

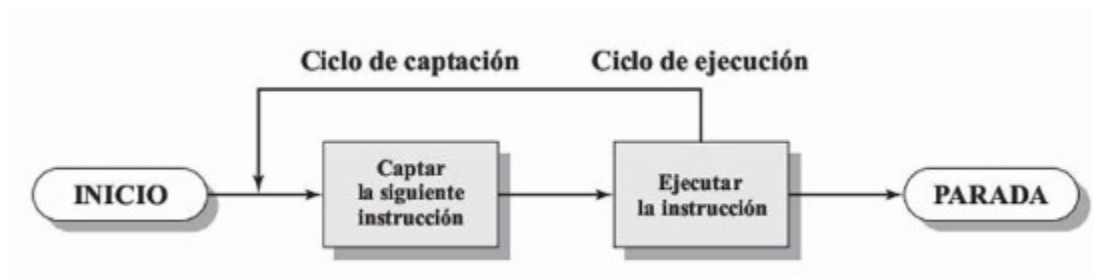


Arquitectura de Computadoras

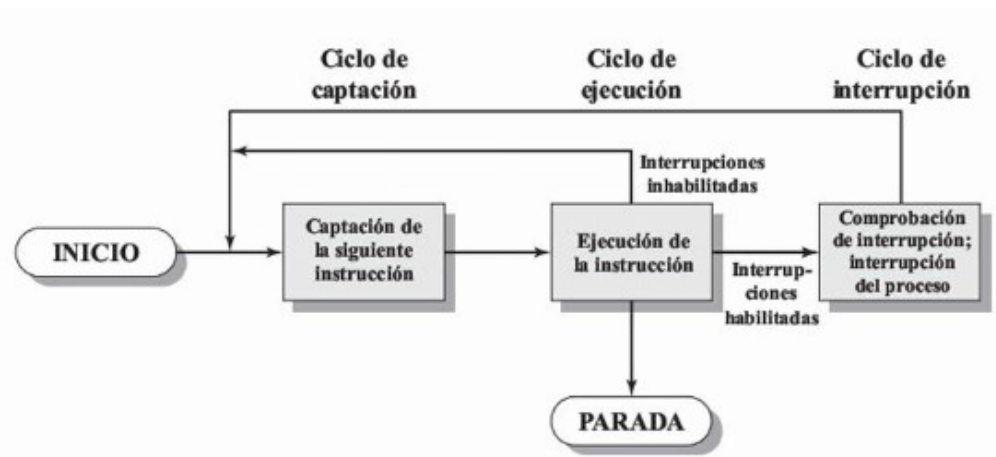
System Buses

Introducción

■ Ciclo de Instrucción básico



■ Ciclo de Instrucción CON INTERRUPCIÓN





Estructuras de Interconexión

- El conjunto de caminos que conectan los distintos módulos de una computadora (CPU, memoria, I/O) se denomina estructura de interconexión
- La estructura de interconexión debe soportar los siguientes tipos de transferencias:
 - Memoria a CPU
 - CPU a Memoria
 - I/O a CPU
 - CPU a I/O
 - I/O a y desde la Memoria (utilizando Direct Memory Access)



Interconexión por Bus

- Un bus es un medio de transmisión compartido pero debe ser utilizado por sólo un dispositivo a la vez
- Contiene tres grupos funcionales de líneas de comunicación
 - Líneas de datos
 - Líneas de direcciones
 - Líneas de control
- Datos
 - El ancho es un elemento clave para determinar el rendimiento del sistema
- Direcciones
 - Identifican la fuente o el destino de los datos que están en el bus
 - El ancho determina la capacidad máxima posible de la memoria del sistema
 - También se utiliza para direccionar puertos de I/O. En general:
 - Los bits de mayor orden seleccionan un módulo particular
 - Los bits de menor orden seleccionan una dirección de memoria o un puerto de I/O



Interconexión por Bus

■ Control

- Controlan el acceso y el uso de las líneas de datos y direcciones
- Algunas líneas típicas son:
 - Lectura de memoria o escritura de memoria
 - Lectura de I/O o escritura de I/O
 - Reconocimiento (ACK) de transferencia
 - Petición y concesión del bus
 - Requerimiento de interrupción y concesión de interrupción
 - Reloj
 - Reset



