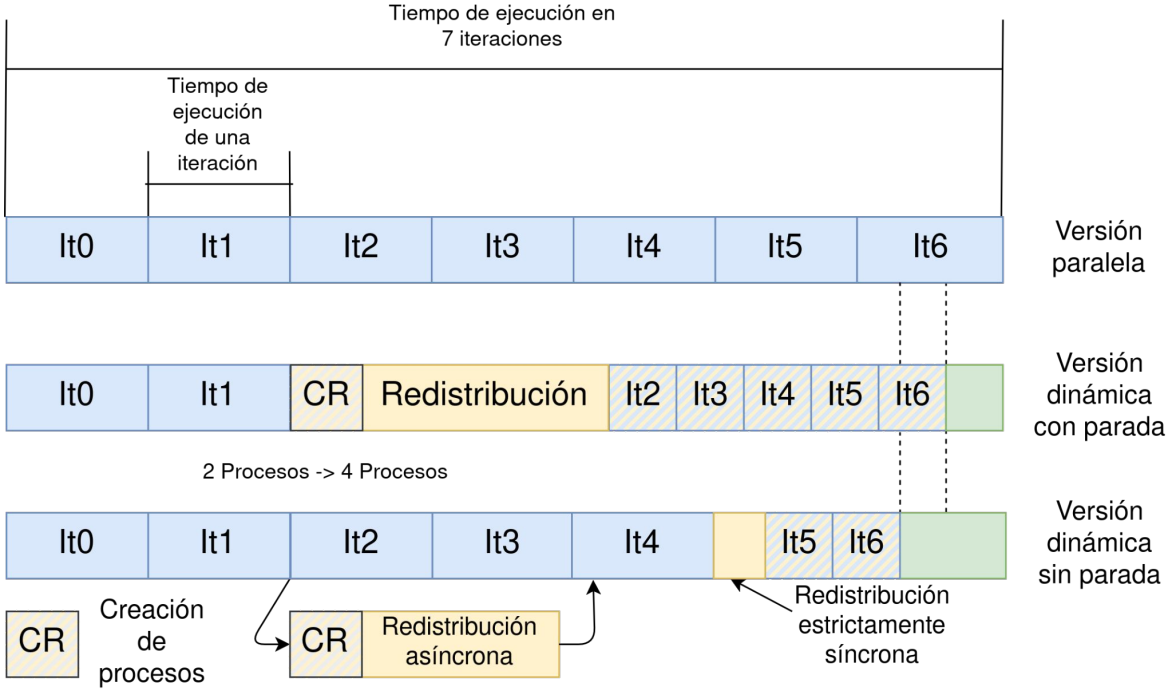
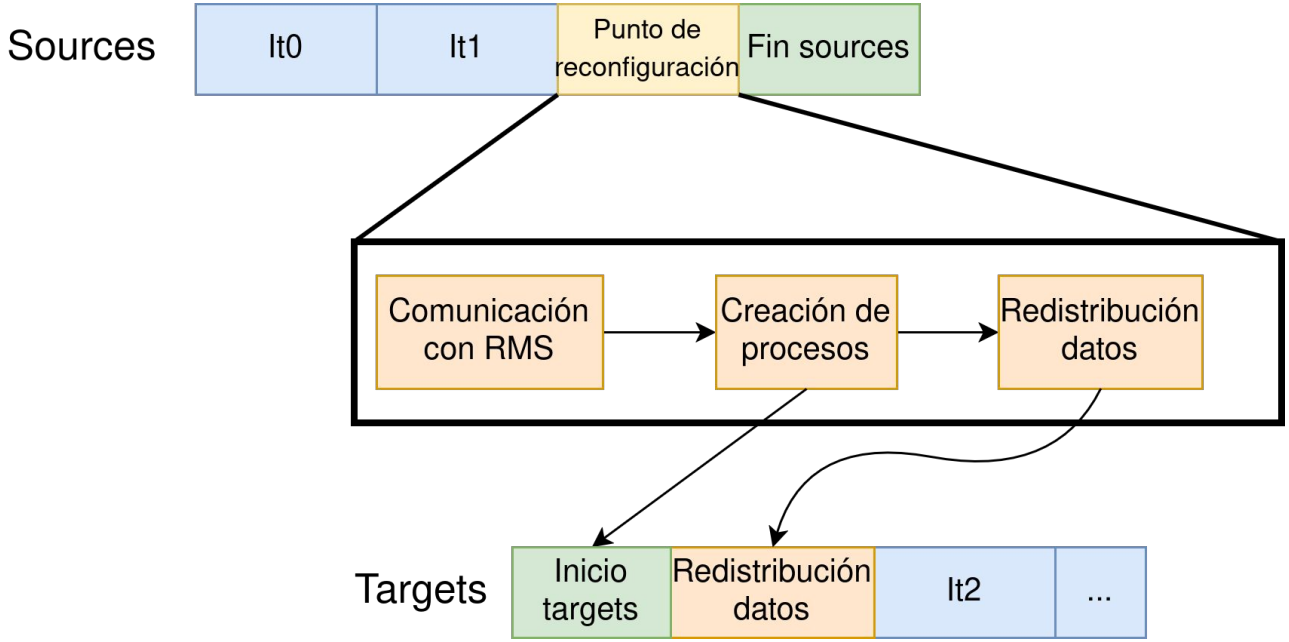
Redimensionamiento Dinámico de Aplicaciones Maleables mediante RMA

