

Sistemas Operativos I

Tema 4

Comunicación y sincronización de procesos



Equipo de Sistemas Operativos DISCA / DSIC

UPV

Comunicación y sincronización de procesos



Objetivos

- Presentar dos alternativas básicas para comunicación entre procesos
 - □ Memoria común
 - Mensajes
- Analizar las condiciones de carrera y estudiar el concepto de seriabilidad.
- Estudiar, en el caso de memoria común, el problema de la sección critica.
- Presentar los criterios de corrección al problema de la sección critica y sus posibles soluciones de una forma estructurada:
 - Soluciones hardware
 - Semáforos
 - Realización de semáforos mediante soluciones hardware.
- Adquirir destreza en la resolución de problemas de sincronización a través de problemas clásicos (productores y consumidores, lectores y escritores, cinco filósofos, etc.)

Sistemas Operativos I (00-01)

Comunicación y sincronización de Procesos

Comunicación y sincronización de procesos



o Contenido

- 1.- Introducción
- 2.- Comunicación por memoria común
- 3.- El problema de la sección crítica
- 4.- Soluciones hardware
- 5.- Semáforos
- 6.- Problemas clásicos de programación concurrente
- 7.- Construcciones lingüísticas
- 8.- Sincronización en POSIX
- 9.- Comunicación por mensajes

o Bibliografia

- A. Silberschatz, J. Peterson , P. Galvin. Sistemas Operativos. Conceptos Fundamentales. 5^a ed. Capítulo 6.
- A. Tanenbaum. Modern Operating Systems. Capítulos 2,11 y 12

Sistemas Operativos I (00-01)

Comunicación y sincronización de Procesos

Comunicación y sincronización de procesos





- 1.- Introducción
- 2.- Comunicación por memoria común
- 3.- El problema de la sección crítica
- 4.- Soluciones hardware
- 5.- Semáforos
- 6.- Problemas clásicos de programación concurrente
- 7.- Construcciones lingüísticas
- 8.- Sincronización en POSIX
- 9.- Comunicación por mensajes

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos

1.Introducción



- o Existe la necesidad de comunicación entre procesos.
- Los procesos requieren con frecuencia comunicación entre ellos (ejemplo: tubos).
- La comunicación entre procesos puede seguir dos esquemas básicos:
 - Comunicación por memoria común
 - Comunicación por mensajes

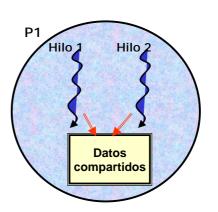
Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos

1.Introducción



- o Comunicación por memoria común
 - La comunicación por memoria común se puede dar en los siguientes casos:
 - Espacio de direcciones único: es el caso de los hilos de ejecución
 - El s.o. crea una zona de memoria accesible a un grupo de procesos
 - Problema de la sección crítica: en un sistema con procesos concurrentes que se comunican compartiendo datos comunes es necesario sincronizar el acceso (lectura, escritura) a los datos compartidos para asegurar la consistencia de los mismos.



Sistemas Operativos I (00-01)

Comunicación y sincronización de Procesos

1.Introducción



o Comunicación por mensajes

- La comunicación por mensajes se da "normalmente" en el siguiente caso:
 - Espacio de direcciones independentes
- Sincronización en la comunicación por mensajes: Cuando dos procesos se comunican vía mensajes se necesitan mecanismos para que el proceso receptor espere (se suspenda hasta) a que el proceso emisor envíe el mensaje y éste esté disponible.
- No aparece el problema de la sección crítica.



Comunicación y sincronización de procesos 🔭



Contenido

1.- Introducción



- 2.- Comunicación por memoria común
- 3.- El problema de la sección crítica
- 4.- Soluciones hardware
- 5.- Semáforos
- 6.- Problemas clásicos de programación concurrente
- 7.- Construcciones lingüísticas
- 8.- Sincronización en POSIX
- 9.- Comunicación por mensajes



- o El problema de los productores y consumidores con buffer acotado
 - **Productor**: proceso que produce elementos (a una cierta velocidad) y los deposita en un buffer.
 - **Consumidor**: proceso que toma elementos del buffer y los consume (a una velocidad probablemente diferente a la del productor)
 - **Buffer**: Estructura de datos que sirve para intercambiar información entre los procesos productores y consumidores. Actúa a modo de depósito para absorber la diferencia de velocidad entre productores y consumidores
 - Ejemplo: buffer de impresora.

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos

9

2.Comunicación por memoria común

Sisterias Operativos i (00-01)

Comunicación y sincronización de Procesos



Productores y consumidores con buffer acotado

```
Variables compartidas:
   var buffer: array[0..n-1]
                    of elemento;
        entrada:=0, salida:=0:0..n-1;
        contador:= 0:0..n;
•proceso productor:
   task productor;
       var item: elemento;
       repeat
       item := producir();
       while contador= n do no-op;
       buffer[entrada] := item;
       entrada := (entrada +1) mod n;
       contador:=contador+1;
       until false
   end productor;
```

```
•proceso consumidor:
    task consumidor;
    var item: elemento;
    repeat
    while contador= 0 do no-op;
    item := buffer[salida]
    salida := (salida +1) mod n;
    contador:=contador-1;
    consumir(item);
    until false
    end consumidor;
```

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos

11

2.Comunicación por memoria común



Condiciones de carrera

- Corrección en programes secuenciales: el programa cumple con sus especificaciones (responde a unos invariantes o reglas de corrección).
- Corrección secuencial no implica corrección concurrente: Un programa que tiene una implementación "secuencialmente correcta" (correcta con un sólo hilo de ejecución) puede presentar problemas cuando se intenta introducir concurrencia en forma de hilos de ejecución.
- Condición de carrera: la ejecución de un conjunto de operaciones concurrentes sobre una variable compartida, deja la variable en un estado inconsistente con las especificaciones de corrección. Además, el resultado de la variable depende de la velocidad relativa en que se ejecutan las operaciones.
- **Peligro potencial**: Las condiciones de carrera pueden presentarse en algún <u>escenario</u>, pero no tienen por qué observarse en todas las posibles trazas de la ejecución del programa.

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos



o Ejemplo de un escenario de condición de carrera

Productor:



contador:= contador+1;

mov reg1, contador;
inc reg1;
mov contador, reg1;

Consumidor:

contador:= contador-1;

mov reg2, contador;
dec reg2;
mov contador, reg2;

Inicialmente: contador =5

| Т | Proceso | Operación | reg1 | reg2 | contador |
|---|---------|--------------------|------|------|----------|
| 0 | Prod. | mov contador, reg1 | 5 | ? | 5 |
| 1 | Prod. | inc reg1 | 6 | ? | 5 |
| 2 | Cons. | mov contador, reg2 | ? | 5 | 5 |
| 3 | Cons. | dec reg2 | ? | 4 | 5 |
| 4 | Cons. | mov reg2, contador | ? | 4 | 4 |
| 5 | Prod | mov reg1, contador | 6 | ? | 6 |



Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos

13

2.Comunicación por memoria común



o Criterio de corrección en programas concurrentes

El criterio de corrección/consistencia más usual para programas concurrentes es:

• **Seriabilidad** (consistencia secuencial): El resultado de la ejecución concurrente de un conjunto de operaciones ha de ser equivalente al resultado de ejecutar secuencialmente cada una de las operaciones, en alguno de los ordenes secuenciales posibles.

Condición de carrera = no seriabilidad: no hay ninguna posible ejecución secuencial de un conjunto de operaciones que de el mismo resultado que la ejecución concurrente.

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos



o Ejemplo de seriabilidad

La ejecución concurrente de op1, op2 y op3 :

se considera correcta si el estado final de los datos compartidos es igual al estado en que quedarían después de alguna de las siguientes ejecuciones secuenciales:

> op1; op2; op3 o op1; op3; op2 o op2; op1; op3 o op2; op3; op1 o op3; op1; op2 o op3; op2; op1.