Interconexión por Bus

- Si un módulo desea enviar datos a otro debe:
 - Obtener el uso del bus
 - Transferir los datos a través del bus
- Si un módulo desea requerir datos de otro debe:
 - Obtener el uso del bus
 - Transferir un requerimiento al otro módulo a través de las líneas de control y direcciones
 - Esperar a que el otro módulo envíe los datos



Configuración física

- Típicamente un bus consiste de
 - Un número de conductores eléctricos en paralelo
 - Cada componente del sistema (normalmente en una o más tarjetas) se conecta a alguna o todas las líneas del bus (slotted conector)
 - El sistema puede ser expandido agregando tarjetas (los slots deben estar previstos)
 - Un componente que falle puede ser reemplazado reemplazando la tarjeta en la que reside

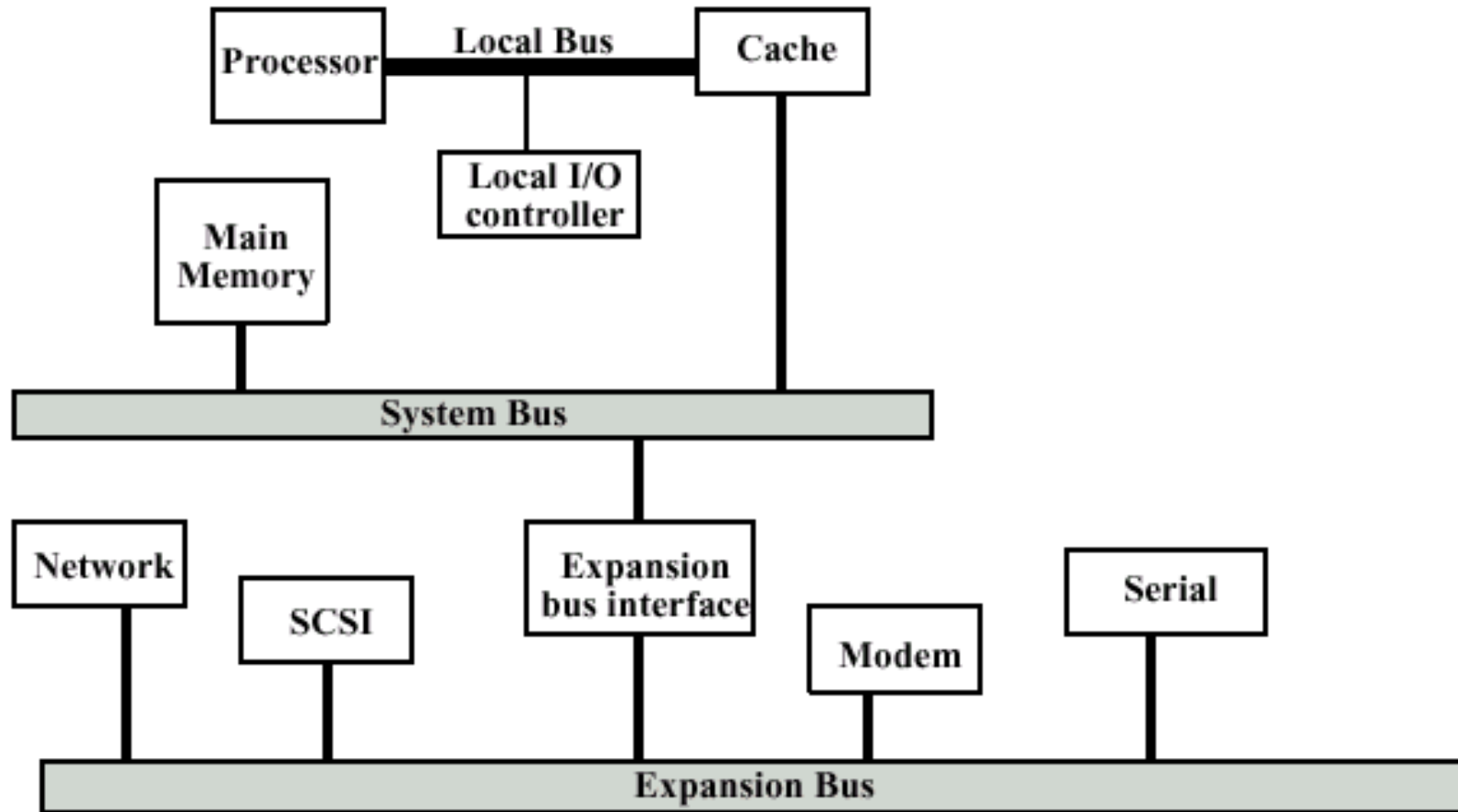


Jerarquía de buses

- A medida que el número de dispositivos conectados a un bus aumenta el rendimiento se degrada debido a que
 - aumenta la demora de propagación (el tiempo que toma a los dispositivos coordinar el uso del bus)
 - el bus puede constituir un cuello de botella a medida que la demanda de transferencia de datos se aproxima a la capacidad del bus
- La solución pasa por establecer una jerarquía de buses
- En una jerarquía tradicional
 - Un bus local conecta la CPU con una memoria cache
 - El bus del sistema conecta la cache, la memoria principal y los módulos de expansión
 - Los buses de expansión (o de I/O) conectan los módulos de expansión (conectados al bus del sistema) y dispositivos de I/O



