APUNTES DE
DISEÑO DE DISEÑO DE CIRCUITOS INTEGRADOS 1

ESCUELA SUPERIOR DE INFORMÁTICA

UNIVERSIDAD COMPLUTENSE DE MADRID

Juan Lanchares Dávila

1. INTRODUCCIÓN

1.1. PERSPECTIVA HISTORICA
1.2. DEL DISEÑADOR DE CIRCUITOS INTEGRADOS
1.3. EL CICLO DE DISEÑO VLSI
1.3.1. FASES DEL CICLO DE DISEÑO
1.4. BREVE HISTORIA DE LAS HERRAMIENTAS CAD
1.5. ESTUDIO LÓGICO DE LOS TRANSISTORES MOS
1.5.1. INTERRUPTOR NMOS
1.5.2. INTERRUPTOR PMOS
1.5.3. INTERRUPTOR CMOS
1.6. LOGICA CMOS
1.6.1. INVERSOR CMOS
1.6.2. PUERTAS NAND CMOS DE M ENTRADAS
1.6.3. PUERTA NOR CMOS DE M ENTRADAS
1.6.4. PRINCIPALES CARACTERISTICAS DE LA LÓGICA CMOS
1.6.5. LÓGICA DE MULTIPLEXORES
1.6.6. ELEMENTOS DE MEMORIA
1.6.6.1. Biestable D
1.6.6.2. Carga por flanco
1.7. PRINCIPALES CARACTERÍSTICAS DE LA LÓGICA CMOS
1.8. TECNOLOGÍA CMOS VS OTRAS TECNOLOGÍAS

2. METODOLOGÍAS DE DISEÑO

2.1. SIMULACIÓN
2.1.1. SIMULACIÓN A NIVEL DE CIRCUITO
2.1.2. SIMULACIÓN DE TIMMING
2.1.3. SIMULACIÓN LÓGICA
2.2. VERIFICACIÓN
2.2.1. VERIFICACIÓN FUNCIONAL (O FORMAL)
2.3. SÍNTESIS DE DISEÑOS
2.3.1. SÍNTESIS DE ARQUITECTURA
2.3.2. SÍNTESIS RTL
2.3.3. SÍNTESIS LÓGICA
2.3.4. SÍNTESIS DE CIRCUITOS
2.4. VALIDACIÓN Y TEST
2.4.1. TEST DE VALIDACIÓN
2.4.2. DISEÑO PARA TESTABILIDAD

3. ESTILOS DE DISEÑO

3.1. EL DISEÑO FULL-CUSTOM

3.2. EL DISEÑO SEMICUSTOM
3.2.1. BIBLIOTECAS DE CELDAS ESTÁNDARES
3.2.1.1. Celdas compiladas
3.2.1.2. generadores de módulos
3.2.2. GATE ARRAYS
3.3. ELECCIÓN DEL ESTILO DE DISEÑO
3.4. ARQUITECTURAS FPGA
3.4.1. ¿QUÉ ES UNA FPGA?
3.4.2. TECNOLOGÍAS FPGA
3.4.2.1. FPGA´s basadas en SRAM
3.4.3. VENTAJAS E INCONVENIENTES

4. TEORÍA DEL TRANSISTOR MOS

4.1. TRANSISTOR MOS
4.2. TRANSISTOR NMOS DE ENRIQUECIMIENTO
4.2.1. DESCRIPCIÓN.
4.2.2. MODO DE OPERACIÓN
4.2.3. REGIONES DE TRABAJO DE UN TRANSISTOR NMOS DE ENRIQUECIMIENTO
4.3. TRANSISTOR PMOS DE ACUMULACION
4.4. TRANSISTORES DE EMPOBRECIMIENTO
4.5. COMPARACION ENTRE PMOS Y NMOS
4.6. POTENCIAL UMBRAL
4.6.1. ECUACIONES DEL POTENCIAL UMBRAL
4.6.1.1. Potencial de capacitor MOS, V_{TMOS}
4.6.1.2. Potencial de Flat-Band
4.7. EFECTO SUSTRATO [EFECTO BODY]
4.8. ECUACIONES BASICAS DE DISPOSITIVOS MOS
4.9. EFECTOS DE SEGUNDO ORDEN
4.9.1. VARIACIONES DEL POTENCIAL UMBRAL
4.9.2. CONDUCCIÓN SUBUMBRAL
4.9.3. SATURACIÓN DE LA VELOCIDAD DE LOS PORTADORES
4.9.4. DEGRADACIÓN DE LA MOVILIDAD
4.9.5. TUNEL FOWLER-NORHEIM
4.9.6. PERFORACION DE CANAL
4.9.7. ELECTRONES CALIENTES. IONIZACION DE IMPACTO

5. LOS INVERSORES MOS

5.1. DEFINICIONES Y PROPIEDADES
5.1.1. RUIDO
5.1.2. MARGENES DE RUIDO
5.2. EL INVESOR CMOS DE CARGA DINÁMICA

5.2.1. REGIONES DE TRABAJO DEL INVERSOR
5.2.2. INFLUENCIA DE LA RELACIÓN βN/βP EN LA CARACTERÍSTICA DE ENTRADA SALIDA
5.2.3. INVERSORES MOS DE CARGA ESTATICA
5.2.4. INVERSOR NMOS GENÉRICO
5.2.5. EL INVERSOR PSEUDONMOS
5.2.6. INVERSOR PSEUDONMOS DE CARGA SATURADA
5.2.7. INVERSOR CON CARGA NMOS DE ENRIQUECIMEINTO
5.2.8. INVERSOR CON CARGA PMOS DE ENRIQUECIMIENTO
5.2.9. INVERSOR DE CARGA NMOS DE EMPOBRECIMIENTO
5.2.10. INVERSOR DE CONEXIÓN A LOGICA TTL
5.3. LA PUERTA DE TRANSMISIÓN
5.3.1. TRANSISTOR DE PASO NMOS
5.3.2. TRANSISTOR DE PASO PMOS
5.4. EL INVERSOR TRIESTATE
5.5. 3.6 TECNOLOGÍA BICMOS
5.5.1. INVERSOR BICMOS

6. TECNOLOGIA DE PROCESOS CMOS

6.1. FABRICACION BASICA DE DISPOSITIVOS
6.1.1. OXIDACIÓN
6.1.2. EPITAXIS, DEPOSICIÓN, IMPLANTACIÓN IONICA Y DIFUSIÓN
6.1.3. FABRICACIÓN DEL AREA ACTIVA
6.1.4. POLISILICIO
6.1.5. PASOS PARA UN PROCESO TIPICO DE FABRICACION DE UNA PUERTA (TRANSISTOR MOS)
6.1.6. TRANSISTORES MOS PARASITOS
6.2. TECNOLOGIA CMOS BASICA
6.2.1. CMOS DE POZO N
6.2.1.1. Polarización de los substratos
6.2.2. CMOS DE POZO P
6.2.3. PROCESO TWIN TUB
6.2.4. SILICIO SOBRE AISLANTE (SOI).
6.3. MEJORAS EN LOS PROCESOS CMOS
6.3.1. MEJORA DE LA RUTABILIDAD
6.3.1.1. Mas de un nivel de metal
6.3.2. MEJORA DEL LAYER DE POLISILICIO
6.3.3. ELEMENTOS PASIVOS DEL CIRCUITO.
6.3.3.1. Resistencias
6.3.3.2. Capacitores
6.4. ROM'S ALTERABLES ELECTRICAMENTE

