



INSTITUTO POLITÉCNICO NACIONAL

Unidad Profesional Interdisciplinaria en Ingeniería y
Tecnologías Avanzadas

Programa académico:
Ingeniería Mecatrónica

Unidad de aprendizaje:
Microcontroladores, Microprocesadores e Interfaz

Tarea 1. Investigación de temas de Arquitectura de Computadoras

Alumna:
Linda Fernández



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

Tabla de Contenido

1. Arquitectura de computadoras	4
1.2 ¿Qué es la arquitectura de computadoras?	4
1.2 Evolución de la tecnología	5
1.3 Ley de Moore	9
1.4 Computadora	10
1.4.1 Funciones, organización y arquitectura	10
1.4.1 Estructura y funcionalidad.....	11
1.4.2 Estructura en alto, medio y bajo nivel	14
1.4.3 Plataformas informáticas	15
1.5 Arquitectura de Hardware	17
1.6 Arquitectura Von Neumann.....	19
1.7 Arquitectura Harvard.....	21
1.8 CPU y sus componentes.....	22
1.9 Microprocesador y sus partes lógicas	24
1.10 ALU, funciones y componentes	26
1.11 Unidad de control y sus componentes	27
1.12 Unidades funcionales.....	29
1.13 Bus y tipos de buses	29
1.14 Fases de ejecución (de una instrucción).....	31
1.15 Memorias, sus jerarquías y sus funciones	33
1.16 Registros y tipos de registros.....	40
1.17 La memoria caché	42
1.18 Norma IEEE754	44
2. Circuitos Integrados	47
2.1 Terminología de los circuitos integrados digitales.....	48
2.2 Características estándar de la serie TTL, CMOS y LVTs.....	52
2.2.1 Series de la tecnología TTL.....	52



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

2.2.2 Series de la tecnología CMOS.....	55
2.2.3 Series de la tecnología LVTs.....	59
2.3 Definir y ejemplificar los Circuitos Integrados:	62
2.3.1 De tres estados.....	62
2.3.2 Buffer	64
2.3.3 De Colector abierto.....	65
2.4 Conmutador bilateral (como el 74LS245).....	66
2.5 Interfaz de CMOS de alto voltaje con TTL	67
2.6 Interfaz de CMOS de bajo voltaje con TTL	67
2.7 Interfaz de TTL con CMOS de bajo voltaje	68
2.8 Interfaz de TTL con CMOS de alto voltaje	68
2.9 Interfaz con tecnologías LVT.	69
3. Memoria	71
3.1 Conceptos generales de la memoria	71
3.2 Tecnologías de memorias ROM (antecedentes y evolución).....	72
3.3 Tecnologías de memorias RAM (antecedentes y evolución)	75
3.4Características principales de las memorias.....	78
4. Referencias	81



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

1. Arquitectura de computadoras

1.2 ¿Qué es la arquitectura de computadoras?

Es el diseño conceptual y la estructura operacional fundamental de un sistema computacional, esto quiere decir que es una descripción funcional de los requerimientos y las implementaciones de diseño de las partes que componen una computadora, sobre todo en la manera en que la unidad central de proceso UCP trabaja internamente y accede a las direcciones de memoria.

La computadora envía y recibe información a través de los periféricos que se adaptan a su infraestructura por medio de los canales. La UCP es la unidad que lleva el procesamiento de la información que la computadora recibe.

La arquitectura también se refiere a los atributos visibles de un sistema, estos tienen un impacto directo en la ejecución lógica de un algoritmo o programa, esta se diferencia de la organización de computadoras pues no se refiere a las unidades funcionales y sus interconexiones.

Algunos atributos arquitectónicos pueden ser el número de bits que representan diferentes tipos de datos, el conjunto de instrucciones, mecanismos y técnicas de direccionamiento de memoria.

Las computadoras son sistemas complejos, actualmente se conforman de millones de componentes electrónicos. Su evolución se caracteriza en un incremento de la velocidad del procesador, una disminución del tamaño de los componentes, un aumento del tamaño de memoria, de la velocidad y de la capacidad E/S.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

Para obtener un aumento de la velocidad del procesador es necesario disminuir el tamaño de los componentes del microprocesador, ya que la distancia entre componentes se reduce y las señales navegan más rápidamente.

Los cambios en la tecnología permiten la introducción de arquitecturas más ricas y potentes.

1.2 Evolución de la tecnología

La evolución de las computadoras es una larga historia que se puede explicar por generaciones según William Stallings. La clasificación de generaciones se basa en la tecnología hardware empleada, esta clasificación se caracteriza por la mayor velocidad, capacidad de memoria y menor tamaño de la generación anterior.

La Segunda Guerra Mundial trajo consigo necesidades computacionales militares, era necesario un computador de propósito general usando tubos de vacío.

El primer computador ENIAC (Electronic Numerical Integrator And Computer) fue diseñado y construido entre 1943 y 1946 bajo la supervisión de John Mauchly y John Presper Eckert en la Universidad de Pennsylvania, se terminó demasiado tarde para ser utilizado en la guerra, su primera misión fue realizar los cálculos complejos con los que se determinó la viabilidad de la bomba de hidrógeno.

El ENIAC era una máquina decimal y no binaria, su memoria consistía en acumuladores capaces de contener un número decimal de diez dígitos. Cada dígito estaba representado por un anillo de diez tubos de vacío. Este computador debía ser



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

programado mediante conmutadores, conectando y desconectando cables manualmente.

Este proceso de programación era una tarea complicada, es por eso que se propone el concepto del programa-almacenado, un proceso donde se presenta el programa y se guarda en la memoria junto a los datos, esta idea se le atribuye a John von Neumann, un matemático asesor del proyecto ENIAC. Esta idea fue desarrollada al paralelo por Alan Turing.

En 1946 von Neumann y algunos compañeros comenzaron el diseño de un nuevo computador de programa-almacenado llamado IAS. Este computador consta de una memoria principal que almacena tanto datos como instrucciones, una unidad aritmicológica (ALU) capaz de hacer operaciones con datos binarios, una unidad de control que interpreta las instrucciones en memoria y provoca su ejecución y por último un equipo de entrada-salida dirigido por la unidad de control.

La mayoría de las computadoras actuales tienen la misma estructura general y funcionamiento que las máquinas de von Neumann.

Para los años 50s ya existían compañías como Sperry e IBM. En 1947 se fundó Eckert-Mauchly Computer Corporation y fabricaban computadoras con fines funcionales. La primera máquina de éxito fue el UNIVAC I, que tenía tanto aplicaciones científicas como comerciales y años después se siguieron construyendo una serie de máquinas sucesoras de la primera.

IBM era en ese entonces el principal fabricante de equipos de procesamiento con tarjetas perforadas. Su primer computador 701 (1953), fungía principalmente aplicaciones científicas. La serie 700/7000 situó a IBM como el fabricante de computadoras dominante.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

Un cambio importante en la construcción de computadoras fue cuando se sustituyeron los tubos de vacío por transistores. Estos elementos más pequeños y baratos disipan menos calor y son usados de la misma manera que los tubos de vacío.

El transistor fue inventado en los laboratorios Bell en 1947 y trajo consigo la revolución electrónica de los años 50s. Las computadoras completamente transistorizadas estuvieron disponibles comercialmente hasta el final de la década. Los transistores definen la segunda generación de computadoras, se introducen también las unidades lógicas y aritméticas y unidades de control más complejas, el uso de lenguajes de programación de alto nivel y se crea un software del sistema para interactuar con la computadora.

Estas computadoras hacían uso de canales de datos, que son módulos de E/S independientes con su propio procesador y conjunto de instrucciones. De esta manera la CPU no ejecuta las instrucciones detalladas de entrada y salida, estas se almacenan en una memoria principal y se ejecutan con el procesador de uso específico para su referente canal de datos. La CPU envía señales de control al canal de datos, así instruye para ejecutar la secuencia de instrucciones en la memoria. Este procedimiento libera a la CPU de la carga de procesamiento. El multiplexor es el punto de conexión central de los canales de datos, la CPU y la memoria. Este dispositivo organiza los accesos a la memoria desde la CPU y los canales de datos, lo que permite que actúen de manera independiente.

La segunda generación se caracterizaba por el uso de transistores simples conocidos como componentes discretos. A través de los años 50s la electrónica se basaba en el uso de dispositivos compuestos por transistores. Generar estos dispositivos resultaba costoso ya por su manera de producción, lo que produjo el surgimiento de problemas en la industria.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

En 1958 comenzó la era de la microelectrónica, debido a la invención del circuito integrado. Entonces los elementos básicos de la computadora pasaron a usar puertas y celdas de memoria. Las puertas son dispositivos que implementan una función lógica o booleana simple, controlan el flujo de información. Las celdas de memoria por su parte son dispositivos que almacenan datos de un bit.

Las computadoras básicas comenzaron a componerse por elementos como el almacén de datos (celdas de memoria), procesamiento de datos (puertas lógicas), transferencias de datos (puertas lógicas) y un control.

Los circuitos integrados simplificaron la arquitectura de las computadoras, ya que los circuitos que antes se tenían ahora se podían reducir e implementar en una pequeña placa de silicio. Con el paso del tiempo era posible encapsular más componentes en pequeños circuitos integrados.

Se estima que en las siguientes generaciones existirán más de 1000 componentes en un solo chip de un circuito integrado.

En el caso de los microprocesadores conforme el tiempo pasaba se volvían más completos. Intel desarrolló el 4004 en 1971, contenía todos los componentes de la CPU en un solo chip. Este procesador es muy primitivo, sin embargo, comenzó con la evolución continua de la capacidad y potencia en los microprocesadores. Después surgieron los procesadores 8008 (1972) y 8080 (1974), este último fue el primer microprocesador de uso general.

En los 80s comenzaron a desarrollarse los microprocesadores de 16 bits. El 8086 de Intel© y posteriormente los Laboratorios Bell y Hewlett-Packard© desarrollaban microprocesadores de un solo chip de 32 bits.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

En la siguiente década, surgen más procesadores cada vez más potentes, compuestos por millones de procesadores, con memorias de capacidad medida el Gigabytes, como lo fueron el 486TM SX (1991), Pentium (1993), Pentium Pro (1995), Pentium II (1997). Estos procesadores contaban con un ancho de bus de hasta 64 bits.

Los procesadores más recientes compiten y apuestan por velocidades de reloj cada vez más grandes, varios núcleos en el procesador y tamaños cada vez más pequeños.

1.3 Ley de Moore

Esta ley fue propuesta por Gordon Moore, cofundador de Intel©, en 1965. Moore se dio cuenta de que el número de transistores integrables en un solo chip se duplicaba cada año y se predecía correctamente que esto continuaría en años futuros.

La ley de Moore tuvo consecuencias en los siguientes años, pues el ritmo continuó por varias décadas. Los costos de los circuitos o chips encapsulados se mantuvieron invariables a pesar del rápido crecimiento en densidad. Podemos interpretar esto como la rápida disminución del precio de la lógica computacional y los componentes electrónicos de las memorias.

Otra consecuencia positiva es que entre menos distancia hay en los componentes y la disminución de la longitud de las interconexiones eléctricas ha incrementado la velocidad operativa, además las computadoras han disminuido su tamaño porque no requieren de elementos tan grandes en su arquitectura, lo que también demanda mucha menos potencia y refrigeración.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

1.4 Computadora

Podemos definir a una computadora como un sistema digital con tecnología microelectrónica que tiene la capacidad de procesar datos a partir de un grupo de instrucciones a través de software. Se compone de microprocesador (CPU), memoria y dispositivos de entrada/salida (E/S) e interconexiones entre ellos.

1.4.1 Funciones, organización y arquitectura

Una computadora se compone por una unidad central de procesamiento (CPU), memoria y las unidades de entrada y salida.

Su organización se refiere a las unidades funcionales e interconexiones, lo cual permite identificar las especificaciones arquitectónicas, a diferencia, la arquitectura de computadoras como se menciona al principio se refiere a los atributos de un sistema que tienen un impacto directo en la ejecución lógica de un programa.

Entre los atributos arquitectónicos podemos encontrar los números de bit usados para representar varios tipos de datos, mecanismos de entrada y salida y técnicas para direccionamiento de memoria, en cambio los de organización son aquellos atributos de hardware invisibles para el/la programador/a, como

El sistema funciona gracias a una fuerte relación entre la arquitectura y la organización, en el intercambio de señales de datos y control entre sus componentes en caso de usar hardware específico y en caso del hardware de uso general acepta datos y señales de control y produce resultados. De esta manera no es necesario reconfigurar el hardware, lo que cambia es la nueva secuencia de códigos, estos códigos son instrucciones y el hardware las interpreta. Este método de programación se denomina software.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

La organización de una computadora debe ser diseñada específicamente para implementar los requerimientos de la arquitectura particular. Los fabricantes ofrecen series de modelos que poseen la misma arquitectura computacional, pero con diferencias en la organización. Es por eso que la serie de modelos tienen precios y características distintas, en la memoria, el procesador, etc. Las arquitecturas sobreviven sobre todo porque no cambian tan frecuentemente como la organización lo hace, ya que esta depende de actualizaciones y su evolución tecnológica es muy rápida.

1.4.1 Estructura y funcionalidad

Para comprender la estructura de las computadoras es necesario reconocer la jerarquía de los sistemas y analizarlos. Un sistema jerárquico se compone de subsistemas interrelacionados que se organiza en una estructura jerárquica hasta alcanzar el nivel mínimo del subsistema elemental.

Cada nivel consiste en componentes y sus interrelaciones, su comportamiento depende de la caracterización abstracta del siguiente nivel más bajo.

La estructura es el modo en que los componentes están interrelacionados y el funcionamiento se basa en la operación de cada componente individual como parte de la estructura.

En la figura 1.1 se ilustran las funciones básicas de la computadora.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

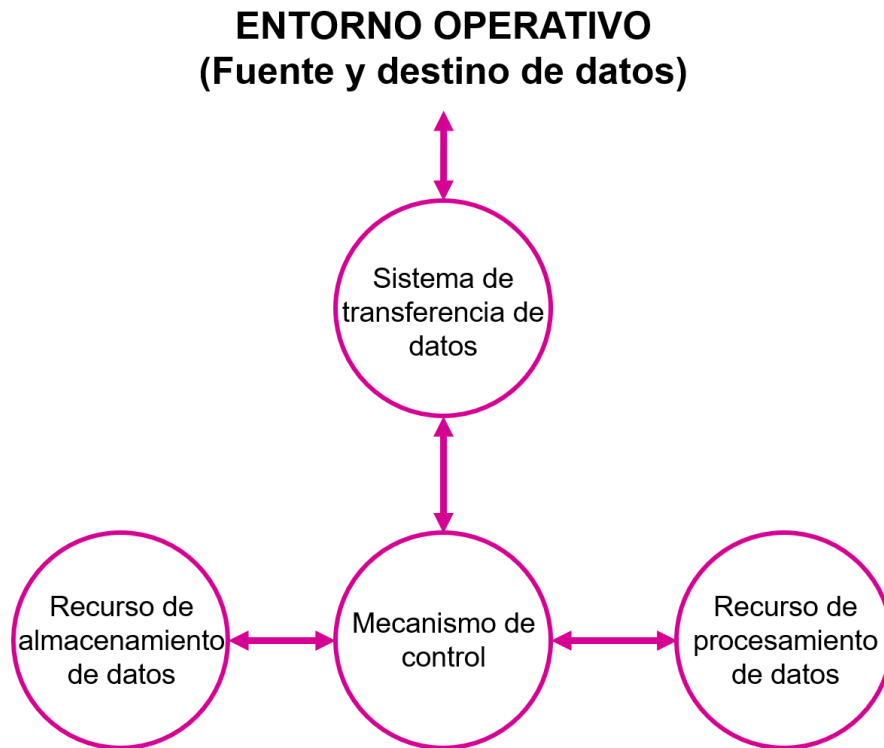


Figura 1.1 Funciones de las computadoras

El funcionamiento se basa en el procesamiento, almacenamiento, transferencia y control de los datos. Estos datos pueden ser de diferentes tipos y su procesamiento depende del rango de requisitos.

