

SISTEMAS OPERATIVOS

Arquitectura de computadores

Erwin Meza Vega
emezav@unicauca.edu.co

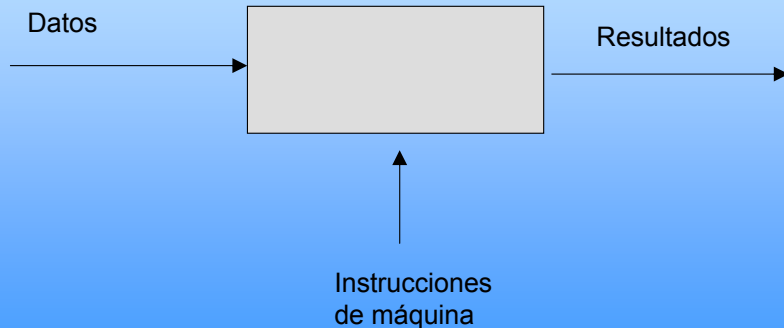
Erwin Meza Vega (emezav@unicauca.edu.co)

Universidad del Cauca

Esta presentación tiene por objetivo mostrar los conceptos generales de la arquitectura de los computadores, necesarios para lograr una mejor comprensión los temas que se van a desarrollar dentro del curso.

Estructura y funcionamiento de los computadores

Visión esquemática



Erwin Meza Vega (emezav@unicauca.edu.co)

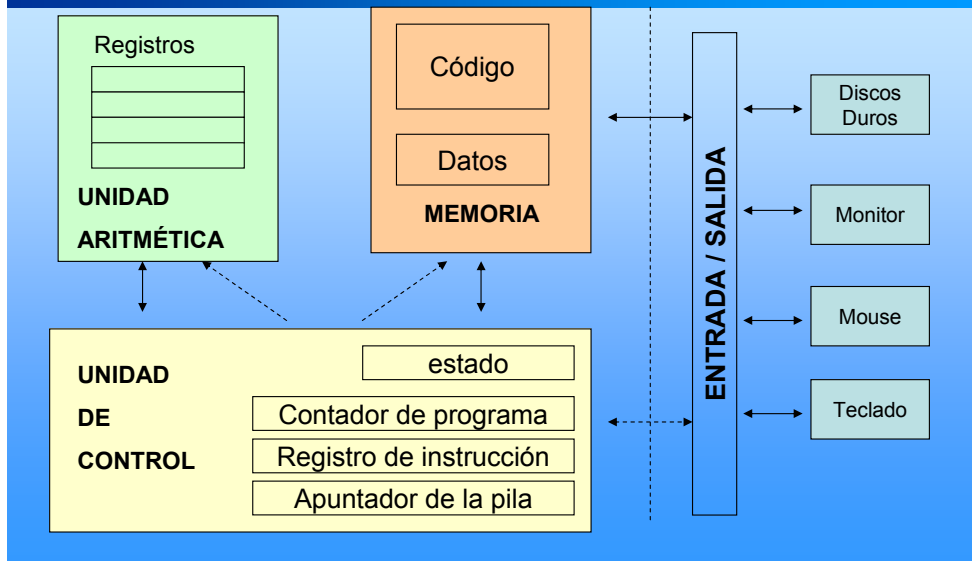
Universidad del Cauca

ESTRUCTURA Y FUNCIONAMIENTO DE LOS COMPUTADORES

Un computador es una máquina que toma como entrada datos e instrucciones propia de su arquitectura, ejecuta una a una las instrucciones, y produce resultados, en forma de datos de salida. De esta forma, la “inteligencia” de los computadores estará dada por la secuencia de instrucciones que éste debe ejecutar.

Las instrucciones de máquina son un conjunto predeterminado de órdenes propias de la máquina, que ha sido especificado por sus fabricantes. Este conjunto de órdenes tiene como objetivo que el computador realice determinadas funciones.

Unidad central de proceso (CPU)



Erwin Meza Vega (emezav@unicauca.edu.co)

Universidad del Cauca

Unidad central de proceso (CPU)

La CPU está compuesta por la unidad aritmética, la unidad de control y la memoria. Adicionalmente se puede contar con dispositivos de entrada/salida, como discos duros, monitor, mouse, etc.

