

# PROGRAMACIÓN PARALELA Y CONCURRENTE

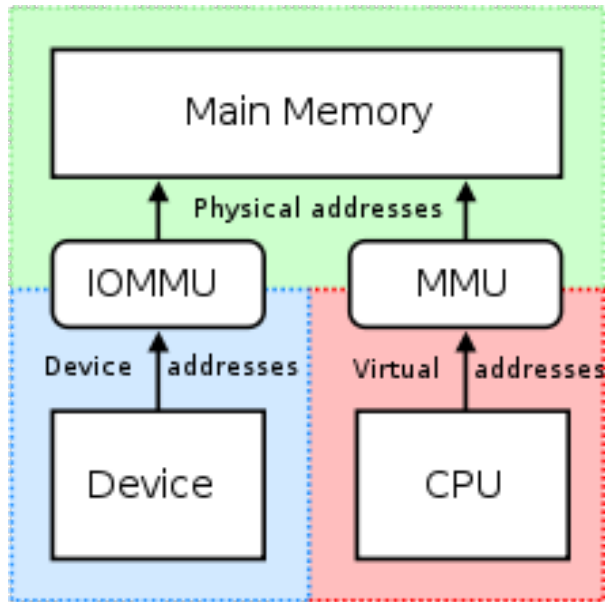
ICC-303





# ARQUITECTURA DE MEMORIA COMPARTIDA / DISTRIBUIDA

## ARQUITECTURA DE MEMORIA COMPARTIDA / DISTRIBUIDA

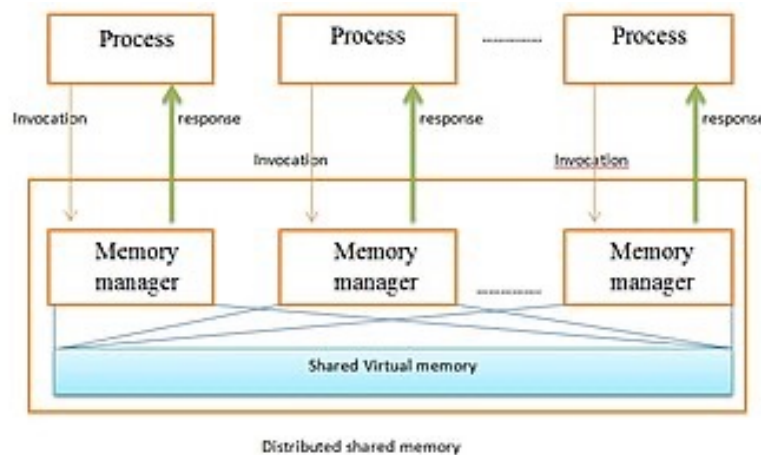


- **Arquitectura de memoria compartida**
  - Una arquitectura de memoria compartida es un sistema en el que todos los procesadores o nodos tienen acceso a un grupo de memoria común. Esto significa que pueden leer y escribir datos en las mismas ubicaciones de memoria y compartir información a través de variables compartidas. La ventaja de esta arquitectura es que simplifica el modelo de programación, ya que los programadores no necesitan gestionar explícitamente la transferencia de datos o la sincronización entre los procesadores. La desventaja es que puede causar problemas de contención y escalabilidad, ya que los procesadores compiten por el acceso a la memoria y el ancho de banda, y el tamaño de la memoria limita la cantidad de procesadores.

# ARQUITECTURA DE MEMORIA COMPARTIDA / DISTRIBUIDA

- **Arquitectura de memoria distribuida**

- Una arquitectura de memoria distribuida es un sistema en el que cada procesador o nodo tiene su propia memoria local y se comunican entre sí mediante el paso de mensajes. Esto significa que solo pueden acceder a sus propias ubicaciones de memoria e intercambiar información mediante el envío y la recepción de mensajes. La ventaja de esta arquitectura es que evita los problemas de contención y escalabilidad de la arquitectura de memoria compartida, ya que los procesadores no interfieren entre sí con el acceso a la memoria y el ancho de banda, y el tamaño de la memoria no limita la cantidad de procesadores. La desventaja es que complica el modelo de programación, ya que los programadores necesitan gestionar explícitamente la distribución de datos y la comunicación entre los procesadores.



