Práctica 9

**Introducción a MPI**

Jordi Blasco Lozano

Computación de alto rendimiento

Grado en Inteligencia Artificial

## Índice:

[Índice: 2](#_Toc194339654)

[1. Introducción 2](#_Toc194339655)

[2. Desarrollo 2](#_Toc194339656)

[Ejercicio 1 - Comunicación punto a punto 2](#_Toc194339657)

[¿Qué hace el programa propuesto? 2](#_Toc194339658)

[¿Qué sucede si cambiamos los tags? 3](#_Toc194339659)

[¿Cómo compilar y ejecutar con mpicc? 3](#_Toc194339660)

## Introducción

## Desarrollo

## Ejercicio 1 - Comunicación punto a punto

Este ejercicio tiene como objetivo afianzar el conocimiento en la comunicación punto a punto utilizando MPI (Message Passing Interface).

## ¿Qué hace el programa propuesto?

El programa inicia con la llamada a MPI\_Init(&argc, &argv), que configura el entorno de MPI y permite que los procesos se comuniquen. Cada proceso obtiene su identificador único (rank) mediante MPI\_Comm\_rank(MPI\_COMM\_WORLD, &rank). Esto es fundamental para asignar roles diferentes a cada proceso según su rango.

En este ejemplo se implementa una comunicación punto a punto entre dos procesos: el proceso 0 y el proceso 1. El proceso 0 es el encargado de enviar un dato (en este caso, el número 42) al proceso 1, que se encargará de recibirlo y posteriormente imprimirlo.

Cuando el proceso 0 detecta que su rank es 0, define una variable dato con el valor 42 y realiza la operación de envío mediante MPI\_Send(&dato, 1, MPI\_INT, 1, 0, MPI\_COMM\_WORLD). Aquí se especifica:

* La dirección del dato a enviar (&dato).
* El número de elementos a enviar (1).
* El tipo de dato (MPI\_INT).
* El proceso destino (1).
* Un tag (0) que actúa como etiqueta para identificar el mensaje.
* El comunicador (MPI\_COMM\_WORLD).

En el proceso 1, al comprobar que su rank es 1, se declara la variable recibida para almacenar el dato entrante. Luego se llama a MPI\_Recv(&recibido, 1, MPI\_INT, 0, 0, MPI\_COMM\_WORLD, MPI\_STATUS\_IGNORE), la cual espera recibir el mensaje del proceso 0. Los parámetros indican:

* La dirección donde se almacenará el dato recibido.
* El número de elementos que se esperan.
* El tipo de dato (MPI\_INT).
* El proceso fuente (0), es decir, de dónde se espera el mensaje.
* El tag (0), que debe coincidir con el tag usado en el envío.
* El comunicador.

Y se ignora el estado del mensaje recibido.

Una vez que el proceso 1 recibe el mensaje, se imprime el dato utilizando printf

## ¿Qué sucede si cambiamos los tags?

Los tags (etiquetas) son fundamentales para emparejar mensajes entre el envío y la recepción. En el programa, tanto el proceso 0 al enviar como el proceso 1 al recibir utilizan el tag 0. Si se cambiara el tag en uno de ellos, por ejemplo, enviando con tag 1 pero intentando recibir con tag 0, el mensaje no se emparejaría correctamente. Esto provocaría que el proceso que llama a MPI\_Recv se quede bloqueado esperando un mensaje con el tag correcto, o se produzca un error en la comunicación.

## ¿Qué ocurre si el proceso 1 ejecuta MPI\_Recv antes de que el proceso 0 envíe el mensaje?

MPI utiliza la comunicación bloqueante en este ejemplo, lo que significa que si el proceso 1 ejecuta MPI\_Recv antes de que el proceso 0 envíe el mensaje, el proceso 1 se bloqueará y esperará hasta que el mensaje llegue. Este comportamiento es el esperado y forma parte de la sincronización de la comunicación punto a punto.

## ¿Qué sucede si cambiamos los tags?

Si el proceso 0 intenta enviar el dato a un proceso que no existe (por ejemplo, especificando un destino con un rank que no se haya iniciado), MPI detectaría un error. Este error puede causar que el programa termine de forma inesperada o que se comporte de manera indefinida, ya que se estaría intentando comunicar con un proceso que no forma parte del comunicador. Es esencial que el rango del proceso destino exista y esté incluido en el comunicador utilizado.

## ¿Cómo compilar y ejecutar con mpicc?

* **Compilación**

Para compilar con nuestro compilador mpicc lo hacemos igual que cualquier otro archivo .c,   
mpicc -o ejercicio1 ejercicio1.c.