Aún cuando el procesamiento de datos está en función la computadora debe ser capaz de almacenar datos. Es necesario guardar temporalmente los datos requeridos para el proceso que se realiza. El almacenamiento puede ser de largo o corto plazo dependiendo de la necesidad de recuperar o no los datos en el futuro.

La computadora también transfiere datos, es la manera de comunicación del sistema. El entorno de operación está formado por dispositivos que son fuentes o receptores



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

de datos. A este proceso se le conoce como entrada-salida o E/S y a los dispositivos se les conoce como periféricos.

En la computadora existe también una unidad de control que gestiona los recursos y dirige los atributos funcionales dependientes de las instrucciones de las funciones anteriores.

La computadora es una entidad que debe interactuar con el entorno, por medio de la transferencia de datos. Todas sus conexiones se clasifican como dispositivos periféricos o líneas de comunicación. La estructura de las computadoras se conforma por cuatro principales componentes.

El primero es la Unidad Central de Procesamiento (CPU), esta controla el funcionamiento y produce el procesamiento de datos. Es conocido como el procesador.

Posteriormente tenemos la memoria principal que se encarga de almacenar los datos a grosso modo.

Los periféricos o líneas de comunicación, otro componente principal, transfiere datos entre la computadora y el entorno.

Por último, el sistema de interconexión que proporciona la comunicación entre la CPU, la memoria principal y los periféricos.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

1.4.2 Estructura en alto, medio y bajo nivel

La Unidad Central de Procesamiento (CPU) estructuralmente está compuesta por una unidad de control, la unidad aritmético-lógica (ALU), los registros y las interconexiones de la CPU.

Podemos entender la computadora por niveles como se muestra en la figura 2.1.

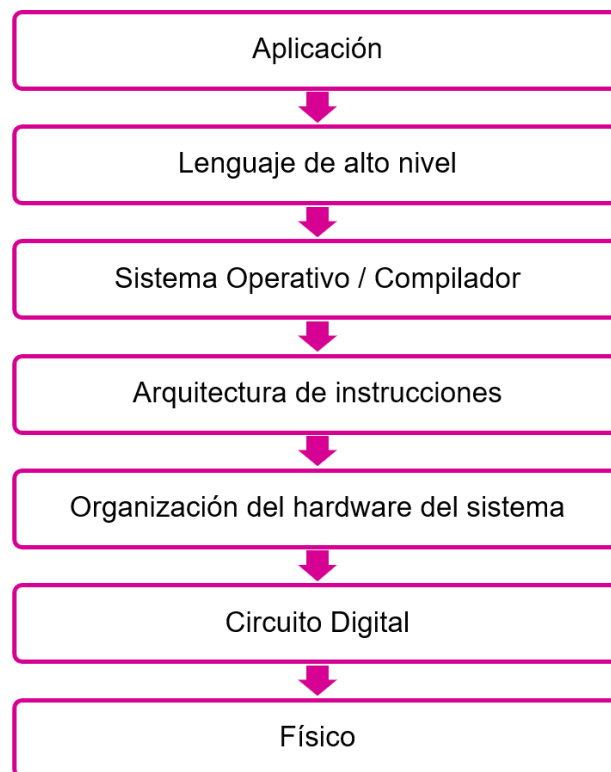


Figura 1.2 Niveles descriptivos de las computadoras

En el nivel de aplicación encontramos el software con el que interactúa el usuario, posteriormente el lenguaje de alto nivel interpreta las instrucciones del usuario. El sistema operativo gestiona la memoria, los procesos, los ficheros y realiza las acciones de compilación, enlazado y ubicación de la memoria. La arquitectura



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

procesa las instrucciones y genera que la organización del sistema realice las interconexiones necesarias para que los circuitos y microchips realicen las funciones lógicas necesarias para cumplir las instrucciones del usuario.

1.4.3 Plataformas informáticas

Las plataformas son los tipos de procesadores u otro hardware en el que se ejecuta el sistema operativo y la combinación del tipo de hardware.

Los sistemas operativos son programas que controlan la ejecución de los programas de aplicación y son la interfaz entre el usuario y el hardware de la computadora. Permite que los recursos se utilicen de forma eficiente.

El usuario final no necesariamente debe conocer la arquitectura de la computadora para poder usarla y eso se debe a la interfaz del sistema operativo.

Este proceso se puede interpretar por capas como lo podemos ver en la figura 1.3.

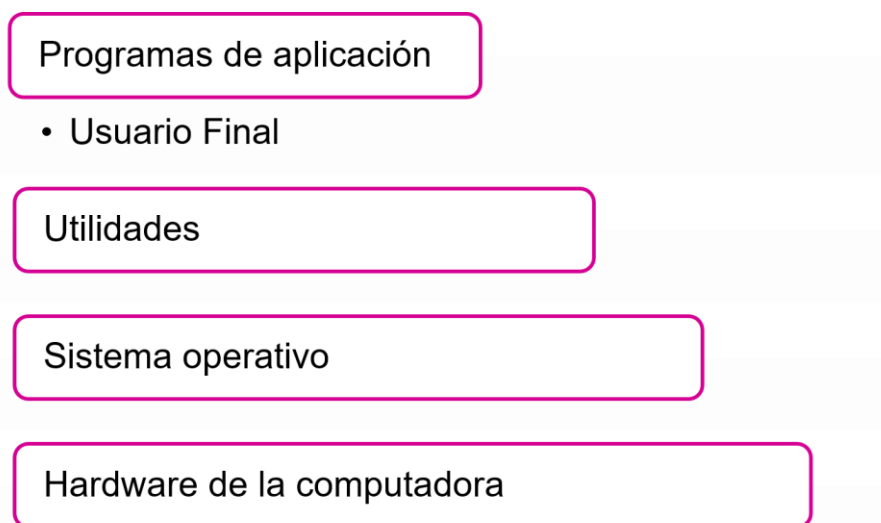


Figura 1.3 Capas de la relación usuario - máquina

El usuario tiene una visión de la computadora en términos de aplicación, esta se utiliza mediante un lenguaje de programación de alto nivel, el sistema operativo interpreta



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

las instrucciones y procesa la información para que el hardware de la computadora ejecute los requerimientos del usuario de aplicaciones. Esto permite que cualquier persona pueda hacer uso de una computadora y reduce la complejidad del empleo de este tipo de tecnologías.

Las computadoras están formadas por recursos para transferir, almacenar y procesar datos, el sistema operativo es el responsable de la administración de los recursos, funciona de la misma forma que el software, sin embargo, este cede el control y depende del procesador para recuperar el control.

Una parte del sistema operativo está en la memoria principal, su núcleo (Kernel) realiza las funciones más frecuentes del sistema operativo, el resto de la memoria se reparte entre los programas y los datos.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

1.5 Arquitectura de Hardware

La computadora está compuesta por un procesador, memoria y las unidades de E/S, todos estos dispositivos físicos conforman la arquitectura de hardware de la computadora.

Todos los dispositivos electrónicos tangibles que permiten el funcionamiento de la computadora se comprenden como hardware.

Se componen de la unidad central de proceso o CPU que se encarga de controlar el sistema de cómputo (hardware y software), desarrolla las operaciones matemáticas y lógicas y controla el envío y recepción de datos de los dispositivos periféricos. Para realizar sus funciones, se compone a su vez de dos dispositivos principales ubicados en la tarjeta madre (motherboard).

La unidad de almacenamiento primario o memoria principal que almacena la información requerida para los procesos.

El microprocesador, que contiene a la unidad de control, la unidad de control de periféricos y la unidad aritmética y lógica.

Unidad de control. Controla el flujo de información al interior de la computadora, reconoce las instrucciones, determina la instrucción a ejecutarse e identifica la ubicación de los datos a procesarse.

Unidad de control de periféricos. Controla la entrada y salida de datos a través de los diferentes dispositivos periféricos.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

Unidad aritmética y lógica. Realiza las operaciones aritméticas, lógicas y de comparación de valores.

La tarjeta madre cuenta además con una serie de ranuras llamadas slots a través de las que se pueden conectar y controlar diversos dispositivos periféricos.

Otro elemento importante del hardware es la memoria principal que se divide en dos medios de almacenamiento primario: memoria RAM y memoria ROM.

En la memoria RAM (Random Access Memory) se almacenan todos los programas que son usados por una aplicación al momento de su ejecución. Esta memoria es volátil, por lo que cuando se apaga el equipo o falla la energía eléctrica se pierde toda la información contenida en ella; lo mismo sucede cuando se sobrescribe información. La memoria RAM comúnmente se puede expandir en unidades de uno o más Megabytes (Mb) dependiendo de la computadora.

La memoria ROM (Read Only Memory) es de sólo lectura, su capacidad no se puede modificar al igual que el contenido. En ella vienen grabados de fábrica algunos programas usados por el sistema para configurarse por sí misma al iniciar una sesión, tales como rutinas de arranque, de diagnóstico y de entrada/salida. El usuario no puede acceder los programas cargados en esta área ya que son exclusivos del sistema.

Los periféricos más usados para comunicar al usuario con su sistema operativo y su procesador son los teclados, monitores, mouse, mousepad, pluma electrónica, micrófono, bocina, scanner, impresora, etc.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

1.6 Arquitectura Von Neumann

Esta arquitectura se basa en los conceptos desarrollados por John von Neumann del Instituto de Estudios Avanzados de Princeton. Von Neumann propone el concepto del programa-almacenado.

Los conceptos clave de esta arquitectura son:

- Los datos y las instrucciones se almacenan en una sola memoria de lectura-escritura.
- Los contenidos de esta memoria se direccionan indicando su posición, sin considerar el tipo de dato contenido en la misma.
- La ejecución se produce siguiendo una secuencia de instrucción tras instrucción.

Von Neumann esboza la estructura en cinco partes:

1. La computadora tendrá que realizar las operaciones aritméticas elementales, esto exige que contenga elementos especializados solo en esas operaciones. Entonces tendrá que existir la función aritmética central **CA (Central Arithmetical)**.
2. El control lógico del dispositivo debe ser realizado eficientemente por un órgano de control central. La exigencia de dispositivos para uso general propone que las instrucciones tanto generales como específicas se almacenen en alguna parte del control central. A la segunda parte específica se le denomina así, **CC (Central Control)**.
3. Cualquier computadora debe realizar secuencias de instrucciones y operaciones largas o complejas que necesitan una cantidad de memoria importante. Estas instrucciones constituyen un material considerable que debe



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

ser almacenado. Lo que nos lleva a la tercera parte específica, la memoria total, **M (Memory)**.

4. El dispositivo necesita transferir información a partir de la función aritmética central, el control central y la memoria. Para esto es necesario la parte específica de entrada **I (Input)**. Que mantiene una entrada y salida mediante el medio de grabación exterior del dispositivo.
5. El dispositivo también debe transferir información por medio de sus partes específicas hasta la salida, a partir de la memoria comunicará con el usuario usando la parte específica **O (output)**.

La unidad de control dirige captando instrucciones de la memoria y ejecuta las instrucciones una a una. En este tipo de arquitectura la computadora usa el mismo dispositivo de almacenamiento para datos e instrucciones conectados mediante un bus de datos y un bus de direcciones. Como se muestra en la figura 1.4.

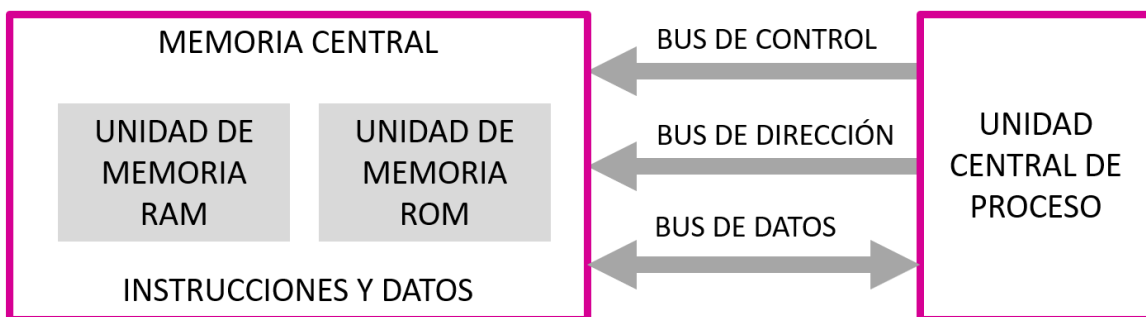


Figura 1.4 Arquitectura Von Neumann



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

1.7 Arquitectura Harvard

Esta es una arquitectura con el almacenamiento y la vía de señal separada. A diferencia de la arquitectura de von Neumann esta posee memorias separadas para las instrucciones y los datos. Tiene dos sistemas de buses, uno para datos y otro para instrucciones. Es una arquitectura más rápida porque permite analizar datos e instrucciones al mismo tiempo.

Como se muestra en la figura 1.5. podemos ver los buses que se direccionan a la memoria ROM, en caso de las instrucciones y los buses de datos se direccionan a la unidad de memoria RAM.

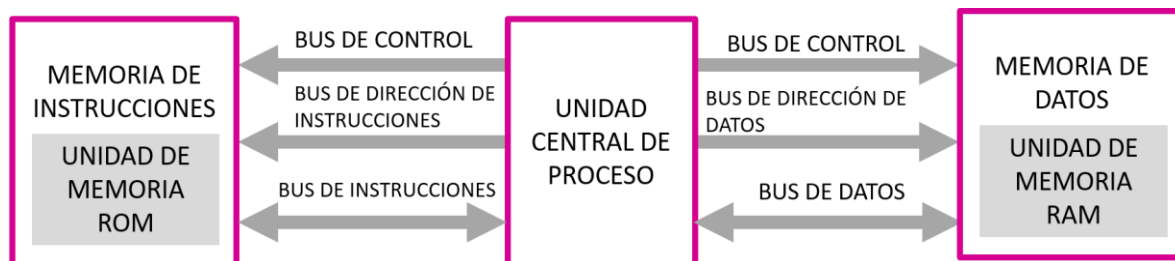


Figura 1.5 Arquitectura Harvard

La ventaja de esta arquitectura en relación con la de von Neumann es el acceso simultáneo a más memoria del sistema y que también ha desarrollado sistemas de memoria caché de CPU. Este tipo de memoria es una parte de la memoria RAM que se almacena en el procesador y su tiempo de respuesta es más pequeño.

Más adelante se explicará más ampliamente los diferentes tipos de memorias y su función.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

1.8 CPU y sus componentes

La CPU (Unidad Central de Procesamiento) es uno de los elementos más complejos e importantes que forma parte de una computadora, las funciones que realiza se resumen de la siguiente manera:

- Interpreta y lleva a cabo la ejecución de las órdenes o instrucciones de los programas.
- Realiza todas las operaciones aritméticas y lógicas necesarias para la manipulación de los datos.
- Controla el correcto funcionamiento del resto de elementos que forman parte de un ordenador.

Una CPU Se compone de muchos circuitos en un chip de silicio, a este chip se le denomina microprocesador. La CPU junto a otros componentes electrónicos se ubican en un tablero de circuitos conocidos como placa madre o motherboard.

Entre los aspectos más importantes y característicos de la CPU sobresale la velocidad, que está determinada por la velocidad de su reloj interno, un dispositivo que produce pulsos eléctricos que sincronizan las operaciones de la computadora. La velocidad de reloj se mide en megahercios o gigahercios y está determinada por la arquitectura del procesador.