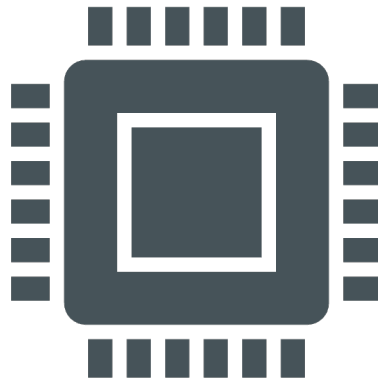
# ARQUITECTURA DE MEMORIA COMPARTIDA / DISTRIBUIDA

- En informática, la memoria compartida distribuida ( DSM ) es una forma de arquitectura de memoria en la que las memorias físicamente separadas se pueden abordar como un único espacio de direcciones compartido . El término "compartido" no significa que haya una única memoria centralizada, sino que el espacio de direcciones es compartido; es decir, la misma dirección física en dos procesadores se refiere a la misma ubicación en la memoria.
- Espacio de direcciones global distribuido ( DGAS ), es un término similar para una amplia clase de implementaciones de software y hardware, en las que cada nodo de un clúster tiene acceso a la memoria compartida además de la memoria privada de cada nodo (es decir, no compartida) memoria.



# ARQUITECTURA DE MEMORIA COMPARTIDA / DISTRIBUIDA

- DSM se puede lograr tanto a través de software como de hardware. Los ejemplos de hardware incluyen circuitos de coherencia de caché y controladores de interfaz de red . Hay tres formas de implementar DSM:
  - Enfoque basado en páginas que utiliza memoria virtual
  - Enfoque de variable compartida que utiliza rutinas para acceder a variables compartidas
  - Enfoque basado en objetos, idealmente accediendo a datos compartidos a través de una disciplina orientada a objetos.



# ARQUITECTURA DE MEMORIA COMPARTIDA / DISTRIBUIDA

## Ventajas

- Se escala bien con una gran cantidad de nodos
- El paso de mensajes está oculto
- Puede manejar bases de datos grandes y complejas sin replicación ni envío de datos a procesos.
- Generalmente más barato que usar un sistema multiprocesador
- Proporciona un gran espacio de memoria virtual
- Los programas son más portátiles debido a las interfaces de programación comunes.
- Proteger a los programadores del envío o recepción de primitivas

## Desventajas

- Generalmente es más lento de acceder que la memoria compartida no distribuida
- Debe proporcionar protección adicional contra accesos simultáneos a datos compartidos.
- Puede incurrir en una penalización de rendimiento.
- Poco control del programador sobre los mensajes reales que se generan.
- Los programadores necesitan comprender los modelos de coherencia para escribir programas correctos.

## ARQUITECTURA DE MEMORIA COMPARTIDA / DISTRIBUIDA

- **Comparación con el paso de mensajes**
- El paso de mensajes es una técnica que se usa para, desde un proceso, invocar de forma abstracta un comportamiento concreto por parte de otro actor (Por ejemplo, ejecutar una función o un programa)

### Paso de mensajes

Las variables deben ordenarse

El costo de la comunicación es obvio

Los procesos están protegidos al tener espacio de direcciones privado

Los procesos deben ejecutarse al mismo tiempo

### Memoria compartida distribuida

Las variables se comparten directamente.

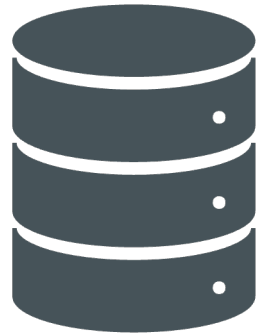
El costo de la comunicación es invisible

Los procesos podrían causar errores al alterar los datos

La ejecución de los procesos puede ocurrir con vidas útiles que no se superpongan



# ARQUITECTURA DE MEMORIA COMPARTIDA / DISTRIBUIDA



Los sistemas de software DSM también tienen la flexibilidad de organizar la región de memoria compartida de diferentes maneras. El enfoque basado en páginas organiza la memoria compartida en páginas de tamaño fijo. Por el contrario, el enfoque basado en objetos organiza la región de memoria compartida como un espacio abstracto para almacenar objetos compartibles de tamaños variables. Otra implementación comúnmente vista utiliza un espacio de tupla, en el que la unidad de intercambio es una tupla.