Si queremos compilar con un compilador de c++ podemos usar el de mpic++,   
mpic++ -o ejercicio1 practica1.cpp

* **Ejecución**

Para la ejecución de los programas lo haremos de la siguiente forma, tanto si es de c como si es de c++ mpirun -np 3 ./ejercicio1, el número que sigue a -np es el número de procesos que se ejecuta en paralelo

## Código para la tarea propuesta

La tarea consiste en modificar el ejemplo para que el proceso 0 envíe un número al proceso 2, y este último lo multiplique por 3 antes de imprimirlo.

#include <mpi.h>

#include <stdio.h>

**int** main(**int** argc, **char**\*\* argv) {

MPI\_Init(&argc, &argv);

**int** rank;

MPI\_Comm\_rank(MPI\_COMM\_WORLD, &rank);

if (rank == 0) {

**int** dato = 42; *// Valor a enviar*

MPI\_Send(&dato, 1, MPI\_INT, 2, 0, MPI\_COMM\_WORLD);

} else if (rank == 2) {

**int** recibido;

*// Recibimos el dato enviado por el proceso 0*

MPI\_Recv(&recibido, 1, MPI\_INT, 0, 0, MPI\_COMM\_WORLD MPI\_STATUS\_IGNORE);

*// Multiplicamos el dato por 3*

**int** resultado = recibido \* 3;

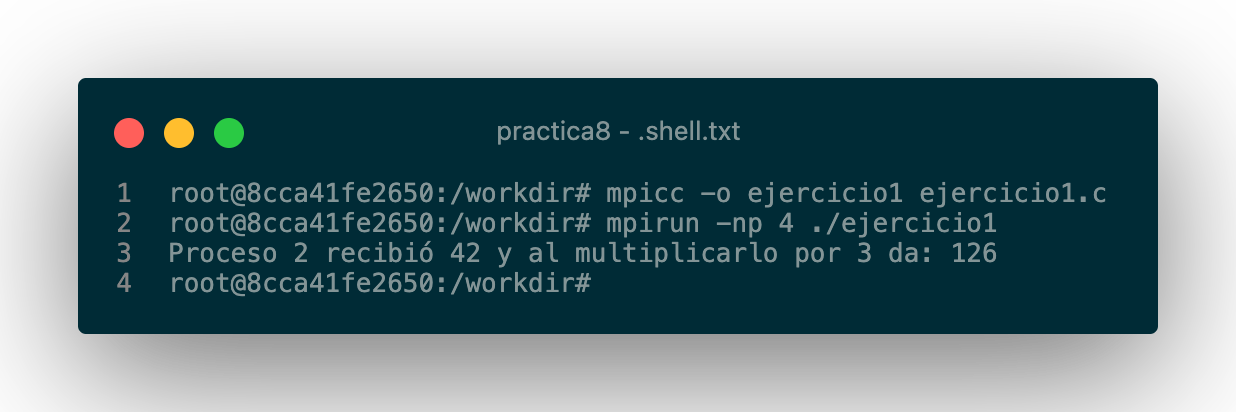
printf("Proceso 2 recibió %d y al multiplicarlo por 3 da: %d\n", recibido,   
 resultado);

}

MPI\_Finalize();

return 0;

}



## Salida

## ¿Cómo funciona?

**Proceso 0:** Envía un número (en este ejemplo, 42) al proceso 2 utilizando MPI\_Send.

**Proceso 2:** Recibe el mensaje con MPI\_Recv, multiplica el valor recibido por 3 y muestra el resultado por pantalla.

**MPI\_Send y MPI\_Recv**

Estas funciones permiten la comunicación bloqueante entre procesos. El proceso emisor envía datos a un proceso receptor especificando, entre otros parámetros, el identificador del proceso destino y un tag para identificar el mensaje. El receptor, a su vez, llama a MPI\_Recv esperando recibir el mensaje con el tag correspondiente.

**Manejo de Procesos**

La estructura condicional (if y else if) en el código permite asignar tareas específicas según el rango del proceso. Esto garantiza que solo el proceso 0 envíe datos y solo el proceso 2 los reciba y procese.

## Ejercicio 2 - Comunicación en cadena

El programa propuesto utiliza la comunicación en cadena para que el dato viaje de un proceso al siguiente hasta alcanzar el último.

## ¿Cómo funciona?

* **Proceso 0**:

Inicia con un dato = 0.

Envía este dato al proceso 1 usando MPI\_Send.