Otro aspecto es el rendimiento, no todo el software se ejecuta a la misma velocidad en todas las CPU. Ciertas aplicaciones se ejecutan más rápido dependiendo del procesador sobre el que se ejecuten.

La CPU se encuentra dividida en dos secciones: la unidad de control (UC) y la unidad aritmético-lógica (ALU).



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

1.8.1 Unidad de control (UC)

La unidad de control es el núcleo del procesador y tiene cuatro funciones principales:

- Leer e interpretar las instrucciones del programa.
- Dirigir las operaciones internas del procesador.
- Controlar el flujo de programas y datos desde o hacia la memoria RAM.
- Controla las operaciones de entrada/salida desde o hacia los periféricos.

La unidad de control a su vez también está formada por distintos componentes necesarios para que se realicen todas las funciones a ella encomendadas.

Registros de instrucciones: Áreas de almacenamiento temporal de alta velocidad y reducido tamaño. Estos registros manejan y almacenan instrucciones a una velocidad unas 10 veces mayor que la de la memoria caché y se usan para una gran variedad de operaciones de procesamiento.

Decodificador: Componente encargado de decodificar e identificar una instrucción y generar una serie de señales de control al resto de elementos con el fin de ejecutar dicha operación.

Reloj: Genera una secuencia de pulsos eléctricos que marca los ritmos de ejecución de cualquier operación.

Unidad aritmético-lógica (ALU)

Unidad encargada de realizar operaciones aritméticas (sumas, productos, etc..) y lógicas (=, <, >) en función de las ordenes o señales de control enviadas por la U.C.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

1.9 Microprocesador y sus partes lógicas

Un microprocesador es un circuito de alta escala de integración (LSI), compuesto de circuitos simples como son los Flip-flops, contadores, registros, decodificadores, comparadores, entre otros. Todos ellos en una placa de silicio, de modo que el microprocesador puede ser considerado un dispositivo lógico de propósito general o universal.

El microprocesador se compone de partes lógicas específicas.

MAR. Este es el registro de direcciones de memoria (memory address register). Debido a que la memoria RAM requiere cierto tiempo para encontrar y posicionarse en una dirección, dicha dirección debe estar siempre en las entradas. Este registro funciona como buffer del PC.

PC. Este es un contador (program counter) que sirve para tener la dirección de memoria de la siguiente instrucción o dato que se debe de leer.

MBR. Este registro es el buffer de datos de la memoria (memory buffer register). Cuando la memoria se posiciona en una dirección, muestra el contenido en su bus de datos, pero estos datos o contenido cambiarán si se cambia la dirección que se pone en las entradas del bus de direcciones de la memoria. Su función también es actuar como buffer para instrucciones de escritura a la memoria.

IR. Registro de instrucciones (instruction register). Cuando el dato contenido en una dirección de memoria es una instrucción se envía a este registro para decodificarlo y ordenar las señales correctas para realizar dicha instrucción.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

ALU. La unidad aritmético-lógica es un dispositivo integrado que puede realizar operaciones aritméticas y operaciones lógicas ya sea con un solo operando o hasta con dos.

ACC. Es un registro acumulador y es de propósito general. En nuestro caso será utilizado para guardar uno de los operandos de la ALU para llevar a cabo operaciones con dos operandos. El segundo se proporcionará directamente a la ALU desde el MBR.

ROP. Este registro contiene el código de operación de la ALU. Esto es, se requiere de un código binario en las entradas de la ALU para seleccionar qué tipo de operación se quiere llevar a cabo.

R. La función de este registro es supervisar el resultado de las operaciones realizadas en la ALU pues está conectado directamente a las salidas.

Generador de tiempos. Se encarga de generar señales de tiempos de acuerdo con la frecuencia de reloj que estemos manejando.

Unidad de control. Genera las señales correctas para realizar las operaciones deseadas por el usuario. Como una analogía tenemos al director de una orquesta.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

1.10 ALU, funciones y componentes

Von Neumann hace hincapié en que una ALU es un requisito fundamental para una computadora porque necesita efectuar operaciones matemáticas básicas: adición, sustracción, multiplicación, y división, como se mencionó anteriormente. Por lo tanto, creyó que era "razonable que una computadora debería contener los órganos especializados para estas operaciones".

La ALU es un circuito digital del microprocesador que realiza realmente las operaciones aritméticas y lógicas con los datos. Se basa en dispositivos lógicos sencillos.

Generalmente las ALU pueden realizar las siguientes operaciones sencillas.

- Operaciones aritméticas de números enteros (adición, sustracción, y a veces multiplicación y división, aunque esto es más complejo)
- Operación lógica de bits (AND, NOT, OR, XOR, XNOR)
- Operación de desplazamiento de bits (Desplazan o rotan una palabra en un número específico de bits hacia la izquierda o la derecha, con o sin extensión de signo). Los desplazamientos pueden ser interpretados como multiplicaciones o divisiones por 2.

Las operaciones más complejas pueden usar un ciclo de reloj, cálculo iterativo, etc. La programación de estas instrucciones eleva el precio de la unidad dependiendo de la complejidad.

La ALU procesa números usando el mismo formato que el resto del circuito digital. Para los procesadores modernos, este formato casi siempre es la representación del número binario de complemento a dos.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

Se compone de Circuitos Operacionales, Registros de Entradas, Registro Acumulador y un Registro de Estados, conjunto de registros que hacen posible la realización de cada una de las operaciones.

La mayoría de las acciones de la computadora son realizadas por la ALU. Esta toma datos de los registros del procesador, posteriormente son procesados y los resultados de esta operación se almacenan en los registros de salida de la ALU. Otros mecanismos mueven datos entre estos registros y la memoria.

Una unidad de control controla a la ALU, al ajustar los circuitos que le señala a la ALU qué operaciones realizar.

1.11 Unidad de control y sus componentes

La unidad de control (UC), es uno de los tres bloques funcionales principales en los que se divide una unidad central de procesamiento (CPU).

Administra las operaciones y maneja todas las señales de control dentro del procesador de una computadora. Permite que la unidad lógica, la memoria y los periféricos de entrada y salida sepan cómo replicar a las órdenes recibidas de un programa. Su función es encaminar todo el flujo de entrada y salida, busca el código de las operaciones de los programas y dirige otras unidades enviando señales de temporización y de control.

Una unidad de control actúa admitiendo los datos de entrada, que transforma en señales de control y que posteriormente son transmitidas al procesador central. El procesador ordena a los diferentes dispositivos qué operaciones deben realizar.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

Para ejecutar una acción, la unidad de control de la CPU debe producir la señal de control necesaria en la secuencia adecuada. La UC se compone de los siguientes elementos.

Unidad integrada. La unidad de control está constituida por compuertas lógicas, flip-flops, circuitos digitales, codificadores y decodificadores, que se encuentran cableados de una forma ya establecida.

Se puede considerar como un aparato de estados, que varía de un estado a otro en cada ciclo de reloj, dependiendo del contenido del registro de instrucciones y las entradas externas.

La salida de este aparato de estados son las señales de control. La secuencia de las operaciones realizadas por esta máquina está determinada por el cableado de los elementos lógicos. Por esta razón se denomina “integrada”.

Señales de control. Son fundamentales para dirigir la ejecución de las operaciones. Son difundidas por circuitos lógicos expresamente diseñados, en los que no se puede modificar el método de generación de las señales sin un cambio físico de la estructura del circuito.

Decodificador de instrucciones. Este está constituido por un conjunto de decodificadores que decodifican los diferentes campos del código de operación de la instrucción. Permite que la unidad de control ingrese al primer estado vinculado con la ejecución de la nueva operación, el cual dura tanto tiempo como permanezcan sin alterarse las señales de temporización y otras señales de entrada.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

Unidad microprogramada. Indica la dirección de un microprograma que está almacenado en el almacén de control.

1.12 Unidades funcionales

Una unidad funcional o unidad de ejecución es una parte de la CPU que realiza las operaciones y cálculos llamados por los programas. Tiene a menudo su propia unidad de control de secuencia, algunos registros, otras unidades internas como una sub-ALU o una unidad de coma flotante, o algunos componentes menores más específicos.

1.13 Bus y tipos de buses

Los buses son un dispositivo que genera la transferencia de la información dentro de la computadora. Todos los componentes internos y elementos periféricos de un ordenador se comunican mediante los buses internos. Los buses pueden clasificarse en tres grupos funcionales como podemos ver en la figura 1.6.

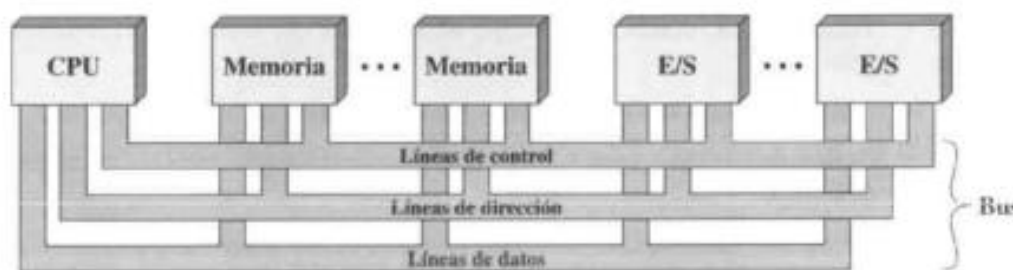


Figura 1.6 Esquema de interconexión mediante un bus. ¹

¹W. Stallings, Organización y arquitectura de computadoras, Madrid: Pearson Educación, 2005. (Figura 3.16).



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

A continuación, se presenta los tipos de buses en la Informática y su función:

Los buses pueden clasificarse según el método de envío de información.

Bus paralelo. Se envían en formato bytes de manera simultánea, con el soporte de diferentes líneas que poseen funciones específicas. En los ordenadores se ha empleado de forma intensa, desde bus del mismo procesador, de los discos duros, tarjetas de video y hasta las impresoras.

El **bus front side** es un tipo de bus de esta clase y posee las siguientes funciones:

- Las **líneas de dirección**, que se encargan de señalar la ubicación de la memoria o el componente que se va a comunicar con el ordenador. La anchura del bus de direcciones determina la máxima capacidad de memoria posible del sistema.
- Las **líneas de control** se encargan de emitir señales entre los componentes, ejemplo de este son los indicadores o señalizadores de estado. Transmiten tanto órdenes como información de temporización entre los módulos del sistema. Las señales de temporización indican la validez de datos y direcciones. Las señales de órdenes especifican las operaciones a realizar, algunas líneas de control son:
 - Escritura en memoria: hace que el dato del bus se escriba en la posición direccionada.
 - Lectura de Memoria: Hace que el dato de la posición direccionada se sitúe en el bus.
 - Escritura E/S: hace que el dato del bus se transfiera a través del puerto de E/S direccionado.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

- Transferencia reconocida: indica que el dato se ha aceptado.
 - Petición de bus: indica que un módulo necesita disponer del control del bus.
 - Cesión de bus: indica que se cede el control del bus a un módulo solicitante.
 - Reloj: se utiliza para sincronizar operaciones.
 - Inicio (reset): pone los módulos conectados en estado inicial.
-
- Las **líneas de datos** se encargan de transmitir los bits al azar. Proporcionan un camino de transmisión entre los módulos del sistema. El conjunto constituido por estas líneas se denomina **bus de datos**, puede incluir entre 32 y cientos de líneas y este número se conoce como la anchura del bus de datos.

Bus serie. Los datos o la información son enviada bit a bit y son restaurados mediante registros. Está compuesto por algunos conductores. Se está utilizando desde hace poco tiempo en buses para discos duros, tarjetas de expansión y procesador.

1.14 Fases de ejecución (de una instrucción)

La función básica que realiza un computador es la ejecución de un programa. Un programa consiste en un conjunto de instrucciones y datos almacenados en la unidad de memoria. La CPU ejecuta todos los programas que se encuentran en la memoria principal. Los programas están formados por un conjunto de instrucciones y en cada una de ellas se distinguen dos fases:

Fase de búsqueda: Consiste en localizar la instrucción a ejecutar dentro de la memoria y llevarla a la UC.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

1. Transferir el contenido del Contador de Programa (CP) al registro de Direcciones (RD).
2. Pasar a registro de Memoria (RM) el dato almacenado en la dirección de memoria indicada por RD.
3. Transferir el dato leído desde el RM al registro de instrucción (RI).
4. Incrementar el valor del Contador de Programa para apuntar a la instrucción siguiente.

Fase de ejecución: Es la realización de las acciones que lleva asociadas la instrucción.

1. Comprende el conjunto de operaciones elementales específicas de la instrucción en curso.
2. Decodificación de la Instrucción. Por ejemplo, la Instrucción ADD. (la instrucción es cambiada por su código máquina correspondiente)
3. Transferencia del campo "CD" de la instrucción en curso (este campo contiene la dirección de memoria en la que se encuentra el operando, y que se encuentra en el registro RI desde que finalizó la fase de búsqueda), al registro RD. Se inicia, por tanto, un proceso análogo al de la fase de búsqueda, pero en este caso para buscar en memoria el operando de la instrucción.

Ejemplo de una fase de ejecución:

Instrucción: Leer A

Fase de búsqueda

1. En el CP se almacena la dirección de memoria de comienzo del programa.
2. La UC genera una orden para que el contenido de CP se envíe al RDM.
3. El selector de memoria localiza la dirección 4 y transfiere su contenido al RIM.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

4. La UC da la orden de transferir el contenido del RIM al RI, en el que se deposita el código de la instrucción a ejecutar.
5. El decodificador de instrucción (DI) analiza el código contenido en el RI y genera las señales de control para ejecutar la instrucción.
6. El CP se incrementa en 1 y apunta a la instrucción siguiente.

Fase de ejecución

7. Una vez conocido el código de operación, la UC establece las conexiones con el dispositivo de entrada para aceptar el dato A.
8. La UC da la orden de que el dato leído se cargue en el RIM y en el RDM se ponga la dirección de memoria donde se almacenará., la 27.
9. El selector de memoria guarda en la dirección de memoria indicada por el RDM el contenido del RIM.

1.15 Memorias, sus jerarquías y sus funciones

Es un dispositivo principal de las computadoras, se utiliza para almacenar datos e instrucciones. Es el espacio de almacenamiento en la computadora donde los datos que van a ser procesados se almacenan, y dependiendo de la arquitectura, también las instrucciones necesarias para procesamiento. La memoria se divide en gran número de piezas pequeñas llamadas células. Cada ubicación o celda tiene una dirección única que varía desde cero hasta el tamaño de la memoria menos uno.

Las memorias pueden distinguirse por su método de acceso que incluye:

- Acceso directo. Tiene asociado un mecanismo de lectura/escritura.
- Acceso secuencial. Se organiza en unidades de datos llamados registros.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

- Acceso aleatorio. La posición direccionable de memoria tiene un único mecanismo de acceso cableado físicamente.
- Asociativa. Es una memoria de tipo aleatoria que hace una comparación de ciertas posiciones de bits dentro de una palabra buscando que coincidan los valores.

La cuestión de tamaño en la memoria, su rapidez y otras prestaciones óptimas con relación a su precio nos permite estructurarla en una jerarquía.

La organización jerárquica de la memoria se basa en una característica denominada principio de localidad que poseen la mayoría de los programas. Esta característica establece que los programas acceden a una porción relativamente reducida del espacio de direcciones en un determinado lapso.

Las memorias deben cumplir las siguientes relaciones:

A menos tiempo de acceso, mayor coste por bit.

A mayor capacidad, menor coste por bit.

A mayor capacidad, mayor tiempo de acceso.

Si en el diseño de la memoria contamos con diferentes componentes aprovechamos la jerarquía que cumple las relaciones anteriores.

