживает модели выполнения SPMD и (начиная с версии 2.0) MPMD.
  - Стандарт представляет собой набор спецификаций подпрограмм (более 120) на языках C, C++ и FORTRAN.
  - Стандарт реализуется разработчиками в виде библиотек подпрограмм для различных аппаратно-программных платформ (кластеры, персональные компьютеры, ..., Windows, Unix/Linux, ...).
  - Коммерческие (IntelMPI, MSMPI и др.) и свободно распространяемые реализации стандарта (MPICH и др.).
  - Текущая версия стандарта: <http://www.mpi-forum.org>.

Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

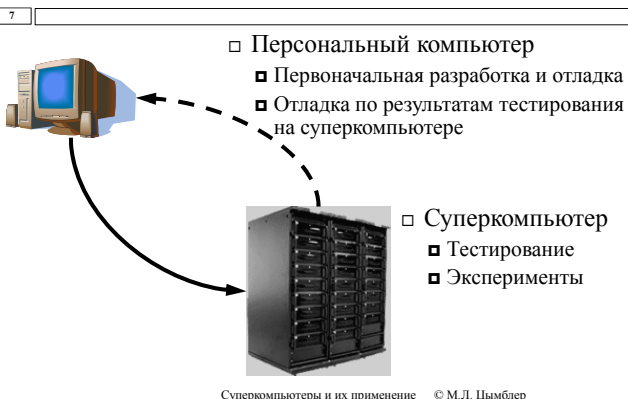
---

---

---

---

## Цикл разработки MPI-программ




---

---

---

---

---

---

---

---

---

---

## MPI-программа

- 8
- *MPI-программа* – множество параллельных взаимодействующих процессов.
  - Процессы порождаются один раз, в момент запуска программы средствами среды исполнения MPI программ\*.
  - Все процессы программы последовательно перенумерованы, начиная с 0. Номер процесса именуется *рангом* процесса.
  - Каждый процесс работает в своем адресном пространстве, каких-либо общих данных нет. Единственный способ взаимодействия процессов – явный обмен сообщениями.

\*Порождение дополнительных процессов и уничтожение существующих возможно только начиная с версии MPI-2.0.

Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

---

---

## Коммуникаторы

- 9
- Для локализации области взаимодействия процессов можно динамически создавать специальные программные объекты – *коммуникаторы*. Один и тот же процесс может входить в разные коммуникаторы.
  - Взаимодействия процессов проходят в рамках некоторого коммуникатора. Сообщения, переданные в разных коммуникаторах, не пересекаются и не мешают друг другу.
  - Атрибуты процесса MPI-программы:
    - ▣ номер коммуникатора;
    - ▣ номер в коммуникаторе (от 0 до  $n-1$ ,  $n$  – число процессов в коммуникаторе).
  - Стандартные коммуникаторы:
    - ▣ MPI\_COMM\_WORLD – все процессы приложения
    - ▣ MPI\_COMM\_SELF – текущий процесс приложения
    - ▣ MPI\_COMM\_NULL – пустой коммуникатор

Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

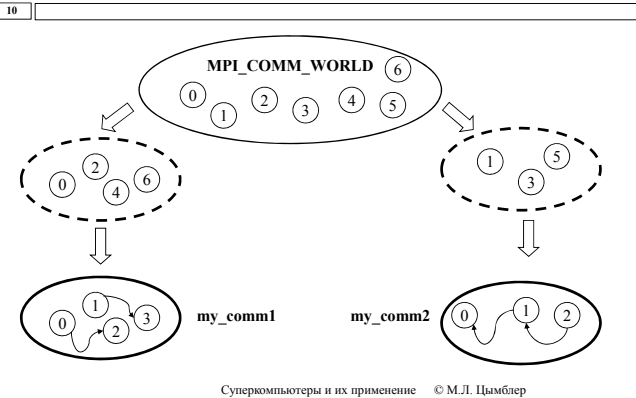
---

---

---

---

## Коммуникаторы




---

---

---

---

---

---

---

---

---

---

## Сообщение

- 11
- *Сообщение* процесса – набор данных стандартного (определенного в MPI) или пользовательского типа.
  - Основные атрибуты сообщения:
    - номер процесса-отправителя (получателя)
    - номер коммуникатора
    - тег (уникальный идентификатор) сообщения (целое число)
    - тип элементов данных в сообщении
    - количество элементов данных
    - указатель на буфер с сообщением

Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

---

---

## Структура MPI-программы

12

```
#include "mpi.h" // Подключение библиотеки

int main (int argc, char *argv[])
{
// Здесь код без использования MPI функций

  MPI_Init(&argc, &argv); // Инициализация выполнения

// Здесь код, реализующий обмены

  MPI_Finalize(); // Завершение

// Здесь код без использования MPI функций

  return 0;
}
```

Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

---

---

## МРІ-функции

13

- Имеют имена вида `MPI_...`
- Возвращают целое число – `MPI_SUCCESS` или код ошибки.
- Простые функции общего назначения:
  - // Количество процессов в коммуникаторе  
`int MPI_Comm_size(MPI_Comm comm, int * size);`
  - // Номер (ранг) процесса в коммуникаторе  
`int MPI_Comm_rank(MPI_Comm comm, int * rank);`
  - // Замер времени (сек.)  
`double MPI_Wtime(void);`
  - // Имя текущего процессорного узла  
`int MPI_Get_processor_name(char * name, int * len);`

Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

## Простая МРІ-программа

14

```
#include "mpi.h"
#include <stdio.h>

int main (int argc, char *argv[])
{
    int num, rank;

    MPI_Init (&argc, &argv);
    MPI_Comm_size (MPI_COMM_WORLD, &num);
    MPI_Comm_rank (MPI_COMM_WORLD, &rank);
    printf("Привет! Я %d-й процесс из %d.\n", rank, num);
    MPI_Finalize();
    return 0;
}
```

```
Привет! Я 1-й процесс из 4.
Привет! Я 0-й процесс из 4.
Привет! Я 4-й процесс из 4.
Привет! Я 3-й процесс из 4.
Привет! Я 2-й процесс из 4.
```

Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

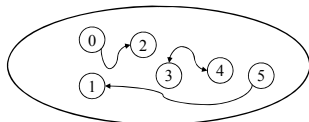
---

---

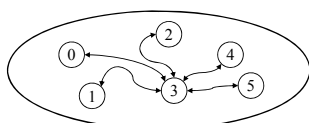
## Виды взаимодействия процессов

15

- *Взаимодействие "точка-точка"* – обмен между двумя процессами одного коммуникатора.



- *Коллективное взаимодействие* – обмен между всеми процессами одного коммуникатора.



Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

## Взаимодействие "точка-точка"

16

- Участвуют два процесса: отправитель сообщения и получатель сообщения.
  - Отправитель должен вызвать одну из функций отправки сообщения и явно указать атрибуты получателя (коммуникатор и номер в коммуникаторе) и тег сообщения.
  - Получатель должен вызвать одну из функций получения сообщения и указать (тот же) коммуникатор отправителя; получатель может не знать номер отправителя и/или тег сообщения.
- Свойства:
  - Сохранение порядка (если P0 передает P1 сообщения A и затем B, то P1 получит A, а затем B).
  - Гарантированное выполнение обмена (если P0 вызвал функцию отправки, а P1 вызвал функцию получения, то P1 получит сообщение от P0).

Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

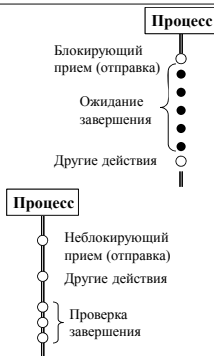
---

---

## Виды коммуникационных функций "точка-точка"

17

- **Блокирующая** функция запускает операцию и возвращает управление процессу только после ее завершения.
  - После завершения допустима модификация отправленного (принятого) сообщения.
- **Неблокирующая** функция запускает операцию и возвращает управление процессу немедленно.
  - Факт завершения операции проверяется позднее (с помощью другой функции).
  - До завершения операции недопустима модификация отправляемого (получаемого) сообщения.



Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

---

---

## Блокирующие vs неблокирующие операции

18

- **Блокирующие операции**
  - Имеют более простую семантику.
  - Относительно легко используются в программе, не требуют дополнительных действий для завершения обмена.
  - Могут снизить быстродействие программы.
  - Могут привести к тупикам.
- **Неблокирующие операции**
  - Имеют более сложную семантику.
  - Относительно трудно используются в программе, требуют дополнительных действий для завершения обмена.
  - Могут повысить быстродействие программы.
  - Не вызывают тупиков.

Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

---

---

## Режимы отправки сообщений "точка-точка"

19

- Стандартный* – операция завершается сразу после отправки сообщения.
- Синхронный* – операция завершается после приема подтверждения от адресата.
- Буферизованный* – операция завершается, как только сообщение копируется в системный буфер для дальнейшей отправки.
- "По готовности"* – операция начинается, если адресат инициализировал прием и завершается сразу после отправки.

Суперкомпьютеры и их применение © М.Л. Цымблер

## Режимы отправки сообщений "точка-точка"

20

- Режим *по готовности* формально является наиболее быстрым, но используется достаточно редко, т.к. обычно сложно гарантировать готовность операции приема.
- Стандартный* и *буферизованный* режимы также выполняются достаточно быстро, но могут приводить к большим расходам ресурсов (памяти), и могут быть рекомендованы для передачи коротких сообщений.
- Синхронный* режим является наиболее медленным, т.к. требует подтверждения приема. В то же время этот режим наиболее надежен, и может быть рекомендован для передачи длинных сообщений.

Суперкомпьютеры и их применение © М.Л. Цымблер

## Коммуникационные функции "точка-точка"

21

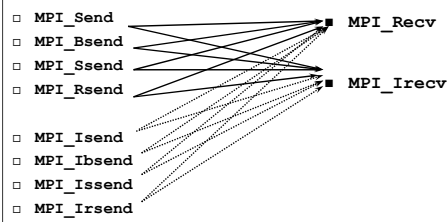
- Отправка: **MPI\_[I][R, S, B]Send**
- Прием: **MPI\_[I]Recv**

Блокирующие			Неблокирующие		
Отправка		Прием	Отправка		Прием
Стандартная	MPI_Send	MPI_Recv	Стандартная	MPI_Isend	MPI_Irecv
Синхронная	MPI_Ssend		Синхронная	MPI_Issend	
Буферизованная	MPI_Bsend		Буферизованная	MPI_Ibsend	
По готовности	MPI_Rsend		По готовности	MPI_Irsend	

Суперкомпьютеры и их применение © М.Л. Цымблер

## Коммуникационные функции "точка-точка"

22



Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

---

---

## Блокирующая стандартная отправка сообщения

23

- **int MPI\_Send**
  - IN void \* buf – указатель на буфер с сообщением
  - IN int count – количество элементов в буфере
  - IN MPI\_Datatype datatype – MPI-тип данных элементов в буфере
  - IN int dest – номер процесса-получателя
  - IN int tag – тег сообщения
  - IN MPI\_Comm comm – коммунитор

Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

---

---

## Блокирующее стандартное получение сообщения

24

- **int MPI\_Recv**
  - OUT void \* buf – указатель на буфер с сообщением
  - IN int count – количество элементов в буфере
  - IN MPI\_Datatype datatype – MPI-тип данных элементов в буфере
  - IN int src – номер процесса-отправителя
  - IN int tag – тег сообщения
  - IN MPI\_Comm comm – коммунитор
  - OUT MPI\_Status\* status – информация о фактически полученных данных (указатель на структуру с двумя полями: source – номер процесса-источника, tag – тег сообщения)

Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

---

---



## Пример программы

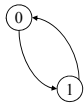
25

```

#include "mpi.h"
#include <stdio.h>
int main(int argc, char *argv[])
{
    int numtasks, rank, dest, src, rc, tag=777;
    char inmsg, outmsg='x';
    MPI_Status Stat;

    MPI_Init(&argc, &argv);
    MPI_Comm_size(MPI_COMM_WORLD, &numtasks);
    MPI_Comm_rank(MPI_COMM_WORLD, &rank);
    if (rank == 0) {
        dest = 1; src = 1;
        rc = MPI_Send(&outmsg, 1, MPI_CHAR, dest, tag, MPI_COMM_WORLD);
        rc = MPI_Recv(&inmsg, 1, MPI_CHAR, src, tag, MPI_COMM_WORLD, &Stat);
    } else
    if (rank == 1) {
        dest = 0; src = 0;
        rc = MPI_Recv(&inmsg, 1, MPI_CHAR, src, tag, MPI_COMM_WORLD, &Stat);
        rc = MPI_Send(&outmsg, 1, MPI_CHAR, dest, tag, MPI_COMM_WORLD);
    }
    MPI_Finalize();
}

```



Суперкомпьютеры и их применение © М.Л. Цымблер

## Неблокирующая стандартная отправка сообщения

26

### □ int MPI\_Isend

- IN void \* buf – указатель на буфер с сообщением
- IN int count – количество элементов в буфере
- IN MPI\_Datatype datatype – MPI-тип данных элементов в буфере
- IN int dest – номер процесса-получателя
- IN int tag – тег сообщения
- IN MPI\_Comm comm – коммуникатор
- OUT MPI\_Request \*request – дескриптор операции (для последующей проверки завершения операции)

Суперкомпьютеры и их применение © М.Л. Цымблер

## Неблокирующее стандартное получение сообщения

27

### □ int MPI\_Irecv

- OUT void \* buf – указатель на буфер с сообщением
- IN int count – количество элементов в буфере
- IN MPI\_Datatype datatype – MPI-тип данных элементов в буфере
- IN int src – номер процесса-отправителя
- IN int tag – тег сообщения
- IN MPI\_Comm comm – коммуникатор
- OUT MPI\_Request \*request – дескриптор операции (для последующей проверки завершения операции)

Суперкомпьютеры и их применение © М.Л. Цымблер

## Завершение неблокирующих обменов

28

- Проверка завершения
  - int MPI\_Test**  
(MPI\_Request \*request, int \*flag, MPI\_Status \*status)
  - int MPI\_Testany (...)
  - int MPI\_Testall (...)
  - int MPI\_Testsome (...)
- Ожидание завершения
  - int MPI\_Wait**  
(MPI\_Request \*request, MPI\_Status \*status)
  - int MPI\_Waitany (...)
  - int MPI\_Waitall (...)
  - int MPI\_Waitsome (...)

Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

---

---

## Пример

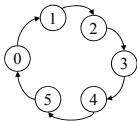
29

```
#include "mpi.h"
#include <stdio.h>
int main(int argc, char *argv[])
{
    int numtasks, rank, next, prev, buf[2], tag1=111, tag2=222;
    MPI_Request reqs[4];
    MPI_Status stats[4];

    MPI_Init(&argc, &argv);
    MPI_Comm_size(MPI_COMM_WORLD, &numtasks);
    MPI_Comm_rank(MPI_COMM_WORLD, &rank);

    prev = rank - 1;
    if (prev < 0) prev = numtasks - 1;
    next = rank + 1;
    if (next > numtasks - 1) next = 0;
    MPI_Irecv(&buf[0], 1, MPI_INT, prev, tag1, MPI_COMM_WORLD, &reqs[0]);
    MPI_Irecv(&buf[1], 1, MPI_INT, next, tag2, MPI_COMM_WORLD, &reqs[1]);
    MPI_Isend(&rank, 1, MPI_INT, prev, tag2, MPI_COMM_WORLD, &reqs[2]);
    MPI_Isend(&rank, 1, MPI_INT, next, tag1, MPI_COMM_WORLD, &reqs[3]);
    MPI_Waitall(4, reqs, stats);

    MPI_Finalize();
}
```



Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

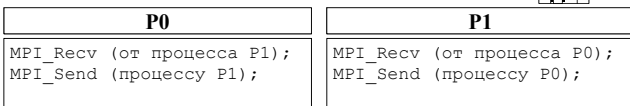
---

---

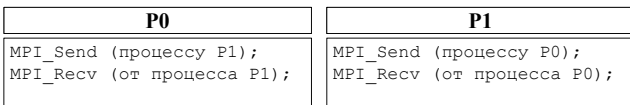
## Тупики (deadlocks)

30

- Гарантированный тупик



- Возможный тупик



Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

---

---

## Разрешение тупиков

31	
<b>P0</b>	<b>P1</b>
MPI_Send (процессу P1); MPI_Recv (от процесса P1);	MPI_Recv (от процесса P0); MPI_Send (процессу P0);
<b>P0</b>	<b>P1</b>
MPI_Send (процессу P1); MPI_Recv (от процесса P1);	MPI_Irecv (от процесса P0); MPI_Send (процессу P0); MPI_Wait
<b>P0</b>	<b>P1</b>
// Совмещенные прием и передача MPI_Sendrecv	// Совмещенные прием и передача MPI_Sendrecv

Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

---

---

---

---

## Совмещение приема и передачи сообщения

32	
<ul style="list-style-type: none"> <li>□ <b>int MPI_Sendrecv</b> <ul style="list-style-type: none"> <li>□ OUT void * sbuf - адрес начала буфера с посылаемым сообщением;</li> <li>□ IN int scount - число передаваемых элементов в сообщении;</li> <li>□ IN MPI_Datatype stype - тип передаваемых элементов;</li> <li>□ IN int dest - номер процесса-получателя;</li> <li>□ IN int stag - идентификатор посылаемого сообщения;</li> <li>□ OUT void * rbuf - адрес начала буфера приема сообщения;</li> <li>□ IN int rcount - число принимаемых элементов сообщения;</li> <li>□ IN MPI_Datatype rtype - тип принимаемых элементов;</li> <li>□ IN int source - номер процесса-отправителя;</li> <li>□ IN int rtag - идентификатор принимаемого сообщения;</li> <li>□ IN MPI_Comm comm - идентификатор коммуникатора;</li> <li>□ OUT MPI_Status status - параметры принятого сообщения.</li> </ul> </li> <li>□ Объединяет в едином запросе отсылку и прием сообщений и гарантирует отсутствие тупиков.</li> <li>□ Принимающий и отправляющий процессы могут являться одним и тем же процессом. Буфера приема и отсылки обязательно должны быть различными. Сообщение, отправленное операцией MPI_Sendrecv, может быть принято обычным образом, и точно также операция MPI_Sendrecv может принять сообщение, отправленное обычной операцией MPI_Send.</li> </ul>	
Суперкомпьютеры и их применение © М.Л. Цымблер	

---

---

---

---

---

---

---

---

---

---

---

---

## Пример

33	
<pre>#include "mpi.h" #include &lt;stdio.h&gt; int rank, numtasks, prev, next, buf[2]; MPI_Status status1, status2;  int main(int argc, char *argv[]) {     MPI_Init(&amp;argc, &amp;argv);     MPI_Comm_size(MPI_COMM_WORLD, &amp;numtasks);     MPI_Comm_rank(MPI_COMM_WORLD, &amp;rank);      prev=rank-1; if (prev &lt; 0) prev=numtasks-1;     next=rank+1; if (next &gt; size-1) next=0;     MPI_Sendrecv(rank, 1, MPI_INTEGER, prev, 777, buf[1], 1, MPI_INTEGER, next,     777, MPI_COMM_WORLD, status2);     MPI_Sendrecv(rank, 1, MPI_INTEGER, next, 888, buf[0], 1, MPI_INTEGER, prev,     888, MPI_COMM_WORLD, status1);     printf("Process %d, prev=%d, next=%d\n", rank, buf[0], buf[1]);      MPI_Finalize();     return 0; }</pre>	
Суперкомпьютеры и их применение © М.Л. Цымблер	