Unidad de control

Se encarga del funcionamiento de todo el conjunto. Su funcionamiento general consiste en:

- Leer una instrucción de la memoria
- Interpretar la instrucción
- Leer los datos de memoria referenciados
- Ejecutar la instrucción, y si es necesario, almacenar el resultado

Registros de control y estado

- Residen en la unidad de control
 - Contador de programa (PC)
 - Cabecera de la pila (SP)
 - Registro de instrucción (IR)
 - Registro de estado

Registros de control y estado

Residen en la unidad de control, dependen de la arquitectura del computador. Los registros principales son:

- Estado: Información producida por alguna de las últimas instrucciones del programa
- Contador de programa (PC – Program Counter): Almacena la dirección de la siguiente instrucción
- Registro de instrucción (IR – Instruction Register): almacena la instrucción que se va a ejecutar
- Apuntador a la pila (SP – Stack Pointer): Maneja la pila en memoria principal

Modelo de programación del computador

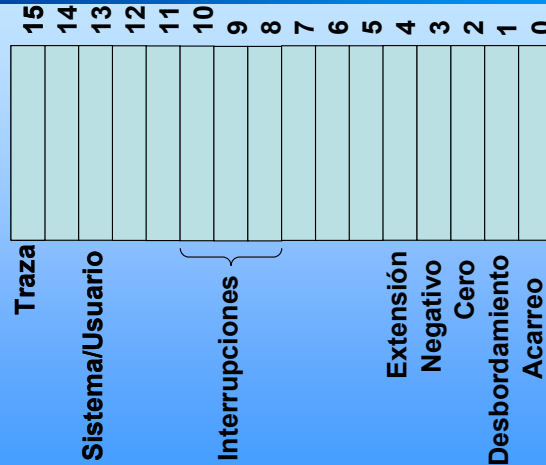
- Elementos de almacenamiento
 - Registros generales
 - PC
 - SP
 - SR
 - RAM
 - Mapa E/S
- Juegos de instrucciones
- Secuencia de funcionamiento

Modelo de programación del computador

El modelo de programación del computador está compuesto de:

- Elementos de almacenamiento: Son aquellos elementos visibles a las instrucciones de máquina, tales como los registros generales, el contador de programa, el apuntador a la pila, el registro de estado, la memoria RAM y el mapa de Entrada/Salida.
- Juegos de instrucciones: Son aquellas operaciones que puede realizar la máquina, unido a los modos de direccionamiento, que especifican cómo se identifican los elementos de almacenamiento que intervienen en las instrucciones.
- Secuencia de funcionamiento: Modo en que se ejecutan las instrucciones de máquina.

Registro de estado



Erwin Meza Vega (emezav@unicauca.edu.co)

Universidad del Cauca

Registro de estado

Almacena la información producida por alguna de las últimas instrucciones del programa. El valor de los bits de este registro puede ser usado para conocer el estado de una operación aritmética o la información sobre la forma en que se debe comportar la máquina. El registro de estado se compone de:

- Bits aritméticos
 - Signo: Almacena el signo de la última operación realizada
 - Acarreo: Se activa si la última suma o resta realizada posee acarreo
 - Cero: Se activa si el resultado de la última operación es cero
 - Desbordamiento: Se activa si la última operación produjo desbordamiento
- Bits de nivel de ejecución: Permiten establecer el nivel de ejecución
- Bits de control de interrupciones: Interrupciones que puede aceptar el procesador

Niveles de ejecución

- Nivel de usuario: Instrucciones de máquina restringidas, acceso a determinados registros
- Nivel de núcleo: Todas las instrucciones, acceso sin restricción

Nivel de usuario

Este nivel es menos permisivo, algunas de las instrucciones de máquina se encuentran restringidas (entrada/salida, gestión de memoria), y sólo se tiene acceso a determinados registros. El nivel de usuario sólo permite modificar los bits 0 a 5 del registro de estado, y el apuntador de la pila del usuario

Nivel de núcleo (kernel)

En este nivel no se tiene restricción en las instrucciones, los registros generales ni el registro de estado.

Secuencia de funcionamiento del computador

Ejecución lineal:

- Lectura de la RAM
- Incrementar PC
- Ejecutar la instrucción

Mecanismos para romper la ejecución lineal

- Instrucciones de salto o bifurcación
- Interrupciones

Secuencia de funcionamiento del computador

El computador funciona de manera secuencial, con los siguientes pasos:

1. Leer de la RAM la instrucción apuntada por el Program Counter (PC)
2. Incrementar el PC
3. Ejecutar la instrucción leída de memoria

Se debe contar con instrucciones que permitan saltar a otra instrucción, diferente a la siguiente en secuencia. Para ellos se puede contar con instrucciones de salto o bifurcación e interrupciones. Las instrucciones de bifurcación permiten que el programa rompa su ejecución y salte a **otro segmento del mismo programa**. Las interrupciones (internas o externas) modifican el valor del PC y lo hacen saltar a **otro programa**.