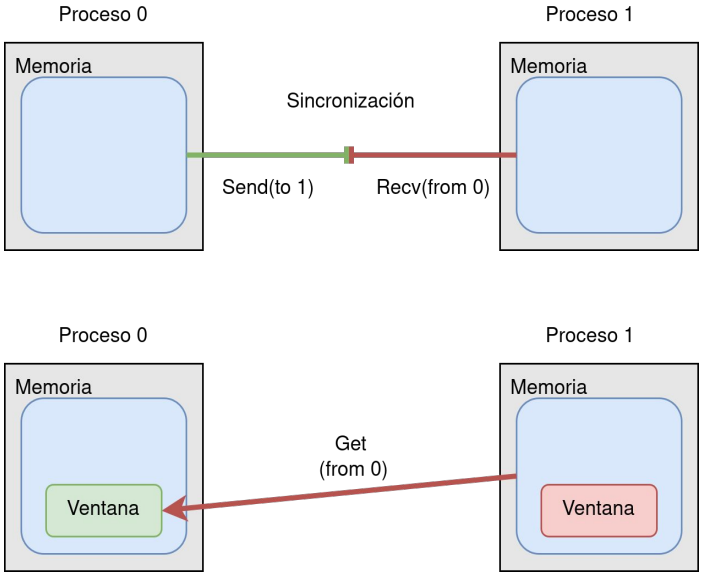
Authors

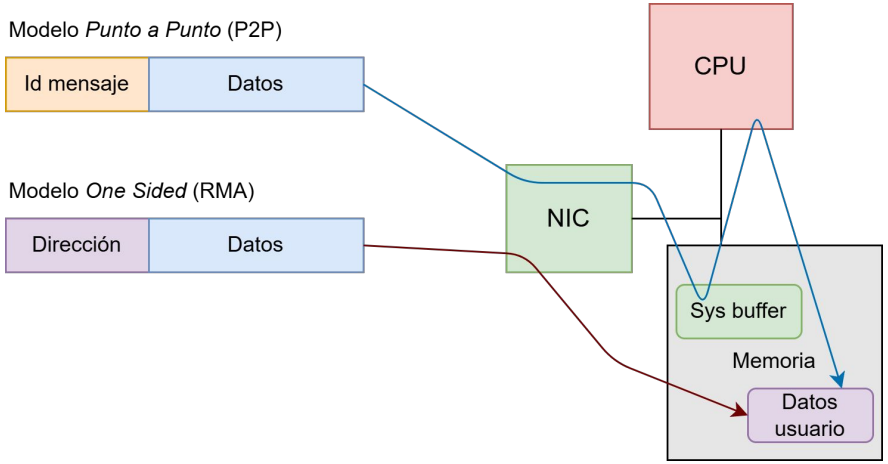
Iker Martín-Álvarez, José I. Aliaga, Maribel Castillo,


Tiempo de ejecución en
7 iteraciones
Tiempo de
ejecución
de una
iteración
It0
It1
It2
It3
It4
It5
It6
Versión
paralela
It0
It1
CR
Redistribución
It2
It3
It4
It5
It6
Versión
dinámica
con parada
2 Procesos -> 4 Procesos
It0
It1
It2
It3
It4
It5
It6
Versión
dinámica
sin parada
CR
Creación
de
procesos
CR
Redistribución
asíncrona
Redistribución
estrictamente
síncrona


Sources
It0
It1
Punto de reconfiguración
Fin sources
Comunicación con RMS
Creación de procesos
Redistribución datos
Targets
Inicio targets
Redistribución datos
It2
...

Uso del modelo de MPI *One-Sided* (RMA)

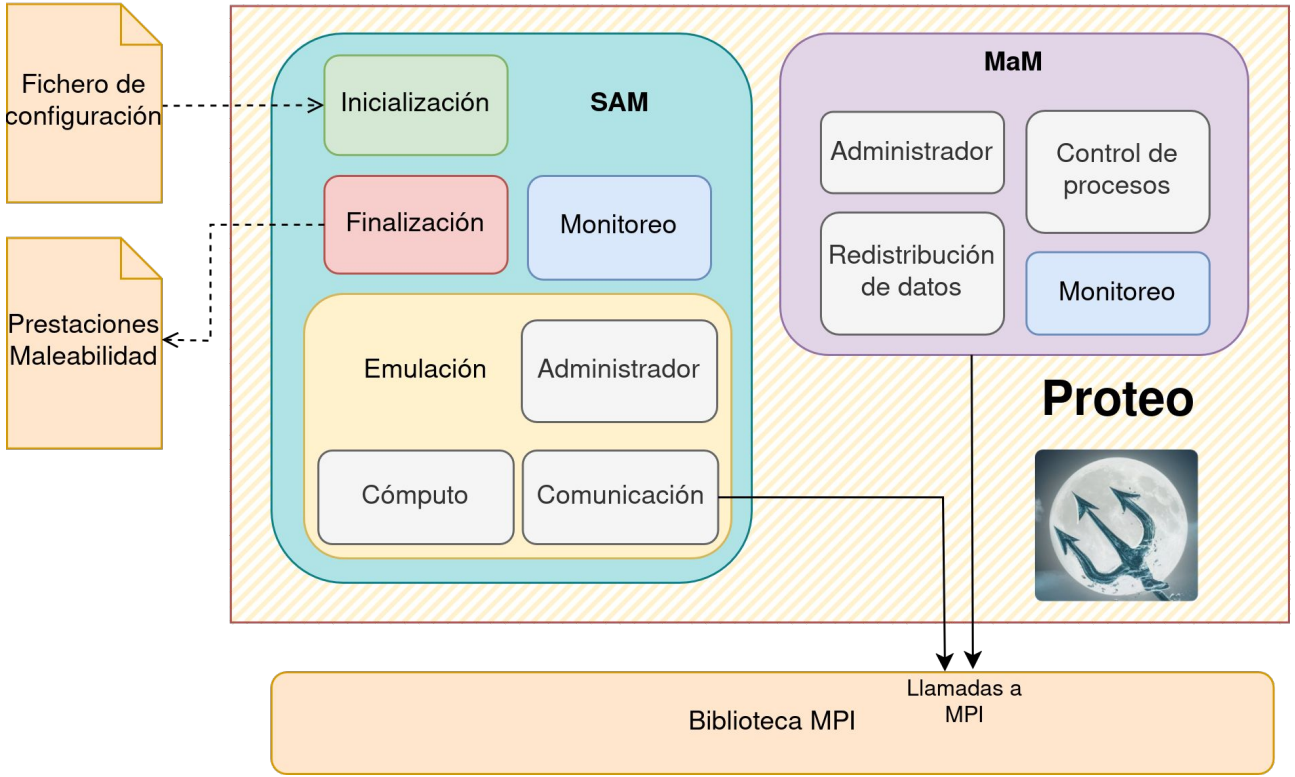




Índice:

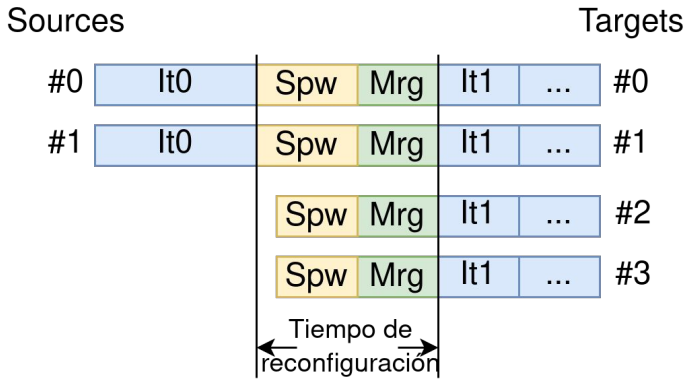
Resumen Proteo

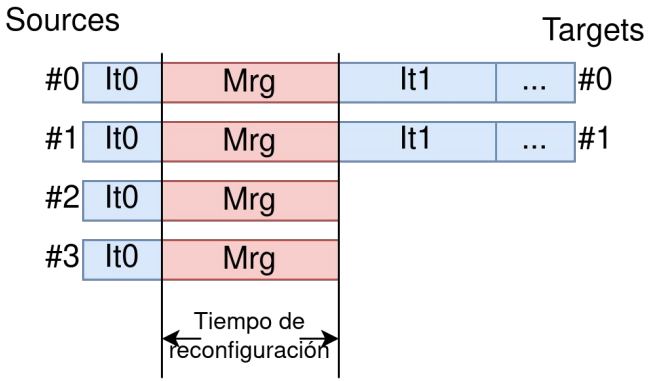
- Framework para emular aplicaciones reales.
- Compuesto por dos módulos:
  - SAM: Realiza la emulación.
  - MaM: Se encarga de la reconfiguración.
- Requiere un archivo de configuración para emular una aplicación real.