6.5. LATCH-UP
6.5.1. ESTUDIO DEL CIRCUITO REALIMENTADO
6.5.2. DISPARO DEL LATCH-UP
6.5.2.1. Ejemplos de carga vertical y horizontal
6.5.3. PREVENCIÓN DEL LATCH-UP
6.5.3.1. Mejoras en los procesos
6.5.3.2. Mejoras en el diseño de Layouts
6.5.3.3. Prevención del Latch-up en los dispositivos de E/S
6.6. REGLAS DE DISEÑO
6.6.1. REPRESENTACIÓN DE LAYERS
6.6.2. ENUMERACIÓN DE ALGUNAS REGLAS
6.7. REGLAS DE DISEÑO DE ES2 PARA LA TECNOLOGÍA DE 0.7 MICRAS ECPD7

7. CARACTERIZACIÓN DE CIRCUITOS

7.1. ESTIMACIÓN DE RESISTENCIAS
7.1.1. RESISTENCIA DE UNA PLANCHA DE MATERIAL
7.2. RESISTENCIA DE REGIONES NO RECTANGULARES
7.3. ESTIMACIÓN DE CAPACIDADES
7.3.1. CARACTERÍSTICAS DEL CAPACITOR MOS
7.3.2. CAPACIDAD DE LOS DISPOSITIVOS MOS
7.3.2.1. Capacidad de puerta
7.3.3. CAPACIDAD DE DIFUSIÓN
7.4. CAPACIDADES DE CONEXIONADO
7.4.1. EFECTO DE LAS MEJORAS TECNOLÓGICAS EN LA CAPACIDAD CONEXIONADO
7.4.2. CAPACIDAD DE LOS CAMPOS LATERALES
7.4.3. LOS CAPACIDADES ENTRE MÚLTIPLES CONDUCTORES CROSS-TALK
7.5. TIEMPO DE RETARDO RC EN LA PROPAGACIÓN DE LA SEÑAL
7.5.1. REDUCCIÓN DE LOS RETARDOS RC

7.5.2. REDUCCIÓN DEL RETARDO DE LA LÍNEA DE RELOJ
7.5.3. CUANDO SE DEBEN TENER EN CUENTA LA LONGITUD DEL HILO
7.6. CARACTERÍSTICAS DE COMUTACIÓN
7.7. MODELOS ANALÍTICOS DEL RETARDO
7.7.1. TIEMPO DE BAJADA
7.7.1.1. Estudio cualitativo
7.7.1.2. Estudio cuantitativo
7.7.2. INVERSOR EQUIVALENTE
8. El **buscando la ganancia equivalente de cada una de las ramas hasta**

8.1.1. CIRCUITOS DE FAN-OUT ELEVADO
8.1.1.1. Etapas de transistores en cascada
8.1.2. SOLUCIÓN AL PROBLEMA DEL FAN-OUT ELEVADO
8.2. DISIPACIÓN DE POTENCIA
8.2.1. POTENCIA ESTÁTICA
8.2.2. POTENCIA DINÁMICA
8.2.2.1. Potencia debido a las capacidades
8.2.2.2. Disipación de corto circuito
8.2.3. RELACIÓN DEL CONSUMO DE POTENCIA CON LA TEMPERATURA

9. LÓGICA COMBINACIONAL ESTÁTICA

9.1. DISEÑO CMOS ESTATICO
9.2. LÓGICA CMOS COMPLEMENTARIA
9.2.1. TÉCNICAS DE DISEÑO PARA EVITAR LOS PROBLEMAS DE FAN-IN ELEVADO
9.3. REGULADORES DE POTENCIAL INTERNOS
9.4. LÓGICA PROPORCIONAL
9.4.1. CARGA A TRAVES DE UNA RESISTENCIA
9.4.2. CARGA MEDIANTE UN TRANSISTOR DE DEPLEXION
9.4.3. CARGA A TRAVÉS PSEUDO-NMOS
9.4.3.1. Ajuste del comportamiento
9.4.3.2. Consideraciones de diseño
9.4.3.3. Modificaciones del dispositivo de carga PMOS
9.4.4. DIFERENTIAL CASCADE VOLTAGE SWITCH (DCVS)
9.5. LÓGICA DE INTERRUPTORES
9.5.1. TRANSISTORES DE PASO
9.5.2. PUERTAS DE TRANSMISION
9.6. LÓGICA DE MULTIPLEXORES
9.6.1. IMPLEMENTACIÓN DE UNA UNIDAD GENERADORA DE FUNCIONES DE DOS ENTRADAS
9.7. CPL COMPLEMENTARY PASS-TRANSISTOR LOGIC

10. LÓGICA COMBINACIONAL DINÁMICA

10.1. PRINCIPIOS
10.2. CARACTERÍSTICAS
10.3. ANÁLISIS DE LOS TIEMPOS DE SUBIDA Y BAJADA
10.4. CORRIENTES DE PERDIDA
10.4.1. DISTRIBUCIÓN DE CARGA
10.4.2. ACOPLAMIENTO DE RELOJ (CLOCK-FEEDTROUGH).
10.5. PUERTAS DINÁMICAS EN CASCADA
10.6. LÓGICA DOMINO
10.6.1. EJEMPLO DE IMPLEMENTACIÓN CON LÓGICA DOMINÓ

10.6.2. PROPIEDADES DE LA LÓGICA DOMINÓ
10.7. LÓGICA DOMINO NP O ZIPPER
10.8. C^2MOS CLOCKED CMOS
10.9. CONSUMO DE POTENCIA EN PUERTAS CMOS
10.9.1. ACTIVIDAD DE CONEXIÓNADO (INTERRUPCION) EN LÓGICA ESTÁTICA
10.9.2. ACTIVIDAD DE INTERRUPCIÓN EN LA LÓGICA DINÁMICA
10.9.3. GLITCHING (FALLOS)
10.9.4. CORRIENTES DE CORTOCIRCUITO
10.10. DISEÑO CMOS DE BAJA POTENCIA
10.10.1. MODIFICACION DEL POTENCIAL DE ALIMENTACION
10.10.2. REDUCCION DE LA CAPACIDAD EFECTIVA
10.11. ELECCIÓN DEL ESTILO LÓGICO