- Ejemplo numérico: sea una variable x con valor inicial 3 sobre la que se pueden realizar dos operaciones:
 - op1: incremento de 8 - op2: multiplicación por 5
- Resultados correctos de op1 || op2 : 55 y 23

Sistemas Operativos I (00-01)

Comunicación y sincronización de procesos 🔭



Contenido

- 1.- Introducción
- 2.- Comunicación por memoria común



- 3.- El problema de la sección crítica
- 4.- Soluciones hardware
- 5.- Semáforos
- 6.- Problemas clásicos de programación concurrente
- 7.- Construcciones lingüísticas
- 8.- Sincronización en POSIX
- 9.- Comunicación por mensajes

3.El problema de la sección crítica



o Concepto de sección crítica

Una sección crítica es una **zona de código** en la que se accede a variables compartidas por varios procesos.

- **Problemas potenciales**: puede introducir condiciones de carrera si no se adoptan las medidas adecuadas.
- Posible solución: sincronizar el acceso a los datos de manera que mientras un proceso ejecuta su sección crítica ningún otro proceso ejecuta la suya (exclusión mútua).

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos

17

3. El problema de la sección crítica



- o Formulación del problema de la sección crítica
 - Sean n procesos compitiendo para acceder a datos compartidos
 - Cada proceso tiene una zona de código, denominada sección crítica, en la que accede a los datos compartidos.
 - Problema: encontrar un protocolo del tipo:



Que satisfaga las tres condiciones siguientes :

Sistemas Operativos I (00-01)

Comunicación y sincronización de Procesos

3. El problema de la sección crítica



o Solución al problema de la sección crítica

Cualquier solución al problema de la sección crítica ha de cumplir tres requisitos:

- Exclusión mútua: si un proceso está ejecutando su sección crítica ningún otro proceso puede estar ejecutando la suya.
- **Progreso**: si ningún proceso está ejecutando su sección crítica y hay otros que desean entrar a las suyas, entonces la decisión de qué proceso entrará a la sección crítica se toma en un tiempo finito y sólo puede ser seleccionado uno de los procesos que desean entrar.
- Espera limitada: Después de que un proceso haya solicitado entrar en su sección crítica, existe un límite en el número de veces que se permite que otros procesos entren a sus secciones críticas.

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos

19

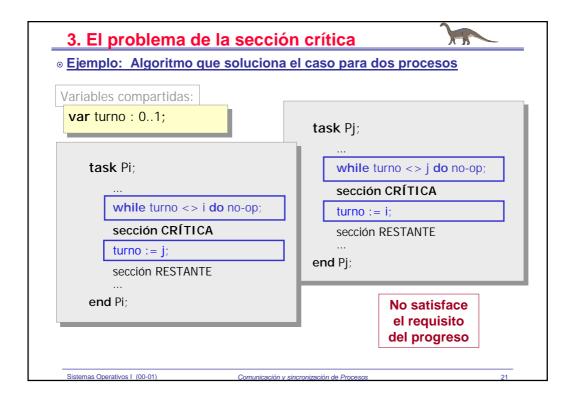
3. El problema de la sección crítica



- o Solución al problema de la sección crítica
 - Supuestos:
 - Los procesos se ejecutan a velocidad no nula.
 - La corrección no ha de depender de hacer suposiciones sobre la velocidad relativa de ejecución de los procesos.

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos





4. Soluciones hardware



o Soluciones hardware al problema de la sección crítica

Las soluciones hardware son soluciones a nivel de instrucciones del lenguaje máquina:

- Deshabilitación de interrupciones
- Instrucción test_and_set atómica (indivisible).
- Instrucción swap atómica.

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos

23

4. Soluciones hardware



o Deshabilitación de interrupciones

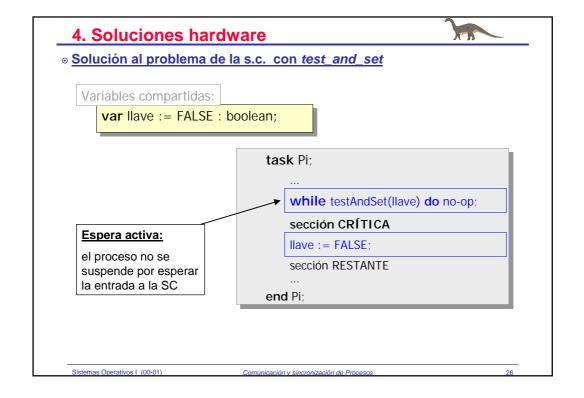
- La deshabilitación de interrupciones se realiza utilizando las instrucciones:
 - □ DI : Deshabilitar interrupciones
 - EI: Habilitar interrupciones
- Se consigue la exclusión mútua inhibiendo los cambios de contexto durante la sección crítica, obligando así a que los procesos se ejecuten de manera atómica.
- Solución únicamente viable al nivel del núcleo del sistema operativo: puesto que no es deseable dejar el control de las interrupciones en manos de los procesos de usuario. Las instrucciones *DI* y *EI* sólo se pueden ejecutar en modo privilegiado.



Sistemas Operativos I (00-01)

Comunicación y sincronización de Procesos

4. Soluciones hardware • Operación test and set atómica _La operación test_and_set permite evaluar y modificar una variable atómicamente en una sola operación de ensamblador. La especificación funcional de esta operación es: function testAndSet(var objetivo: boolean): boolean; begin testAndSet := objetivo; objetivo := TRUE; end atómicamente!



4. Soluciones hardware



o Solución a la s.c. con test and set y espera limitada

La solución anterior no satisface el requisito de la espera limitada.

Variables compartidas:

```
var esperando := FALSE : array[0..n-1] of boolean;
    cerradura := FALSE : boolean;
```

Sistemas Operativos I (00-01

Comunicación v sincronización de Procesos

27

4. Soluciones hardware



o Solución a la s.c. con test and set y espera limitada

```
task Pi;
var j: 0..n-1;
llave: boolean;
...

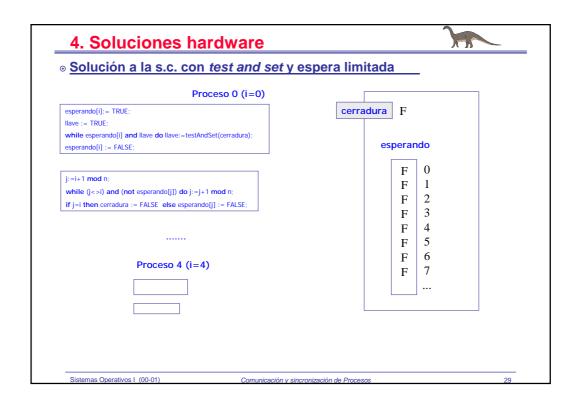
esperando[i]:= TRUE;
llave := TRUE;
while esperando[i] and llave do llave:=testAndSet(cerradura);
esperando[i] := FALSE;

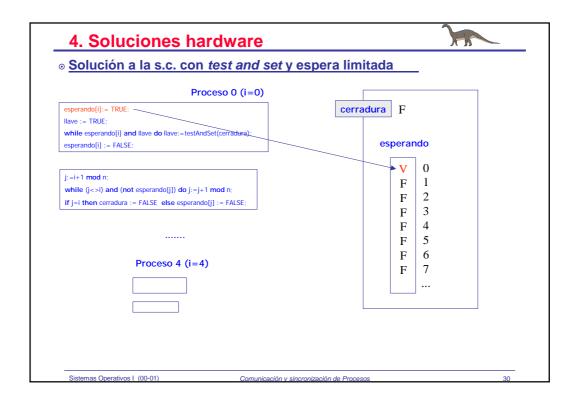
sección CRÍTICA

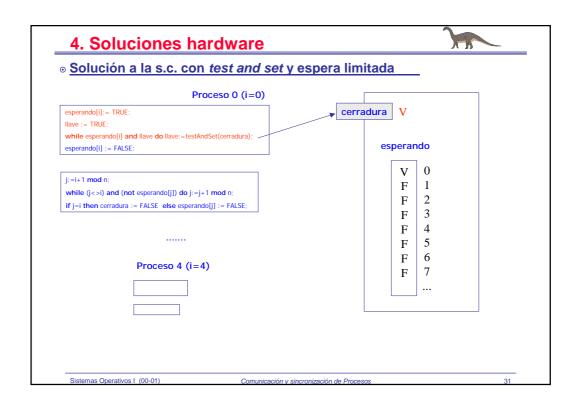
j:=i+1 mod n;
while (j<>i) and (not esperando[j]) do j:=j+1 mod n;
if j=i then cerradura := FALSE else esperando[j] := FALSE;
sección RESTANTE
end Pi;
```