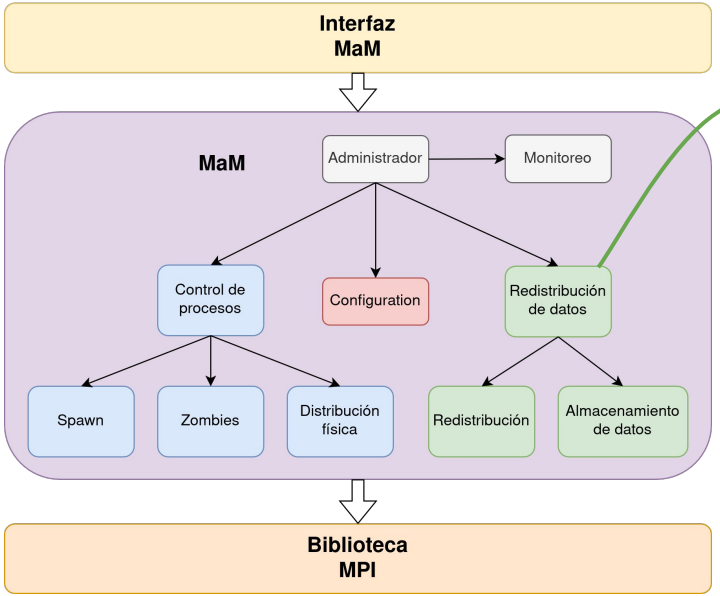
Repositorio Proteo

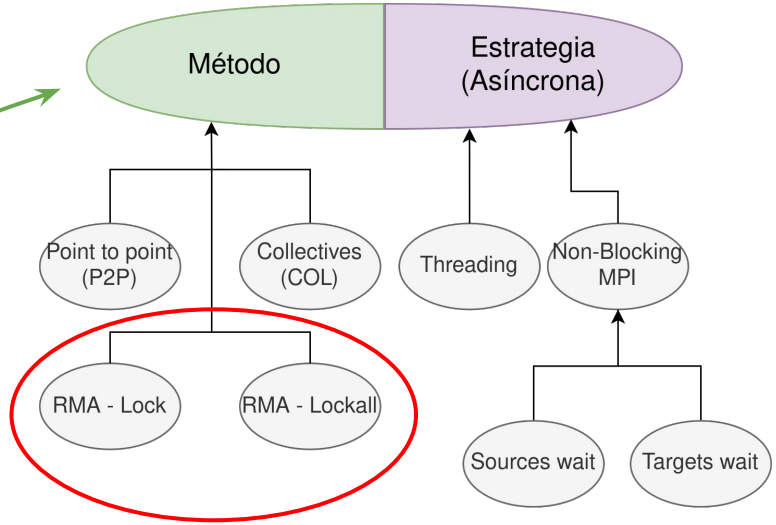
Creación de procesos - Método Merge





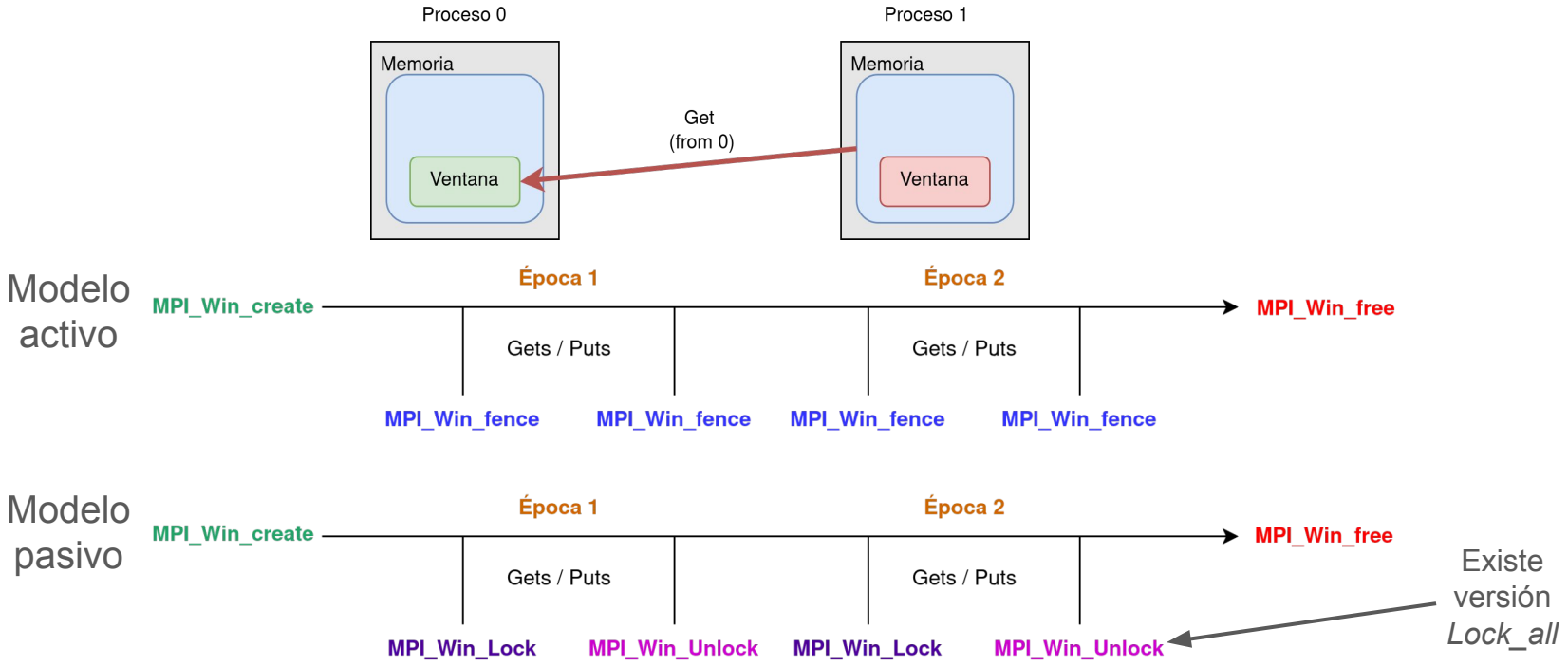
Métodos de redistribución de datos



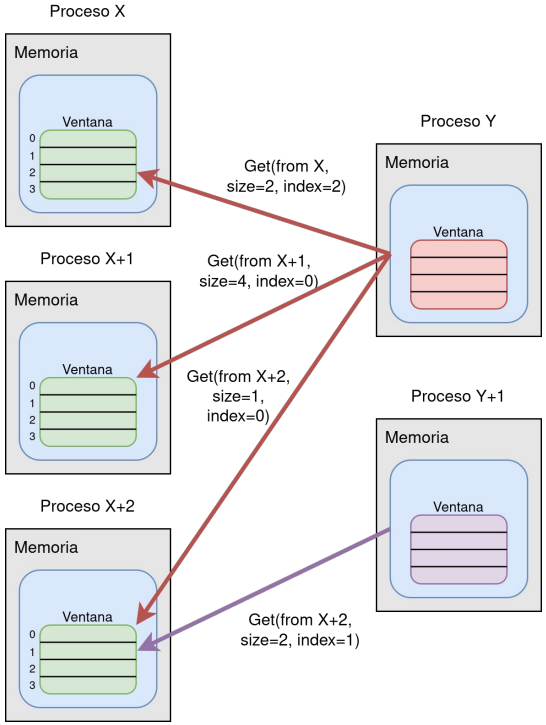


Índice:

Bases


Proceso 0
Proceso 1
Memoria
Memoria
Ventana
Ventana
Get
(from 0)
Modelo activo
MPI_Win_create
Época 1
Época 2
MPI_Win_free
Gets / Puts
Gets / Puts