11.DISEÑO SECUENCIAL

11.1. TIPOS DE SISTEMAS SECUENCIALES
11.2. TIEMPOS RELEVANTES EN LA CARGA DE UN DISPOSITIVO
11.3. ELEMENTOS DE MEMORIA
11.4. PIPELINE CON REGISTROS Y CON LATCHES
11.4.1. CON REGISTROS:
11.4.2. CON LATCHES
11.5. TIEMPOS DE SET-UP Y DE HOLD
11.6. ESTRUCTURAS DE MEMORIA DE UNA FASE DE RELOJ
11.6.1. REGISTRO ESTÁTICO DE UNA FASE DE RELOJ
11.6.2. OTROS REGISTROS ESTÁTICOS
11.6.2.1. Latch estático c^2MOS
11.6.2.2. Latch con solo una puerta de transmisión
11.6.2.3. Celda RAM estática
11.6.2.4. Registros con Set y Reset asíncronos
11.6.2.5. registro con set y reset síncrono:
11.6.3. REGISTROS DINÁMICOS
11.6.3.1. Registro de desplazamiento dinámico
11.6.3.2. Registro C^2MOS
11.7. ESTRUCTURAS LÓGICAS DE UNA FASE DE RELOJ
11.8. LÓGICA NORA
11.8.1. ESTRUCTURA NORA N-P-C2MOS
11.8.2. LÓGICA NOR MIXTA
11.8.3. TRUE SINGLE-PHASE CLOCKED LOGIC (TSPCL)
11.8.4. SPLIT- OUTPUT
11.9. DOS FASES DE CK

11.9.1. ESTRUCTURAS DE MEMORIA DE DOS FASES DE CK
11.10. ESTRUCTURAS LÓGICAS DE DOS FASES DE RELOJ
11.11. ESTUDIO DEL RENDIMIENTO DE LOS CIRCUITOS SECUENCIALES. CLOCK SKEW (DESVIACIONES DE RELOJ)
11.11.1. RETARDO MAXIMO DE PROPAGACION.
11.11.2. SINCRONIZACION DE UNA FASE CK.
11.11.3. SINCRONIZACIÓN DE DOS FASES DE RELOJ.
11.12. SINCRONIZACIÓN DE SISTEMA MEDIANTE PLL (PHASE LOCKED LOOP)

12.DISEÑO DE SUBSISTEMAS DE MEMORIA

12.1. MEMORIAS DE LECTURA/ESCRITURA.
12.2. CELDAS RAM
12.2.1. CELDA DE 6 TRANSISTORES
12.2.2. CELDA DE 4 TRANSISTORES Y CARGA RESISTIVA.
12.2.3. CELDA DE 5 TRANSISTORES
12.2.4. CELDA DE 4 TRANSISTORES (DINAMICA)
12.2.5. CELDA DE TRES TRANSISTORES
12.2.6. CELDA DE UN TRANSISTOR
12.3. LECTURA DE MEMORIAS ESTÁTICAS
12.3.1. LECTURA CON PRECARGA
12.3.2. OTRO CIRCUITO DE PRECARGA MEDIANTE TRANSISTORES N
12.4. RAM SIN PRECARGA (CARGA ESTATICA)
12.5. CIRCUITOS PERIFERICOS
12.5.1. SENSORES AMPLIFICADORES
12.5.2. AMPLIFICADOR DIFERENCIAL
12.5.3. AMPLIFICADOR “CURRENT-MIRROR”
12.5.4. OTRO AMPLIFICADOR
12.6. AMPLIFICADORES DE UNA ENTRADA

1. INTRODUCCIÓN

1.1 PERSPECTIVA HISTORICA

1925 Lilienfeld da los principios básicos de los transistores MOS

- 1935 O´heil da una estructura similar a los MOS actuales
- 1947 Bell inventa el transistor
- 1949 Aparecen los transistores bipolares
  - De las dos tecnologías aparecidas es la bipolar la que se desarrolla con más fuerza
- 1960 Aparece la **tecnología bipolar TTL:**
  - Es la primera familia lógica que realmente tiene éxitos como circuito integrado.
  - Su principal ventaja frente a otras era su gran densidad de integración y de hecho supuso la primera gran revolución dentro del los circuitos integrados
  - Hasta los años 80 esta tecnología era la que dominaba el mercado
  - Pero tenia un importante problema: el gran consumo de potencia limitaba la capacidad de integración de los circuitos.
- Esta es la razón de que se continuara a lo largo de todos estos años en las tecnologías MOS.
- El gran problema de los circuitos MOS desde que en 1925 se enunciaron sus principio era la gran dificultad tecnológica para su fabricación.
- La primera tecnología MOS que se utilizo fue la CMOS pero otra vez la gran dificultad tecnología de su fabricación hizo desistir de ello y se empezó a utilizar la PMOS.
- La segunda revolución del circuito integrado apareció cuando 1970 Intel fabrico el primer microprocesador 4004 y en 1974 el 8080 totalmente en tecnología NMOS, cuyas principales características eran ser más rápida que la PMOS.
- De manera paralela en 1970 aparece la primera memoria semiconductora de gran densidad (1K)
- El principal problema de la tecnología NMOS era el gran consumo de potencia que tenía.

- En la actualidad la tecnología más usada (80-90% de los circuitos) es la CMOS cuyas principal característica es el bajo consumo de potencia y la robustez.
- Existen en la actualidad otras tecnologías:
    - BiCMOS que combina tecnología MOS y bipolar que se usa para memorias de alta velocidad o Gate arrays
    - ECL tecnología bipolar de alto rendimiento
    - Arseniuro de galio

ν **DENSIDADES Y FRECUENCIAS**

- En 1960 Moore predijo que el número de transistores que se podría integrar en un circuito crecería exponencialmente con el tiempo.
- La ley de Moore que se ha cumplido hasta el momento
- En la actualidad :
    - se integran millones de transistores
    - frecuencias de 200mhz

1.2 DEL DISEÑADOR DE CIRCUITOS INTEGRADOS

- En sus inicios el diseño de un circuito consistía en el estudio individualizado de cada transistor buscando su optimización y su perfecta localización en el entorno.
- En la actualidad un diseño puede incluir varios millones de transistores, luego su tratamiento individualizado es imposible. Esto da lugar a la aparición de metodologías de diseño rígidas y estrategias que son automatizables mediante herramientas CAD.
- En lugar de una aproximación individualizada el circuito se aborda de una manera jerárquica. Es decir, el circuito se ve como una colección de módulos, siendo un módulo un conjunto de módulos o de celdas.
    - ∗ Estas celdas se reutilizan tanto como sea posible para reducir el esfuerzo de diseño
- La jerarquización del estudio del un circuito integrado da lugar al concepto de abstracción.
- En cada nivel de diseño los detalles internos de un módulo complejo pueden abstraerse y sustituirse por un modelo de comportamiento o caja negra.
    - ∗ Esta caja negra contiene toda la información para poder conectar el módulo en el nivel de jerarquía superior.
    - ∗ Ventaja: en lugar de tener que trabajar con miles de transistores el diseñador trabaja con módulos sencillos caracterizados por un pequeño conjunto de parámetros.