En la figura 1.7 podemos entender la jerarquía típica de la memoria y como podemos satisfacer las prestaciones requeridas por el diseñador de la memoria, con relación a costo, capacidad y tiempo de acceso.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

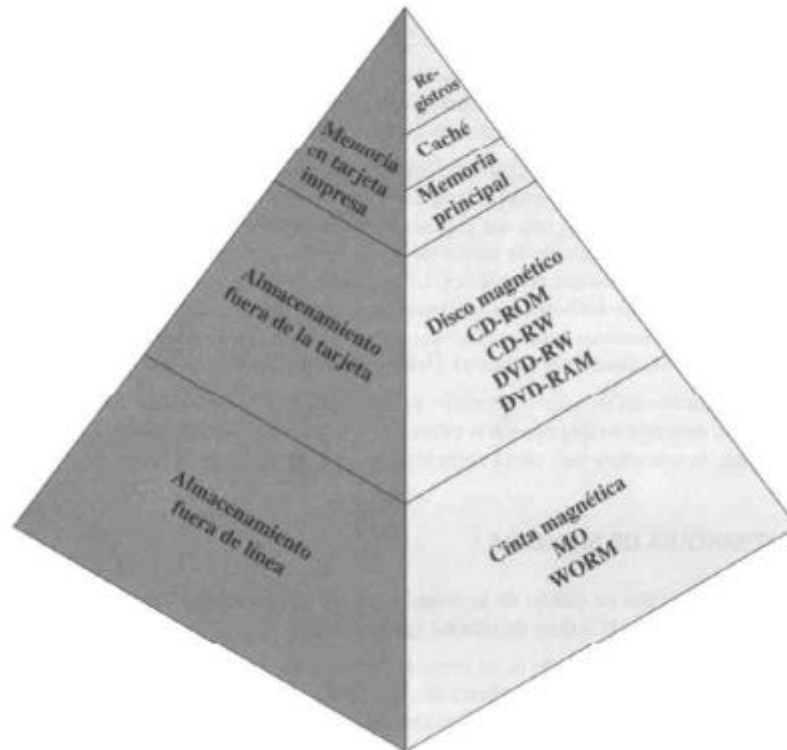


Figura 1.7 Jerarquía de memoria²

El principio de localidad posee dos variantes:

- Localidad temporal. Si un dato o instrucción es referenciado en determinado momento, es común que vuelva a ser referenciado poco tiempo después.
- Localidad espacial. Cuando un dato o instrucción es referenciado en determinado momento, es común que los datos o instrucciones con direcciones cercanas también sea accedidos poco tiempo después.

El principio de localidad es aplicado por distintos actores dentro de un sistema:

² W. Stallings, Organización y arquitectura de computadoras, Madrid: Pearson Educación, 2005. (Figura 4.1).



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

- Registros (memoria): lo aplica el compilador ó el programador de bajo nivel.
- Memoria cache (disco): lo aplica el hardware.
- Memoria (disco): lo aplica el sistema operativo(memoria virtual) o el programador.

Cuando se habla de jerarquía de memoria se emplea la siguiente terminología.

Hit. Suceso que ocurre cuando un objeto de información se encuentra en el lugar de la jerarquía donde se lo está buscando.

Miss. Suceso que ocurre cuando el elemento de información no es encontrado en el lugar de la jerarquía donde se lo está buscando. En este caso hay que buscarlo en un nivel jerárquico inferior.

Hit Rate. Es la tasa de acierto de encontrar un elemento de información en el lugar jerárquico en que se busca.

Miss Rate. Es la tasa de fallos en encontrar un elemento de información en el lugar buscado.

Hit Time. Tiempo de acceso promedio en el nivel de jerarquía considerado.

Miss Penalty. Tiempo de acceso promedio adicional requerido para acceder al elemento de información en el nivel de jerarquía inferior.

Tiempo promedio de acceso a memoria. Si consideramos el nivel de jerarquía “memoria caché” podemos establecer que el tiempo promedio de acceso a memoria es: $\text{Hit Time} + \text{Miss Rate} * \text{Miss Penalty}$.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

La velocidad de acceso a los elementos de información, incluyendo las instrucciones de los programas condiciona la capacidad de proceso de las computadoras. Por eso el rendimiento del sistema de memoria tiene un impacto significativo sobre el rendimiento genera.

Los sistemas de memoria tienen distintos parámetros que tienen relación con su capacidad de movilizar elementos de información: Tiempo de Acceso. Tiempo que transcurre entre que la dirección se presenta estable y los datos pueden ser manipulados en forma confiable.; Tiempo de Ciclos. Tiempo mínimo que debe pasar entre un acceso y el siguiente.; Tasa de Transferencia. Velocidad de movimiento de datos de la memoria.

La **memoria caché** es una memoria de tamaño reducido, de alta velocidad que se ubica entre la memoria principal y la CPU. Utiliza el principio de localidad mantiene copias de los bloques de memoria principal más accedidos, de manera que cuando la CPU requiere una palabra que se encuentre en los bloques de almacenados en la memoria cache es satisfecho por este, con un tiempo de acceso mucho menor que si fuera desde la memoria principal.

Desde la memoria caché se accede a la memoria principal en bloques de k palabras y se almacena en una línea de la caché.

Cuando la CPU requiere una palabra de memoria ocurre un “hit” en la memoria de mayor jerarquía y por eso se satisface desde esa memoria. Pero si se dispara un “miss” se trae desde la memoria el bloque que contiene la palabra buscada para escribirlo en la línea de memoria caché y entregarla a la CPU. Lo más probable es que ocurra un “hit” por la estrategia de mover bloques.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

En el diseño de un sistema de memoria caché hay que tomar una serie de decisiones sobre distintos aspectos que tienen impacto sobre su rendimiento (porcentaje de “hits” frente al total de accesos).

Tamaño. El punto de equilibrio depende fuertemente del tipo de programa que se ejecuten en el sistema, existirá una cierta cantidad de memoria de la cual el incremento del rendimiento obtenido no compensa el costo adicional. La memoria caché se coloca directamente en el “chip” del procesador, con lo cual los transistores disponibles son finitos y compite con los demás sub-sistemas que deben ser construidos. Se utilizan tres niveles de caché que son el L1(de 10 KB a 20 KB), el L2(de 128 KB a 512 KB) y el L3(de 4 MB a 12 MB).

Política de Escritura. En las operaciones de escritura existen dos estrategias, write through la cual implica que de existir un “hit” la escritura se hace en la memoria caché y en la memoria principal mientras que en write back si ocurre un “hit” la escritura se realiza solamente en la memoria caché, aunque presenta la complejidad de determinar si es necesario actualizar la memoria principal.

Tamaño del bloque. Una línea más grande por el principio de localidad la probabilidad de un “hit” tras un “miss” aumenta pero más cantidad de líneas puede sufrir entre cambios de contexto en ambientes de multiprogramación.

Función de Correspondencia. Determina como se asocia una cierta posición de memoria con su posible ubicación en la memoria caché está formada por un cierto conjunto de líneas y en cada línea se puede almacenar un bloque de memoria. Con lo cual la función de correspondencia establece para cada bloque de memoria cuales son las líneas posibles de ser utilizado en la cache para almacenarlo, la relación



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

puede ser desde que cualquier línea de caché puede recibir a un bloque dado en memoria hasta que solo una línea determinada pueda recibir un bloque dado.

Algoritmo de Sustitución.

La **memoria principal** contiene sólo los datos y las instrucciones sobre qué equipo está trabajando actualmente. Tiene una capacidad limitada y datos se pierde cuando se apaga el suministro eléctrico. Generalmente se compone del dispositivo semiconductor. Estos recuerdos no son tan rápidos como los registros. Instrucción necesaria para procesar los datos residen en la memoria principal. Se divide en dos subcategorías RAM y ROM.

Características de la memoria principal

- Son memorias de los semiconductores
- Es conocida como memoria principal.
- Memoria generalmente volátil.
- Datos se pierden en caso de alimentación conmutada libres
- Es la memoria de trabajo de la computadora.
- Más rápido que la memoria secundaria.
- Una computadora no puede funcionar sin memoria primaria.

La **memoria secundaria** también se conoce como memoria externa o no volátil. Es más lenta que la memoria principal. Estos se utilizan para almacenar información permanentemente. CPU no accede directamente estos recuerdos en cambio son accesibles mediante rutinas de entrada / salida. Contenido de memorias secundarias primero se transfiere a la memoria principal, y luego CPU puede acceder a él. Por ejemplo: disco, CD-ROM, DVD etc.

Característica de memoria secundaria



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

- Estos son recuerdos magnéticos y ópticos
- Es conocida como memoria de respaldo.
- Es una memoria no volátil.
- Los datos se almacenan permanentemente, aunque el poder es conmutada libres.
- Se utiliza para el almacenamiento de datos en una computadora.
- Una computadora puede funcionar sin memoria secundaria.

1.16 Registros y tipos de registros

Anteriormente se mencionó la jerarquía de memoria, en los niveles más altos de la jerarquía la memoria es más rápida, más pequeña y costosa. Dentro del procesador hay un conjunto de registros que funciona como un nivel de memoria por encima de la memoria principal y de la caché en la jerarquía. Los registros del procesador son de dos tipos.

Registros visibles por el usuario

Permiten al programador de ensamblador minimizar las referencias a memoria principal por medio de la optimización del uso de registros.

Así mismo estos se clasifican en las siguientes categorías.

Registro de uso general. Pueden ser asignados por el programador a diversas funciones. Usualmente dentro del repertorio de instrucciones ortogonales a la operación. Cualquier registro de uso general puede contener el operando para cualquier código de operación.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

Registro de datos. Son usados únicamente para contener datos y no se pueden emplear en el cálculo de la dirección de un operando.

Registro de dirección. Están dedicados a un modo de direccionamiento particular, como lo pueden ser:

- Punteros de segmento. Contiene la dirección de la base del segmento.
- Registros índice. Se usan para direccionamiento indexado.
- Puntero de pila: Apunta a la cebera del direccionamiento a pila.

Registro de códigos de condición. También conocidos como indicadores, son bits fijados por el hardware del procesador como resultado de alguna operación. Normalmente forman parte de un registro de control.

Registros de control y de estado

Son usados por la unidad de control para controlar el funcionamiento del procesador y por programas privilegiados del sistema operativo para controlar la ejecución de programas.

Son esenciales cuatro registros para la ejecución de una instrucción.

- **Contador de programa (PC).** Contiene la dirección de la instrucción a captar.
- **Registro de instrucción (IR).** Contiene la instrucción captada más reciente.
- **Registro de dirección de memoria (MAR).** Contiene la dirección de una posición de memoria.
- **Registro intermedio de memoria (MBR).** Contiene la palabra de datos a escribir en memoria o la palabra leída recientemente.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

1.17 La memoria caché

Se conoce como memoria caché o memoria de acceso rápido a uno de los recursos con los que cuenta una CPU (Central Processing Unit) para almacenar temporalmente los datos recientemente procesados en un búfer especial, es decir, en una memoria auxiliar.

La memoria caché opera de modo similar a la Memoria Principal del CPU, pero con mayor velocidad a pesar de ser de mucho menor tamaño. Su eficacia provee al microprocesador de tiempo extra para acceder a los datos más frecuentemente utilizados, sin tener que rastrearlos a su lugar de origen cada vez que sean necesarios.

Esta memoria alterna se sitúa entre el CPU y la Memoria RAM (Random Access Memory) proveen de un empuje adicional en tiempo y ahorro de recursos al sistema.

Existen varios tipos de memoria caché, como los siguientes:

Caché de disco. Es una porción de memoria RAM asociada a un disco particular, en donde se almacenan los datos de reciente acceso para agilizar su carga.

Caché de pista. Similar a la RAM, este tipo de memoria caché sólida empleada por supercomputadores es potente, pero costosa.

Caché de Web. Se ocupa de almacenar los datos de las páginas Web recientemente visitadas, para agilizar su carga sucesiva y ahorrar ancho de banda. Este tipo de caché a su vez puede funcionar para un solo usuario (privada), varios usuarios a la



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

vez (compartida) o en conjunto para toda la red administrada por un servidor (en pasarela).

cuando accedemos a un dato cualquiera en nuestro sistema computarizado, se crea de inmediato una copia de los datos más relevantes del mismo en la memoria caché, de modo que los accesos siguientes a dicha información la tengan a mano y no deban rastrearla hacia su lugar de origen.

Accediendo a la copia y no al original, se ahorra tiempo de procesamiento y por ende velocidad, ya que el microprocesador no debe acudir todo el tiempo a la memoria principal. Se trata de una copia de trabajo constantemente actualizada de los datos de más frecuente utilización.

La liberación de memoria caché cumple con dos propósitos fundamentales, que son:

- Eliminar datos viejos o innecesarios, como archivos viejos o procesos que no volveremos a necesitar pero que están allí almacenados “por si acaso” para acelerar su ejecución.
- Acelerar y agilizar el sistema al brindarle nuevo espacio libre para copiar los datos en uso actual, acortando los tiempos de procesamiento.

Dicha labor de mantenimiento debe hacerse con cierta periodicidad, que sin embargo no debería ser exagerada, pues estaríamos impidiendo que la memoria caché cumpla con su cometido.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

1.18 Norma IEEE754

IEEE establece el estándar para aritmética en coma flotante (IEEE 754), esta es la norma o estándar técnico para computación en coma flotante, establecida en 1985 por el Instituto de Ingenieros Eléctricos y Electrónicos (IEEE). La norma surge como la necesidad en las diversas implementaciones de coma flotante que las hacían difíciles de usar de forma fiable y portátil. Muchas unidades de coma flotante de hardware utilizan ahora el estándar IEEE 754.

El estándar define:

- Formatos aritméticos: conjuntos de datos de coma flotante binarios y decimales, que consisten en números finitos, incluidos los ceros con signo y los números desnormalizados o subnormales, infinitos y valores especiales "no numéricos"-
- Formatos de intercambio: codificaciones (cadenas de bits) que se pueden utilizar para intercambiar datos de coma flotante de forma eficiente y compacta.
- Reglas de redondeo: propiedades que deben satisfacerse al redondear los números durante las operaciones aritméticas y las conversiones.
- Operaciones: operaciones aritméticas y otras (como funciones trigonométricas) en formatos aritméticos.
- Manejo de excepciones: indicaciones de condiciones excepcionales, tales como división por cero, desbordamiento, etc.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

La norma también es conocida como IEC 60559:1989, Binary floating-point arithmetic for microprocessor systems (originalmente el número de referencia era IEC 559:1989).¹

La versión actual, IEEE 754-2008 publicada en agosto de 2008, incluye casi todo el estándar IEEE 754-1985 original y el estándar IEEE para aritmética de coma flotante independiente de la base (IEEE 854-1987).

El estándar IEEE 745 define cinco formatos básicos que se denominan por su base numérica y el número de bits utilizados en su codificación de intercambio. Existen tres formatos básicos binarios de coma flotante (codificados con 32, 64 o 128 bits) y dos formatos básicos de coma flotante decimal (codificados con 64 o 128 bits). Los formatos binary32 y binary64 son los formatos simple y doble del estándar original IEEE 754-1985. Una implementación conforme debe implementar completamente al menos uno de los formatos básicos.

El estándar también define los formatos de intercambio, que generalizan estos formatos básicos.⁶ Para los números binarios, se requiere la convención de bit de entrada. La siguiente tabla resume los formatos de intercambio más pequeños (incluidos los más básicos).

El estándar de coma flotante IEEE especifica formatos de precisión extendidos y extensibles, que se recomiendan para permitir una mayor precisión que la proporcionada por los formatos básicos.⁶ Un formato de precisión extendido amplía un formato básico utilizando más precisión y más rango de exponentes.

Para un formato extendido con una precisión entre dos formatos básicos, el rango de exponentes debe ser tan grande como el del siguiente formato básico más amplio.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

Por ejemplo, un número binario de 64 bits de precisión extendida debe tener un exponente máximo de al menos 16383. El formato extendido x87 de 80 bits cumple este requisito.