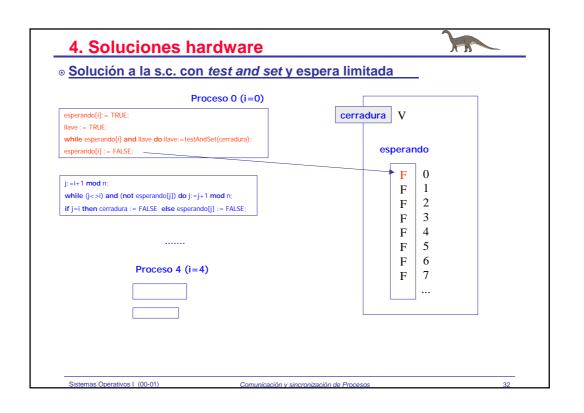
Sistemas Operativos I (00-01)

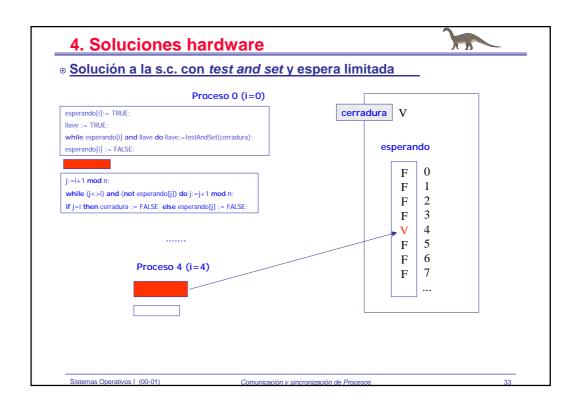
Comunicación y sincronización de Procesos

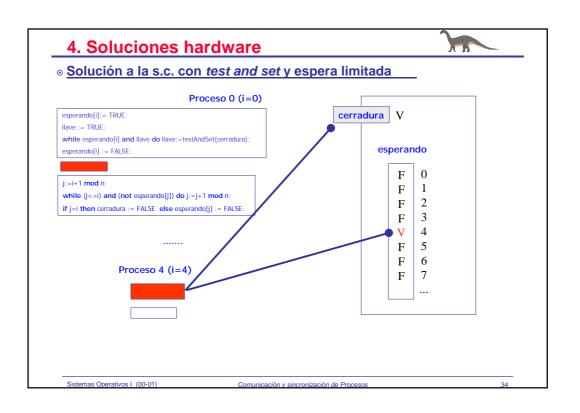


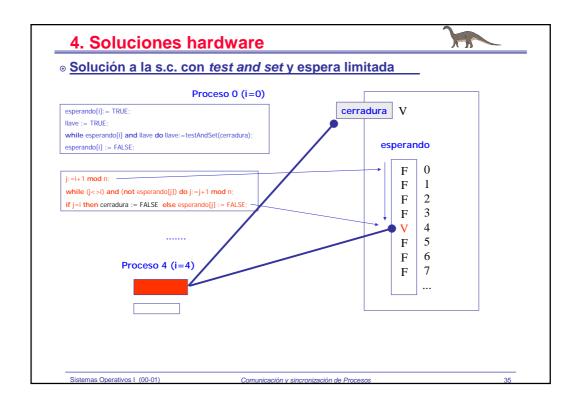


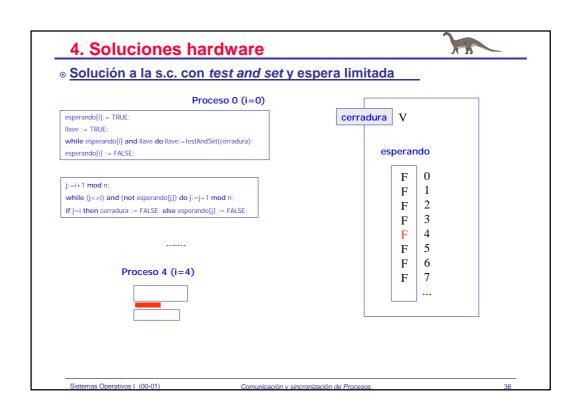


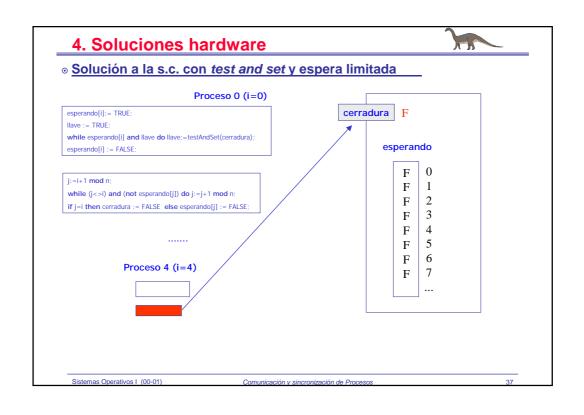


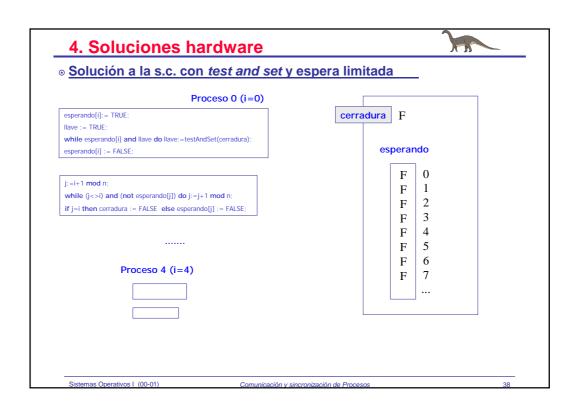












```
4. Soluciones hardware

• Operación swap atómica

_La operación swap permite intercambiar atómicamente dos variables en una sola operación de ensamblador.

La especificación funcional de esta operación es:

function swap(var a,b: boolean);
var temp: boolean;
begin
temp:= a;
a:=b;
b:= temp;
end

Sistemas Operativos I (00-01)

Comunicación y sincronización de Procesos

39
```



Comunicación y sincronización de procesos 🔭



- 1.- Introducción
- 2.- Comunicación por memoria común
- 3.- El problema de la sección crítica
- 4.- Soluciones hardware



- 5.- Semáforos
- 6.- Problemas clásicos de programación concurrente
- 7.- Construcciones lingüísticas
- 8.- Sincronización en POSIX
- 9.- Comunicación por mensajes

Sistemas Operativos I (00-01)

5. Semáforos



o Definición de semáforo

- Tipo de datos que toma valores enteros y sobre el que se definen las siguientes operaciones atómicas:
 - □ S: semaforo(N); (* semáforo S con valor inicial N >=0 *)
 - □ **P**(S) S:=S-1; if S<0 then esperar(S);</pre>

atómicamente!

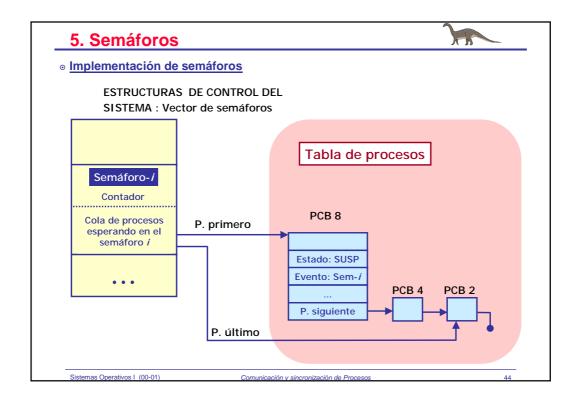
□ **V**(S) S:=S+1; if S<=0 then despertar(S);</pre>

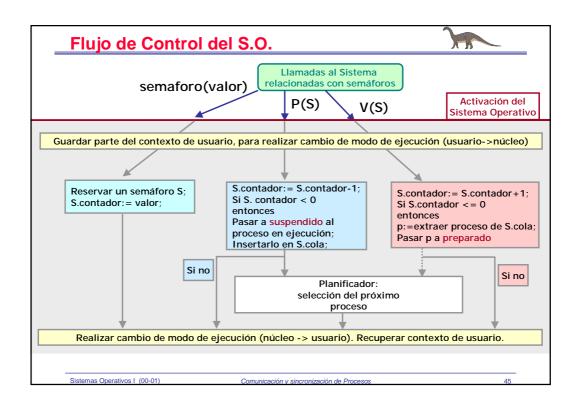
atómicamente!

