

Memoria compartida distribuida y paralela

6.2. Modelos de consistencia

- La duplicidad de los bloques compartidos aumenta el rendimiento, pero produce un problema de consistencia entre las diferentes copias de la página en caso de una escritura.
- Si con cada escritura es necesario actualizar todas las copias, el envío de las páginas por la red provoca que el tiempo de espera aumente demasiado, convirtiendo este método en impracticable..

Para solucionar este problema se proponen diferentes modelos de consistencia, que establezcan un nivel aceptable de acercamiento tanto a la consistencia como al rendimiento.

Nombramos algunos modelos de consistencia, del más fuerte al más débil:

- consistencia estricta

Cualquier lectura a una localidad de memoria x regresa el valor guardado por la operación de escritura más reciente en x .

- secuencial

El resultado de cualquier ejecución es el mismo que si las operaciones de todos los procesos fueran ejecutadas en algún orden secuencial, y las operaciones de cada proceso individual aparecen en esta secuencia en el orden especificado por su programa.

- causal

Las escrituras potencialmente relacionadas de forma causal son vistas por todos los procesos en el mismo orden. Las escrituras concurrentes pueden ser vistas en un orden diferente en máquinas diferentes

- PRAM

Las escrituras realizadas por un proceso son recibidas por los otros procesos en el orden en que son realizadas, pero las escrituras de procesos diferentes pueden verse en un orden diferente por procesos diferentes.

- del procesador

Para cada posición de memoria x existe un acuerdo local acerca del orden de las escrituras en x . Las escrituras en diferentes posiciones no tienen que ser vistas en el mismo orden por los diferentes procesos.

- débil

- 1.- *Los accesos a las variables de sincronización son secuencialmente consistentes.*
- 2.- *No se permite realizar un acceso a una variable de sincronización hasta que las escrituras anteriores hayan terminado en todas partes.*
- 3.- *No se permite realizar un acceso a los datos (L/E) hasta realizar todos los accesos anteriores a las variables de sincronización*

- de liberación

- 1.- *Antes de realizar un acceso ordinario a una variable compartida, deben terminar con éxito todas las adquisiciones anteriores del proceso en cuestión.*
- 2.- *Antes de permitir la realización de una liberación, deben terminar las (L/E) anteriores del proceso.*

3.- *Los accesos de adquisición y liberación deben ser consistentes con el procesador (no se pide la consistencia secuencial).*

- de entrada

1.- *No se permite realizar un acceso de adquisición a una variable de sincronización con respecto de un proceso hasta que se realicen todas las actualizaciones de los datos compartidos protegidos con respecto de ese proceso*

2.- *Antes de permitir la realización de un acceso en modo exclusivo a una variable de sincronización por un proceso, ningún otro proceso debe poseer la variable de sincronización, ni siquiera en modo exclusivo*

3.- *Después de realizar un acceso en modo exclusivo a una variable de sincronización, no se puede realizar el siguiente acceso en modo no exclusivo de otro proceso a esa variable de sincronización hasta haber sido realizado con respecto del propietario de esa variable.*

6.3. Administración de memoria distribuida y paralela

Se refiere en hacer que los recursos locales y remotos de una manera efectiva y fácil. Esta localización de recursos debe ser transparente para el usuario.

- Los recursos deben estar disponibles de la siguiente manera:
- **Migración de datos:**
- Los datos son traídos al lugar del sistema donde son necesitados, pueden llegar a ser desde un archivo (local o remoto) o hasta el contenido de una memoria física.
- **Características:**
- 1. Si un proceso actualiza un conjunto de datos, la localización original debe de ser también actualizada.
- 2. Se utiliza normalmente sistemas de archivos distribuidos que son componentes que implementan un sistema común de archivos disponibles para todas las computadoras autónomas del sistema. Su objetivo principal es proveer la misma capacidad funcional de acceso a los archivos sin importar su localización dentro de la red. A esto se le conoce como transparencia.
- 3. Otro de los elementos utilizados para la migración de datos es la administración de memoria compartida distribuida. Provee de un espacio de direccionamiento virtual que es compartido entre todas las computadoras del sistema distribuido. La mayor complejidad de implementar memoria compartida distribuida es el mantener la consistencia de los datos y reducir los retardos en el acceso de los mismos.
- **Migración de Cálculos:**
- Los cálculos procesados son llevados hacia otra localización y puede ser eficiente bajo ciertas circunstancias, por ejemplo: Cuando se requiere información de un directorio remoto es más eficiente enviar el mensaje solicitando la información necesaria y recibéndola de regreso.
- **Planeación Distribuida:**
- Los procesos son transferidos de una computadora a otra dentro del sistema distribuido, esto es que un proceso puede ser ejecutado en una computadora diferente de donde fue originado.
- Este proceso de relocalización puede ser deseable si la computadora donde se origino se encuentra sobrecargada o no posee los recursos necesarios.

- La planeación distribuida es la responsable de todos los procesos distribuidos que se lleven a cabo entre computadoras sean efectuados con criterio y transparencia para obtener el máximo rendimiento.

6.4. Modelos y algoritmos para el control de memoria distribuida

Módulo de Estación de Trabajo

Este modelo permite la administración de los procesos ejecutándose en diferentes equipos a través de la implementación de un esquema particular para cada equipo donde es el único responsable de controlar y manejar sus propios procesos.

Módulo de Pila de Procesadores

Este modelo consiste en la utilización de varios microprocesadores conectados en arreglo los cuales se pueden asignar en forma dinámica a los procesos que lo requieren.

Módulo Híbrido

Consiste en proporcionar una integración entre la estación de trabajo y la pila de procesadores para combinar la ventaja de los dos esquemas. El trabajo interactivo se lleva a cabo en las estaciones de trabajo, mientras que los procesos no interactivos se ejecutan en el arreglo de microprocesadores. Es la solución mas cara pero mas eficiente.