Los formatos de intercambio están destinados al intercambio de datos de coma flotante utilizando una cadena de bits de longitud fija para un formato dado. Para el intercambio de números de coma flotante binaria, se definen formatos de intercambio de longitudes de 16 bits, 32 bits, 64 bits y cualquier múltiplo de 32 bits mayor o igual a 128. El formato de 16 bits está destinado al intercambio o almacenamiento de pequeños números.

El IEEE754 va más allá de la simple definición de un formato.



2. Circuitos Integrados

Los Circuitos Integrados (CI) encapsulan muchos más circuitos en un encapsulado pequeño, por lo que el tamaño total de casi cualquier sistema digital se reduce. El costo se reduce en forma sustancial debido al ahorro que representa la producción en masa de grandes volúmenes de dispositivos similares. Algunas de las otras ventajas no son tan evidentes.

Los CIs han hecho que los sistemas digitales sean más confiables al reducir el número de interconexiones externas de un dispositivo a otro. Ahora la mayoría de las conexiones son internas en los CIs, en donde están protegidas de malas soldaduras, quebraduras o cortos en las rutas de conexión en un circuito impreso (PCB), y demás problemas físicos. Los CIs también han reducido en forma dramática la cantidad de energía eléctrica necesaria para realizar una función dada, debido a que sus circuitos en miniatura, por lo general, requieren menos energía que sus contrapartes discretas.

Como todo, los CI también tienen sus desventajas, como que no pueden manejar corrientes o voltajes muy grandes, ya que el calor generado en espacios tan pequeños provocaría que la temperatura aumentara más allá de los límites aceptables. Además, no pueden implementar con facilidad ciertos dispositivos eléctricos tales como los inductores, los transformadores y los capacitores grandes.

El principal uso de los CIs es para realizar operaciones de circuitos de baja energía, a lo cual se le conoce, por lo general, como procesamiento de información. Las operaciones que requieren altos niveles de energía o dispositivos que no pueden integrarse se siguen manejando mediante componentes discretos.

Las diversas familias lógicas difieren en cuanto a los componentes principales que utilizan en sus circuitos. TTL y ECL utilizan transistores bipolares como su elemento



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

principal en el circuito; PMOS, NMOS y CMOS utilizan transistores MOSFET unipolares como su componente principal.

2.1 Terminología de los circuitos integrados digitales

La mayor parte de la nomenclatura y la terminología están estandarizadas. A continuación, se definen los términos más útiles.

Parámetros de voltaje y corriente

$V_{IH}(\text{mín})$. Voltaje de entrada en nivel alto. El mínimo nivel de voltaje requerido para un 1 lógico en una entrada. El circuito lógico no aceptará como nivel ALTO cualquier voltaje que se encuentre por debajo de este nivel.

$V_{IL}(\text{máx})$. Voltaje de entrada en nivel bajo. El máximo nivel de voltaje requerido para un 0 lógico en una entrada. El circuito lógico no aceptará como nivel BAJO cualquier voltaje que se encuentre por encima de este nivel.

$V_{OH}(\text{mín})$. Voltaje de salida en nivel alto. El mínimo nivel de voltaje en la salida de un circuito lógico en el estado 1, bajo condiciones de carga específicas.

$V_{OL}(\text{máx})$. Voltaje de salida en nivel bajo. El máximo nivel de voltaje en la salida de un circuito lógico en el estado 0, bajo condiciones de carga específicas.

I_{IH} . Corriente de entrada en nivel alto. La corriente que fluye hacia una entrada cuando se le aplica a ésta un voltaje específico de alto nivel.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

I_{IL}. Corriente de entrada en nivel bajo. La corriente que fluye hacia una entrada cuando se le aplica a ésta un voltaje específico de bajo nivel.

I_{OH}. Corriente de salida en nivel alto. La corriente que fluye desde una salida en el estado 1 lógico, bajo condiciones de carga específicas.

I_{OL}. Corriente de salida en nivel bajo. La corriente que fluye desde una salida en el estado 0 lógico, bajo condiciones de carga específicas.

Capacidad de carga

En general, se requiere que la salida de un circuito lógico controle varias entradas lógicas. Algunas veces todos los CIs en el sistema digital son de la misma familia lógica, pero muchos sistemas tienen una mezcla de varias familias lógicas.

La capacidad de carga se define como el número máximo de entradas lógicas que puede controlar una salida de manera confiable. La capacidad de carga depende de la naturaleza de los dispositivos de entrada que se conectan a una salida. A menos que se especifique una familia lógica distinta como el dispositivo de carga, se asume que la capacidad de carga se refiere a la acción de cargar dispositivos de la misma familia que la salida reforzadora.

Tiempos de propagación

Una señal lógica siempre experimenta un retraso al pasar a través de un circuito. Los dos tiempos de propagación se definen de la siguiente manera:

t_{PLH}. El tiempo que tarda la compuerta en cambiar del estado de 0 lógico al de 1 lógico (de BAJO a ALTO).



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

t_{PHL} . El tiempo que tarda la compuerta en cambiar del estado de 1 lógico al de 0 lógico (de ALTO a BAJO).

Requerimientos de energía

Todo CI requiere una cierta cantidad de energía eléctrica para operar. Esta energía se suministra mediante uno o más voltajes de fuente de energía conectados a las terminales de energía en el chip, las cuales se etiquetan como V_{CC} (para TTL) o V_{DD} (para los dispositivos MOS). La cantidad de energía que requiere un CI se determina con base en la corriente I_{CC} (o I_{DD}) que consume de la fuente V_{CC} (o V_{DD}), y la energía real es el producto ($I_{CC} \times V_{CC}$).

En algunos circuitos lógicos I_{CCH} e I_{CCL} tendrán valores distintos. Para estos dispositivos, la corriente promedio se calcula con base en la suposición de que las salidas de las compuertas están en BAJO la mitad del tiempo, y en ALTO la otra mitad del tiempo.

$$I_{CC}(\text{prom}) = \frac{I_{CCH} + I_{CCL}}{2}$$

Inmunidad al ruido

Los campos eléctricos y magnéticos dispersos pueden inducir voltajes en los conductores que conectan los circuitos lógicos. Estas señales espurias indeseables se conocen como ruido y algunas veces pueden hacer que el voltaje en la entrada de un circuito lógico caiga por debajo de $V_{IH}(\text{mín})$ o se eleve por encima de $V_{IL}(\text{máx})$, lo cual podría producir una operación impredecible. La inmunidad al ruido de un circuito lógico se refiere a la habilidad de ese circuito para tolerar el ruido sin producir cambios falsos en el voltaje de salida. A una medida cuantitativa de la inmunidad al ruido se le conoce como margen de ruido.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

Acción de suministro de corriente y drenado de corriente

Las familias lógicas pueden describirse de acuerdo con la forma en que fluye la corriente entre la salida de un circuito lógico y la entrada de otro. Se ilustran en la figura 2.1 (a).

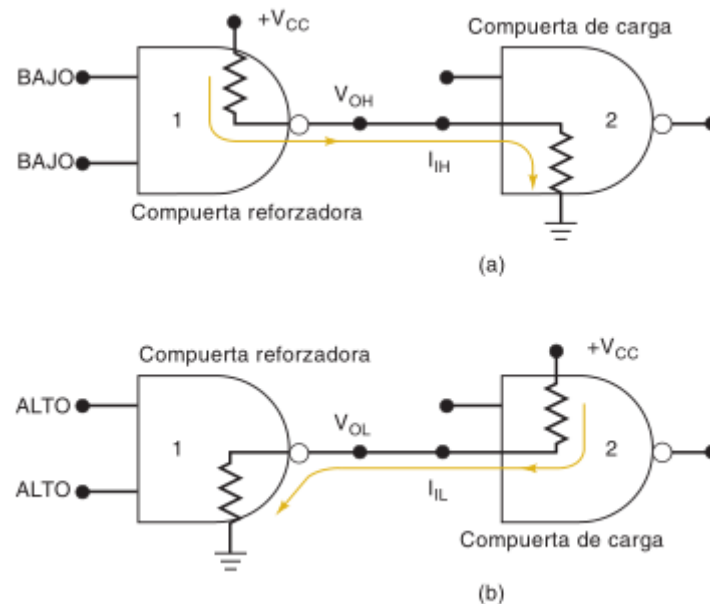


Figura 2.1. Comparación de suministro y drenado de corriente.³

Cuando la salida de la compuerta 1 se encuentra en el estado ALTO, suministra una corriente I_{IH} a la entrada de la compuerta 2, la cual actúa en esencia como una resistencia a tierra. Por ende, la salida de la compuerta 1 actúa como un suministro de corriente para la entrada de la compuerta 1. La acción de drenado de corriente se ilustra en la figura 2.1 (b). Aquí el circuito de entrada de la compuerta 2 se representa como una resistencia enlazada a V_{CC} . La terminal positiva de una fuente de energía. Cuando la salida de la compuerta 1 cambie a su estado BAJO, la corriente fluirá en la

³ R. J. Tocci, G. L. Moss y N. S. Widmer, Sistemas Digitales. Principios y Aplicaciones, Pearson Educación, 2007. Figura 8-5.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

dirección que se muestra desde el circuito de entrada de la compuerta 2, de vuelta a través de la resistencia de la compuerta 1 hasta tierra.

La distinción entre suministro de corriente y drenado de corriente es importante, y se volverá más aparente a medida que examinemos las diversas familias lógicas.

Encapsulados de CI

Los desarrollos y avances en los circuitos integrados continúan a un ritmo acelerado. Lo mismo se aplica a los encapsulados de CI. Existen diversos tipos de encapsulados, los cuales difieren en cuanto al tamaño físico, a las condiciones ambientales y de consumo de energía bajo las cuales el dispositivo puede operar en forma confiable, y a la manera en que el encapsulado de CI se monta en el circuito impreso.

2.2 Características estándar de la serie TTL, CMOS y LVTs

2.2.1 Series de la tecnología TTL

Estas series originales de dispositivos y sus descendientes en la familia TTL han tenido una tremenda influencia sobre las características de todos los dispositivos lógicos en la actualidad. Los dispositivos TTL aún se utilizan como lógica de “pegamento” que conecta los dispositivos más complejos en los sistemas digitales. También se utilizan como circuitos de interfase para dispositivos que requieren una operación con alta corriente. A pesar de que la familia TTL bipolar ya prácticamente no es utilizada, comenzaremos nuestra discusión sobre los CIs lógicos con los dispositivos que dieron forma a la tecnología digital.

El circuito lógico TTL básico es la compuerta NAND. Figura 2.2 (a). Las características de las entradas TTL provienen de la configuración de múltiple emisor (unión de diodos) del transistor Q_1 . La polarización directa de cualquiera (o ambas) de estas uniones de diodos encenderá a Q_1 . Sólo cuando todas las uniones se polaricen en

Tarea 1. Investigación de temas de Arquitectura de Computadoras.

forma inversa se apagará el transistor. Este transistor con entrada de múltiple emisor puede tener hasta ocho emisores para una compuerta NAND de ocho entradas.

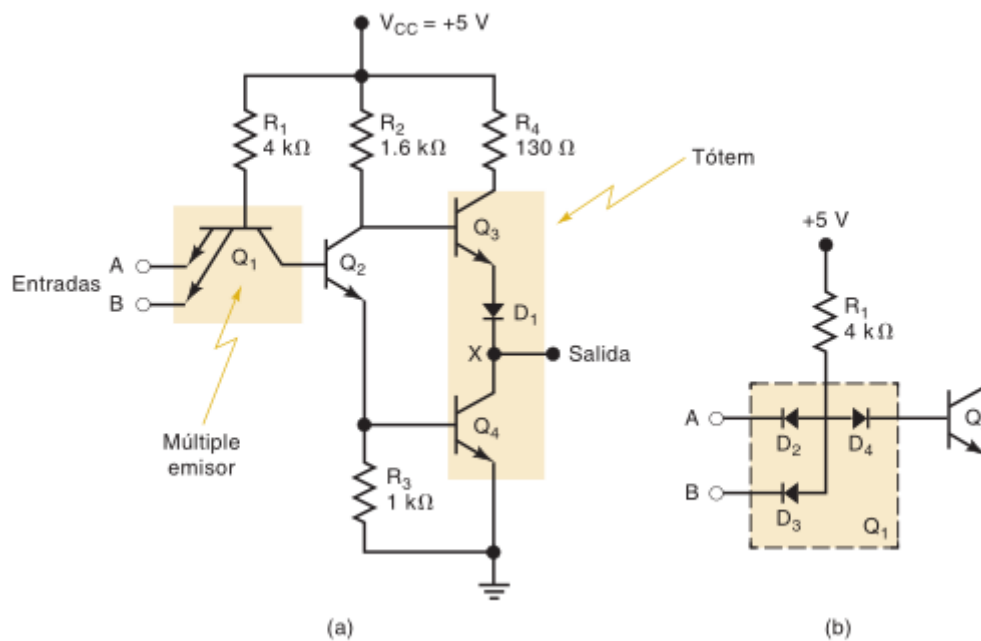


Figura 2.2. Compuerta NAND TTL⁴

Operación del circuito: estado BAJO

Aunque el circuito se ve en extremo complejo, podemos simplificar su análisis de cierta forma si utilizamos el equivalente en diodos del transistor de múltiple emisor Q_1 , como se muestra en la figura 2.2(b). Los diodos D_2 y D_3 representan las dos uniones E-B de Q_1 , y D_4 es la unión colector-base (C-B). En el siguiente análisis utilizaremos esta representación para Q_1 .

⁴ R. J. Tocci, G. L. Moss y N. S. Widmer, Sistemas Digitales. Principios y Aplicaciones, Pearson Educación, 2007. Figura 8-7.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

Operación del circuito: estado ALTO

Esta situación puede producirse si se conecta una de las entradas (o ambas) a un nivel BAJO. Aquí, la entrada B está conectada a tierra. Esto producirá la polarización directa de D_3 , de forma que la corriente fluirá de la terminal de alimentación de 5V a través de R_1 y D_3 , y de la terminal B a tierra. El voltaje positivo a través de D_3 retendrá el punto Y a un valor aproximado a los 0.7 V. Este voltaje no es suficiente para polarizar en forma directa a D_4 y la unión E-B de Q_2 lo necesario como para la conducción.

Acción de drenado de corriente

Una salida TTL actúa como un drenador de corriente en el estado BAJO, ya que recibe corriente de la entrada de la compuerta que está controlando.

Acción de suministro de corriente

En el estado ALTO, una salida TTL actúa como un suministro de corriente.

Todos los circuitos TTL tienen una estructura similar. Las compuertas NAND y AND utilizan entradas de transistor con múltiple emisor o uniones con varios diodos; las compuertas NOR y OR utilizan transistores de entrada separados. En cualquier caso, la entrada será el cátodo (región N) de una unión P-N, por lo que un voltaje de entrada en ALTO desconectará la unión y sólo fluirá una pequeña corriente de fuga (I_{IH}).

Comparación de las características de las series TTL

La figura 2.3 proporciona los valores típicos para algunas de las características más importantes de cada una de las series TTL. Todas las clasificaciones de rendimiento, excepto la clasificación de frecuencia máxima de reloj, son para una compuerta NAND en cada serie. La frecuencia de reloj máxima se especifica como la máxima frecuencia que puede utilizarse para conmutar un flip-flop J-K. Esto nos proporciona una medida



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

útil del intervalo de frecuencia a través del cual puede operarse cada una de las series de CIs.

