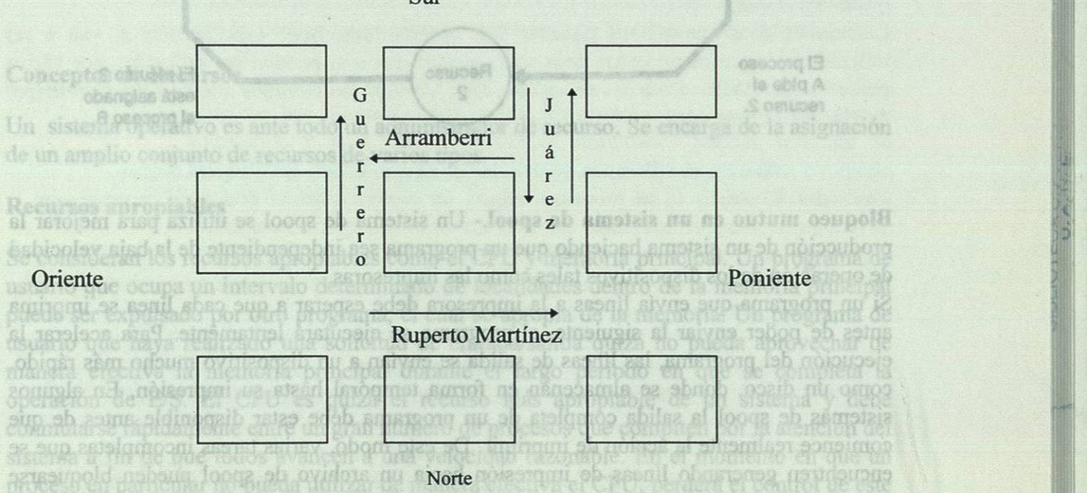
Unidad III
Bloqueo Mutuo

Objetivo de esta unidad.- Durante el desarrollo de esta unidad, se analizarán los bloqueos de los procesos.
El alumno deberá comprender: ¿Qué es un bloqueo?, las condiciones que lo propician, las áreas de investigación de los bloqueos y temas asociados.

Bloqueos

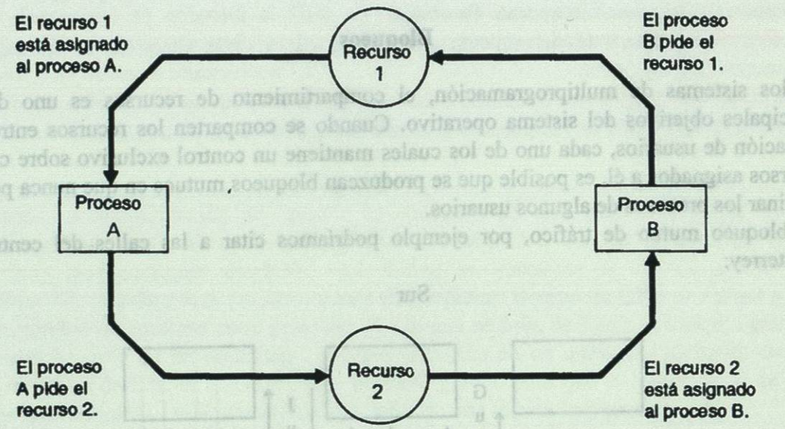
En los sistemas de multiprogramación, el compartimiento de recursos es uno de los principales objetivos del sistema operativo. Cuando se comparten los recursos entre una población de usuarios, cada uno de los cuales mantiene un control exclusivo sobre ciertos recursos asignados a él, es posible que se produzcan bloqueos mutuos en que nunca podrán terminar los procesos de algunos usuarios.
Un bloqueo mutuo de tráfico, por ejemplo podríamos citar a las calles del centro de Monterrey:



El tráfico se detiene por completo, de poco o nada sirven los semáforos controladores del tráfico, siendo necesario la intervención del agente de tránsito para solucionar el embrollo alejando lenta y cuidadosamente los autos y camiones que circulan por el área congestionada. El tráfico comienza a fluir normalmente, no sin antes haber provocado molestias, movilizaciones y una considerable pérdida de tiempo.

La mayor parte de los bloqueos mutuos de los sistemas operativos se presentan a causa de una competencia normal por los recursos dedicados (recursos que sólo pueden ser utilizados por un usuario a la vez, o sea, recursos reutilizables en serie).