ν Esta filosofía de diseño ha provocado la aparición de **herramientas CAD**, sin las cuales no sería posible la complejidad de los circuitos actuales. Estas herramientas incluyen

- Simulaciones lógicas y eléctricas
- Generación de layout
- Síntesis
- Verificación

- Para evitar el rediseño y reverificación se usan módulos de memoria y aritméticos puertas básicas ya diseñadas e incluidas en las bibliotecas de celdas estándar.
- Incluso el pentium las usa

- El análisis anterior conduce a pensar que con herramientas CAD y el diseño modular la existencia de diseñadores de circuitos integrados con conocimientos profundos del funcionamiento electrónico de los diseños queda trasnochado.
    - ¿Que necesidad existe de conocer el comportamiento de los de los transistores o de las capacidades y resistencias parásitas?
    - ¿Que necesidad existe de conocer entidades inferiores a puertas y módulos?
- A continuación se dan algunos razones de peso:
    - Todavía hay que diseñar e implementar las bibliotecas de módulos, dado que el importante avance de las tecnologías hace que sean inservible las celdas de una tecnología para la siguiente.
    - Crear el modelo de una celda o módulo requiere un conocimiento de la operación interna que realiza.
    - El diseño basado en librerías trabaja correctamente cuando las ligaduras no son excesivamente fuertes.
        - Ejemplo es el circuito ASIC; el objetivo una solución integrada de rápido funcionamiento.
        - Este no es el caso de muchos diseños, como los microprocesadores, que fuerzan el diseño hasta los limites de la tecnología, en estos casos los módulos se hacen a medida
    - El modelo de abstracción es solo válido hasta ciertos grados

1.3 EL CICLO DE DISEÑO VLSI

- Como otros productos de ingeniería la manufactura de un circuito integrado se compone de
  - Diseño del producto
  - Fabricación
  - Test
- En lo que respecta al ciclo de diseño los objetivos que de be cumplir
  - Alcanzar un diseño óptimo del producto
  - Ciclo de diseño corto
- La principal característica del ciclo de diseño de un CI es que en la actualidad esta en su mayor parte automatizado.
- Debido al gran número de componentes y de detalles que requiere el proceso de fabricación, el diseño no tendría sentido sin las herramientas automáticas
- Efectos de la automatización:
  - Mejores optimizaciones de rendimiento, área, potencia
  - Ciclos de diseño cortos
  - Bajo coste por unidad
  - Introducción en muchos campos de la ciencia y de la vida diaria

1.3.1 FASES DEL CICLO DE DISEÑO

ν Se puede ver el ciclo de diseño de un circuito integrado como una sucesión de fases en cada una de las cuales se transforma la representación del sistema.

ν **FASES**

- Especificación del sistema
- Diseño funcional
- Diseño lógico
- Diseño circuito
- Diseño físico

ν **ESPECIFICACIÓN DEL SISTEMA**

- **QUE**
- Descripción de alto nivel
- Factores a tener en cuenta
    - Rendimiento
    - Funcionalidad a implementar
    - Dimensiones físicas
- Elección de la tecnología de fabricación
- Técnicas de diseño

ν **DISEÑO FUNCIONAL**

- **CÓMO**
- Primera descomposición en módulos
- Estructura del sistema
- Se consideran aspectos de comportamiento
- El resultado de la fase es un diagrama de relación entre unidades funcionales

ν **DISEÑO LÓGICO**

- Se obtienen y comprueban las expresiones booleanas
- Se pueden representar los módulos mediante ecuaciones booleanas que se pueden optimar

ν **DISEÑO DEL CIRCUITO**

- **QUIÉN**
- Desarrollo de una representación del circuito basada en el diseño lógico
- Las expresiones lógicas se convierten en una representación del circuito teniendo en cuanta las especificaciones de potencia y velocidad
- Comportamiento eléctrico de las partes críticas

ν **DISEÑO FÍSICO**

- La representación de cada componente se convierte en representación física
- Layout
- Los detalles del layout dependen de las reglas de diseño

- Es uno de los pasos más complejos
- El diseño se verifica para comprobar que el layout cumple las especificaciones DRC y extracción eléctrica
- DRC.- verifica que el circuito cumple las reglas de fabricación
- extracción verifica la funcionalidad del circuito, generando la funcionalidad del circuito a partir del layout

ν **FABRICACIÓN**

- Preparación de la oblea
- Deposición y difusión de materiales en la oblea según la descripción del layout
- tamaño típico de la oblea es de 10 cm

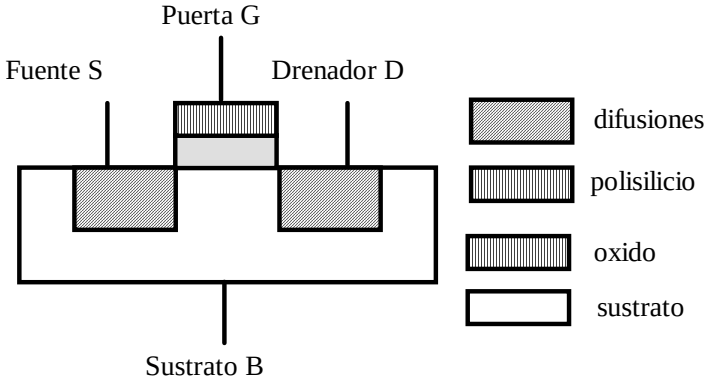
ν **ENCAPSULADO**

- Después de la fabricación la oblea se corta en dados
- cada circuito(dado) se encapsula y se prueba
- El encapsulado es el encargado de eliminar el calor que genera la disipación del circuito.
- El ciclo de diseño conlleva iteraciones tanto dentro de un paso como entre pasos

1.4 BREVE HISTORIA DE LA HERRAMIENTAS CAD

- 1950-1965 diseño manual
- 1965-1975
  - Editores de layout
  - Rutadores automáticos para PCB
  - Algoritmos de partición eficientes
- 1975-1985
  - Herramientas de ubicación automático
  - Fases de diseño bien definidas
  - Importante desarrollo teórico en todas las fases del diseño
- 1985-hoy
  - Ubicación y rutado
  - Algoritmos paralelos de diseño físico
  - Desarrollo de la teoría de grafos
  - Optimación combinatoria de layout.

1.5 ESTUDIO LÓGICO DE LOS TRANSISTORES MOS



- Se realiza una aproximación cualitativa al comportamiento lógico de los MOS
- Un dispositivo MOS esta fabricado de los siguientes layers
  - ∗ Difusión cuya misión es transmitir la señal
  - ∗ Polisilicio. Señal de control
  - ∗ Oxido aislante
  - ∗ Metal hilado
- Los recursos pueden ser
  - ∗ NMOS portador mayoritario son negativos e^-
  - ∗ PMOS portadores mayoritarios positivos h^+

ν CONTACTOS

- Gate .Puerta. es el contacto que controla la corriente que fluye por el transistor. Se implementa en Polisilicio
- Fuente (S) y Drenador (D), son los terminales del interruptor. Físicamente equivalentes. El nombre depende de la dirección del flujo
- Sustrato o Body (B) .No se tiene en cuenta en esta aproximación

ν Modo de funcionamiento. Al aplicar un potencial en la puerta se crea un canal entre las difusiones de la misma polaridad que estas a través de la cual fluyen los portadores mayoritarios.