* **Procesos intermedios (1 hasta size-2):**

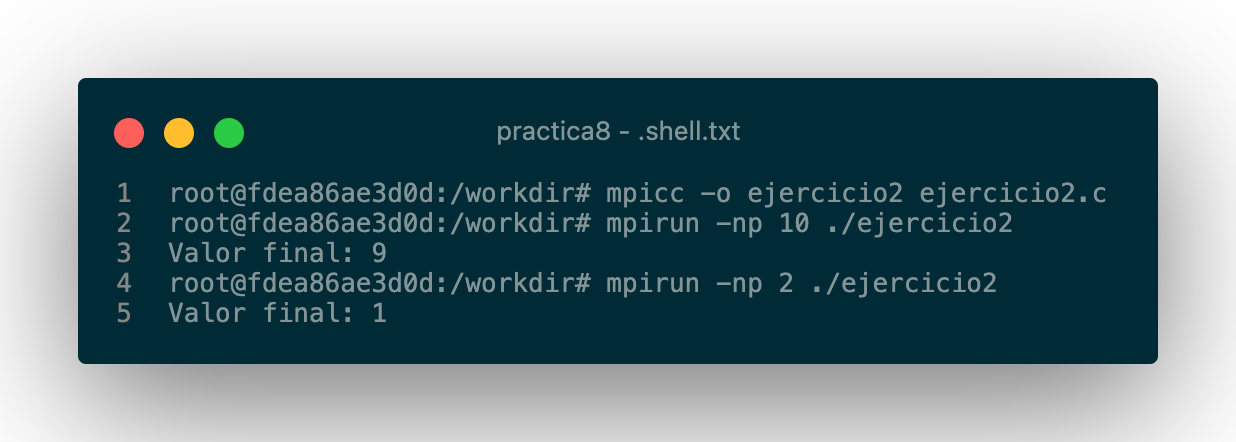
Cada proceso recibe el dato del proceso anterior mediante MPI\_Recv.

Se incrementa el valor recibido (con recibido++).

Envía el nuevo valor al siguiente proceso con MPI\_Send.

* **Proceso final (rank == size - 1):**

Recibe el dato, lo incrementa y, al no tener a quién enviarlo (ya que es el último), imprime el valor final.



## ¿Qué sucede si un proceso omite el envío?

Si algún proceso no realiza el envío (por ejemplo, si no llama a MPI\_Send después de incrementar el dato), la cadena se rompe. Los procesos posteriores se quedarán bloqueados en MPI\_Recv, ya que nunca recibirán el mensaje esperado.

**¿Qué ocurre si ejecutamos el programa con solo un proceso?**

Si se ejecuta el programa con un solo proceso, éste tendrá rank 0 y se ejecutará únicamente el bloque de código correspondiente al proceso 0, que intenta enviar el dato al proceso 1. Como no existe el proceso 1, se generará un error o un comportamiento indefinido. Por ello, se requiere que el programa se ejecute con al menos dos procesos para que la comunicación en cadena tenga sentido.

**¿Cómo podríamos hacer que el proceso 0 también imprima el dato inicial?**

Para que el proceso 0 imprima el dato inicial, se puede incluir una instrucción de impresión (printf) antes de enviar el dato. Esto permitiría ver el valor original que se está transmitiendo en la cadena.

## Código para la tarea propuesta

La tarea consiste en modificar el programa para que:

* El dato inicial sea 10 (en lugar de 0).
* Cada proceso incremente el dato en 2 (en lugar de 1) antes de enviarlo al siguiente.

#include <mpi.h>

#include <stdio.h>

**int** main(**int** argc, **char**\*\* argv) {

MPI\_Init(&argc, &argv);

**int** rank, size;

MPI\_Comm\_rank(MPI\_COMM\_WORLD, &rank);

MPI\_Comm\_size(MPI\_COMM\_WORLD, &size);

if (rank == 0) {

**int** dato = 10;

MPI\_Send(&dato, 1, MPI\_INT, 1, 0, MPI\_COMM\_WORLD);

} else {

**int** recibido;

MPI\_Recv(&recibido, 1, MPI\_INT, rank - 1, 0, MPI\_COMM\_WORLD,

MPI\_STATUS\_IGNORE);

recibido += 2;

if (rank < size - 1) {

MPI\_Send(&recibido, 1, MPI\_INT, rank + 1, 0, MPI\_COMM\_WORLD);

} else {

printf("Valor final: %d\n", recibido);

}

}

MPI\_Finalize();

return 0;