	74	74S	74LS	74AS	74ALS	74F
Clasificaciones de rendimiento						
Tiempo de propagación (ns)	9	3	9.5	1.7	4	3
Disipación de potencia (mW)	10	20	2	8	1.2	6
Frecuencia máxima de reloj (MHz)	35	125	45	200	70	100
Capacidad de carga (misma serie)	10	20	20	40	20	33
Parámetros de voltaje						
$V_{OH}(\text{mín})$	2.4	2.7	2.7	2.5	2.5	2.5
$V_{OL}(\text{máx})$	0.4	0.5	0.5	0.5	0.5	0.5
$V_{IH}(\text{mín})$	2.0	2.0	2.0	2.0	2.0	2.0
$V_{IL}(\text{máx})$	0.8	0.8	0.8	0.8	0.8	0.8

Figura 2.3. Comparación de las características de las series TTL⁵

2.2.2 Series de la tecnología CMOS

La tecnología MOS básica de un electrodo metálico sobre un aislante de óxido, sobre un sustrato semiconductor. Los transistores de la tecnología MOS son transistores de efecto de campo, a los cuales se les conoce como MOSFETs. Esto significa que el campo eléctrico en el lado del electrodo metálico del aislante de óxido tiene un efecto sobre la resistencia del sustrato. La mayoría de los CIs digitales MOS están contruidos en su totalidad a partir de MOSFETs y de ningún otro componente.

Las principales ventajas del MOSFET son que su fabricación es relativamente simple y económica; es pequeño y consume muy poca energía. La fabricación de los CIs MOS es equivalente a casi una tercera parte de la fabricación de los CIs bipolares (TTL, ECL, etc.). Además, los dispositivos MOS ocupan mucho menos espacio e n un

⁵ R. J. Tocci, G. L. Moss y N. S. Widmer, Sistemas Digitales. Principios y Aplicaciones, Pearson Educación, 2007. Tabla 8-6.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

chip que los transistores bipolares; lo que es más importante: los CIs digitales MOS, por lo general, no utilizan los elementos tipo resistencia de los CIs, que ocupan la mayor parte del área del chip de los CIs bipolares.

Hasta ahora hay dos tipos generales de MOSFETs: de depleción y de enriquecimiento. Los CIs digitales MOS utilizan sólo MOSFETs de enriquecimiento, por lo que sólo consideraremos este tipo de MOSFETs en el siguiente análisis. Lo que es más, sólo nos preocuparemos por la operación de estos MOSFETs en los interruptores de encendido/apagado.

La figura 2.4 muestra los símbolos esquemáticos para los MOSFETs de enriquecimiento de canal N y canal P, en donde la dirección de la flecha indica ya sea el canal P o el canal N.

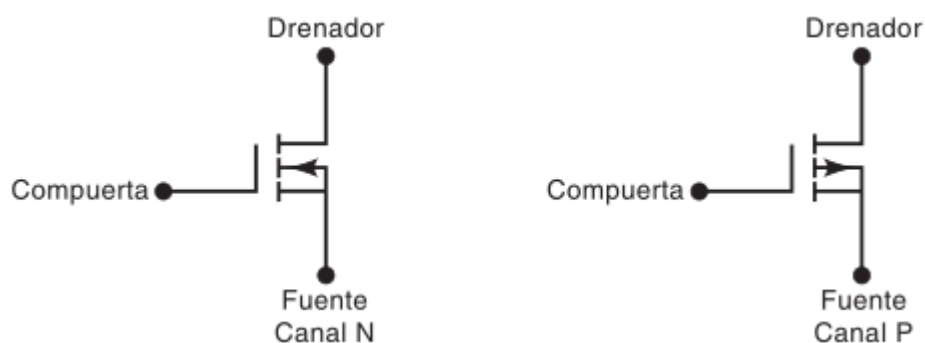


Figura 2.4. Símbolos esquemáticos para MOSFET.⁶

Inversor CMOS

La figura 2.5 muestra el circuito para el INVERSOR CMOS básico. Para este diagrama y los siguientes. El INVERSOR CMOS tiene dos MOSFETs en serie, de manera que el dispositivo de canal P tiene conectada su fuente a V_{DD} (un voltaje positivo) y el

⁶ R. J. Tocci, G. L. Moss y N. S. Widmer, Sistemas Digitales. Principios y Aplicaciones, Pearson Educación, 2007. Figura 8-19.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

dispositivo de canal N tiene su fuente conectada a tierra. Las compuertas de los dos dispositivos se conectan entre sí para formar una entrada común. Los drenadores de los dos dispositivos se conectan entre sí para formar la salida común.

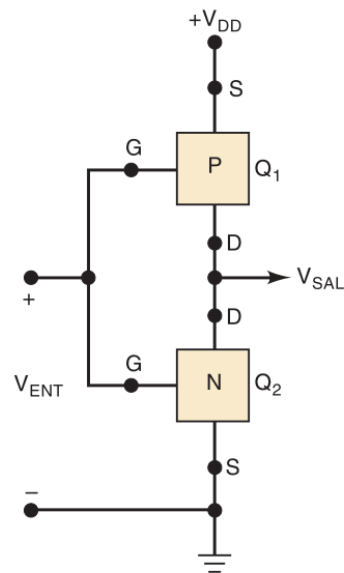


Figura 2.5. Inversor CMOS básico.⁷

Compuerta NAND CMOS

Pueden construirse otras funciones lógicas si se modifica el INVERSOR básico. Una compuerta NAND que se forma al agregar un MOSFET de canal P en paralelo y un MOSFET de canal N en serie al INVERSOR básico.

Para analizar este circuito, es útil tener en cuenta que una entrada de 0 V enciende su correspondiente P-MOS y apaga su correspondiente N-MOS, y viceversa, para una entrada de V_{DD} . Así, podemos ver que la única ocasión en que se producirá una salida en BAJO será cuando las entradas A y B estén ambas en ALTO (V_{DD}) para encender ambos N-MOSFETs, con lo cual se proporciona una resistencia baja de la terminal de salida a tierra. Para todas las demás condiciones de entrada, por lo

⁷ R. J. Tocci, G. L. Moss y N. S. Widmer, Sistemas Digitales. Principios y Aplicaciones, Pearson Educación, 2007. Figura 8-22.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

menos, habrá un P-MOS encendido mientras que al menos un N-MOS estará apagado. Esto produce una salida en ALTO.

Compuerta NOR CMOS

Para formar una compuerta NOR CMOS se agrega un P-MOS en serie y un N-MOS en paralelo al INVERSOR básico, como se muestra en la figura 8-24. Una vez más, este circuito puede analizarse si se toma en cuenta que un nivel BAJO en cualquiera de las entradas encenderá su P-MOS correspondiente y apagará su N-MOS correspondiente, y viceversa, para una entrada en ALTO. Dejaremos al lector la opción de verificar que este circuito opere como una compuerta NOR. Para formar compuertas AND y OR CMOS se combinan las compuertas NAND y NOR con INVERSOR.

FF SET-RESET CMOS

Dos compuertas NOR o NAND CMOS pueden acoplarse en cruz para formar un latch SET-RESET simple. Se utilizan circuitos de compuertas adicionales para convertir el latch SET-RESET básico en flip-flops sincronizados por reloj en D y en J-K.

Los CIs CMOS no sólo proporcionan todas las mismas funciones lógicas están disponibles en TTL, sino también varias funciones de propósito especial que no proporciona TTL.

Características de las series CMOS

Compatibilidad de terminales. Dos CIs tienen terminales compatibles cuando sus configuraciones de terminales son iguales. Por ejemplo, la terminal 7 en ambos CIs es TIERRA, la terminal 1 en ambos es una entrada para el primer INVERSOR, y así en lo sucesivo.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

Equivalente funcional. Dos CIs son equivalentes funcionales si las funciones lógicas que realizan son idénticas. Por ejemplo, ambos contienen cuatro compuertas NAND de dos entradas, o ambos contienen seis flip-flops D con disparo en el flanco positivo del reloj.

Compatible eléctricamente. Dos CIs son compatibles eléctricamente cuando pueden conectarse en forma directa uno con el otro, sin necesidad de tomar medidas especiales para asegurar su operación correcta.

Las entradas CMOS nunca se deben dejar desconectadas. Todas las entradas CMOS deben conectarse ya sea a un nivel de voltaje fijo (0 V o V_{DD}) o a otra entrada.

2.2.3 Series de la tecnología LVTs

Los fabricantes de CIs están continuamente buscando formas de colocar dispositivos semiconductores (diodos, resistencias, transistores, etc.) lo más junto posible en un chip; es decir, buscan incrementar la densidad del chip. Esta mayor densidad tiene cuando menos dos beneficios principales. En primer lugar, permite encapsular más circuitos en el chip; en segundo lugar, como los circuitos están más cerca unos de otros, el tiempo para que las señales se propaguen de un circuito a otro disminuirá, con lo cual se mejorará la velocidad de operación del circuito en general.

También hay desventajas en cuanto a tener una mayor densidad de chip. Cuando los circuitos se colocan más juntos, el material aislante que separa un circuito de otro es más estrecho. Esto reduce la cantidad de voltaje que puede soportar el circuito antes de que se produzca un rompimiento dieléctrico. Al aumentar la densidad del chip se aumenta la disipación de potencia total del chip, con lo cual se puede elevar la



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

temperatura del chip por encima del máximo nivel permitido para una operación confiable.

Familia CMOS

La serie 74VLC (CMOS de bajo voltaje) contiene el surtido más amplio de las conocidas compuertas SSI y funciones MSI de las familias de 5 V, junto con muchos dispositivos de interfase de bus tales como búferes, latches, reforzadores, etcétera. Esta serie puede manejar niveles lógicos de 5-V en sus entradas, por lo que puede realizar conversiones de sistemas de 5-V a sistemas de 3-V. Siempre y cuando el control de corriente se mantenga lo suficientemente bajo como para mantener el voltaje de salida dentro de los límites aceptables, la serie 74VLC también podrá manejar entradas TTL de 5 V. Los requerimientos de entrada de V_{IH} de piezas CMOS de 5-V tales como la serie 74HC/AHC no permiten que los dispositivos LVC los controlen.

En la actualidad, la serie 74ALVC (CMOS avanzado de bajo voltaje) ofrece el rendimiento más alto. Los dispositivos de esta serie están diseñados principalmente para aplicaciones de interfase de bus en las que se utilice sólo lógica de 3.3 V.

La serie 74LV (Bajo voltaje) ofrece tecnología CMOS y muchas de las compuertas SSI y funciones lógicas MSI comunes, junto con algunos búferes octales populares, latches y flip-flops. Está diseñada para operar sólo con otros dispositivos de 3.3 V.

La serie 74AVC (CMOS avanzado de muy bajo voltaje) se introdujo teniendo en mente los sistemas del mañana. Está optimizada para sistemas de 2.5 V, pero puede operar con voltajes de alimentación con valores tan bajos como 1.2 V o de hasta 3.3 V. Este amplio intervalo de voltaje de alimentación la hace útil en sistemas con voltajes mixtos.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

Tiene tiempos de propagación de menos de 2 ns, lo cual hace de esta serie un digno rival de los dispositivos bipolares 74AS. Tiene muchas de las características de interfase de bus de las series BiCMOS que la harán útil en futuras generaciones de estaciones de trabajo de bajo voltaje, PCs, redes y equipo de telecomunicaciones.

La serie 74AUC (CMOS avanzado de ultra bajo voltaje) está optimizada para operar a niveles lógicos de 1.8 V.

La serie 74AUP (Ultra baja potencia avanzada) es la serie lógica de menor potencia de la industria y se utiliza en aplicaciones portátiles, operadas por baterías.

La serie 74CBT (Tecnología de interconexión cruzada) ofrece circuitos de interfase de bus de alta velocidad que pueden conmutar con rapidez cuando se habilitan, y no cargan el bus cuando se deshabilitan.

La serie 74CBTLV (Tecnología transversal de bajo voltaje) es el complemento de 3.3 V para la serie 74CBT.

La serie 74GTLP (Lógica de transceptor de disparo Plus) está hecha para aplicaciones de segundo plano en paralelo de alta velocidad.

La serie 74SSTV (Lógica terminada en series de acoplamiento) es útil en los sistemas de memoria avanzada de alta velocidad de las computadoras modernas.

La serie Interruptor TS (Interruptor de señal TI) está hecha para aplicaciones de señales mixtas y ofrece algunas soluciones de conmutación analógica y digital, y de multiplexeo.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

La serie 74TVC (Circuito de sujeción de traducción de voltaje) se utiliza para proteger las entradas y salidas de los dispositivos sensibles contra el exceso de voltaje en las líneas de bus.

Familia BiCMOS

La serie 74LVT (Tecnología BiCMOS de bajo voltaje) contiene piezas BiCMOS que están diseñadas para aplicaciones de interfase de bus de 8 bits y de 16 bits. Al igual que en la serie LVC, las entradas pueden manejar niveles lógicos de 5 V y sirven como un traductor de 5 V a 3 V. Como los niveles de salida [V_{OH} (mín) y V_{OL} (máx)] son equivalentes a los niveles TTL, son totalmente compatibles eléctricamente con TTL. La tabla 8-10 compara las diversas características.

La serie 74ALVT (Tecnología BiCMOS avanzada de bajo voltaje) es una mejora en relación con la serie LVT. Ofrece la operación con 3.3 V o 2.2 V a 3 ns y sus terminales son compatibles con las de las series ABT y LVT existentes. También está diseñada para aplicaciones de interfase de bus.

2.3 Definir y ejemplificar los Circuitos Integrados:

2.3.1 De tres estados

La configuración triestado es un tercer tipo de circuitos de salida que se utilizan en las familias TTL y CMOS. Aprovechan la operación de alta velocidad del arreglo de salida puesta en alto/puesta en bajo, al tiempo que permite conectar las salidas juntas para compartir un conductor común. Se llama triestado porque permite tres posibles los estados de salida: ALTO, BAJO y alta impedancia (Hi-Z).



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

El estado Hi-Z es una condición en la que tanto el transistor de puesta en alto como el de puesta en bajo se APAGAN, de forma que la terminal de salida sea una alta impedancia tanto para tierra como para el voltaje de alimentación V . La figura 2.6 ilustra estos tres estados para un circuito inversor simple.

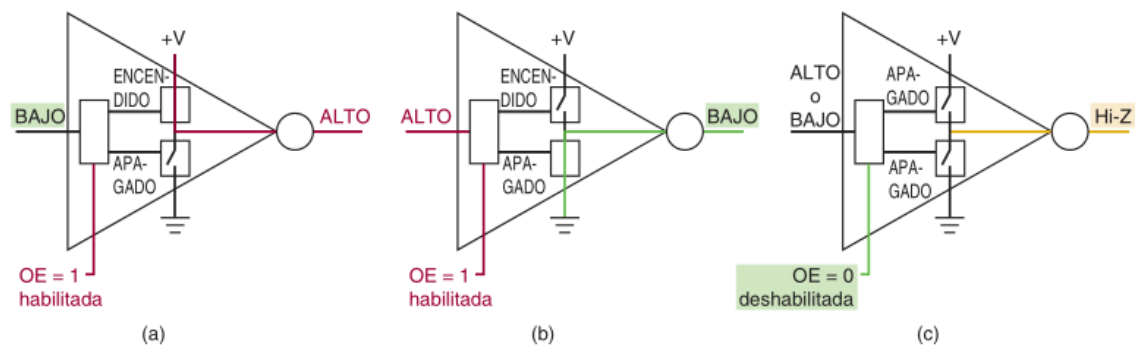


Figura 2.6. Las tres posibles condiciones de una salida triestado.⁸

Los dispositivos con salidas triestado tienen una entrada de habilitación. A menudo esta entrada se etiqueta como E (habilita) o como OE (habilita salida). Cuando $OE=1$, como se muestra en la figura 2.6 (a) el circuito opera como INVERSOR normal, ya que el nivel lógico ALTO en OE habilita la salida. Ésta se encontrará en ALTO o en BAJO, dependiendo del nivel de entrada. Cuando $OE=0$, como se muestra en la figura 2.6(c), la salida del circuito se deshabilita. Cambia a su estado Hi-Z con ambos transistores en el estado no conductor. En este estado, la terminal de salida es en esencia un circuito abierto (no está conectada a nada).