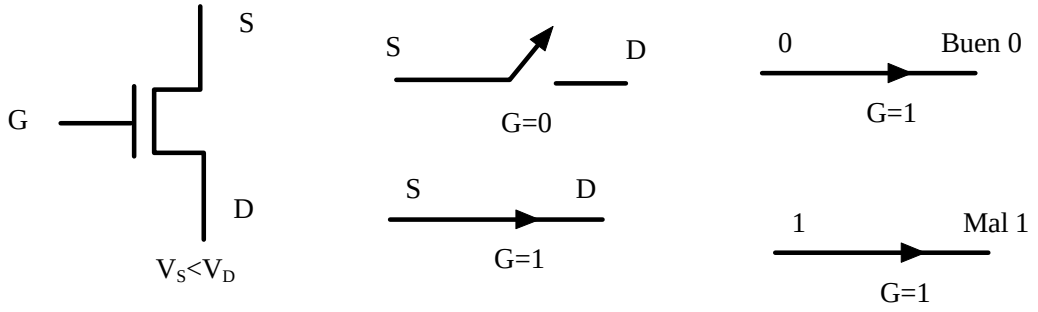
1.5.1 SUPOSICIONES PARA EL ESTUDIO DEL TRANSITOR MOS

ν Simplificando un transistor se puede ver como un interruptor sencillo

- Uno lógico , suele ser un valor comprendido entre 1.5 y 15 voltios. Tambien se le llama alimentación y se le suele representar por Vdd
- Cero lógico. Suele tener el valor 0 voltios. También se le llama tierra y se le representa por Gnd
- Por convenio la corriente la proporciona el Vdd y la elimina el Gnd es decir el sentido de los portadores positivos (de más a menos).
- La dureza mide la capacidad de suministrar o eliminar corriente . Esta dureza puede variar. Las salidas siempre tendrán mayor dureza que las entradas
    - Vdd y Gnd suministran la mayor dureza

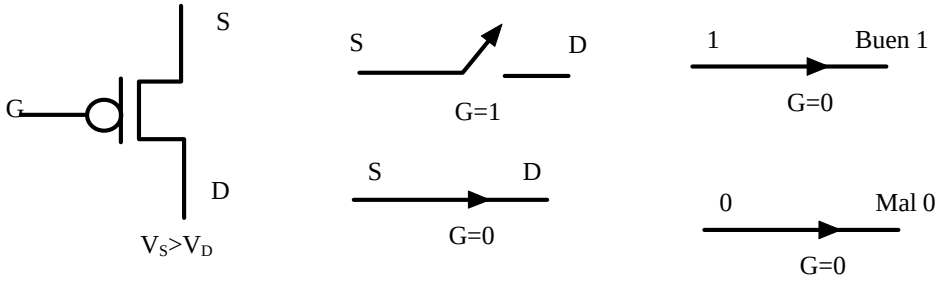
1.5.2 INTERRUPTOR NMOS

- Los Portadores mayoritarios son las cargas negativas
- Cuando Gate =1 la fuente y el drenador se unen mediante un cana de tipo N, luego el transistor conduce
- Cuando traslada un cero de la fuente al drenador es un interruptor perfecto
- Cuando traslada un uno se degenera la señal



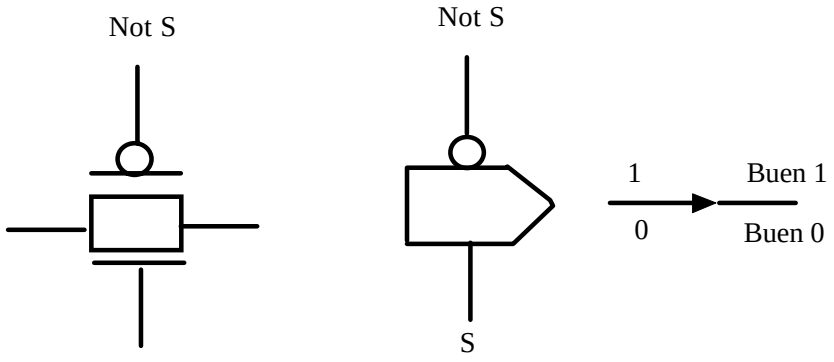
1.5.3 INTERRUPTOR PMOS

- Los portadores mayoritarios son las cargas positivas
- Cuando Gate =0 la fuente y el drenador se unen mediante un canal P luego el transistor conduce
- Cuando traslada un uno de la fuente a un drenador es un interruptor perfecto
- Cuando traslada un cero se degenera la señal



1.5.4 INTERRUPTOR CMOS

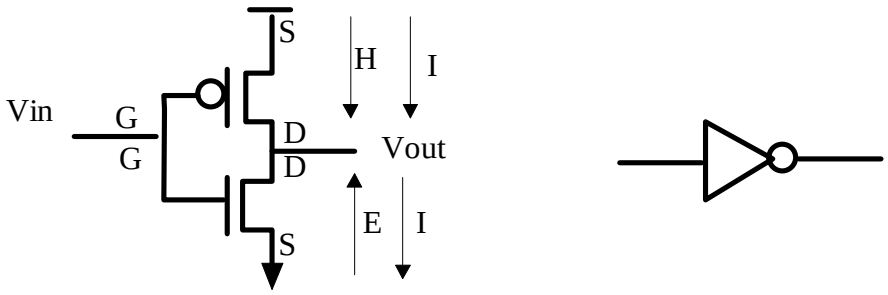
- Combinación en paralelo de un interruptor NMOS y un interruptor PMOS
- Transmite bien tanto el cero como el uno
- La señal de control del NMOS es la señal de control del PMOS complementada
- Otros nombres son:
    - puerta de transmisión
    - puerta de paso



1.6 LOGICA CMOS

1.6.1 INVERSOR CMOS

- Interruptor PMOS en serie con un interruptor NMOS
- PMOS se le llama de pull up genera el uno
- al NMOS de pull down genera el cero
- la salida se toma de la unión de los interruptores