Interrupciones

- Señales que llegan a la unidad de control
 - Excepciones
 - Interrupciones de reloj
 - Interrupciones de E/S
 - Excepciones de Hardware
 - Instrucción TRAP
- Pueden manejar prioridades

Las interrupciones son uno de los mecanismos que tienen los modelos de computación para interrumpir el funcionamiento secuencial de la CPU, y de esta manera permitir que se ejecute otro programa.

Las excepciones se pueden producir por desbordamientos, errores de entrada/salida, direccionamiento errado.

Las interrupciones de reloj se ejecutan en un intervalo de tiempo determinado.

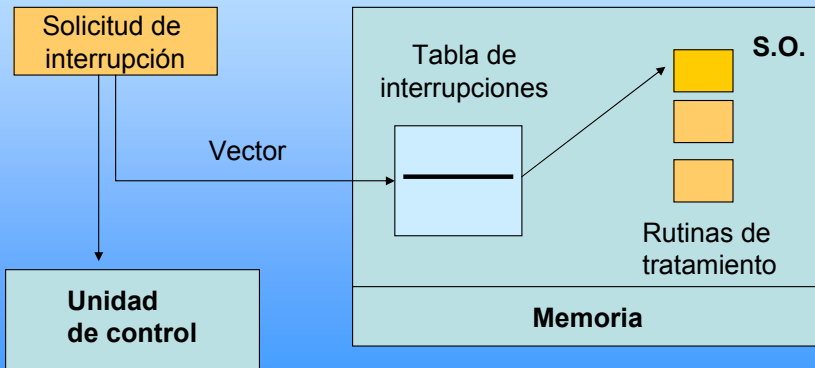
Las interrupciones de Entrada/Salida permiten avisar a la CPU de la ocurrencia de un evento.

Las excepciones de hardware se producen cuando ocurre un error de memoria, o corte de corriente

La instrucción TRAP permite que un programa lance una interrupción

De otro lado, las interrupciones pueden manejar prioridades, con el fin de distinguirlas en orden de relevancia.

Esquema de una interrupción



Erwin Meza Vega (emezav@unicauca.edu.co)

Universidad del Cauca

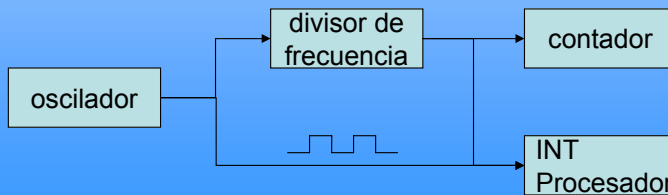
A nivel físico, una interrupción es una señal que se envía a la unidad de control. El agente que solicita la interrupción debe proporcionar un vector que especifica la dirección de la rutina de tratamiento. Este proceso se lleva a cabo cuando termina la ejecución de la instrucción de máquina actual, y consiste en las siguientes operaciones:

- Salvar algunos registros (Estado y pc)
- Pasar a modo núcleo
- Cargar un nuevo valor en el PC

El nuevo valor cargado en el PC corresponde a la dirección que ha sido proporcionada por el agente que solicitó la interrupción, con esto se activa la rutina de tratamiento. Luego de ejecutar dicha rutina, se recupera el estado almacenado y se pasa al modo usuario.

RELOJ

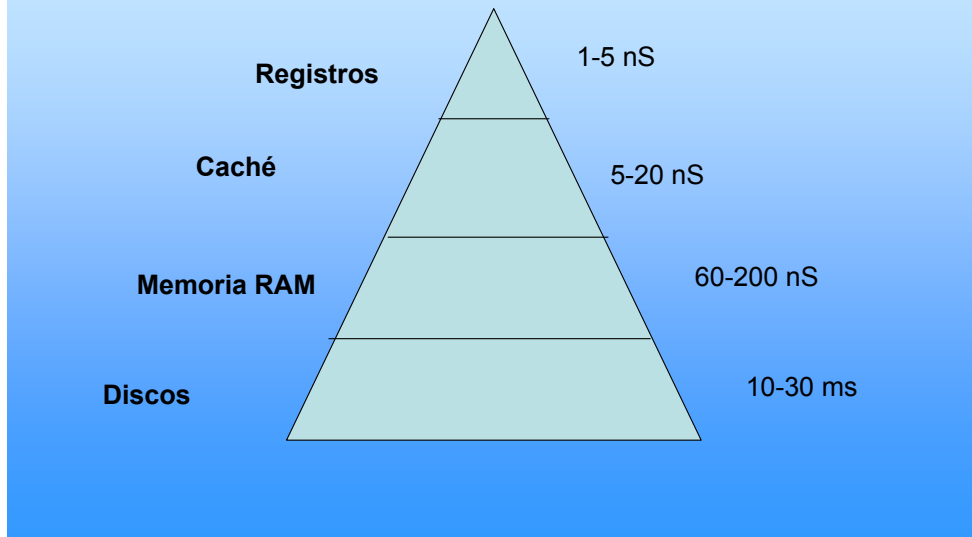
- Señal que gobierna el ritmo de ejecución de las instrucciones de la máquina
- Generador de interrupciones periódicas
- Contador de fecha y hora