⁸ R. J. Tocci, G. L. Moss y N. S. Widmer, Sistemas Digitales. Principios y Aplicaciones, Pearson Educación, 2007. Figura 8-35.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

2.3.2 Buffer

Las aplicaciones de las salidas de colector/drenador abierto que hemos descrito se utilizaban más en los primeros días de los circuitos lógicos de lo que se emplean actualmente. Un uso más común de estos circuitos ahora es como un búfer/reforzador. Un búfer o reforzador es un circuito lógico que está diseñado para tener una mayor capacidad de corriente y/o voltaje de salida que un circuito lógico ordinario.

Estos dispositivos permiten que un circuito de salida más débil controle una carga pesada. Los circuitos de colector/drenador abierto ofrecen cierta flexibilidad única como búfer/reforzadores.

Un búfer triestado es un circuito que se utiliza para controlar el paso de una señal lógica de la entrada a la salida. Algunos búferes triestados también invierten la señal a medida que va pasando. A los circuitos de la figura 8-35 se les puede llamar búferes triestado inversores.

Dos de los CIs de búfer triestado que se utilizan comúnmente son el 74LS125 y el 74LS126. Ambos contienen cuatro búferes triestado no inversores, como los que se muestran en la figura 8-36. El 74LS125 y el 74LS126 difieren sólo en el estado activo de sus entradas HABILITAR (ENABLE). El 74LS125 permite que la señal de entrada A llegue a la salida cuando $E=0$, mientras que el 74LS126 pasa la entrada cuando $E=1$.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

2.3.3 De Colector abierto

Una solución al problema de compartir un conductor común entre compuertas es remover el transistor de puesta en alto activo del circuito de salida de cada compuerta. De esta forma, ninguna de las compuertas tratará siquiera de asegurar un nivel ALTO lógico. Las salidas TTL que están modificadas de esta forma se llaman salidas de colector abierto. Los circuitos CMOS de salida que se modifican de esta forma se llaman salidas de drenador abierto. La salida se toma en el drenador del MOSFET de puesta en bajo de canal N, el cual es un circuito abierto (es decir, no está conectado a ningún otro circuito).

El equivalente de TTL se llama salida de colector abierto ya que el colector del transistor inferior en el tótem se conecta en forma directa a la terminal de salida y en ninguna otra parte. La estructura de colector abierto elimina los transistores de puesta en alto Q_3 , D_1 y R_4 . En el estado BAJO de salida, Q_4 está ENCENDIDO (tiene corriente de base y en esencia es un corto entre colector y emisor); en el estado ALTO de salida, Q_4 está APAGADO (no tiene corriente de base y es en esencia un circuito abierto entre colector y emisor). Como este circuito no tiene una forma interna de elevar la salida a ALTO, el diseñador del circuito debe conectar una resistencia de puesta en alto externa R_P a la salida.

Cuando Q_4 está ENCENDIDO, lleva el voltaje de salida a BAJO. Cuando Q_4 está APAGADO, R_P lleva la salida de la compuerta a ALTO.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

2.4 Conmutador bilateral (como el 74LS245)

Un circuito CMOS especial que no tiene su contraparte en TTL o en ECL es la compuerta de transmisión o interruptor bilateral, el cual actúa en esencia como un interruptor de un polo, un tiro y se controla mediante un nivel lógico de entrada.

La figura 2.7(a) es el arreglo básico para el interruptor bilateral. Consiste de un PMOSFET y un NMOSFET en paralelo, de manera que puedan cambiarse ambas polaridades del voltaje de entrada. La entrada CONTROL y su inverso se utilizan para encender (cerrado) y apagar (abierto) el interruptor. Cuando el CONTROL está en ALTO, ambos MOSFETs están apagados y el interruptor está abierto. En teoría, este circuito opera como un relevador electromecánico. No obstante, en la práctica no es un cortocircuito perfecto cuando el interruptor está cerrado; la resistencia R_{ENC} del interruptor es, por lo general, de $200\ \Omega$. En el estado abierto, la resistencia del interruptor es muy grande, por lo general de $10e12\ \Omega$, que para la mayoría de los fines se puede considerar como circuito abierto. El símbolo en la figura 2.7 (b) se utiliza para representar el interruptor bilateral.

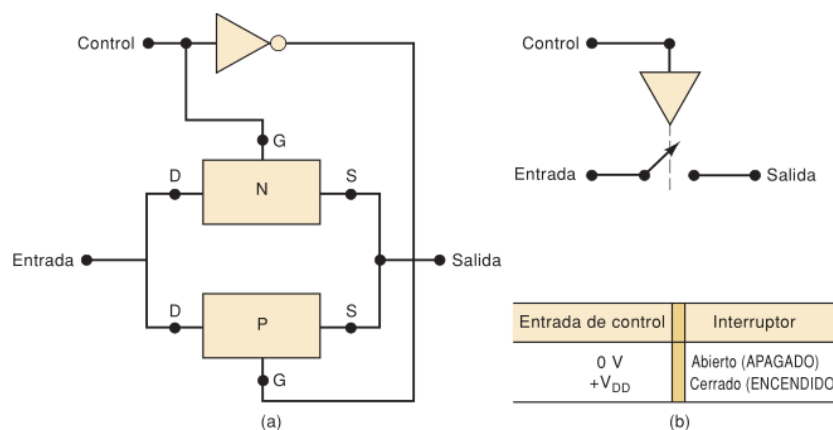


Figura 2.7. Interruptor bilateral CMOS (compuerta de transmisión).⁹

⁹ R. J. Tocci, G. L. Moss y N. S. Widmer, Sistemas Digitales. Principios y Aplicaciones, Pearson Educación, 2007. Figura 8-43.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

2.5 Interfaz de CMOS de alto voltaje con TTL

Antes de considerar el problema de integrar las salidas CMOS con las entradas TTL, será útil revisar las características de salida de CMOS para los dos estados lógicos.

La figura 2.8 muestra el circuito de salida equivalente en el estado ALTO. La R_{ENC} del P-MOSFET conecta la terminal de salida a V_{DD} (recuerde que el N-MOSFET está apagado). En consecuencia, el circuito de salida CMOS actúa como una fuente de V_{DD} con una resistencia de R_{ENC} . Por lo general el valor de R_{ENC} varía de 100 a 1000 Ω .

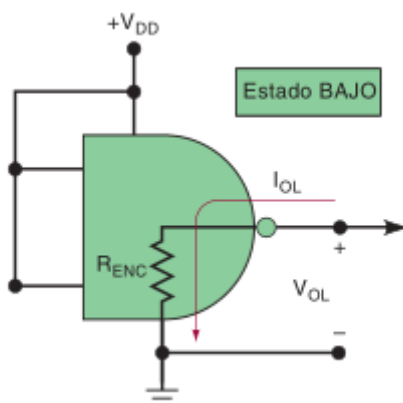


Figura 2.8 Interfaz de CMOS de alto voltaje con TTL¹⁰

2.6 Interfaz de CMOS de bajo voltaje con TTL

La figura 2.9 muestra el circuito de salida equivalente en el estado BAJO. La R_{ENC} del N-MOSFET conecta la terminal de salida a tierra (recuerde que el PMOSFET está apagado). Por ende, la salida CMOS actúa como una resistencia baja a tierra; es decir, actúa como colector de corriente.

¹⁰ R. J. Tocci, G. L. Moss y N. S. Widmer, Sistemas Digitales. Principios y Aplicaciones, Pearson Educación, 2007. Figura 8-48 (a).



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

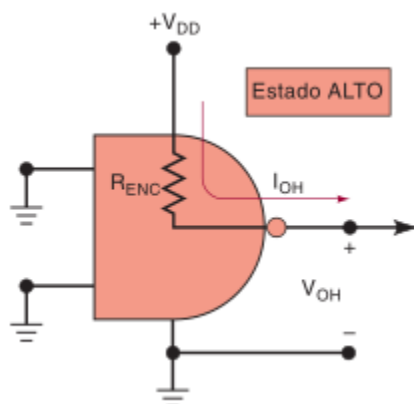


Figura 2.9 Interfaz de CMOS de alto voltaje con TTL¹¹

2.7 Interfaz de TTL con CMOS de bajo voltaje

Las salidas CMOS pueden suministrar con facilidad el suficiente voltaje (V_{OH}) como para satisfacer el requerimiento de entrada de TTL en el estado ALTO (V_{IH}). Las salidas CMOS pueden suministrar más de la suficiente corriente (I_{OH}) como para cumplir con los requerimientos de corriente de entrada de TTL (I_{IH}). Por lo tanto, no se necesita ninguna consideración especial para el estado ALTO.

2.8 Interfaz de TTL con CMOS de alto voltaje

Las entradas TTL tienen una corriente de entrada relativamente alta en el estado BAJO, cuyo intervalo es de 100 A a 2 mA. Las series 74HC y 74HCT pueden drenar hasta 4 mA, por lo que no tendrían problema para manejar una carga TTL individual de cualquier serie. No obstante, la serie 4000B es mucho más limitada. Su capacidad de I_{OL} en nivel bajo no es suficiente como para manejar incluso una entrada de las series 74S o 74AS. La serie 74AHC tiene un control de salida comparable a de la serie 74LS.

¹¹ R. J. Tocci, G. L. Moss y N. S. Widmer, Sistemas Digitales. Principios y Aplicaciones, Pearson Educación, 2007. Figura 8-48 (b).



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

2.9 Interfaz con tecnologías LVT.

En ciertas situaciones, el V_{OH} del reforzador es un poco menor que lo que requiere la carga para reconocerlo como un nivel ALTO. Hablamos sobre esta situación cuando vimos cómo integrar salidas TTL con entradas CMOS de 5 V. El único componente de interfase necesario era una resistencia de puesta en alto de 10-k, la cual provoca que el voltaje de salida TTL se impulse por encima del nivel de 3.3 V cuando la salida está en ALTO.

Cuando existe la necesidad de un desplazamiento más sustancial en el voltaje debido a que el reforzador y la carga operan con distintos voltajes de la fuente de energía, se requiere un circuito de interfase conocido como traductor de nivel de voltaje. Un ejemplo es un dispositivo CMOS de bajo voltaje (1.8 V) que maneja una entrada CMOS de 5 V.

El reforzador puede producir un máximo de sólo 1.8 V como nivel ALTO, y la compuerta de carga requiere 3.33 V para un nivel ALTO. Necesitamos una interfase que pueda aceptar niveles lógicos de 1.8 V y que los traduzca a niveles de CMOS de 5 V. La manera más simple de lograr esto es mediante un búfer que tenga un colector abierto, tal como el 74LVC07 que se muestra en la figura 2.10(a).

La resistencia de puesta en alto se conecta a la alimentación de 5 V, mientras que la fuente de energía para el búfer de interfase es de 1.8 V. Otra solución es utilizar un circuito traductor de nivel de voltaje de alimentación dual tal como el 74AVC1T45, como se muestra en la figura 2.10(b). Este dispositivo utiliza dos voltajes distintos de la fuente de poder, uno para las entradas y otro para las salidas, y traduce entre los dos niveles.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

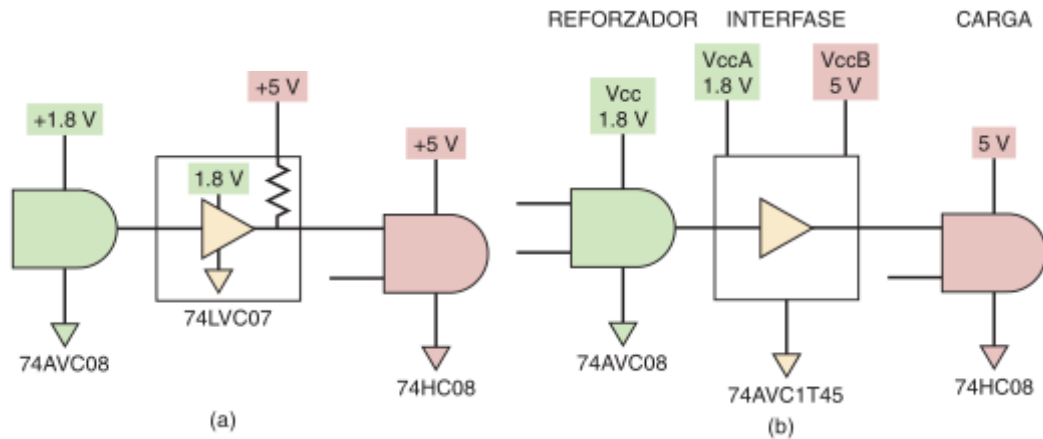


Figura 2.10. (a) Colector abierto con elevación a un alto voltaje (b) Uso de un traductor de nivel de voltaje. ¹²

¹² R. J. Tocci, G. L. Moss y N. S. Widmer, Sistemas Digitales. Principios y Aplicaciones, Pearson Educación, 2007. Figura 8-50.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

3. Memoria

Como ya mencioné anteriormente, la memoria es un dispositivo principal de las computadoras, se utiliza para almacenar datos e instrucciones. Es el espacio destinado al almacenamiento en la computadora donde los datos que van a ser procesados se almacenan, y dependiendo de la arquitectura, también las instrucciones necesarias para procesamiento.

3.1 Conceptos generales de la memoria

Localidad

La base para la mejora de prestaciones de una memoria de dos niveles es un principio conocido como localidad de las referencias. Establece que las referencias a memoria tienden a formar agrupaciones. A lo largo de un período de tiempo largo las agrupaciones en uso cambian, pero durante periodos cortos el procesador trabaja fundamentalmente con agrupaciones fijas de referencias a memoria.

La localidad espacial se refiere a la tendencia durante la ejecución a involucrar múltiples posiciones de memoria que estén agrupadas. La localidad espacial refleja la tendencia del procesador a acceder a las instrucciones secuencialmente y también la tendencia de los programas a acceder a posiciones de datos consecutivas.

La localidad temporal hace referencia a la tendencia del procesador a acceder a posiciones de memoria que han sido utilizadas recientemente.

Funcionamiento

La memoria del nivel superior M1 es más pequeña y rápida, pero más costosa (por bit) que la del nivel inferior M2. M1 se utiliza como almacén temporal para una parte del contenido de la otra más grande. Cuando se hace una referencia a memoria, se



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

intenta acceder al elemento en M1. Si tiene éxito, entonces tiene lugar un acceso rápido. Si no, se copia un bloque de posiciones de memoria de M2 a M1.

3.2 Tecnologías de memorias ROM (antecedentes y evolución)

Memoria de Solo Lectura (ROM), contiene un patrón permanente de datos que no puede alterarse. Este tipo de memoria no es volátil, no requiere fuente de alimentación para mantener memorizados los valores de los bits. Es posible leer una ROM, pero no se pueden escribir nuevos datos en ella. Sus aplicaciones potenciales son:

- Subrutinas de biblioteca para funciones de uso frecuente.
- Programas del sistema.
- Tablas de funciones.

Su alternativa económica es la ROM programable PROM, no es volátil y solo puede grabarse una vez. El proceso de escritura se lleva a cabo eléctricamente y puede realizarlo el programador.

La memoria de sobre-todo lectura es útil para aplicaciones en las que las operaciones de lectura son bastante frecuentes, más que las de escritura. Existen tres formas comunes de este tipo de memorias: EPROM, EEPROM y la memoria flash.

La memoria de solo – lectura programable y borrrable ópticamente EPROM, se lee y escribe eléctricamente como la PROM, sin embargo, antes de la escritura, todas las celdas de almacenamiento deben primero borrarse a la vez, mediante la exposición del chip encapsulado a radiación ultravioleta. Su alternativa es la EEPROM, que se borra eléctricamente y no ópticamente.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

La memoria flash se introdujo a mediados de 1980, se denominan así por la velocidad con la que puede reprogramarse. Esta memoria no permite borrar a nivel de byte.