Cada proceso espera que el otro libere un recurso que no liberará hasta que el otro libere su recurso, etc. Esta espera circular es característica de los sistemas en bloqueo mutuo. Dado que la tenaz retención de los recursos puede provocar un bloqueo mutuo, este a veces recibe el nombre de abrazo mortal.



Bloqueo mutuo en un sistema de spool.- Un sistema de spool se utiliza para mejorar la producción de un sistema haciendo que un programa sea independiente de la baja velocidad de operación de los dispositivos tales como las impresoras.
Si un programa que envía líneas a la impresora debe esperar a que cada línea se imprima antes de poder enviar la siguiente, el programa se ejecutará lentamente. Para acelerar la ejecución del programa, las líneas de salida se envían a un dispositivo mucho más rápido, como un disco, donde se almacenan en forma temporal hasta su impresión. En algunos sistemas de spool la salida completa de un programa debe estar disponible antes de que comience realmente la acción de imprimir. De este modo, varias tareas incompletas que se encuentren generando líneas de impresión hacia un archivo de spool pueden bloquearse mutuamente si el espacio se termina antes de que acabe cualquiera de las tareas.
La recuperación de un sistema de un bloqueo mutuo de esa naturaleza puede requerir el inicio completo del sistema con la consecuente pérdida de todo trabajo realizado hasta entonces. Si se bloquea de manera que el operador pueda tomar el control, existe la posibilidad de realizar una recuperación menos drástica mediante la desactivación o muerte de una o más tareas hasta disponer del espacio de spool suficiente para que puedan completarse los otros trabajos.
Cuando un programador de sistemas genera un sistema operativo, especifica el espacio para los archivos de spool. Una manera de reducir la posibilidad de un bloqueo mutuo en los sistemas de spool es reservar un espacio mucho mayor para los archivos spool que el considerado indispensable. Si el espacio es insuficiente no siempre es posible esta solución.

Una solución más común es colocar un impedimento en la entrada del spooler para que no acepte más tareas en el momento en que los archivos de spool se acerquen a un *umbral de saturación* que podría ser, un 75% del espacio total por ejemplo. Esto reduciría el rendimiento del sistema que es el precio que hay que pagar por reducir la posibilidad del bloqueo mutuo.
En cualquier sistema que mantenga los procesos en espera mientras se les asigna un recurso o se toman decisiones de planificación, la programación de un proceso puede postergarse indefinidamente mientras otro recibe la atención del sistema. Esta situación se le conoce como Aplazamiento indefinido, Postergación indefinida o Inanición, y es tan peligrosa como un bloqueo mutuo.
Puede ocurrir debido a predisposiciones en las políticas de planificación de recursos del sistema. Cuando los recursos se planifican por prioridad, es posible que un proceso dado espere en forma indefinida un recurso porque siguen llegando otros procesos con mayor prioridad. En algunos sistemas, el aplazamiento indefinido se evita aumentando la prioridad del proceso mientras espera, a esto se le conoce como envejecimiento; en algún momento la prioridad de este superará la prioridad de otros procesos entrantes y el proceso en espera será atendido.

Conceptos de Recursos

Un sistema operativo es ante todo un administrador de recurso. Se encarga de la asignación de un amplio conjunto de recursos de varios tipos.

Recursos apropiables

Se consideran los recursos apropiables como el CPU y memoria principal. Un programa de usuario que ocupa un intervalo determinado de localidades dentro de la memoria principal puede ser expulsado por otro programa, el cual se apropia de la memoria. Un programa de usuario que haya realizado una solicitud de entrada/salida quizá no pueda aprovechar de manera efectiva la memoria principal durante el largo periodo en que se completa la operación de E/S. El CPU es quizá el recurso más apropiable de un sistema y debe conmutarse rápidamente entre un gran número de procesos que compiten por la atención del sistema a fin de que todos avancen a una velocidad razonable. En el momento en que un proceso en particular no pueda utilizar de manera efectiva el CPU, perderá el control de este en favor otro proceso.
La apropiación es de enorme importancia para el éxito de un sistema de computo multiprogramado.

Recursos no apropiables

No pueden arrebatarse al proceso al que han sido asignados.Por ejemplo, las unidades de disco, unidades de cinta que estan asigandas a un proceso en particular por periodos de varios minutos u horas.

Algunos recursos se pueden *compartir* entre varios procesos y otros están dedicados a un solo proceso. Las unidades de disco están dedicadas a un solo proceso, pero a menudo contienen archivos pertenecientes a varios procesos. La memoria principal y el CPU son compartidos por muchos procesos; aunque un solo CPU normalmente sólo puede pertenecer a un proceso a la vez, la conmutación de un CPU entre varios procesos crea la ilusión de un compartimiento simultáneo.