MPI_Win_fence
MPI_Win_fence
MPI_Win_fence
MPI_Win_fence
Modelo pasivo
MPI_Win_create
Época 1
Época 2
MPI_Win_free
Gets / Puts
Gets / Puts
MPI_Win_Lock
MPI_Win_Unlock
MPI_Win_Lock
MPI_Win_Unlock
Existe versión Lock_all

Métodos propuestos

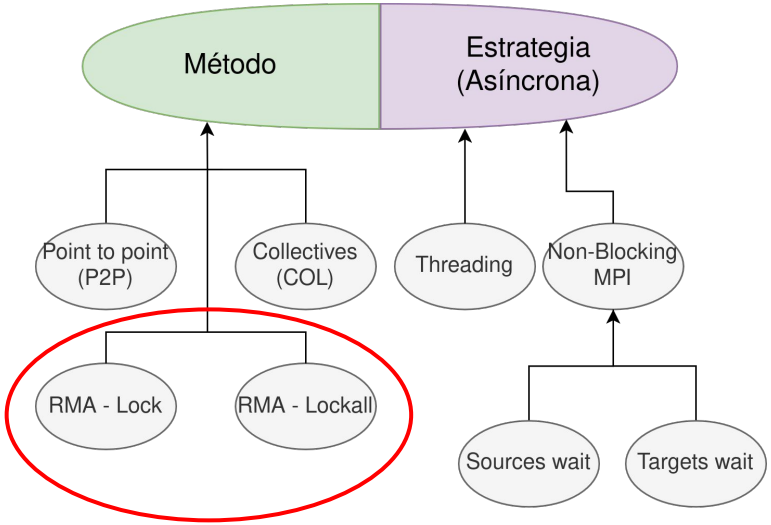


Método 1:

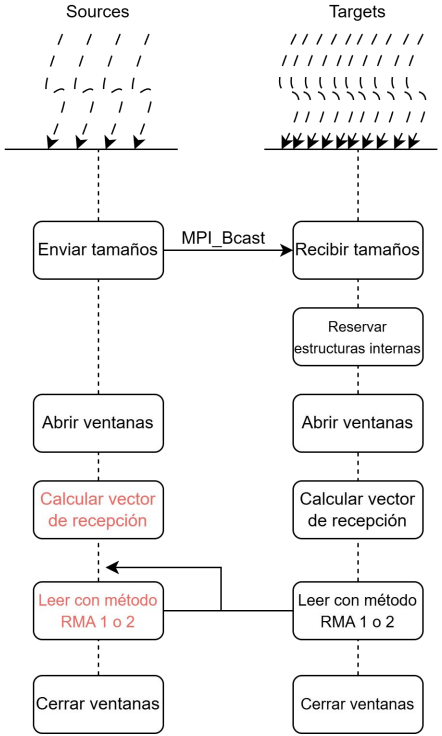
1. Bucle Lock(x) + Get(x)
2. Bucle Unlocks

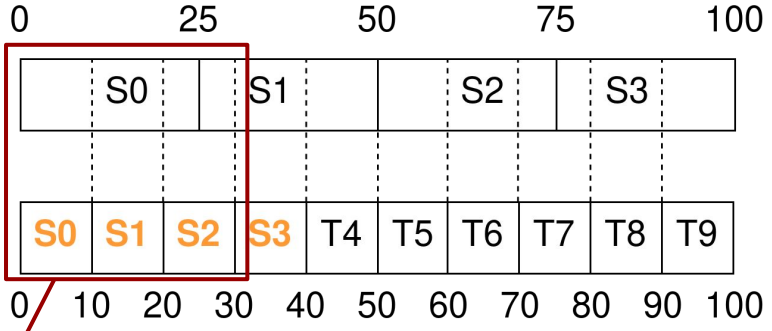
Método 2:

1. Lock_all
2. Bucle Gets
3. Unlock_all

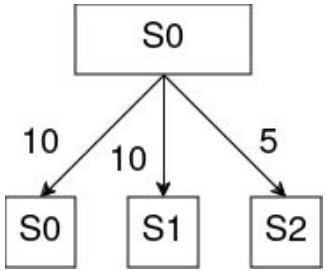


Merge - Pasos





Distribución de 100 elementos de 4 *sources* a 10 *targets.*



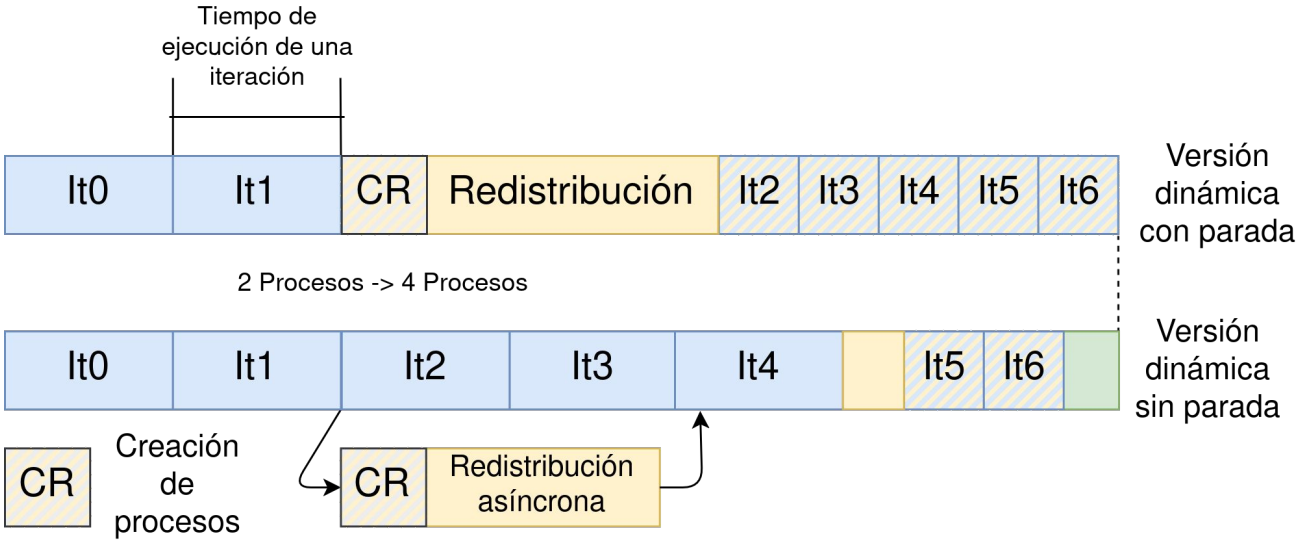
Distribución de elementos desde el *source* 0 a los *targets* 0, 1 y 2.

Índice:

Ventaja asíncrona

3 Estrategias para redistribuir datos:

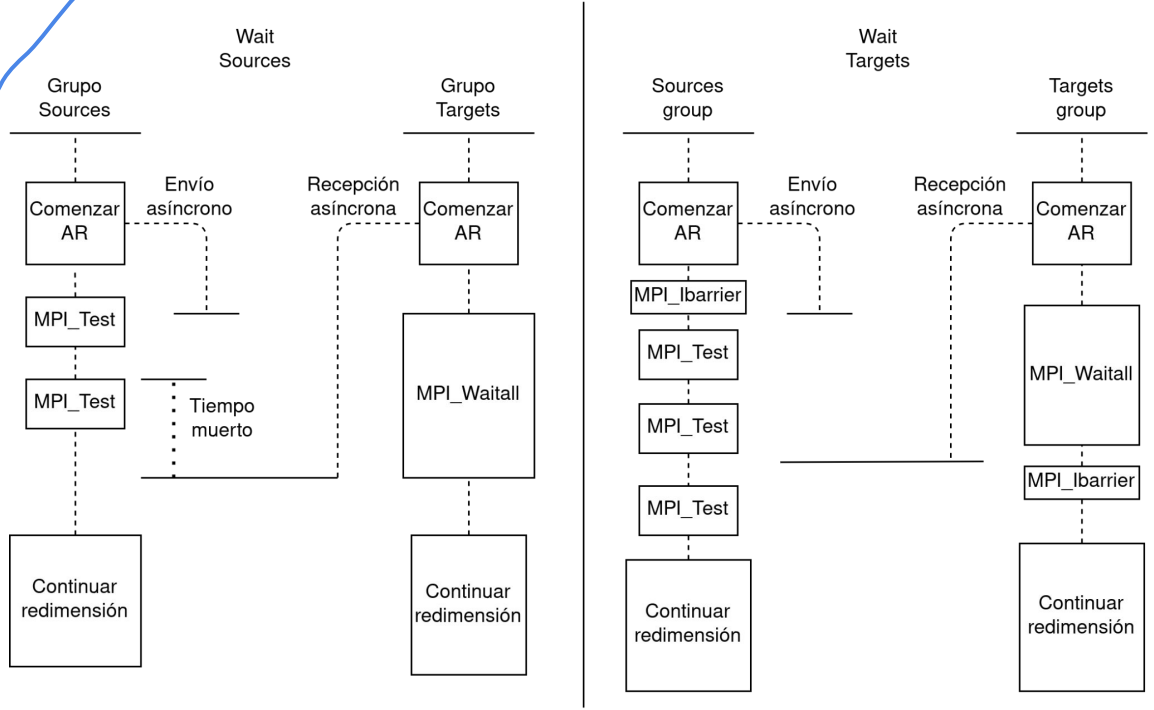
- Non-Blocking
  - Sources Wait
  - Targets Wait
- Threading



Estrategias asíncronas

3 Estrategias para redistribuir datos:

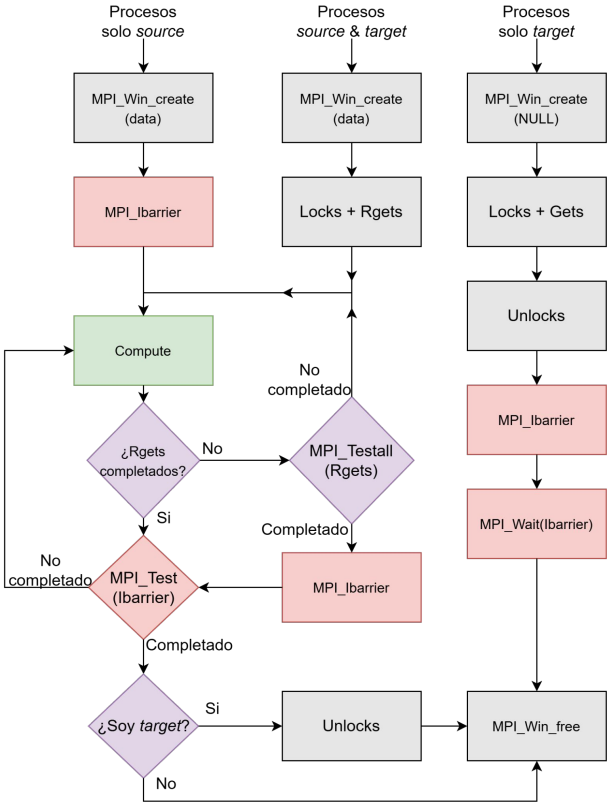
- **Non-Blocking - Sources Wait**
- **Non-Blocking - Targets Wait**
- Threading



Flujo para *One-Sided*

Pasos para modelo no bloqueante:

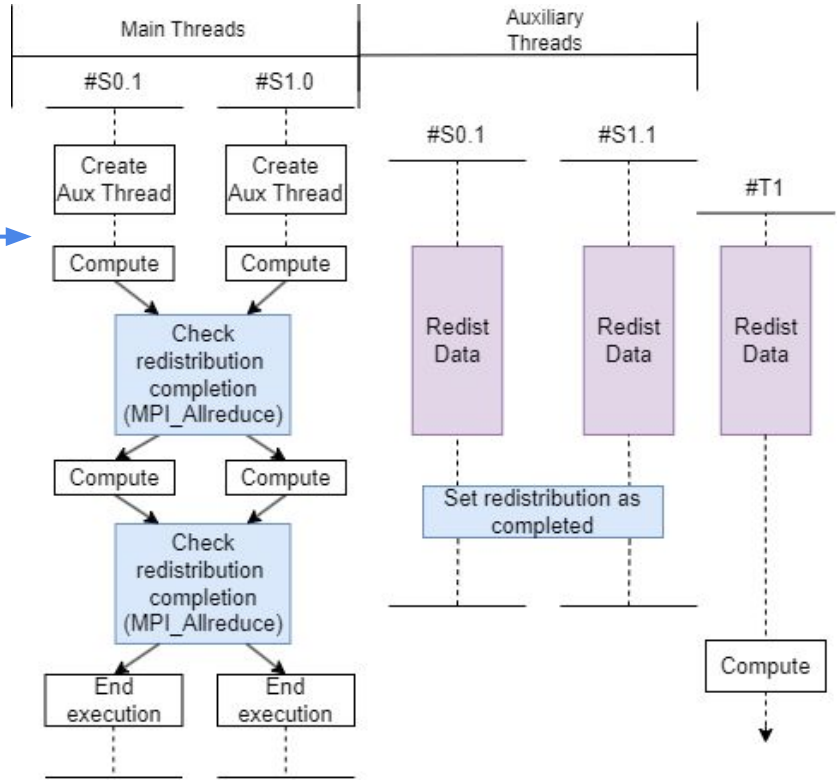
- Inicialización
- Comenzar lecturas
- Comprobar lecturas locales
- Comprobar lecturas globales
- Finalizar



Estrategias asíncronas

3 Estrategias para redistribuir datos:

- Non-Blocking - Sources Wait
- Non-Blocking - Targets Wait
- **Threading**



Índice:

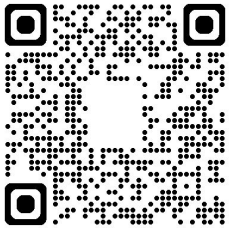
Descripción de los experimentos

Descripción de máquina:

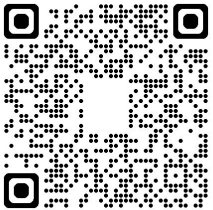
- Nodos cálculo (x8): 2x Intel Xeon 4210 (10-core)
- Total de 160 *cores*
- *Red Infiniband EDR 100GB/s*
- MPICH 4.2.0 CH4:OFI - fi_provider=verbs

Descripción de la emulación:

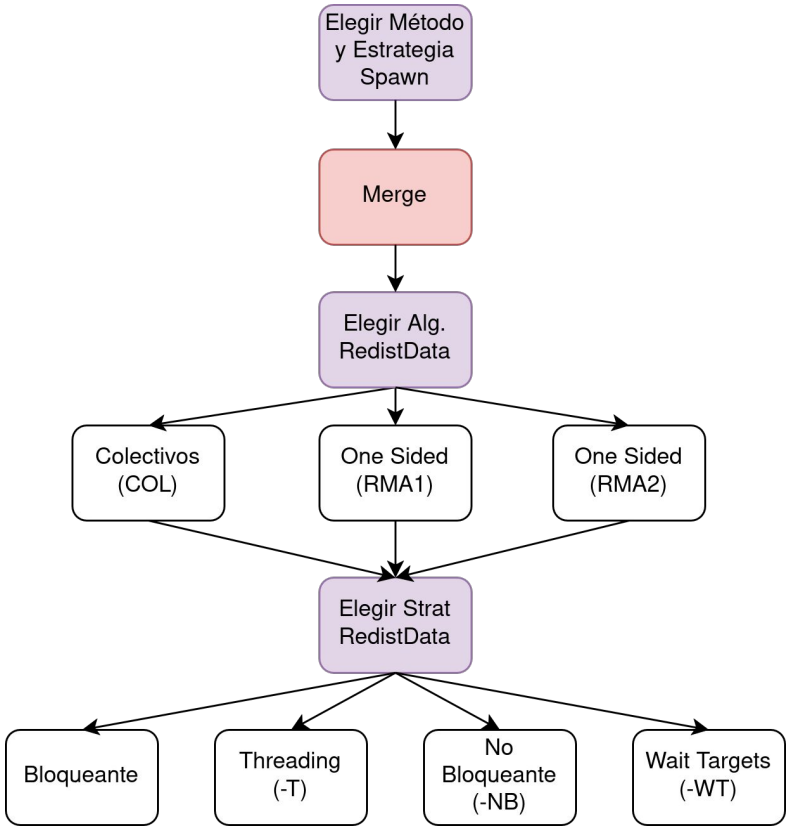
- Proteo emula la aplicación CG
- Una reconfiguración por ejecución.
- Redistribución de 64GB.
- Mediana de 20 ejecuciones.