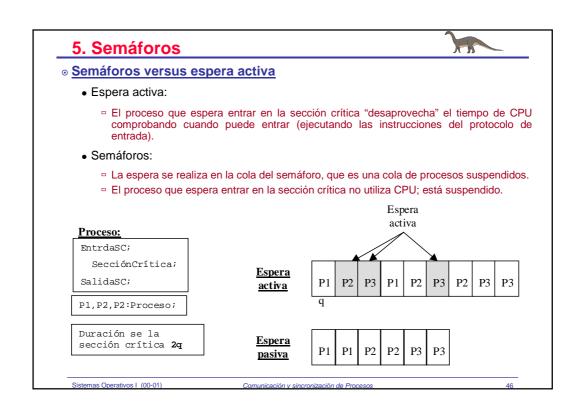
- · La operación esperar(S) suspende al proceso que ha invocado P() y lo introduce
- en una cola de espera asociada a S. • La operación despertar(S) extrae un proceso
- de la cola de espera asociada a S y lo activa.

Semáforo contador → → cola









5. Semáforos



o Semáforos: ejemplos de utilización

_Un semáforo es un mecanismo de sincronización de uso general:

- Para conseguir exclusión mútua.
- Para forzar relaciones de precedencia, como por ejemplo:
 - El proceso Pj debe ejecutar B después de que el proceso Pi haya ejecutado A.

var sinc: semaforo(0);

task Pi; ... A; V(sinc); ... task Pj; ... P(sinc) B; ...

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos

47

5. Semáforos



- o Semáforos: problemas derivados de una utilización incorrecta
 - Interbloqueos: Hay un conjunto de procesos en el que todos esperan (indefinidamente) un evento que sólo otro proceso del conjunto puede producir. *Ejemplo:*

var s1: semaforo(1);
s2: semaforo(1);

task P1; P(s1); P(s2); P(s1); V(s1); V(s2); task P2; P(s2);

$$\label{eq:bounds} \begin{split} & \underline{\text{Pscenario:}} \\ & \text{P1} \rightarrow \text{P(s1);} \\ & \text{P2} \rightarrow \text{P(s2);} \\ & \text{P2} \rightarrow \text{P(s1);} \\ & \text{P1} \rightarrow \text{P(s2);} \\ & \text{Sistema en situación} \\ & \text{de interbloqueo} \dots \end{split}$$

Sistemas Operativos I (00-01)

Comunicación y sincronización de Procesos

Comunicación y sincronización de procesos 🔭



- 1.- Introducción
- 2.- Comunicación por memoria común
- 3.- El problema de la sección crítica
- 4.- Soluciones hardware
- 5.- Semáforos



- 6.- Problemas clásicos de programación concurrente
- 7.- Construcciones lingüísticas
- 8.- Sincronización en POSIX
- 9.- Comunicación por mensajes

Sistemas Operativos I (00-01

Comunicación y sincronización de Procesos

10

6. Problemas clásicos de concurrencia



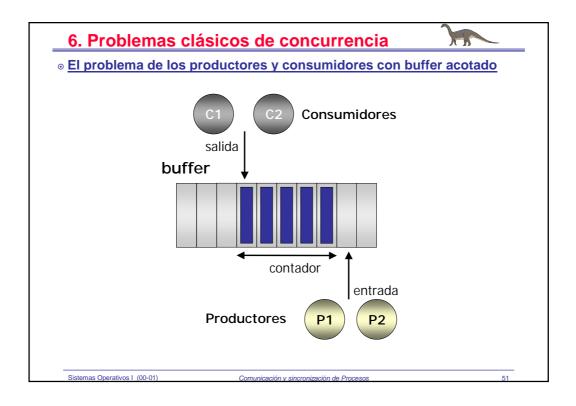
o Problemas clásicos de programación concurrente

_Existe una serie de problemas clásicos que se utilizan como "banco de pruebas" de los diferentes mecanismos de sincronización entre procesos:

- Los productores y los consumidores
- Los lectores y los escritores
- Los cinco filósofos
- El barbero dormilón
- _

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos



6. Problemas clásicos de concurrencia



- o Los productores y consumidores
- o Variables compartidas

```
var buffer: array[0..n-1] of elemento;
```

entrada:=0, salida:=0, contador:= 0 : 0..n-1;

lleno: semaforo(0); vacio: semaforo(n), mutex: semaforo(1);

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos

6. Problemas clásicos de concurrencia



o Los productores y consumidores

```
task productor;
var item: elemento;
repeat
    item := producir();
    P(vacio);
    P(mutex);
    buffer[entrada] := item;
    entrada := (entrada +1) mod n;
    contador:=contador+1;
    V(mutex);
    V(lleno);
until false
end productor;
```

```
task consumidor;
  var item: elemento;
  repeat
    P(lleno);
    P(mutex);
    item := buffer[salida]
    salida := (salida +1) mod n;
    contador:=contador-1;
    V(mutex);
    V(vacio);
    consumir(item);
  until false
end productor;
```

Sistemas Operativos I (00-01)

omunicación v sincronización de Procesos

53

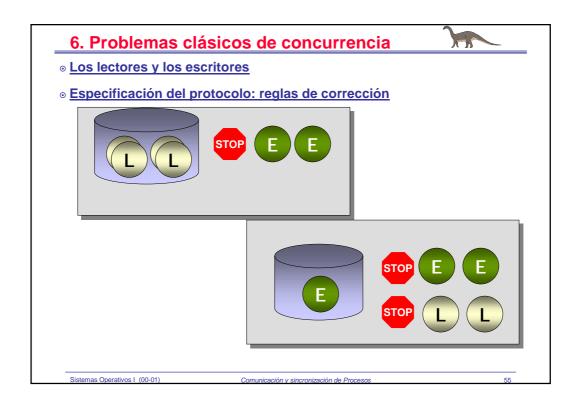
6. Problemas clásicos de concurrencia



- Los lectores y los escritores
- o Problema de los lectores y los escritores
 - Hay un conjunto de datos comunes (un fichero, una base de datos, etc.)
 - Los procesos lectores, acceden a los datos en modo de sólo lectura
 - Los procesos escritores, acceden a los datos en modo lectura-escritura
- o Especificación del protocolo: reglas de corrección
 - Diversos lectores pueden leer concurrentemente.
 - Los escritores se excluyen mútuamente entre ellos
 - Los escritores se excluyen mútuamente con los lectores.