Antecedentes

El tipo más simple de ROM en estado sólido es de la misma antigüedad que la propia tecnología semiconductor. Las puertas lógicas combinacionales pueden usarse en conjunto para indexar una dirección de memoria de n bits en valores de m bits de tamaño (una tabla de consultas). Con la invención de los circuitos integrados se desarrolló la máscara ROM. La máscara ROM consistía en una cuadrícula de líneas formadas por una palabra y líneas formadas por un bit seleccionadas respectivamente a partir de cambios en el transistor. De esta manera podían representar una tabla de consultas arbitraria y un lapso de propagación deducible.

En las máscaras ROM los datos están físicamente codificados en el mismo circuito, así que solo se pueden programar durante la fabricación. Esto acarrea serias desventajas:

1. Solo es económico comprarlas en grandes cantidades, ya que el usuario contrata fundiciones para producirlas según sus necesidades.
2. El tiempo transcurrido entre completar el diseño de la máscara y recibir el resultado final es muy largo.
3. No son prácticas para I+D por el hecho de que los desarrolladores necesitan cambiar el contenido de la memoria mientras refinan un diseño.
4. Si un producto tiene un error en la máscara, la única manera de arreglarlo es reemplazando físicamente la ROM por otra.

Los desarrollos posteriores tomaron en cuenta estas deficiencias, así pues se creó la memoria de solo lectura programable (PROM). Inventada en 1956, permitía a los



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

usuarios modificarla solo una vez, alterando físicamente su estructura con la aplicación de pulsos de alto voltaje. Esto eliminó los problemas 1 y 2 antes mencionados, ya que una compañía podía pedir un gran lote de PROMs vacías y programarlas con el contenido necesario elegido por los diseñadores. En 1971 se desarrolló la memoria de solo lectura programable y borrable (EPROM) que permitía reiniciar su contenido exponiendo el dispositivo a fuertes rayos ultravioleta. De esta manera erradicaba el punto 3 de la anterior lista. Más tarde, en 1983, se inventó la EEPROM, resolviendo el conflicto número 4 de la lista ya que se podía reprogramar el contenido mientras proveyese un mecanismo para recibir contenido externo (por ejemplo, a través de un cable serial). En medio de la década de 1980 Toshiba inventó la memoria flash, una forma de EEPROM que permitía eliminar y reprogramar contenido en una misma operación mediante pulsos eléctricos miles de veces sin sufrir ningún daño.

Todas estas tecnologías mejoraron la versatilidad y flexibilidad de la ROM, pero lo hicieron a expensas de un alto incremento del costo por chip. Por eso las máscaras ROM se mantuvieron como la solución económica durante bastante tiempo. Esto fue así aproximadamente hasta el año 2000, cuando el precio de las memorias reprogramables hubo descendido lo suficiente como para comenzar a desplazar a las ROM no reprogramables del mercado.

El producto más reciente es la memoria NAND, otra vez desarrollada por Toshiba. Los diseñadores rompieron explícitamente con las prácticas del pasado, afirmando que enfocaba "ser un reemplazo de los discos duros", más que tener el tradicional uso de la ROM como una forma de almacenamiento primario no volátil. En 2007, NAND ha avanzado bastante en su meta, ofreciendo un rendimiento comparable al de los discos duros, una mejor tolerancia a los shocks físicos, una miniaturización



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

extrema (como por ejemplo memorias USB y tarjetas de memoria MicroSD), y un consumo de potencia mucho más bajo.

3.3 Tecnologías de memorias RAM (antecedentes y evolución)

Las memorias de acceso aleatorio acceden a palabras individuales directamente mediante lógica de direccionamiento cableada interna. La más común es denominada **memoria de acceso aleatorio (RAM)**. Esta memoria es volátil, debe estar siempre alimentada y si su alimentación se interrumpe pierda los datos. Esta memoria se puede interpretar en el procesador con DRAM (dinámica) o SRAM (estática).

La DRAM está hecha con celdas que almacenan los datos como cargas eléctricas en condensadores.

La SRAM es un dispositivo basado en los mismos elementos que usa el procesador. Los valores binarios se almacenan utilizando configuraciones de Flip-Flops.

Una celda de memoria RAM dinámica es más simple que una estáticas, por lo tanto son más densas y más económicas.

Antecedentes

Uno de los primeros tipos de memoria RAM fue la memoria de núcleo magnético, desarrollada entre 1949 y 1952 y usada en muchos computadores hasta el desarrollo de circuitos integrados a finales de los años 60 y principios de los 70. Esa memoria requería que cada bit estuviera almacenado en un toroide de material ferromagnético de algunos milímetros de diámetro, lo que resultaba en dispositivos con una capacidad de memoria muy pequeña. Antes que eso, las computadoras usaban relés



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

y líneas de retardo de varios tipos construidas para implementar las funciones de memoria principal con o sin acceso aleatorio.

En 1969 fueron lanzadas una de las primeras memorias RAM basadas en semiconductores de silicio por parte de Intel con el integrado 3101 de 64 bits de memoria y para el siguiente año se presentó una memoria DRAM de 1024 bytes, referencia 1103 que se constituyó en un hito, ya que fue la primera en ser comercializada con éxito, lo que significó el principio del fin para las memorias de núcleo magnético. En comparación con los integrados de memoria DRAM actuales, la 1103 es primitiva en varios aspectos, pero tenía un desempeño mayor que la memoria de núcleos.

En 1973 se presentó una innovación que permitió otra miniaturización y se convirtió en estándar para las memorias DRAM: la multiplexación en tiempo de las direcciones de memoria. MOSTEK lanzó la referencia MK4096 de 4096 bytes en un empaque de 16 pines,¹ mientras sus competidores las fabricaban en el empaque DIP de 22 pines. El esquema de direccionamiento² se convirtió en un estándar de facto debido a la gran popularidad que logró esta referencia de DRAM. Para finales de los 70 los integrados eran usados en la mayoría de las computadoras nuevas, se soldaban directamente a las placas base o se instalaban en zócalos, de manera que ocupaban un área extensa de circuito impreso. Con el tiempo se hizo obvio que la instalación de RAM sobre el impreso principal, impedía la miniaturización, entonces se idearon los primeros módulos de memoria como el SIPP, aprovechando las ventajas de la construcción modular. El formato SIMM fue una mejora al anterior, eliminando los pines metálicos y dejando unas áreas de cobre en uno de los bordes del impreso, muy similares a los de las tarjetas de expansión, de hecho los módulos SIPP y los primeros SIMM tienen la misma distribución de pines.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

A finales de los 80 el aumento en la velocidad de los procesadores y el aumento en el ancho de banda requerido, dejaron rezagadas a las memorias DRAM con el esquema original MOSTEK, de manera que se realizaron una serie de mejoras en el direccionamiento como las siguientes:

Módulos formato SIMM de 30 y 72 pines, los últimos fueron utilizados con integrados tipo EDO-RAM.

Fast Page Mode RAM (FPM-RAM) fue inspirado en técnicas como el Burst Mode usado en procesadores como el Intel 486.3 Se implantó un modo direccionamiento en el que el controlador de memoria envía una sola dirección y recibe a cambio esa y varias consecutivas sin necesidad de generar todas las direcciones. Esto supone un ahorro de tiempos ya que ciertas operaciones son repetitivas cuando se desea acceder a muchas posiciones consecutivas. Funciona como si deseáramos visitar todas las casas en una calle: después de la primera vez no sería necesario decir el número de la calle únicamente seguir la misma. Se fabricaban con tiempos de acceso de 70 ó 60 ns y fueron muy populares en sistemas basados en el 486 y los primeros Pentium.

Extended Data Output RAM (EDO-RAM) fue lanzada al mercado en 1994 y con tiempos de accesos de 40 o 30 ns suponía una mejora sobre FPM, su antecesora. La EDO, también es capaz de enviar direcciones contiguas pero direcciona la columna que va a utilizar mientras que se lee la información de la columna anterior, dando como resultado una eliminación de estados de espera, manteniendo activo el búfer de salida hasta que comienza el próximo ciclo de lectura.

Burst Extended Data Output RAM (BEDO-RAM) fue la evolución de la EDO-RAM y competidora de la SDRAM, fue presentada en 1997. Era un tipo de memoria que usaba generadores internos de direcciones y accedía a más de una posición de



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

memoria en cada ciclo de reloj, de manera que lograba un 50 % de beneficios, mejor que la EDO. Nunca salió al mercado, dado que Intel y otros fabricantes se decidieron por esquemas de memoria sincrónicos que si bien tenían mucho del direccionamiento MOSTEK, agregan funcionalidades distintas como señales de reloj.

3.4 Características principales de las memorias

Volatilidad de la información

La memoria volátil requiere energía constante para mantener la información almacenada. La memoria volátil se suele usar solo en memorias primarias. La memoria RAM es una memoria volátil, ya que pierde información en la falta de energía eléctrica.

La memoria no volátil retendrá la información almacenada incluso si no recibe corriente eléctrica constantemente, como es el caso de la memoria ROM. Se usa para almacenamientos a largo plazo y, por tanto, se usa en memorias secundarias, terciarias y fuera de línea.

La memoria dinámica es una memoria volátil que además requiere que periódicamente se refresque la información almacenada, o leída y reescrita sin modificaciones.

Accesibilidad secuencial o aleatoria a información

Acceso aleatorio, significa que se puede acceder a cualquier localización de la memoria en cualquier momento en el mismo intervalo de tiempo, normalmente pequeño.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

Acceso secuencial, significa que acceder a una unidad de información tomará un intervalo de tiempo variable, dependiendo de la unidad de información que fue leída anteriormente.

Habilidad para cambiar la información

Las memorias de lectura/escritura o memorias cambiables permiten que la información se reescriba en cualquier momento. Una computadora sin algo de memoria de lectura/escritura como memoria principal sería inútil para muchas tareas. Las computadoras modernas también usan habitualmente memorias de lectura/escritura como memoria secundaria.

La memoria de sólo lectura (Read-Only Memory, ROM) retiene la información almacenada en el momento de fabricarse.

La memoria de escritura única lectura múltiple (Write Once Read Many, WORM) permite que la información se escriba una sola vez en algún momento tras la fabricación.

Las memorias de escritura lenta y lectura rápida son memorias de lectura/escritura que permite que la información se reescriba múltiples veces pero con una velocidad de escritura mucho menor que la de lectura. Un ejemplo son los CD-RW.

Direccionamiento de la información

En la memoria de localización direccionable, cada unidad de información accesible individualmente en la memoria se selecciona con su dirección de memoria numérica.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

En las memorias de sistema de archivos, la información se divide en archivos informáticos de longitud variable y un fichero concreto se localiza en directorios y nombres de archivos «legible por humanos

En las memorias de contenido direccionable (content-addressable memory), cada unidad de información legible individualmente se selecciona con un valor hash o un identificador corto sin relación con la dirección de memoria en la que se almacena la información.

Capacidad de memoria

Minicomputadoras: se caracterizan por tener una configuración básica regular que puede estar compuesta por un monitor, unidades de disquete, disco, impresora, etc. Su capacidad de memoria varía de 16 a 256 KiB.

Macrocomputadoras: son aquellas que dentro de su configuración básica contienen unidades que proveen de capacidad masiva de información, terminales (monitores), etc. Su capacidad de memoria varía desde 256 a 512 KiB, también puede tener varios megabytes o hasta gigabytes según las necesidades de la empresa.

Microcomputadores y computadoras personales: con el avance de la microelectrónica en la década de los 70 resultaba posible incluir todos los componente del procesador central de una computadora en un solo circuito integrado llamado microprocesador. Ésta fue la base de creación de unas computadoras a las que se les llamó microcomputadoras.



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

4. Referencias

- [1] W. Stallings, Organización y arquitectura de computadoras, Madrid: Pearson Educación, 2005.
- [2] J. B. Vazquez Gomez, “Introducción a la informática”, Estado de México: Red Tercer Milenio, 2012.
- [3] Leandro Alega “Plataforma Informática”. [En línea]. Enlace: alegsa.com.ar/Dic/plataforma.php [Último acceso: 17 Octubre 2020].
- [4] Académicos UNAM, “Introducción a la computación: Unidad I”. [En línea]. Enlace: bunam.unam.mx/intComputacion/Unidad_1/c02u1t02p01.html [Último acceso: 17 Octubre 2020].
- [5] R. J. Tocci, G. L. Moss y N. S. Widmer, Sistemas Digitales. Principios y Aplicaciones, Pearson Educación, 2007.
- [6] J. B. Vazquez Gomez, “Arquitectura de computadoras I”, Estado de México: Red Tercer Milenio, 2012.
- [7] Escuela Politécnica Nacional CIECFIE, “Conceptos Adicionales de Arquitectura de computadoras”. Enlace: ciecfie.epn.edu.ec/wss/VirtualDirectories/80/pag_personales/PChico/Materiales_Micros/Conceptos%20adicionales.pdf [Último acceso: 21 Octubre 2020].
- [8] Docencia FCA UNAM, “Estructura de un ordenador”. Enlace: docencia.fca.unam.mx/~rcastro/estructura_compu.pdf [Último acceso: 21 Octubre 2020].
- [9] J Savage. G. Vázquez. UNAM FI, “DISEÑO DE MICROPROCESADORES – UNAM” Enlace: profesores.fib.unam.mx/luist/archivos/MicroprocesadoresSavage.pdf [Último acceso: 18 Octubre 2020].
- [10] X. Sierra. “Microprocesadores: lógica y diseño”. Enlace: monografias.com/trabajos14/micros/micros.shtml#DISE%C3%91O [Último acceso: 18 Octubre 2020].
- [11] H. Corvo. “Unidad de control: componentes y funciones”. Enlace: lifeder.com/unidad-de-control/ [Último acceso: 20 Octubre 2020].
- [12] Wikipedia “Unidad funcional”. Enlace: es.wikipedia.org/wiki/Unidad_funcional [Último acceso: 20 Octubre 2020].
- [13] Y. Agüero “Tipos de Buses en la Informática y su función”. Enlace: <https://tecnoinformatic.com/c-informatica-basica/tipos-de-buses/> [Último acceso: 19 Octubre 2020].



Tarea 1. Investigación de temas de Arquitectura de Computadoras.

- [14] P. Tumero “Ciclos de Instrucción, búsqueda y ejecución”. Enlace: monografias.com/docs110/ciclo-instruccion-busqueda-y-ejecucion/ciclo-instruccion-busqueda-y-ejecucion.shtml [Último acceso: 20 Octubre 2020].
- [15] M.E. Raffino. “Memoria Caché”. Enlace: concepto.de/memoria-cache/ [Último acceso: 22 Octubre 2020].
- [16] EDM. “Jerarquía de la memoria”. Enlace: estructuradealmacenamientodememoria.wordpress.com/2016/12/07/jerarquia-de-la-memoria/ [Último acceso: 21 Octubre 2020].
- [17] Tutorials Point “Memoria”. Enlace: tutorialspoint.com/es/computer_fundamentals/computer_memory.html [Último acceso: 21 Octubre 2020].
- [18] IEEE “IEEE Std 754-2008: IEEE Standard for Floating-Point Arithmetic”. IEEE Inc. 2008.
- [19] Wikipedia “Memoria de acceso aleatorio: Historia”. Enlace: wikipedia.org/wiki/Memoria_de_acceso_aleatorio#Historia [Último acceso: 23 Octubre 2020].
- [20] Wikipedia “Memoria de solo lectura: Historia”. Enlace: wikipedia.org/wiki/Memoria_de_solo_lectura#Historia [Último acceso: 23 Octubre 2020].
- [21] Wikipedia “Memoria”. Enlace: wikipedia.org/wiki/Memoria [Último acceso: 23 Octubre 2020].