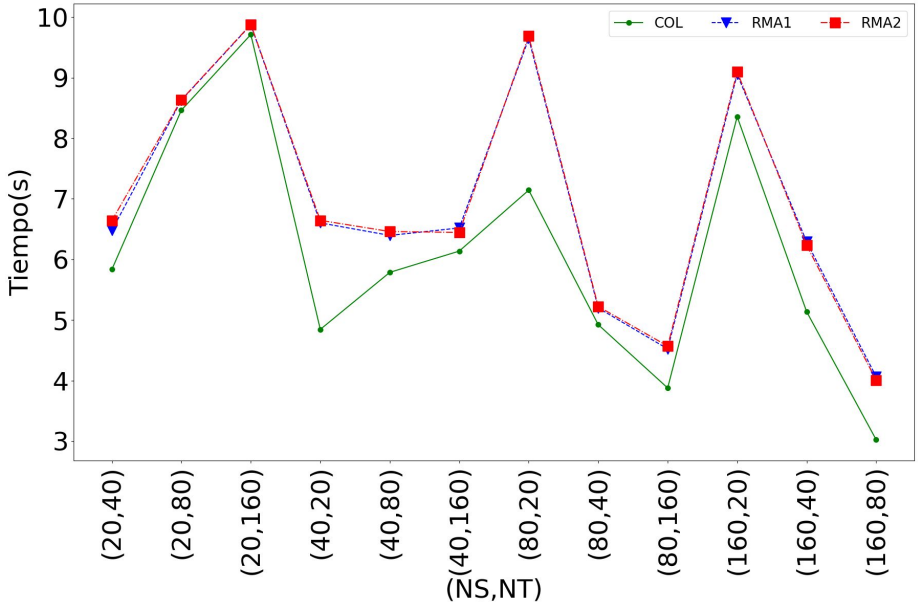
Dataset generado

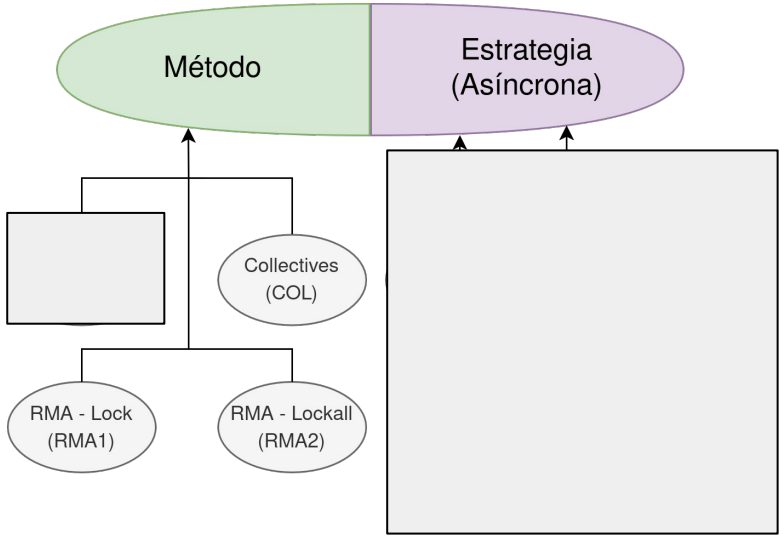


Repositorio Proteo

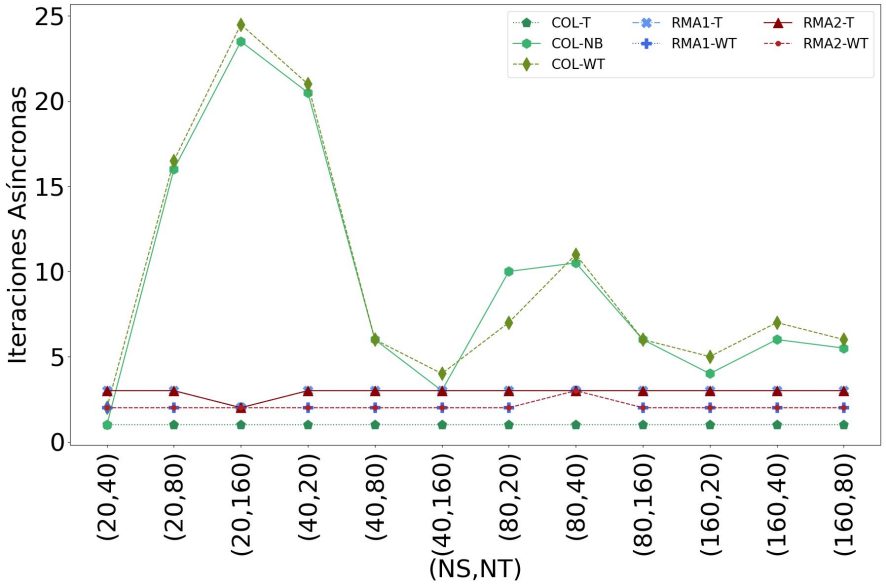


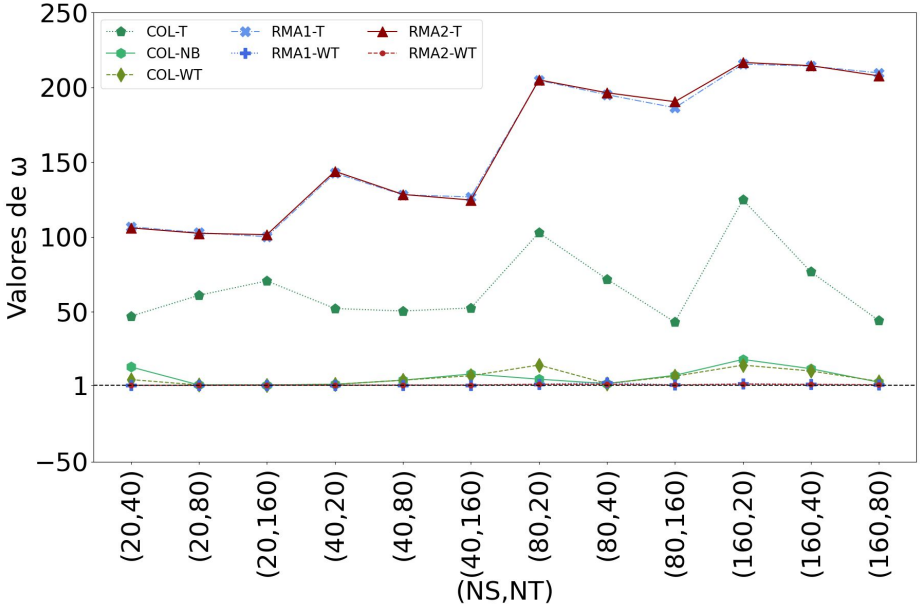
Pruebas síncronas - Tiempos de redistribución de datos