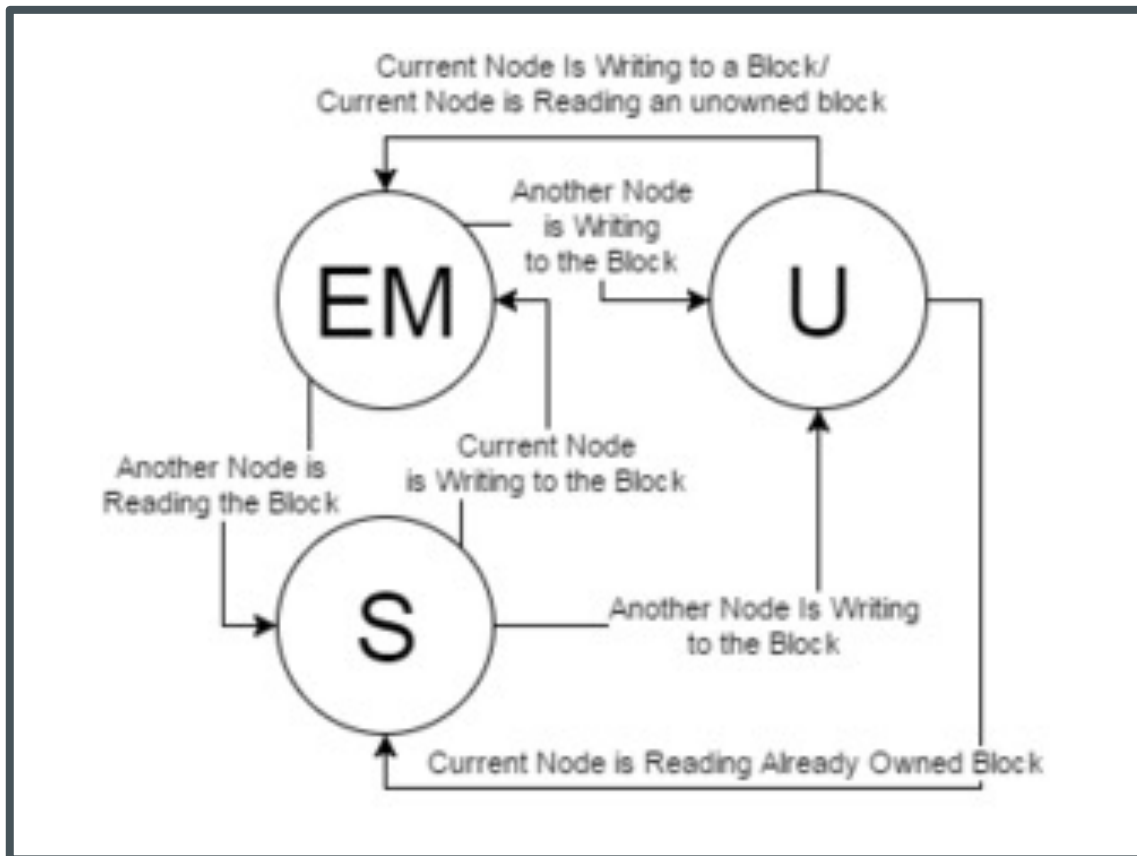


La arquitectura de memoria compartida puede implicar separar la memoria en partes compartidas distribuidas entre los nodos y la memoria principal; o distribuir toda la memoria entre nodos. Un protocolo de coherencia, elegido de acuerdo con un modelo de coherencia, mantiene la coherencia de la memoria.

# ARQUITECTURA DE MEMORIA COMPARTIDA / DISTRIBUIDA

- **Coherencia de la memoria del directorio**
- La coherencia de la memoria es necesaria para que el sistema que organiza el DSM pueda rastrear y mantener el estado de los bloques de datos en los nodos de las memorias que componen el sistema. Un directorio es uno de esos mecanismos que mantiene el estado de los bloques de caché que se mueven por el sistema.