---

---

---

---

---

---

---

---

---

---

---

---

## Получение сообщений

34

- "Джокеры"
  - ▣ ранг процесса `MPI_ANY_SOURCE`
  - ▣ тег сообщения `MPI_ANY_TAG`
- Информация об ожидаемом сообщении
  - ▣ С блокировкой
 

```
int MPI_Probe(int src, int tag,
              MPI_Comm comm, MPI_Status * status);
```
  - ▣ Без блокировки
 

```
int MPI_Iprobe(int src, int tag,
               MPI_Comm comm, int * flag, MPI_Status *
               status);
```
  - ▣ Количество элементов в принятом сообщении
 

```
int MPI_Get_count(MPI_Status * status,
                  MPI_Datatype type, int * cnt);
```



Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

---

---

## Коллективные операции

35

- Прием и/или передачу выполняют одновременно *все* процессы коммутатора.
- Коллективная функция имеет большое количество параметров, часть которых нужна для приема, а часть для передачи. При вызове в разных процессах та или иная часть игнорируется.
- Значения *всех* параметров коллективных функций (за исключением адресов буферов) должны быть идентичными во всех процессах.
- MPI назначает теги для сообщений автоматически.

Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

---

---

## Барьерная синхронизация

36

- `int MPI_Barrier(MPI_Comm comm)`
  - ▣ Обеспечивает *синхронизацию* процессов – одновременное достижение процессами указанной точки вычислений.
  - ▣ Должна вызываться всеми процессами коммутатора.
  - ▣ Продолжение вычислений любого процесса произойдет по окончании выполнения функции `MPI_Barrier` всеми процессами коммутатора.



Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

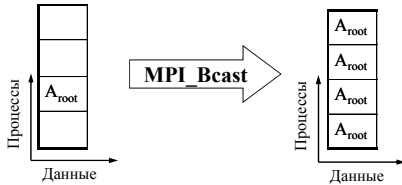
---

---

## Широковещательная рассылка

37

- **int MPI\_Bcast**
  - IN/OUT void \*buf – указатель на буфер с сообщением (отправляемым – для процесса с рангом root, принимаемым – для остальных процессов)
  - IN int count – количество элементов в буфере
  - IN MPI\_Datatype datatype – MPI-тип данных элементов в буфере
  - IN int root – ранг процесса, выполняющего рассылку данных
  - IN MPI\_Comm comm – коммутатор



Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

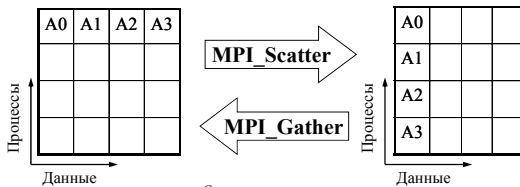
---

---

## Коллективный прием сообщения

38

- Сборка элементов данных из буферов всех процессов в буфере процесса с рангом root
- int MPI\_Gather**(void \* sbuf, int scount, MPI\_Datatype stype, void \* rbuf, int rcount, MPI\_Datatype rtype, int root, MPI\_Comm comm);
- Рассылка элементов данных из буфера процесса с номером rootRank в буфера всех процессов (обратная к MPI\_Gather)
- int MPI\_Scatter**



Суперкомпьютеры и их применение 38

---

---

---

---

---

---

---

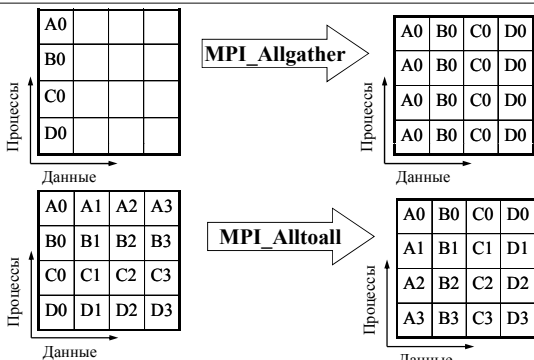
---

---

---

## Широковещательные прием и передача

39



Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

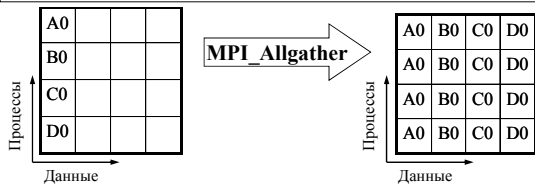
---

---

---

## Широковещательный прием

40



```
int MPI_Allgather(void *sbuf, int scount, MPI_Datatype stype,
void *rbuf, int rcount, MPI_Datatype rtype, MPI_Comm comm)
```

Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

---

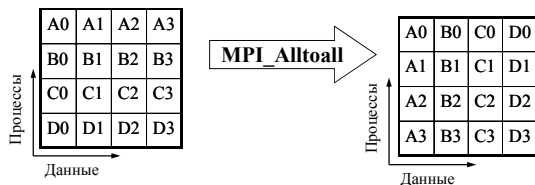
---

## Широковещательная передача

41

```
int MPI_Alltoall(void *sbuf, int scount, MPI_Datatype stype,
void *rbuf, int rcount, MPI_Datatype rtype, MPI_Comm comm)
```

- **sbuf**, **scount**, **stype** - параметры передаваемых сообщений,
- **rbuf**, **rcount**, **rtype** - параметры принимаемых сообщений
- **comm** - коммуникатор, в рамках которого выполняется передача данных



Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

---

---

## Глобальные операции над данными

42

- Выполнение `count` независимых глобальных операций `op` над соответствующими элементами массивов `sbuf`.  
Результат выполнения над  $i$ -ми элементами `sbuf` записывается в  $i$ -й элемент массива `rbuf` процесса с рангом `root`.

```
int MPI_Reduce(void * sbuf, void * rbuf, int
count, MPI_Datatype type, MPI_Op op, int
root, MPI_Comm comm);
```

- Глобальные операции:
  - `MPI_MAX`, `MPI_MIN`
  - `MPI_SUM`, `MPI_PROD`
  - ...
  - `MPI_Op_Create()`

Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

---

---



## Пользовательские типы данных

46

### □ Создание типа "массив"

```
int MPI_Type_contiguous(int count,
MPI_Datatype oldtype, MPI_Datatype *newtype);
```

```
#define N 100
int A[N];
MPI_Datatype MPI_INTARRAY100;
...
MPI_Type_contiguous(N, MPI_INT, & MPI_INTARRAY100);
MPI_Type_commit(&MPI_INTARRAY100);
...
MPI_Send(A, 1, MPI_INTARRAY100, ... );
// то же, что и MPI_Send(A, N, MPI_INT, ... );
...
MPI_Type_free( &intArray100 );
```

Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

---

---

---

---

## MPI-2: нововведения

47

### □ Динамическое порождение процессов

- Разрешается создание и уничтожение процессов по ходу выполнения программы.
- Предусмотрен специальный механизм, позволяющий устанавливать связь между только что порожденными процессами и уже работающей частью MPI-программы.
- Имеется возможность установить связь между двумя приложениями даже в том случае, когда ни одно из них не было инициатором запуска другого.

### □ Одностороннее взаимодействие процессов

- Обмен сообщениями по схеме Put/Get вместо традиционной схемы Send/Receive. Активной стороной может быть один процесс (при обмене не требуется активность отправителя либо получателя).

### □ Параллельный ввод-вывод

- Специальный интерфейс для работы процессов с файловой системой.

### □ Расширенные коллективные операции

- Во многие коллективные операции добавлена возможность взаимодействия между процессами, входящими в разные коммутаторы.
- Многие коллективные операции внутри коммутатора могут выполняться в режиме, при котором входные и выходные буфера совпадают.

### □ Интерфейс для C++

Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

---

---

---

---

## Заключение

48

### □ Модель передачи сообщений для параллельного программирования в системах с распределенной памятью

### □ Модели SPMD и MPMD запуска параллельных программ

### □ Стандарт Message Passing Interface (MPI)

### □ Основные понятия и функции MPI

Суперкомпьютеры и их применение © М.Л. Цымблер

---

---

---

---

---

---

---

---

---

---

---

---