Sistemas Operativos I (00-01)

Comunicación y sincronización de Procesos



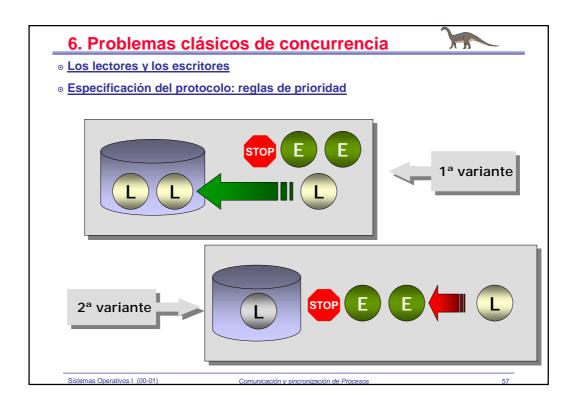
6. Problemas clásicos de concurrencia

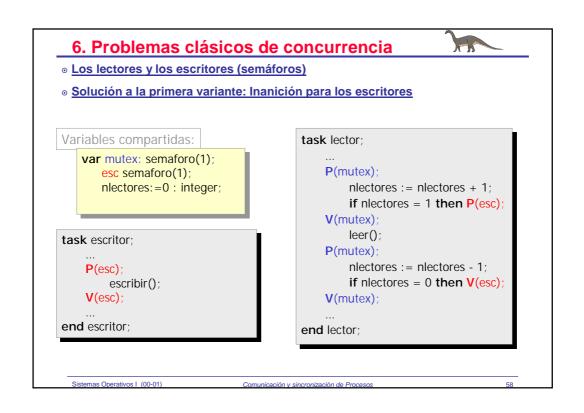


- o Loslectores y escritores
- o Especificación del protocolo: reglas de prioridad
 - Primera variante de los Lectores-Escritores: (Prioridad a los lectores)
 - □ Si hay lectores leyendo y
 - escritores esperando, entonces
 - un nuevo lector tiene preferencia sobre el escritor que espera.
 - → Hay inanición para los escritores, si no paran de llegar lectores
 - Segunda variante de los Lectores-Escritores: (Prioridad a los escritores)
 - □ Si hay escritores esperando,
 - □ éstos siempre tienen preferencia sobre nuevos lectores que lleguen
 - → Hay inanición para los lectores, si no paran de llegar escritores
 - Tercera variante, sin inanición para ningún tipo de proceso

Sistemas Operativos I (00-01)

Comunicación y sincronización de Procesos





Comunicación y sincronización de procesos 🔭



- 1.- Introducción
- 2.- Comunicación por memoria común
- 3.- El problema de la sección crítica
- 4.- Soluciones hardware
- 5.- Semáforos
- 6.- Problemas clásicos de programación concurrente



- 7.- Construcciones lingüísticas
- 8.- Sincronización en POSIX
- 9.- Comunicación por mensajes

Sistemas Operativos I (00-01)

Comunicación y sincronización de Procesos

59

7. Construcciones lingüísticas



- o Con el fin de simplificar la programación de aplicaciones concurrentes, algunos lenguajes de programación proporcionan herramientas (construcciones lingüísticas) que facilitan la sincronización entre las tareas y proporcionan un modelo de programación uniforme.
 - Monitores
 - Tipos protegidos en ADA
 - Métodos synchronized en Java

Soluciones a nivel de lenguaje de programación

Soluciones a nivel de sistema operativo

Soluciones hardware

Sistemas Operativos I (00-01)

Comunicación y sincronización de Procesos

7.1 Monitores



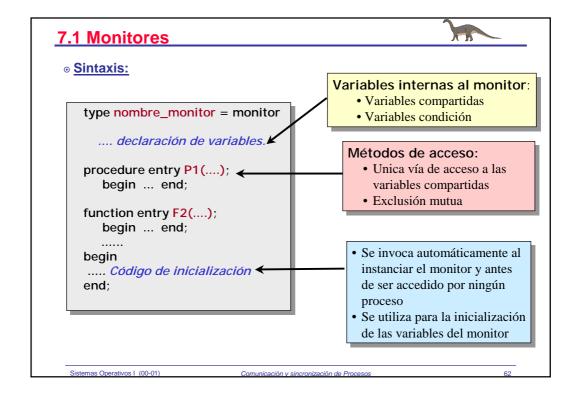
Un monitor es un tipo de datos que encapsula datos y operaciones.

Proporciona:

- Exclusión mútua.
- · Sincronización.

Sistemas Operativos I (00-01

Comunicación v sincronización de Procesos



7.1 Monitores



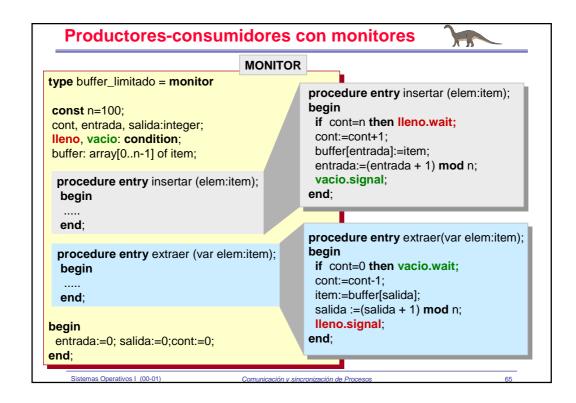
- o Variables de tipo condición:
 - Para definir esquemas complejos de sincronización, se pueden definir las variables del tipo *condition*.
 - □ var x: condition;
 - Sobre las variables *condition* se pueden realizar las siguientes operaciones:
 - x.wait; Causa la suspensión (en una cola asociada a x) del proceso que invocó la operación.
 - □ **x.signal**; Se reanuda (si existe) un proceso suspendido en la cola asociada a x.
 - x.awaited; Indica el número de procesos suspendidos en la cola asociada a x.

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos

63

7.1 Monitores o Productores-consumidores con monitores Monitor PROGRAMA PRINCIPAL B:buffer_limitado; task Consumidor; task Productor: var: elemento:item; var: elemento:item; repeat repeat B.extraer(elemento);Producir un elementoProcesar elemento B.insertar(elemento); until false: until false; end Consumidor; end Productor;



7.1Monitores



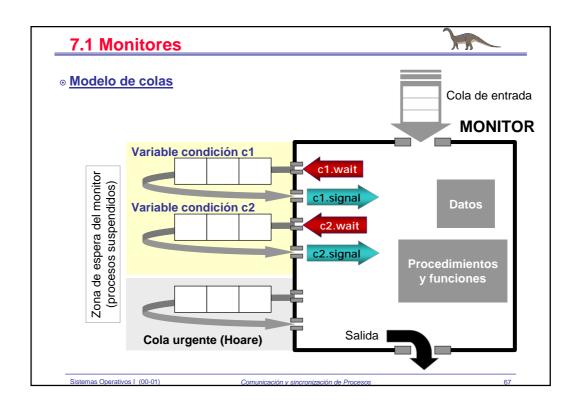
Variantes

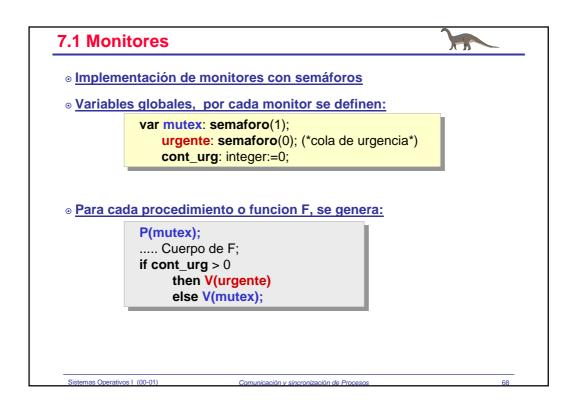
Existen diferentes variantes en la definición de un monitor según resuelvan el siguiente problema:

- Un proceso P ejecuta x.signal y activa a otro proceso Q, → Potencialmente
 P y Q pueden continuar su ejecución dentro del monitor. ¿Cómo se garantiza la ejecución en exclusión mutua en el monitor?
 - Modelo de Hoare: El proceso que invoca la operación x.signal (P) se suspende (en una cola de "urgencia") de forma que el proceso reanudado por la operación x.signal (Q) pasa a ejecutarse dentro del monitor.
 - Modelo de Lampson y Redell: El proceso P continúa ejecutándose y Q se ejecutará cuando el monitor quede libre.
 - Modelo de Brinch-Hansen: La operación x.signal ha de ser la última operación que un proceso invoca antes de salir del monitor.