Siendo

H la corriente de huecos

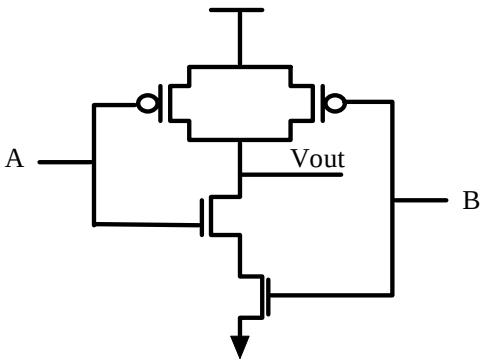
E la corriente de electrones

I la intensidad convencional

1.6.2 PUERTAS NAND CMOS DE M ENTRADAS

- Estructura muy similar a la del inversor CMOS
- **Árbol de pull up**
  - Que proporciona el uno
  - M transistores en paralelo
  - Conectados a Vdd y a la salida
- **Árbol de pull down que**
  - Proporciona el cero
  - M transistores en serie conectado a Gnd y a la salida

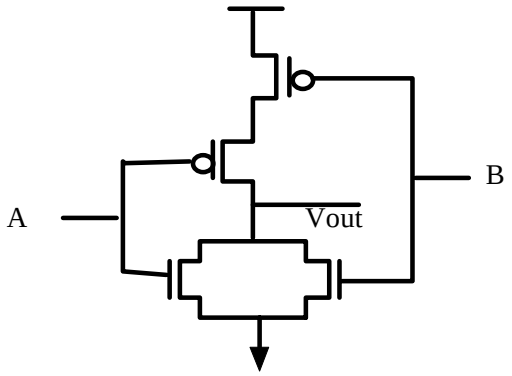
a b	**Vout**
0 0	1
0 1	1
1 0	1
1 1	0



1.6.3 PUERTA NOR CMOS DE M ENTRADAS

- Arbol de pull down: transistores N en paralelo conectadas a la tierra y al Vout
- Arbol de pull up: transistores P en serie conectados a Vdd y Vout

a b	Vout
0 0	1
0 1	0
1 0	0
1 1	0



1.6.4 PRINCIPALES CARACTERISTICAS DE LA LÓGICA CMOS

Para cualquier entrada siempre existe un camino desde la tierra o la alimentación a Vout, esto indica que los Potenciales de salida son totales, es decir máxima dureza. A este tipo de lógicas se les llama totalmente restaurada

Esta característica también tiene como efecto que no sea una lógica proporcional. Es decir no hace falta que los transistores guarden relaciones de tamaño entre ello para que alcance la salida los valores lógicos correctos. Esto simplifica el diseño, y también tendrá una importante consecuencia y es que los márgenes de ruido serán grandes.

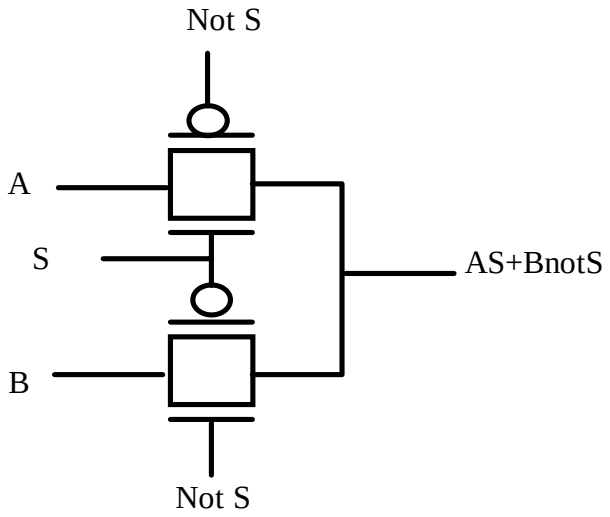
Con la señal estable no existe nunca un camino directo entre Vdd y Gnd (lo contrario de lo que ocurre en estilos de diseño con un solo tipo de canal NMOS o PMOS o con los bipolares o arseniuro de galio. Esto da lugar a baja disipación de potencia estática que es la principal característica de los circuitos CMOS.

- Memorias densas y de baja disipación.
- El potencial necesario para conectar una puerta es un porcentaje fijo de Vdd
    - A este potencial se le llama $V_T = 0{,}2$ Vdd
- Densidad del circuito inferior al de otras lógicas, 2n transistores por cada n entradas
- Layout: da lugar a estilos de layout regulares y automatizables
- Su robustez asegura el buen funcionamiento final del sistema.
- Si las ligadura de sincronización son flojas se puede relajar mucho su estudio.
- Existe una gran cantidad de herramientas CAD disponibles para su diseño.
- Ejemplo de diseño de una puerta cmos: ab+cd

1.6.5 LÓGICA DE MULTIPLEXORES

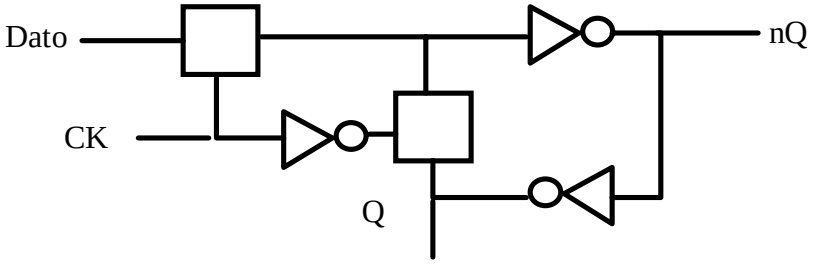
- Se lleva a cabo con puertas de paso CMOS
- Los multiplexores son claves en los elementos de memoria CMOS y en las estructuras de manipulación de datos
- La función lógica que implementa un multiplexor es:

$$y=\Sigma_{i=0\rightarrow 2^{N}-1} Xi.Mi$$

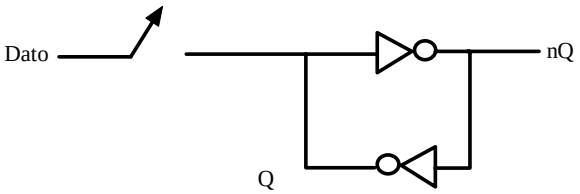


1.6.6 ELEMENTOS DE MEMORIA

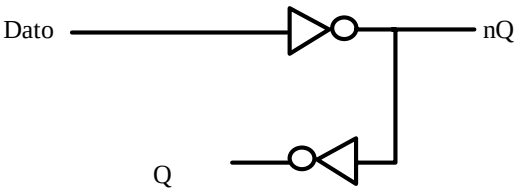
1.6.6.1 Biestable D



- Carga por nivel
- Se compone de un multiplexor de dos entradas y de dos inversores
  - Siendo D la entrada de datos
  - CK la señal de reloj que carga el dato
  - Q la salida de datos
- Cuando CK=0 se establece una realimentación entre los inversores lo que produce que el estado actual de Q quede almacenado.