}

## 

## Salida

## Ejercicio 3 - Comunicación colectiva

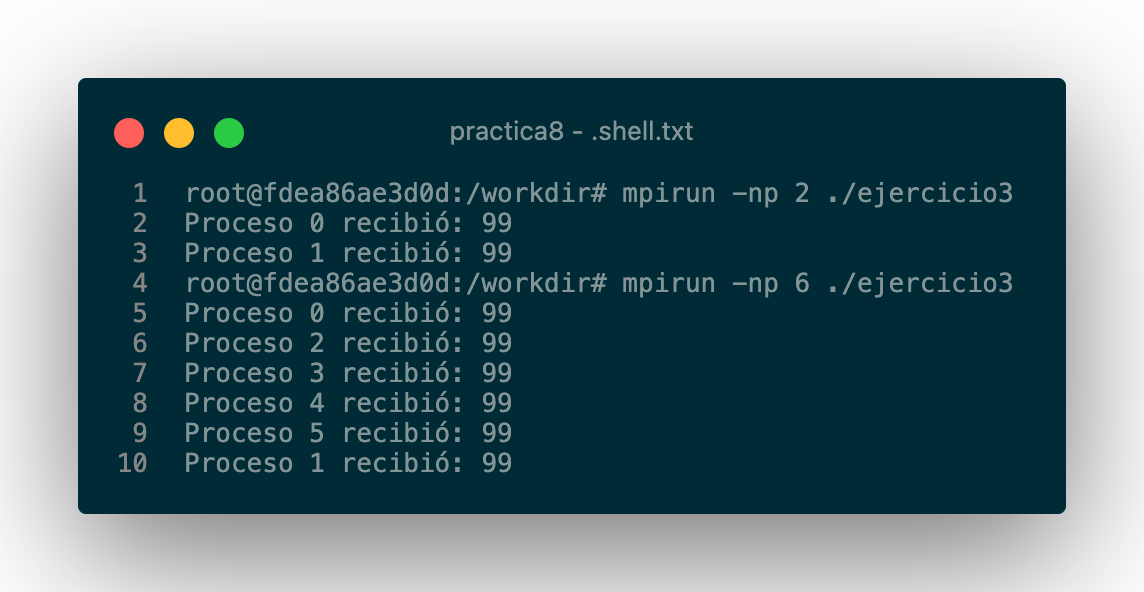
En el ejercicio se usa MPI\_Bcast para difundir un dato desde un proceso raíz a todos los procesos que participan en el comunicador.

## ¿Cómo funciona?

* El proceso con rank 0 inicializa la variable dato con el valor 99.
* Se llama a MPI\_Bcast(&dato, 1, MPI\_INT, 0, MPI\_COMM\_WORLD), lo que permite que todos los procesos, sin importar su rank, reciban el valor 99.

Cada proceso imprime el dato recibido.

La ventaja de usar MPI\_Bcast es que se simplifica la difusión del dato a todos los procesos sin necesidad de múltiples envíos individuales.



## ¿Qué pasa si el root es diferente?

Si se cambia el proceso root (el que inicializa el dato) por otro (por ejemplo, no el 0), el proceso designado es el que debe inicializar el dato. Todos los procesos recibirán el valor desde ese proceso. Es importante que el proceso root tenga el dato inicial correcto, ya que es el que se transmite a los demás.

## ¿Qué problema evitaríamos usando MPI\_Bcast en lugar de MPI\_Send?

Usar MPI\_Bcast evita tener que realizar múltiples llamadas a MPI\_Send para enviar el dato a cada proceso individualmente. Con MPI\_Bcast se garantiza que todos los procesos reciben el mismo dato de manera sincronizada y se simplifica el código, además de reducir el riesgo de errores en la comunicación.

## ¿Qué ocurre si un proceso no participa en la operación colectiva?

En una operación colectiva, todos los procesos que pertenecen al comunicador deben participar. Si un proceso no participa en la llamada a MPI\_Bcast, se pueden generar bloqueos o comportamientos indefinidos, ya que la operación espera la participación de todos los procesos.

## Código para la tarea propuesta

La tarea consiste en modificar el programa para que:

* El proceso 2 sea el root de la difusión (inicializa el dato).
* El proceso 2 envíe un número a todos los procesos mediante MPI\_Bcast.
* Una vez recibido el dato, cada proceso lo incrementa en 1 antes de imprimirlo.