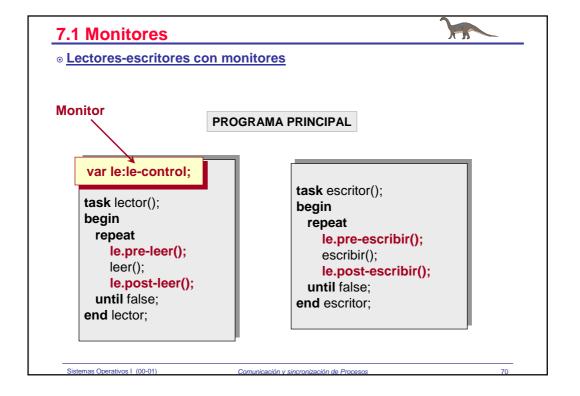
Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos

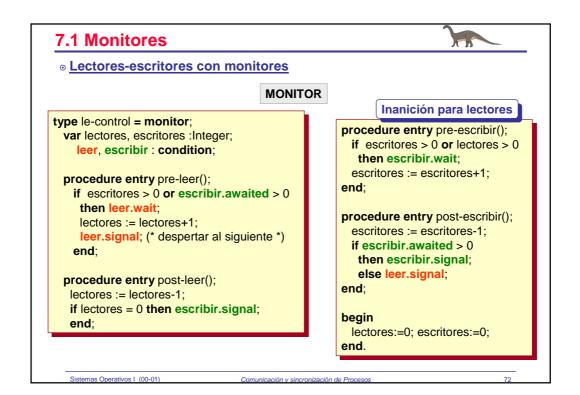




```
7.1 Monitores
 o Implementación de monitores con semáforos
 o Para cada variable condición X, se define:
                      var x_sem: semaforo(0);
                          x_cont: integer:=0;
                                               Operación X.signal
     Operación X.wait
                                    if x cont > 0
   x_cont:= x_cont+1;
                                       then begin
   if cont_urg > 0
                                          cont_urg:= cont_urg + 1;
       then V(urgente)
                                          V(x_sem);
       else V(mutex);
                                          P(urgente);
   P(x_sem);
                                          cont_urg:= cont_urg - 1;
   x_cont:= x_cont-1;
                                       end;
  Sistemas Operativos I (00-01)
```



```
7.1 Monitores
 Lectores-escritores con monitores
                                       MONITOR
                                                           Inanición para escritores
type le-control = monitor;
                                                      procedure entry pre-escribir();
 var lectores, escritores :Integer;
                                                        if escritores > 0 or lectores > 0
    leer, escribir: condition;
                                                         then escribir.wait:
                                                        escritores := escritores+1;
 procedure entry pre-leer();
                                                      end:
   if escritores > 0 then leer.wait;
    lectores := lectores+1;
                                                      procedure entry post-escribir();
    leer.signal; (* despertar al siguiente *)
                                                        escritores := escritores-1;
   end;
                                                        if leer.awaited > 0
                                                         then leer.signal:
 procedure entry post-leer();
                                                         else escribir.signal;
   lectores := lectores-1;
   if lectores = 0 then escribir.signal;
                                                      end;
   end:
                                                      begin
                                                        lectores:=0; escritores:=0;
                                                      end.
   Sistemas Operativos I (00-01)
```



7.1 Monitores



o Lectores-escritores con monitores

MONITOR

```
type le-control = monitor;
  var lectores, escritores :Integer;
  leer, escribir : condition;

procedure entry pre-leer();
  if escritores > 0 or escribir.awaited > 0
    then leer.wait;
    lectores := lectores+1;
    leer.signal; (* despertar al siguiente *)
  end;

procedure entry post-leer();
  lectores := lectores-1;
  if lectores = 0 then escribir.signal;
  end;
```

Sin Inanición

```
procedure entry pre-escribir();
  if escritores > 0 or lectores > 0
    then escribir.wait;
  escritores := escritores+1;
end;

procedure entry post-escribir();
  escritores := escritores-1;
  if leer.awaited > 0
    then leer.signal;
    else escribir.signal;
end;

begin
  lectores:=0; escritores:=0;
end
```

Sistemas Operativos I (00-01)

Comunicación y sincronización de Procesos

73

7.1 Monitores



o Monitores: espera condicional

Cuando se produce una llamada a x.signal, ¿qué proceso se activa?

- El monitor permite una variante de la operación wait sobre una variable condición para que, cuando se produzca la operación signal, se <u>active a los</u> procesos suspendidos por orden de prioridad.
- o Se define la siguiente construcción:
 - x.wait(c)
 - donde:
 - c es una expresión entera que se evalúa cuando se realiza la operación wait.
 - El valor de c se almacena junto al proceso suspendido.
 - Al invocar la operación signal, se activa el proceso que tenga mayor prioridad (menor valor numérico de c).

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos

7.1 Monitores



o Monitores espera condicional

- Asignación de recursos basada en tiempo de uso.
 - De aquellos procesos que se encuentren esperando por el recurso ocupado, se reanudará aquel que lo vaya a usar durante menos tiempo.

```
type asignacion-recurso = monitor
 var ocupado: boolean;
      x: condition;
 procedure entry adquirir(tiempo: integer);
 begin
    if ocupado then x.wait(tiempo);
    ocupado:= true;
 end:
 procedure entry liberar;
 begin
    ocupado:= false;
    x.signal;
 end;
begin
   ocupado:= false;
end.
```

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos

75

7.2 Objetos protegidos



- o Objetos protegidos de ADA: declaración
- Un objeto protegido encapsula datos que son accedidos simultáneamente por varias tareas.
- Las tareas acceden a los datos del objeto utilizando operaciones protegidas, las cuales pueden ser funciones, procedimientos y entradas.

```
protected [type]
   Nombre[(Discriminantes)]
is

procedure Pnombre(Params);
function Fnombre(Params) return Un_Tipo;
entry Enombre(Params);

private
   Datos_Protegidos : Su_Tipo;
end Nombre;
```

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos

7.2 Objetos protegidos



- o Objetos protegidos de ADA
- o La especificación tiene dos partes:
 - Una pública, donde se declaran las operaciones protegidas.
 - Una privada donde se declaran los datos protegidos. La parte privada no es visible por los clientes del objeto protegido.

```
especificación
protected [type]
   Nombre[(Discriminantes)]
  procedure Pnombre(Params);
  function Fnombre(Params) return Un_Tipo;
  entry Enombre(Params);
private
 Datos_Protegidos : Su_Tipo;
end Nombre;
```

Sistemas Operativos I (00-01)

7.2 Objetos protegidos



- o Objetos protegidos de ADA: declaración
- o En el cuerpo del objeto se implementan las operaciones protegidas.

```
protected body Nombre
is
                  cuerpo
end Nombre;
```

 ADA permite declarar tipos protegidos o instancias individuales.

```
Objeto_1: Nombre;
Objeto_2: Nombre;
                   instanciación
```

private

protected [type] Nombre[(Discriminantes)] procedure Pnombre(Params); function Fnombre(Params) return Un_Tipo; entry Enombre(Params);

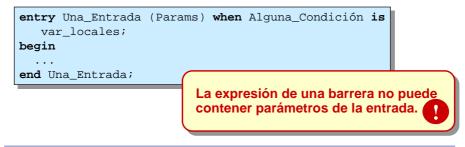
especificación

Datos_Protegidos : Su_Tipo; end Nombre;

7.2 Objetos protegidos



- o Objetos protegidos de ADA: control de concurrencia
- Las **funciones** permiten consultar el objeto protegido y, por tanto, pueden ejecutarse simultáneamente varias funciones.
- Cuando una tarea está ejecutando un procedimiento o una entrada del objeto protegido, otra tarea interesada en acceder al objeto protegido queda detenida hasta que éste queda libre.
- Las entradas tienen asociadas una condición lógica denominada barrera (o guarda). Si en el momento de llamar al objeto la barrera es cierta, la tarea ejecuta la entrada. En caso contrario, la tarea queda detenida en dicha entrada.