Los datos y programas son sin duda recursos que es necesario controlar y asignar. En los sistemas de multiprogramación, varios usuarios utilizar al mismo tiempo un programa editor, siendo un desperdicio de memoria tener una copia del editor para cada programa, en lugar de ello se lleva a la memoria una sola copia del código y se hacen varias copias de los datos, una para cada usuario. El código no debe cambiar pues muchas personas pueden utilizarlo al mismo tiempo.El código no debe cambiar, pues pueden estarlo usando gran cantidad de personas al mismo tiempo. *El código que no se puede modificar mientras se usa se llama reentrante.* El código reentrante puede ser compartido simultáneamente por varios procesos.

El código que se puede modificar pero que vuelve a su forma original cada vez que se utiliza se llama reutilizable en serie. El reutilizable en serie solo puede ser utilizado por un proceso a la vez.

Cuando se dice que ciertos recursos son compartidos, debe especificarse si van a ser utilizados por varios procesos de manera simultánea o si pueden ser utilizados por varios procesos, pero sólo uno a la vez, siendo estos últimos los recursos que tienden a participar en bloqueos mutuos.

Cuatro condiciones necesarias para que exista un bloqueo mutuo

Exclusión Mutua.- Los procesos exigen un control exclusivo de los recursos que necesitan

Espera.- Los procesos mantienen la posesión de los recursos ya asignados a ellos mientras esperan recursos adicionales.

No apropiación.- Los recursos no pueden arrebatarse a los procesos a los que están asignados hasta que no termine su ejecución o su utilización.

Espera circular.- Existe una cadena circular de procesos en la cual cada proceso tiene uno o más recursos que son requeridos por el siguiente proceso en la cadena.

La existencia de un bloqueo mutuo implica que se han dado todas y cada una de las cuatro condiciones.

Áreas de investigación en los bloqueos mutuos

Prevención.- En la prevención de un bloqueo mutuo vamos a ajustar el sistema para eliminar toda la posibilidad de que ocurra un bloqueo mutuo pero esto puede declinar el aprovechamiento de los recursos o bien empobrecerlos.

Técnicas para evitar.- En las técnicas para evitar el bloqueo mutuo el objetivo es imponer condiciones menos restrictivas que en la prevención para tratar de obtener un mejor aprovechamiento de los recursos. No implica ajustar previamente el sistema para eliminar todas las posibilidades de que se produzca aquel sino que permite la posibilidad de que ocurra el bloqueo mutuo pero se esquiva cuando esta a punto de suceder.

Detección.- Los métodos de detección del bloqueo mutuo se utilizan en sistemas que permiten la ocurrencia de los bloqueos mutuos ya sea de manera voluntaria e involuntaria. El objetivo es determinar si ha ocurrido un bloqueo mutuo y saber exactamente cuales son los procesos y recursos implicados en él, una vez que se determina, el bloqueo mutuo puede eliminarse del sistema.

Recuperación.- Los métodos de recuperación ante un bloqueo mutuo sirven para eliminar los bloqueos mutuos de un sistema para que pueda seguir trabajando y para que los procesos implicados puedan terminar su ejecución y liberen los recursos. Esto viene a ser un problema complejo ya que la recuperación se viene logrando en el mejor de los casos eliminando completamente uno o varios procesos. Después, se inician de nuevo los procesos eliminados, perdiéndose la mayor o todo el trabajop previo realizado por el proceso.

1er Área de Investigación en los Bloqueos Mutuos: Prevención

La técnica más popular más popular empleada por los diseñadores para tratar el bloqueo mutuo es la prevención

J.W. Havender llegó a la conclusión de que si falta una de la cuatro condiciones necesarias no puede haber bloqueo mutuo y para negarlas sugiere lo siguiente:

1. Cada proceso deberá pedir todos sus recursos al mismo tiempo y no podrá continuar su ejecución hasta que no reciba todos.
2. Si a un proceso que tiene ciertos recursos se le niegan los demás, este proceso deberá liberar los recursos y, en caso necesario pedirlos de nuevo junto con los recursos adicionales.
3. Se impondrá un ordenamiento lineal de los tipos de recursos en todos los procesos, esto es, si a un proceso le han asignado un tipo de recursos especifico, en lo sucesivo solo podrá pedir aquellos recursos que siguen en el ordenamiento.