#include <mpi.h>

#include <stdio.h>

**int** main(**int** argc, **char**\*\* argv) {

MPI\_Init(&argc, &argv);

**int** rank;

MPI\_Comm\_rank(MPI\_COMM\_WORLD, &rank);

**int** dato;

if (rank == 2) dato = 99; *// El root inicializa el dato*

MPI\_Bcast(&dato, 1, MPI\_INT, 2, MPI\_COMM\_WORLD); *// Difusión a todos*

dato += 1;

printf("Proceso %d recibió: %d\n", rank, dato);

MPI\_Finalize();

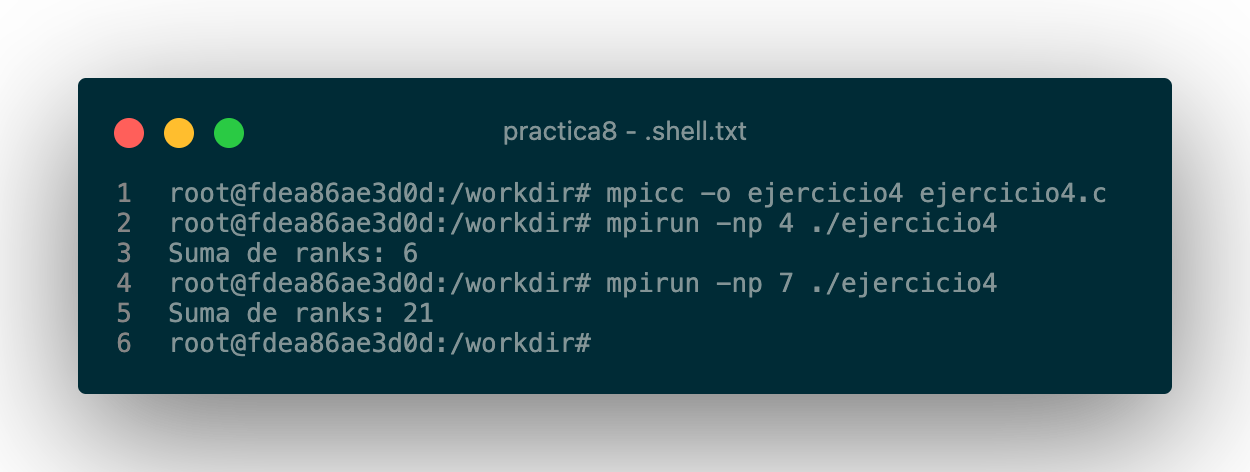
return 0;

}

## Salida

## Ejercicio 4 - Reducción de datos

El código propuesto suma los rangos (identificadores) de todos los procesos. Para ello, cada proceso conoce su propio rank y se utiliza la función MPI\_Reduce con el operador MPI\_SUM para acumular estos valores en el proceso root (en este caso, el proceso 0).



**Corroboración de los datos:**

* procesos (0,1,2,3) suma = 6
* procesos (0,1,2,3,4,5,6,) = 21

## ¿Qué diferencia hay entre MPI\_Gather y MPI\_Reduce?

Mientras que MPI\_Gather recopila datos de todos los procesos y los almacena en un único arreglo en el proceso root (sin aplicar operaciones aritméticas o lógicas), MPI\_Reduce combina los datos de todos los procesos aplicando una operación (por ejemplo, suma, máximo, mínimo, etc.) para obtener un único resultado en el proceso root. En nuestro ejemplo, se usa MPI\_Reduce para sumar los ranks.

## ¿Qué pasaría si omitimos el proceso root en MPI\_Reduce?

Si se omite el proceso root, o se especifica un valor incorrecto, el resultado combinado no se almacenará correctamente en un proceso en particular, lo que puede ocasionar que el proceso no tenga acceso al resultado final. Además, la llamada a MPI\_Reduce requiere que se especifique un proceso root válido para que la operación se realice de forma correcta.

## ¿Qué otras operaciones podrías realizar en lugar de suma?

Además de la suma (MPI\_SUM), MPI\_Reduce permite realizar otras operaciones, tales como el máximo (MPI\_MAX), el mínimo (MPI\_MIN), el producto (MPI\_PROD), la operación lógica AND (MPI\_LAND), entre otras. En el ejercicio que nos ocupa, vamos a utilizar MPI\_MAX para calcular el máximo de los ranks.

## Código para la tarea propuesta

La tarea consiste en modificar el programa para que utilizando MPI\_Reduce con el operador MPI\_MAX, se calcule el máximo de los ranks y, además, se haga visible este resultado en todos los procesos. Para lograr que todos los procesos conozcan el valor máximo obtenido, se puede realizar una reducción en el proceso root y luego utilizar MPI\_Bcast para difundir el resultado a todos los procesos.