Una jerarquía de buses tradicional

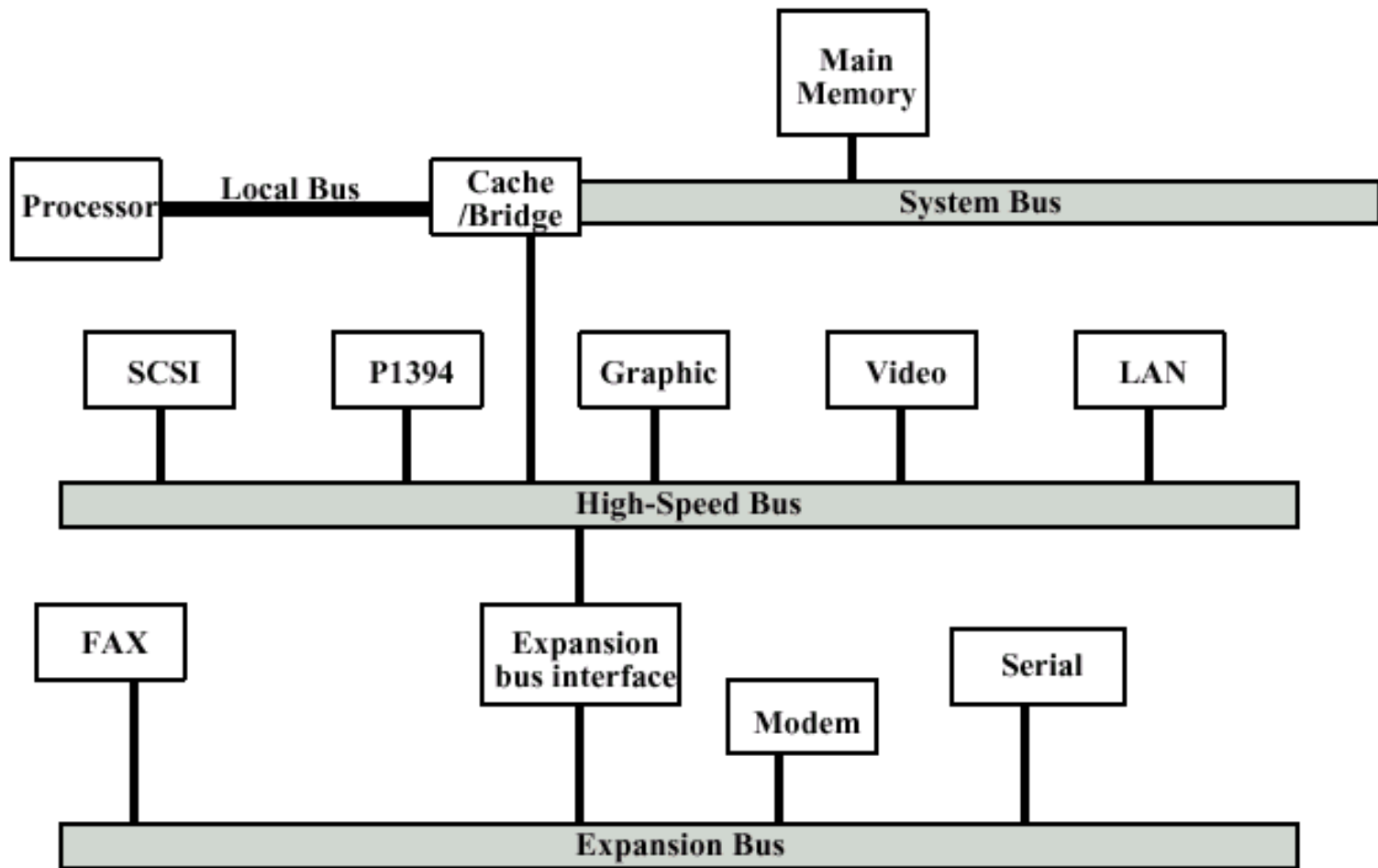




Jerarquía de buses

- Aún la jerarquía tradicional falla a medida que aumenta el rendimiento de los dispositivos de I/O
- Una solución consiste en incorporar un bus de alta velocidad
 - específicamente diseñado para soportar dispositivos de I/O de alta velocidad
 - coloca los dispositivos de alta demanda en una integración estrecha con el procesador al tiempo que los mantiene independientes del procesador
 - los cambios en la arquitectura del procesador no afectan al bus de alta velocidad y viceversa
- Suele denominársela arquitectura de entrepiso

Una jerarquía de buses para alto rendimiento





Elementos de diseño – Tipos de bus

- Dedicado: cada línea está asignada en forma permanente a una función o a un subconjunto de componentes
 - Cuando el bus está dedicado a un subconjunto de componentes se utilizan múltiples buses y hay un módulo adaptador que los conecta y resuelve las disputas al nivel más alto
- Multiplexado en el tiempo: las líneas se utilizan para diferentes funciones en diferentes momentos
 - Se usan menos líneas, lo que significa ahorro de espacio y de costo
 - Se requieren circuitos más complejos en cada módulo
 - Hay una reducción potencial del rendimiento



Elementos de diseño – Método de arbitraje

Determina quien puede utilizar el bus en un momento dado

- Centralizado: un dispositivo denominado controlador del bus o árbitro asigna el tiempo del bus
- Distribuido: cada módulo conectado al bus contiene lógica de control de acceso e interactúa con los otros para establecer quien utiliza el bus
- Ambos métodos designan un dispositivo (CPU o un procesador de I/O) como maestro, quien puede iniciar la transferencia con otro dispositivo, denominado esclavo