El concepto de reloj se puede aplicar en tres ámbitos:

- Señal que gobierna el ritmo de la CPU, y determina la frecuencia de ejecución (Velocidad)
- Temporizador que envía interrupciones periódicas al procesador
- Acumulador de segundos, a partir del cual se puede calcular la fecha y la hora.

Jerarquía de memoria



Erwin Meza Vega (emezav@unicauca.edu.co)

Universidad del Cauca

Jerarquía de memoria

La memoria se puede dividir de acuerdo a su velocidad de acceso, en donde se tiene:

- Registros del procesador: Dispositivos de almacenamiento temporal, cuya velocidad de acceso varía entre 1 a 5 nano-segundos Su capacidad varía de 64 a 256 bytes.
- Memoria caché: Integrada al procesador, es un dispositivo de almacenamiento temporal, cuya velocidad de acceso varía entre 1 a 5 nano-segundos. Su capacidad varía de 8 kB a 2 MB.
- Memoria RAM: Almacenamiento temporal, su velocidad varía de 60 a 200 nano-segundos. Su capacidad varía de 64 MB a 64 GB.
- Discos: Almacenamiento permanente, su velocidad de acceso varía de 10 a 30 mili-segundos, y su capacidad de almacenamiento varía de 50 MB a 809 GB.

Memoria Virtual

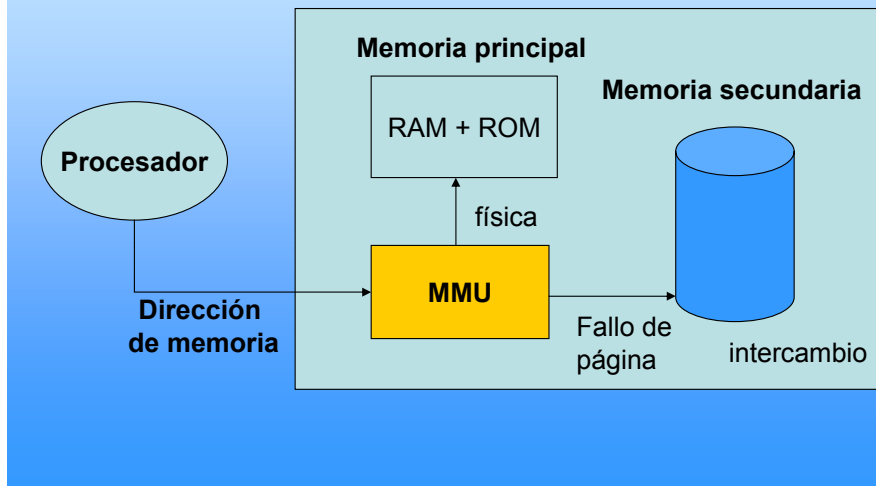
- El S.O. debe dar la impresión de tener memoria ilimitada
- Utiliza dos componentes de la jerarquía de memoria
 - Principal (RAM)
 - Respaldo (Discos)
- La gestiona el S.O. con ayuda del hardware

Memoria Virtual

Debido a la necesidad de ejecutar varios procesos al tiempo, se requiere mayor cantidad de memoria. Sin embargo, el computador posee una memoria RAM limitada; para solucionar este problema se puede hacer uso de otro componente de memoria (discos) para que el sistema la utilice como memoria RAM.

El manejo de memoria virtual es transparente para el usuario, ya que es administrada por el sistema operativo con ayuda del Hardware. De esta forma, el sistema da la impresión de tener una cantidad ilimitada de RAM.

Memoria virtual



Erwin Meza Vega (emezav@unicauca.edu.co)

Universidad del Cauca

Memoria Virtual

La memoria total se divide en páginas, que son cargadas en un *espacio de memoria virtual*, compuesto por la RAM, la ROM y en un dispositivo de almacenamiento secundario. De esta forma, las direcciones generadas por las instrucciones de máquina se refieren al espacio virtual, que puede estar compuesto por un conjunto de páginas en memoria principal y otras en memoria secundaria.