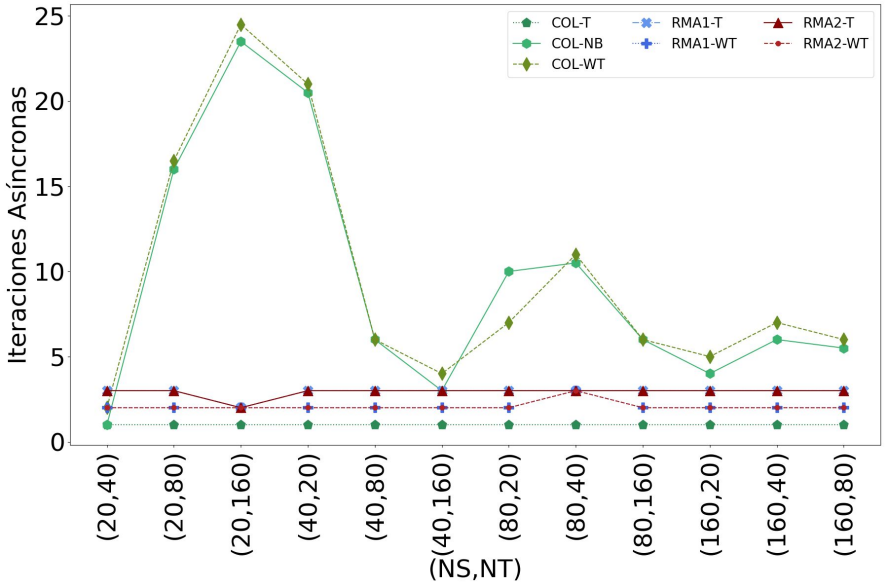


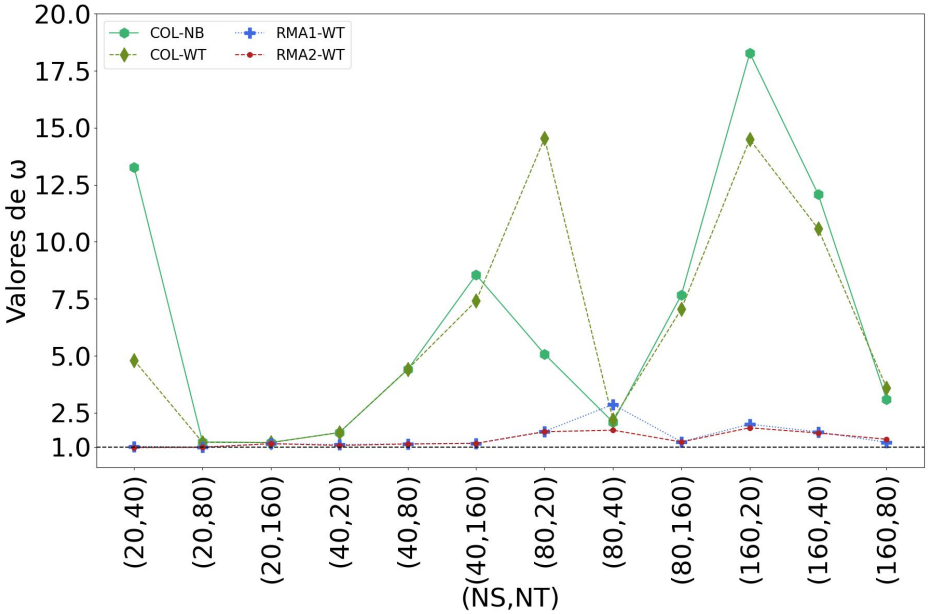
Pruebas asíncronas - Parámetros internos





Pruebas asíncronas - Parámetros internos

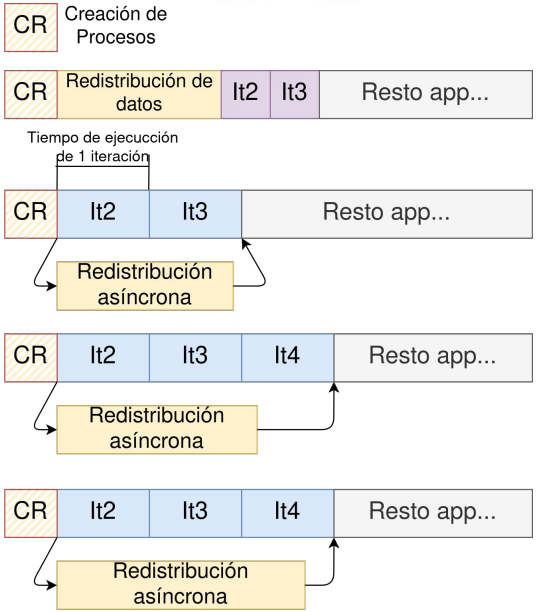


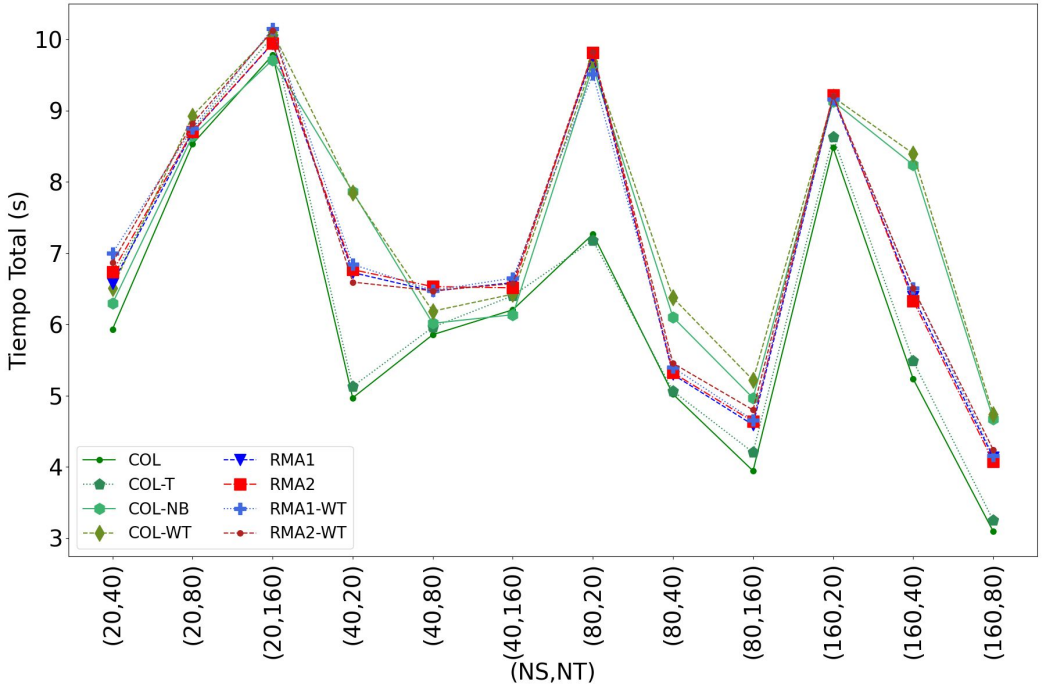


Pruebas asíncronas - Tiempos redistribución de datos

$$T_{total}^{Bl} = T_{redis}^{Bl} + T_{it}^{NT} * \min_{variantes}(N_{it}^{NS \to NT})$$

$$T_{total}^{SP} = T_{redis}^{SP}$$





Índice:

1. Objetivo
2. Proteo
3. Implementación RMA
4. Implementación asíncrona
5. Resultados
6. Conclusiones

MaM se ha ampliado con las operaciones *One-Sided* de MPI para redistribuir datos

El modelo *One-Sided* asíncrono para redistribuir no impacta en las prestaciones de la aplicación

El mismo no obtiene mejores prestaciones que operaciones colectivas debido a la inicialización de las ventanas

Redimensionamiento Dinámico de Aplicaciones Maleables mediante RMA

Authors

Iker Martín-Álvarez (martini@uji.es),

José I. Aliaga, Maribel Castillo,