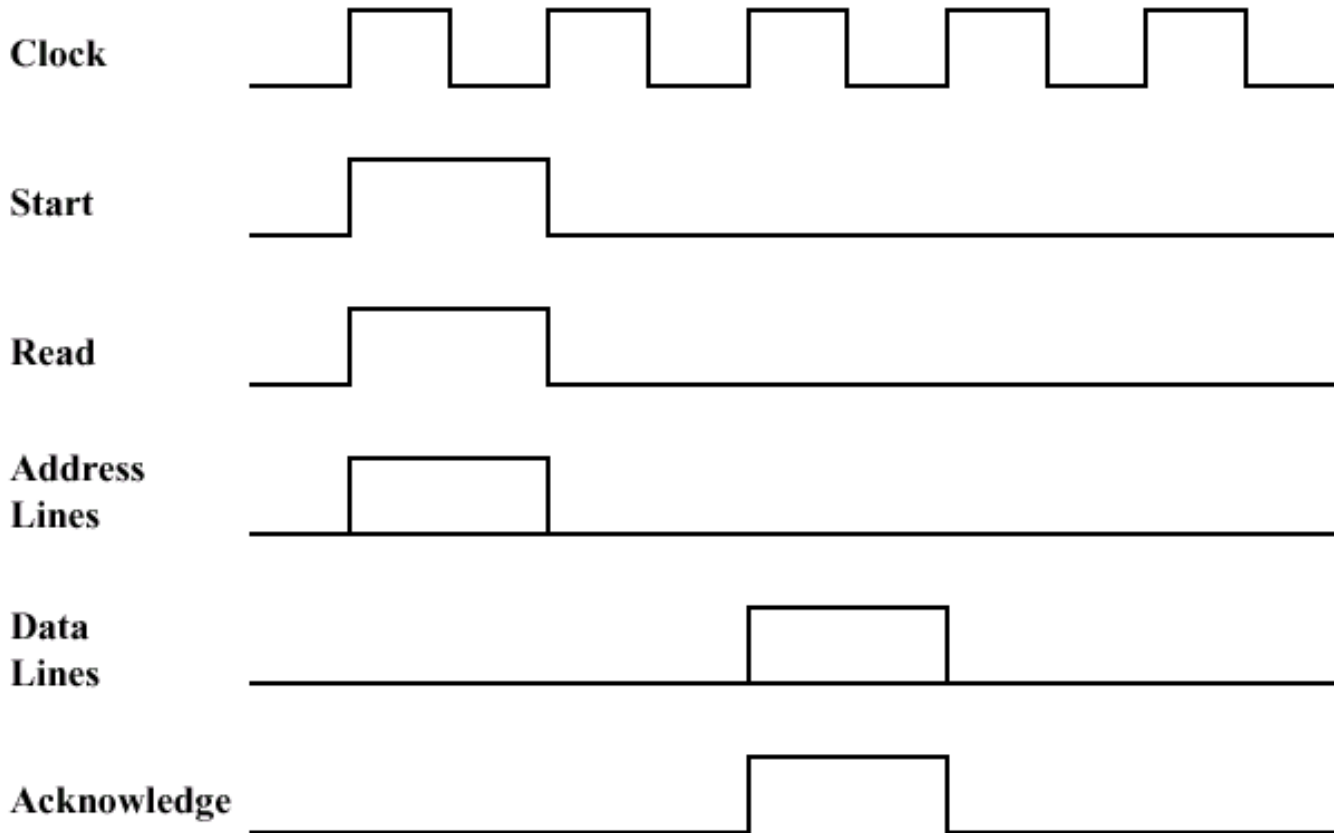


Elementos de diseño – Timing

Se refiere a la forma en que los eventos son coordinados

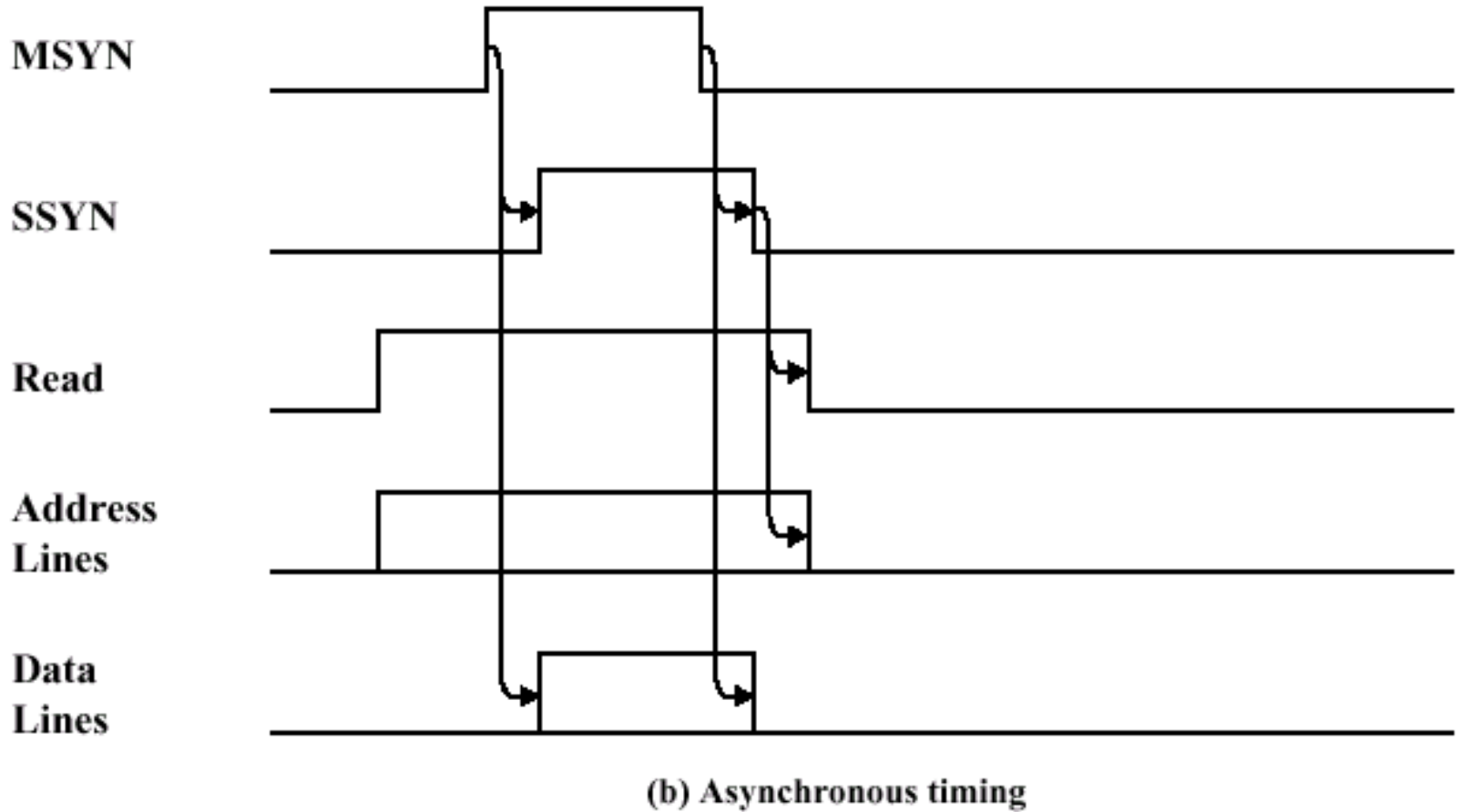
- Sincrónico: La ocurrencia de eventos en el bus está determinada por un reloj que produce una secuencia de unos y ceros alternados de igual duración
 - El reloj es accesible a todos los dispositivos conectados al bus
 - La secuencia de un 1 y un 0 constituye un ciclo de bus (o ciclo de reloj)
 - Los eventos se inician al comienzo de un ciclo de reloj
- Asincrónico: La ocurrencia de un evento sigue a y depende de la ocurrencia de un evento previo
 - Permite tomar ventaja de progresos en el rendimiento de los dispositivos y que una mezcla de dispositivos lentos y rápidos –que utilicen nuevas y viejas tecnologías– compartan el bus
 - Es más difícil de implementar y probar que timing sincrónico

Elementos de diseño - Timing Sincrónico



(a) Synchronous timing

Elementos de diseño - Timing Asincrónico





Elementos de diseño - Otros

- Ancho del bus: se refiere a la cantidad de líneas que componen el bus
- Tipos de transferencia
 - Read (slave to master)
 - Write (master to slave)
 - Read-modify-write: un read es seguido inmediatamente por un write a la misma dirección. La dirección es informada una sola vez -al comienzo de la operación-. La operación es indivisible a fin de prevenir el acceso de otros dispositivos maestros al bus y se utiliza básicamente para proteger la memoria compartida en un sistema multiprogramado
 - Read-after-write: una operación indivisible consistente de un write seguido inmediatamente por un read de la misma dirección. Se utiliza para chequeo.
 - Block transfer: consiste de un ciclo de dirección seguido por n ciclos de datos. El primer item de datos va a la dirección especificada y los siguientes a las direcciones subsecuentes.