Sistemas Operativos I (00-01)

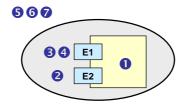
Comunicación v sincronización de Procesos

79

7.2 Objetos protegidos



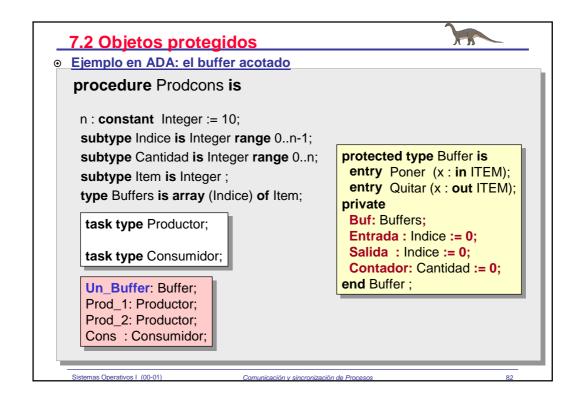
- o Objetos protegidos de ADA: control de concurrencia
- Las barreras se revaluan cuando finaliza la ejecución de una entrada o un procedimiento. Si alguna barrera es ahora cierta se elige una tarea detenida por dicha barrera y reanuda su ejecución.
- Las tareas detenidas en barreras tienen prioridad cuando son despertadas frente a aquellas que están esperando fuera del objeto protegido.



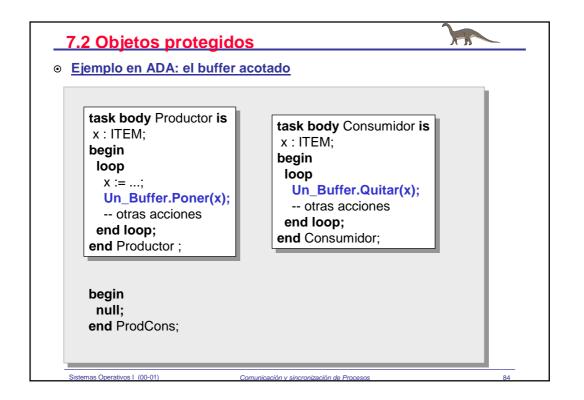
Sistemas Operativos I (00-01)

Comunicación y sincronización de Procesos

```
7.2 Objetos protegidos
o Construcción de semáforos con objetos protegidos de ADA
                            especificación
                                                               cuerpo
protected type Semaforo (Inicial:integer:=1)
                                             protected body Semaforo is
                                              entry P when contador>0 is
 entry P;
                                              begin
 procedure V;
                                                 contador:=contador-1;
private
                                              end P;
 contador:Integer:=Inicial;
end Semaforo;
                                              procedure V is
                                              begin
   utilización
                                                 contador:=contador+1;
                                              end V;
   Un_semaforo: Semaforo(5);
                                             end Semaforo;
   Un_semaforo.P;
   Un_semaforo.V;
```



```
7.2 Objetos protegidos
o Ejemplo en ADA: el buffer acotado
             protected body Buffer is
               entry Poner (x : in Item) when Contador < n is
               begin
                Buf(entrada) := x;
                Entrada := (Entrada + 1) mod n;
                Contador := Contador + 1;
               end Poner;
               entry Quitar (x : out Item) when Contador > 0 is
               begin
                x:= Buf(Salida);
                Salida := (Salida + 1) mod n;
                Contador := Contador - 1;
               end Quitar;
             end Buffer;
```



Comunicación y sincronización de procesos 🔭



- 1.- Introducción
- 2.- Comunicación por memoria común
- 3.- El problema de la sección crítica
- 4.- Soluciones hardware
- 5.- Semáforos
- 6.- Problemas clásicos de programación concurrente
- 7.- Construcciones lingüísticas



- 8.- Sincronización en POSIX
- 9.- Comunicación por mensajes

Sistemas Operativos I (00-01)

Comunicación y sincronización de Procesos

95

8. Sincronización en POSIX



POSIX ofrece dos primitivas de sincronización de hilos: los *mutex* (o cerrojos) y las variables condición.

Mutex (cerrojos)

- Tienen dos estados posibles:
 - □ Abierto → ningún hilo es el propietario
 - □ Cerrado → el hilo que lo cierra es el propietario. Un mutex no puede tener dos hilos propietarios simultáneamente.
- Cuando un hilo intenta cerrar un *mutex* que ya ha sido cerrado, se suspende hasta que el hilo propietario lo abra.
- Sólo el hilo propietario puede abrir el mutex.
- Creación e inicialización de los mutex:
 - p_thread_mutex_t mut = PTHREAD_MUTEX_INITIALIZER;

Sistemas Operativos I (00-01)

Comunicación y sincronización de Procesos





| Llamada | Descripción |
|----------------------------|--|
| pthread_mutex_lock (mut) | Si el mutex <i>mut</i> estaba abierto, lo cierra y el hilo que lo cierra es el propietario del mutex. Si estaba cerrado, suspende(*) al hilo hasta que se abra el mutex. |
| pthread_mutex_unlock (mut) | Abre el mutex <i>mut</i> . Solo el hilo propietario del mutex puede abrirlo (***). |
| pthread_mutex_trylock(mut) | Igual que pthread_mutex_lock pero en lugar de suspender al hilo si el semáforo estaba cerrado, retorna immediatamente con un código de error. |

- (*) Hay tipos de mutex, que le permiten al propietario cerrarlo varias veces sin bloquearse. Con este tipo de mutex, el propietario debe abrirlo tantas veces como lo hubiera cerrado para dejarlo finalmente abierto.
- (**) Hay sistemas, como Redhat 5.1, que permiten que otros hilos abran el mutex, sin embargo este funcionamiento no es portable.

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos

97

8. Sincronización en POSIX



o Variables condición en POSIX

Las variables condición son un tipo abstracto de datos con tres operaciones básicas:

- wait: suspende al proceso que la invoca en la condición.
- signal: activa un proceso suspendido en la condición
- broadcast: activa a todos los procesos suspendidos en la condición

Las variables condición siempre deben estar asociadas a un mutex.

Además existen métodos para limitar el tiempo que un hilo está bloqueado en una variable condición.

oCreación e inicialización de atributos:

p_thread_cond_t c1 = PTHREAD_COND_INITIALIZER;

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos

8. Sincronización en POSIX



o Variables condición en POSIX

| Llamada | Descripción |
|--|--|
| pthread_cond_wait(cond, mut) | De forma atómica, realiza la operación p_thread_mutex_unlock sobre <i>mut</i> , y bloquea al hilo en la variable condición <i>cond</i> . Cuando es despertado, el hilo cierra nuevamente el mutex <i>mut</i> (realizando la operación p_thread_mutex_lock. |
| pthread_cond_signal(cond) | Despierta a uno de los hilos que están bloqueados en la variable condición. Si no hay hilos bloqueados, no sucede nada. |
| pthread_cond_broadcast(cond) | Despierta todos los hilos bloqueados sobre cond. |
| pthread_cond_timedwait (cond, mut,duracion) | Igual que pthread_cond_wait pero si antes de duracion no se ha despertado al hilo, la llamada finalizará con un código de error. Al despertar, el hilo que invoca la llamada, vuelve a cerrar el mutex mut. |

stemas Operativos I (00-01) Comu

89

8. Sincronización en POSIX



©El productor consumidor en POSIX (i)

Variables globales:

#define n 10

int entrada, salida, contador;

int buffer[n];

pthread_cond_t vacio = PTHREAD_COND_INITIALIZER;

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos

```
8. Sincronización en POSIX
 oEl productor consumidor en POSIX (i)
void *productor(void * arg){
                                       void *consumidor(void * arg){
  for (i=0; i<100; i++) Insertar(i);
                                           for (i=0; i<100; i++)
                                                     printf("%d ", Extraer());
   pthread_exit(0);
                                           pthread_exit(0);
             main() {
               pthread_t th_a, th_b;
               entrada = salida = contador = 0;
               pthread_create(&th_a, NULL,productor,NULL);
               pthread_create(&th_b, NULL,consumidor,NULL);
               pthread_join(th_a, NULL);
               pthread_join(th_b, NULL); exit(0);
```

```
8. Sincronización en POSIX
⊙El productor consumidor en POSIX (i)
               Insertar (int dato) {
                pthread_mutex_lock(&mutex);
                while (contador >=n) pthread_cond_wait(&lleno, &mutex);
                buffer[entrada] = dato; entrada = (entrada+1) % n;
                contador = contador + 1;
                pthread_cond_broadcast(&vacio);
                pthread_mutex_unlock(&mutex);
              int Extraer () {
                int dato;
                pthread_mutex_lock(&mutex);
                while (contador == 0) pthread_cond_wait(&vacio, &mutex);
                dato = buffer[salida]; salida= (salida+1) % n;
                contador=contador - 1;
                pthread_cond_broadcast(&lleno);
                pthread_mutex_unlock(&mutex);
                return dato;
```

Comunicación y sincronización de procesos 🔭



- 1.- Introducción
- 2.- Comunicación por memoria común
- 3.- El problema de la sección crítica
- 4.- Soluciones hardware
- 5.- Semáforos
- 6.- Problemas clásicos de programación concurrente
- 7.- Construcciones lingüísticas
- 8.- Sincronización en POSIX



9.- Comunicación por mensajes

Sistemas Operativos I (00-01

Comunicación y sincronización de Procesos

93

9. Comunicación por mensajes



o Sistemas de mensajes

- Los sistemas de mensajes permiten la comunicación entre procesos con espacios de direcciones distintos, bien sean locales o remotos.
- Operaciones básicas (proporcionadas por el Sistema Operativo):
 - send(dst,mensaje): enviar un mensaje a un destino.
 - receive(org,mensaje): recibir un mensaje de un origen.



Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos



- Según el modo de nombrar origen y destino:
 - Comunicación directa: proceso a proceso.
 - □ Comunicación indirecta: basada en puertos o buzones.
- Según la capacidad del enlace:
 - □ Comunicación síncrona: el enlace tiene capacidad = 0.
 - Comunicación asíncrona: el enlace tiene capacidad > 0.

Sistemas Operativos I (00-01

Comunicación v sincronización de Procesos

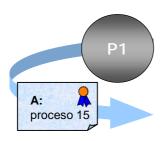
95

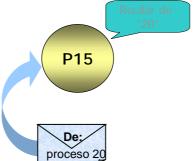
9. Comunicación por mensajes



o Comunicación directa

- Los mensajes se direccionan a procesos.
- Los procesos necesitan conocer sus identificadores: las direcciones dst y org son identificadores de procesos.
- **org** puede ser un comodín, ANY. En este caso se recibe de cualquier proceso y ANY toma el valor del proceso emisor.





Sistemas Operativos I (00-01)

Comunicación y sincronización de Procesos

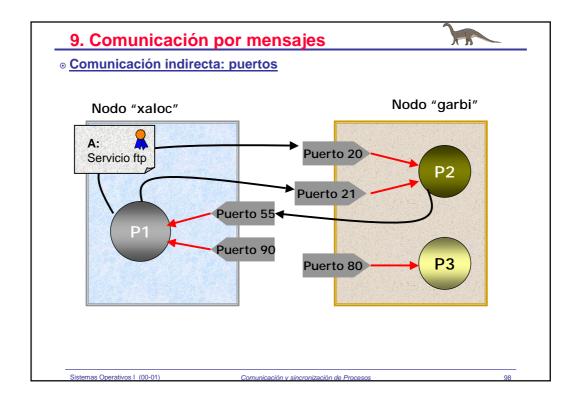


o Comunicación indirecta: puertos o buzones

- Los mensajes se envían a puertos: las direcciones dst y org son identificadores de puertos.
- Un puerto tiene un identificador único en el sistema.
- El protocolo TCP/IP utiliza puertos "bien conocidos" que sirven para invocar servicios de red.
 - ftp: port 20 y 21
 - www: port 80
 - ntp: port 123

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos





- o Comunicación indirecta: puertos y buzones
 - **Puertos:** permiten que varios procesos envíen, pero sólo puede haber un proceso asociado para recibir.
 - **Buzones**: permiten que varios procesos se asocien para recibir. El sistema operativo selecciona qué proceso recibe el mensaje.

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos

99

9. Comunicación por mensajes



- Operaciones sobre puertos
 - ¿Qué proceso puede recibir del puerto? Alternativas:
 - a) El propietario: El proceso que lo crea es el propietario del puerto y es el único que puede recibir o destruir el puerto. Cuando el proceso propietario finaliza, se destruyen todos los puertos que posee. Son necesarias las siguientes operaciones:
 - create(port): crea un puerto.
 - destroy(port): destruye un puerto.
 - b) El que se vincula al puerto: El puerto pertenece al S.O. y su existencia es independiente de los procesos que lo utilizan. Cuando un proceso se vincula a un puerto, es el único que puede recibir de él.
 - bind(puerto): vincula un proceso con un puerto.

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos

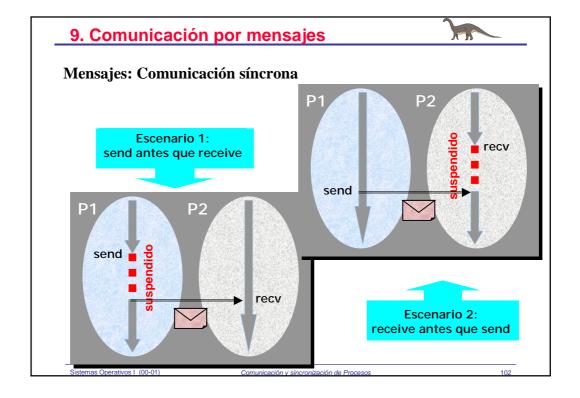


o Comunicación síncrona

- La capacidad del enlace es cero: El enlace no almacena mensajes; no pueden existir mensajes enviados y pendientes de entrega.
- Cita o rendez-vous: La comunicación sólo tiene lugar cuando emisor y receptor han invocado sus operaciones respectivas send y receive.
- Implementación de la cita o rendez-vous: El primer proceso que llega a la cita se suspende esperando a que el otro llegue.

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos





o Comunicación asíncrona

La capacidad del enlace es N: pueden haber mensajes enviados y pendientes de entrega

- Variantes:
 - Capacidad limitada: N está acotado.
 - Capacidad ilimitada: N no está acotado.
- Sincronización:
 - □ Receive de enlace vacío:
 - Opción bloqueante: suspender al proceso que la invoca.
 - Opción no bloqueante : devolver un error al proceso que la invoca.
 - Send a un enlace lleno:
 - Opción bloqueante: suspender al proceso que la invoca.
 - Opción no bloqueante : devolver un error al proceso que la invoca o perder el mensaje.

Sistemas Operativos I (00-01)

Comunicación y sincronización de Procesos

103

9. Comunicación por mensajes



- El problema de los Productores y Consumidores con mensajes:
 - Los espacios de direcciones son separados. No hay variables comunes.
 - Versión con:
 - Comunicación directa.
 - Comunicación síncrona o asíncrona (send y receive bloqueantes)

```
task productor;
var item : elemento;
  repeat
    item := producir();
    send( consumidor, item );
  until false;
end productor;
```

```
task consumidor;
var item : elemento;
repeat
    receive( productor, item );
    consumir(item);
until false;
end consumidor;
```

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos



o El problema de los Productores y Consumidores con mensajes:

• Versión con:

- Comunicación directa.
- □ Comunicación asíncrona, con enlace de capacidad N (receive bloqueante, send no bloqueante)

```
task productor;
var item : elemento;
    m : mensaje;
    repeat
    item := producir();
    receive( consumidor, m );
    construir_msj(m, item);
    send( consumidor, m );
    until false;
end productor;
```

```
task consumidor;
var item : elemento;
    m : mensaje;
    i: integer;
    for i:=1 to N do send( productor, m );
    repeat
        receive( productor, m );
    item := extraer_msj(m);
    consumir(item);
    until false;
end consumidor;
```

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos

105

9. Comunicación por mensajes



- o Mensajes: el modelo cliente-servidor
 - Modelo de comunicación especialmente adecuado para sistemas distribuidos, basado en asociar un proceso servidor al recurso que se desa compartir.
 - Existen dos patrones de comportamiento diferenciados:
 - Servidores: Gestionan un recurso y ofrecen servicios relacionados con el recurso. Reciben peticiones de los clientes, las ejecutan en su nombre y responden a los clientes.
 - Clientes: Envían mensajes de petición a los servidores y esperan respuesta.

Sistemas Operativos I (00-01)

Comunicación v sincronización de Procesos