Sin embargo, para que un programa pueda ser ejecutado, debe residir en memoria principal, por esto se debe realizar un intercambio de páginas, en el cual se llevan a la memoria secundaria aquellas páginas que no están siendo utilizadas actualmente y se cargan en memoria principal las páginas que el programa necesita.

Los espacios virtual y físico se dividen en páginas virtuales y páginas de intercambio. Cuando un programa solicita una dirección de memoria (en el espacio virtual) que no se encuentra en memoria principal, se produce un fallo de página que debe ser atendido por el sistema operativo, para cargar la página correspondiente a memoria principal.

Entrada/Salida

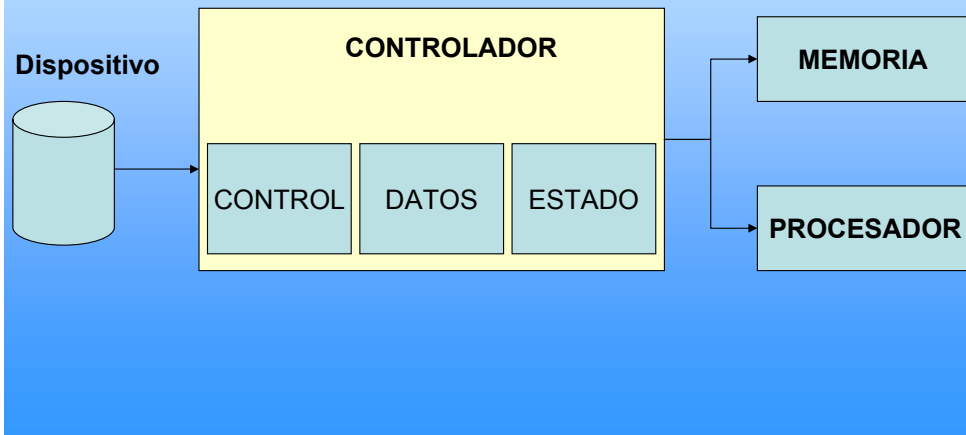
- Intercambio de información entre los dispositivos periféricos, la memoria y los registros del procesador.
- Dispositivos periféricos
 - Bloque: Discos duros, CDs, DVDs, Red
 - Carácter: Terminales
- En los dispositivos de bloques la E/S se realiza a nivel de sector
 - 256, 512 o 1024 bytes.

Entrada y salida

Se puede definir como el intercambio de información entre los dispositivos periféricos, la memoria y los registros del procesador. Las operaciones de entrada/salida se pueden realizar a nivel de bloque y a nivel de carácter.

- Dispositivos de bloque: Cada vez que se realiza una operación de E/S, se intercambia un conjunto de bloques de información
- Dispositivos de carácter: Cada operación intercambia un carácter

Dispositivos de Entrada/Salida



Erwin Meza Vega (emezav@unicauca.edu.co)

Universidad del Cauca

Dispositivos de Entrada/Salida

Todo dispositivo de entrada/salida debe tener además una pieza de software/hardware llamada controlador, que posee registros de control, datos y estado. Según su configuración, un dispositivo de entrada/salida puede leer o escribir directamente en la RAM o en los registros del procesador o lanzar interrupciones para informar que existen datos. El sistema operativo también puede monitorear constantemente los registros de estado del dispositivo para verificar si existe un cambio.

Entrada/salida y concurrencia

- Operación de E/S
 - Escribir la orden en los registros del controlador
 - Leer / escribir los datos
 - Fin de operación
- Tipos de E/S
 - No concurrente: E/S programada
 - Concurrente : Interrupciones, Acceso directo a memoria (DMA)

Entrada/salida y concurrencia

De acuerdo a la concurrencia, las operaciones de E/S se pueden dividir en:

- No concurrente: El sistema operativo es el encargado de monitorear los eventos que ocurren en los dispositivos de entrada/salida (espera activa). No es conveniente porque se consume tiempo de procesador
- Concurrente
 - Interrupciones: EL procesador envía una orden al controlador y se dedica a otra tarea. Cuando el dispositivo tiene datos disponibles, genera una interrupción. El procesador lanza una rutina de interrupción, encargada de leer el dato y almacenarlo en memoria.
 - Acceso directo a memoria (Direct Memory Access – DMA): El controlador del dispositivo almacena el dato directamente en memoria y genera una interrupción.

Referencias

- CARRETERO Jesús, y otros. Sistemas Operativos. Una visión aplicada. Editorial McGraw-Hill.
- TANENBAUM, Andrew. Modern Operating Systems. 2nd edition. Prentice Hall

Créditos

Autor: Erwin Meza Vega
Universidad del Cauca
emezav@unicauca.edu.co