# ARQUITECTURA DE MEMORIA COMPARTIDA / DISTRIBUIDA



## ■ Coherencia de la memoria del directorio

### • Estados

- Un DSM básico rastreará al menos tres estados entre nodos para cualquier bloque determinado en el directorio. Habrá algún estado para dictar el bloque como no almacenado en caché (U), un estado para dictar un bloque como de propiedad exclusiva o propiedad modificada (EM) y un estado para dictar un bloque como compartido (S). A medida que los bloques ingresan a la organización del directorio, pasarán de U a EM (estado de propiedad) en el nodo inicial. El estado puede pasar a S cuando otros nodos comienzan a leer el bloque.
- Existen dos métodos principales para permitir que el sistema rastree dónde se almacenan en caché los bloques y en qué condiciones en cada nodo. La solicitud-respuesta centrada en el hogar utiliza el hogar para atender solicitudes y controlar estados, mientras que la centrada en el solicitante permite que cada nodo impulse y administre sus propias solicitudes a través del hogar.

# ARQUITECTURA DE MEMORIA COMPARTIDA / DISTRIBUIDA

## Coherencia de la memoria del directorio

### Solicitud y respuesta centradas en el hogar

- En un sistema centrado en el hogar, el DSM evitará tener que manejar carreras de solicitud-respuesta entre nodos al permitir que solo se realice una transacción a la vez hasta que el nodo principal haya decidido que la transacción ha finalizado, generalmente cuando el nodo principal haya recibido todas las respuestas. respuesta del procesador a la solicitud. Un ejemplo de esto es el modo de fuente doméstica QPI de Intel. Las ventajas de este enfoque son que es simple de implementar pero su estrategia de solicitud-respuesta es lenta y amortiguada debido a las limitaciones del nodo principal.

### Solicitud y respuesta centradas en el solicitante

- En un sistema centrado en el solicitante, el DSM permitirá que los nodos se comuniquen entre sí a voluntad a través del hogar. Esto significa que varios nodos pueden intentar iniciar una transacción, pero esto requiere consideraciones adicionales para garantizar la coherencia. Por ejemplo: cuando un nodo está procesando un bloque, si recibe una solicitud para ese bloque de otro nodo, enviará un NAck (Reconocimiento negativo) para informarle al iniciador que el nodo de procesamiento no puede cumplir con esa solicitud de inmediato. Un ejemplo de esto es el modo de búsqueda de fuente QPI de Intel. Este enfoque es rápido pero no previene naturalmente las condiciones de carrera y genera más tráfico de autobuses.

# ARQUITECTURA DE MEMORIA COMPARTIDA / DISTRIBUIDA

- **Modelos de consistencia**

- El DSM debe seguir ciertas reglas para mantener la coherencia sobre cómo se ve el orden de lectura y escritura entre los nodos, lo que se denomina modelo de coherencia del sistema .
- Supongamos que tenemos  $n$  procesos y  $m_i$  operaciones de memoria para cada proceso  $i$  , y que todas las operaciones se ejecutan secuencialmente. Podemos concluir que  $(m_1 + m_2 + \dots + m_n)!$  /  $(m_1! m_2! \dots m_n!)$  son posibles entrelazamientos de las operaciones. El problema con esta conclusión es determinar la exactitud de las operaciones entrelazadas. La coherencia de la memoria para DSM define qué entrelazados están permitidos.

# ARQUITECTURA DE MEMORIA COMPARTIDA / DISTRIBUIDA

## ■ Replicación

- Hay dos tipos de algoritmos de replicación. Lectura de replicación y escritura de replicación. En la replicación de lectura, varios nodos pueden leer al mismo tiempo, pero solo un nodo puede escribir. En la replicación de escritura, varios nodos pueden leer y escribir al mismo tiempo. Las solicitudes de escritura son manejadas por un secuenciador. La replicación de datos compartidos en general tiende a:
  - Reducir el tráfico de la red
  - Promover un mayor paralelismo
  - Resulta en menos errores de página
- Sin embargo, preservar la coherencia y la coherencia puede resultar más difícil.