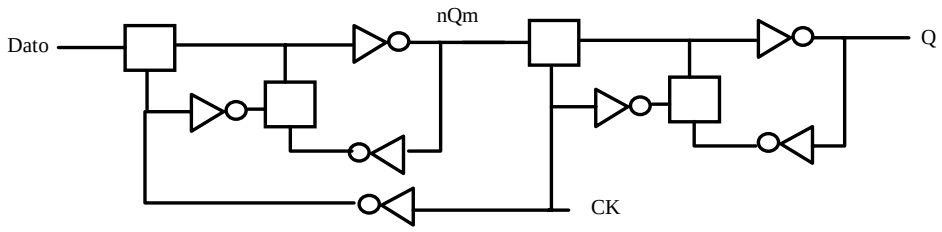


- la entrada D se ignora
- Cuando CK=1 cualquier cambio en D se trasmite a Q

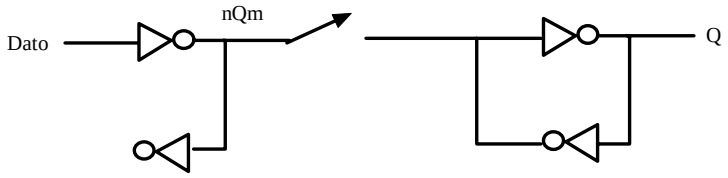


- Se interrumpe el camino de realimentación
- Principal inconveniente es que el dato que se quiere guardar debe estar estable durante todo el ciclo de reloj
- Si se utiliza en circuitos realimentados genera realimentaciones indeseadas.
- la solución es la carga por flanco

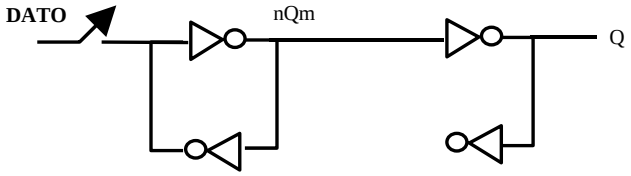
1.6.6.2 Carga por flanco



- Combinando dos biestables de carga por nivel se consigue uno de carga por flanco. Por convenio al primer biestable se le llama maestro y al segundo esclavo
- Su principal ventaja es que la entrada y la salida están siempre aisladas. Evita realimentaciones indeseadas.
- Mientras CK =0→ nqm=f(d) pero la salida q está aislada de la entrada y por lo tanto el dato está almacenado.



- Justo en el instante de transición de 0 a 1 el maestro deja de hacer un muestreo y carga el dato que aparece a la puerta del esclavo como nqm



1.7 PRINCIPALES CARACTERÍSTICAS DE LA LÓGICA CMOS

- Lógica totalmente restaurada(salidas a Vdd ó a Gnd).
- Los tiempos de transición son del mismo orden.
- Memorias densas y de baja disipación.
- Las puertas de transmisión manejan los dos valores lógicos correctamente.
- La disipación de potencia estática es prácticamente nula.
- El potencial necesario para conectar una puerta es un porcentaje fijo de Vdd
    - A este potencial se le llama V_T = 0,2 Vdd
- Densidad del circuito inferior al de otras lógicas, 2n transistores por cada n entradas
- Layout: da lugar a estilos de layout regulares y automatizables
- Su robustez asegura el buen funcionamiento final del sistema.
- Si las ligadura de sincronización son flojas se puede relajar mucho su estudio.
- Existe una gran cantidad de herramientas CAD disponibles para su diseño.

1.8 TECNOLOGÍA CMOS VS OTRAS TECNOLOGÍAS

- Las tecnologías utilizadas habitualmente son:
    - CMOS
    - BIPOLAR
    - ARSENIURO DE GALIO
    - BICMOS
- El arseniuro de galio es la tecnología más rápida, siguiéndole la bipolar y la CMOS.
- La tecnología CMOS es la de más alta densidad y menor consumo por puerta.
- La CMOS se puede usar para diseños analógicos, pero da mejor rendimiento la bipolar, se suele usar por ser la mas barata, cuando las necesidades analógicas son poco elevadas
- Sus costes de diseño son los más bajos debido a la gran cantidad de herramientas automáticas y al desarrollo de las celdas estándar.
- La tecnología BiCMOS es una combinación de tecnología bipolar y tecnología CMOS que se usa para circuitos con señales DC y AC.
- La tecnología CMOS es la mas utilizada.

2. METODOLOGÍAS DE DISEÑO

La enorme complejidad que han alcanzado los circuitos integrados en la actualidad a forzado la aparición de un sin fin de herramientas automáticas cuyo principal objetivo es hacer la tarea de diseñador más fácil y conseguir circuitos más fiables. Estas herramientas se pueden clasificar en tres grandes grupos:

- Análisis y verificación que examinan el comportamiento de un circuitos y ayudan a determinar si se ciñe a las especificaciones.
- Síntesis e implementación.- ayudan al diseñador a generar el esquemático o el layout
- testabilidad sirven para validar el funcionamiento del circuitos una vez diseñado.

2.1 SIMULACIÓN

Si duda las herramientas que inicialmente mas ayudan al diseñador de circuitos son las herramientas de simulación.

La simulación consiste en introducir en las entradas de circuito unas señales determinadas y para comprobar la forma de las señales de salida, de manera que si el circuito fuera correcto deberían coincidir con la funcionalidad implementada.

Una de las principales ventajas de la simulación es su controlabilidad, observabilidad y facilidad de uso

2.1.1 SIMULACIÓN A NIVEL DE CIRCUITO

Una de las primeras herramientas de simulación que existieron y una de las más utilizadas en la actualidad es el SPICE, desarrollado en la Universidad de Berkeley. Es un simulador del comportamiento electrónico de los circuitos cuya principal ventaja es la gran exactitud de sus análisis.

En general los simuladores eléctricos se basan en la resolución de matrices de ecuaciones que relacionan voltajes, corrientes y resistencias. Como ya se ha indicado su principal ventaja es su gran exactitud. Como contrapartida tiene el inconveniente de que debe tener en cuenta gran cantidad de pequeños de talles de los transistores, así como los efectos secundarios de estos, lo que la hacen inservibles para los circuitos de cierto grado de complejidad.

El tiempo de simulación es proporcional a N^M donde N es el número de dispositivos no lineales del circuito y M puede se encuentra entre 1 y 2

Se usa para verificar en detalles pequeños trozos de circuito

Hacen un tratamiento continuo de comportamiento del diseño, es decir suponiendo que el circuito se encuentra alimentado entre 0v y 5v estudia las transiciones en el tiempo que toma la salida del circuito al pasar de 0 a 5 o de 5 a 0 pasando por todos y cada uno de los posibles potenciales.

De todos modos no se debe asumir implícitamente la exactitud total del rendimiento que predicen estas herramientas debido a diversos factores como son:

Inexactitud de los modelos de los parámetros de los modelos MOS

Uso de modelos MOS inapropiados

Inexactitud de las resistencias y de las capacidades parásitas

En la actualidad todas las herramientas relacionadas con SPICE proporcionan diferentes niveles de modelo según el grado de exactitud que se desee obtener en la simulación. Los modelos mas simples están pensados para acelerar los procesos de simulación, mientras que los modelos mas complejos se usan para simulaciones mas exactas.

El problema de la inexactitud de las capacidades y resistencias parásitas solo se puede solucionar una vez realizado el layout del circuito.

La herramienta de CADENCE proporciona una herramienta de simulación llamada SPECTRE.

2.1.2 SIMULACIÓN DE TIMMING

Realiza una simulación menos completa y exacta que realizada por los simuladores eléctricos, pero tiene la ventaja de tener tiempos de ejecución dos ordenes de magnitud menores.

En lugar de resolver matrices de ecuaciones resuelve ecuaciones más sencillas o aplica tablas de look-up

Al ser bastante menos complejos que las simulaciones eléctricas se pueden usar para simular circuitos de mayor complejidad complejos.

Margen de error del 10-20%

2.1.3 SIMULACIÓN LÓGICA

No trabajan con transistores sino con puertas lógicas

Como hemos visto debido a la gran cantidad de información que manejan la simulación eléctrica es ineficaz para circuitos de cierta complejidad.

Esta es la razón de que haya aparecido simuladores lógicos. La diferencia entre ambos es que en el simulador lógico se trabajan exclusivamente con los valores Booleanos 0 y 1 para determinar el comportamiento de los circuitos.

Debido a que solo nos interesan los valores 0 o 1 finales que toman los circuitos y a la elevada abstracción de información que realiza la herramienta, puesto que la descripción booleana de cada puerta es perfectamente conocida, las simulaciones lógicas son mucho menos complejas y por lo tanto mucho más rápidas que las simulaciones eléctricas, pero también menos exactas, en cuanto a comportamiento eléctrico. La simulación lógica se utiliza para simular circuitos de cierta complejidad, y comprobar su funcionalidad.

A estos simuladores se le pueden añadir información de timming y retardo de las puertas que utilizan para la simulación, además de cálculos estadísticos calculados estadísticamente en función del número de puertas lógicas que puede formar el circuito.

El retardo de estas puertas se suele calcular según la siguiente expresión:

Tpuerta=Tintrinseco+Cload·Tload:

siendo

Tpuerta el retardo por puerta

Tintrinseco el retardo intrínseco de la puerta (no debido a la carga)

Cload- la capacidad de carga

Tload el retardo por unidad de carga

Los simuladores lógicos que incluyen estas informaciones son muy exactos para lógicas bien caracterizadas como la CMOS.

En la simulación los valores de los parámetros del diseño , tales como márgenes de ruido, propagación del retardo o energía disipada se determina aplicando un conjunto de vectores de excitación en la entrada y sacando conclusiones de los vectores de salida, esta opción es muy flexible pero tiene la gran desventaja que depende enormemente del los vectores que se elijan, de tal manera que las conclusiones extraídas pueden ser falsa. En definitiva no se tiene en cuenta las estructura del circuito.

En cualquier caso los resultados de la simulación no garantizan que la funcionalidad implementada sea la correcta, solo garantiza que se cumple para los vectores de excitación probados. Ejemplo una simulación eléctrica no detecta una distribución de carga si no se mete la secuencia correcta .

La solución podría ser simular todos los vectores posibles pero esto es imposible para sistemas complejos.

2.2 VERIFICACIÓN

La verificación intenta extraer los parámetros del sistema directamente de la descripción del circuito. Por ejemplo el camino critico de un circuito puede extraerse al observar un esquemático o una descripción del mismo. Tienen la ventaja de que no dependen de la elección de un determinado vector de excitación. El problema es que necesitan una comprensión del estilo de diseño elegido.

Las herramientas de verificación deben analizar los circuitos, luego deben incluir gran cantidad de información sobre los estilos y tecnologías de diseño y modos de sincronización. Esto hace que una herramienta de verificación diseñada para circuitos con el estilo de diseño CMOS, no sea valida para circuitos con el estilo de diseño NP ZIPPER.

- **Verificación eléctrica:**

Dado el esquemático de transistores de un circuito es sencillo comprueba que se cumplen un con junto de reglas. Por ejemplo para un estilo de diseño C^2MOS comprueba que el número de inversiones entre dos inversores C2MOS es impar. En cambio para un estilo pseudoNmos, comprueba que existe una relación correcta entre el tamaño del canal del transistor P y los canales de los transistores n, y de esta manera unos márgenes de ruido correctos.

El sentido común y el conocimiento de los estilos de diseño ayudan a diseñar grandes conjuntos de reglas que ayudan a verificar los problemas.

Verificadores de timming.

Según se va haciendo el circuito más complejo es más difícil determinar cual es el camino critico del diseño, información que se venía obteniendo con los simuladores de timming.

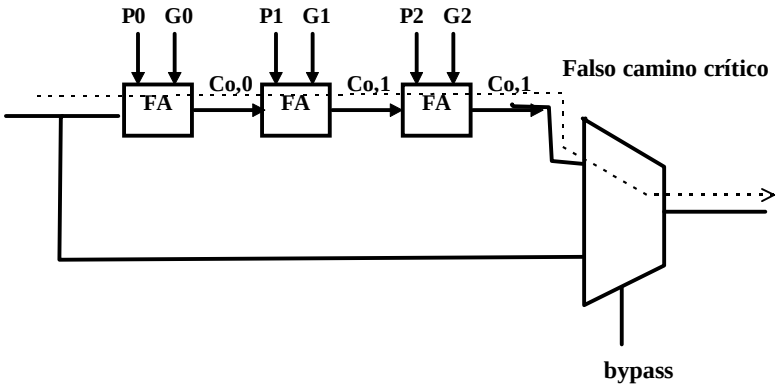
Una posible solución podría ser ejecutar una simulación extensivacon la herramienta ,que llevaría mucho tiempo. Pero ni siquiera en este caso podríamos asegurar que los vectores de excitación recorren el camino crítico.

Un verificador de timing recorre la red eléctrica y ordena todos los caminos en función de su retardo. Este retardo se puede calcular de múltiples maneras. Por ejemplo sustituyendo los transistores por sus retardo RC. Muchos simuladores obtienen mediante este método el camino crítico y posteriormente realizan una simulación pata determinar con mayor exactitud su retardo.

Un problema habitual en estos verificadores es determinar falsos caminos críticos que no se van a recorrer jamas.

Ejemplo:

sumador con carry bypass. Este sumador se supone implementado mediante un generador de carry, La señal de entrada al primer generador de carry se propaga a la salida cuando p0=p1=p2=1. Haciendo la señal de bypass=p0·p1·p2 acelero el proceso luego el camino critico que atraviesa todos los generadores de acarreo es falso porque nunca se recorre.



VERIFICACIÓN FUNCIONAL (O FORMAL)

Cualquier componente de un sistema independientemente de su nivel de jerarquía se puede describir en función de sus entradas y de un estado interno.

Si se unen todos estos componentes se conseguiría describir el sistema. La verificación funcional consiste en comparar la especificación resultante con la especificación inicial del sistema. Aun que no idénticas estas dos descripciones necesitan ser equivalentes para que el circuito sea equivalente.

Esta herramienta es el sueño de todo diseñador de circuitos. Comprobar que el el diseño cumple las especificaciones propuestas.

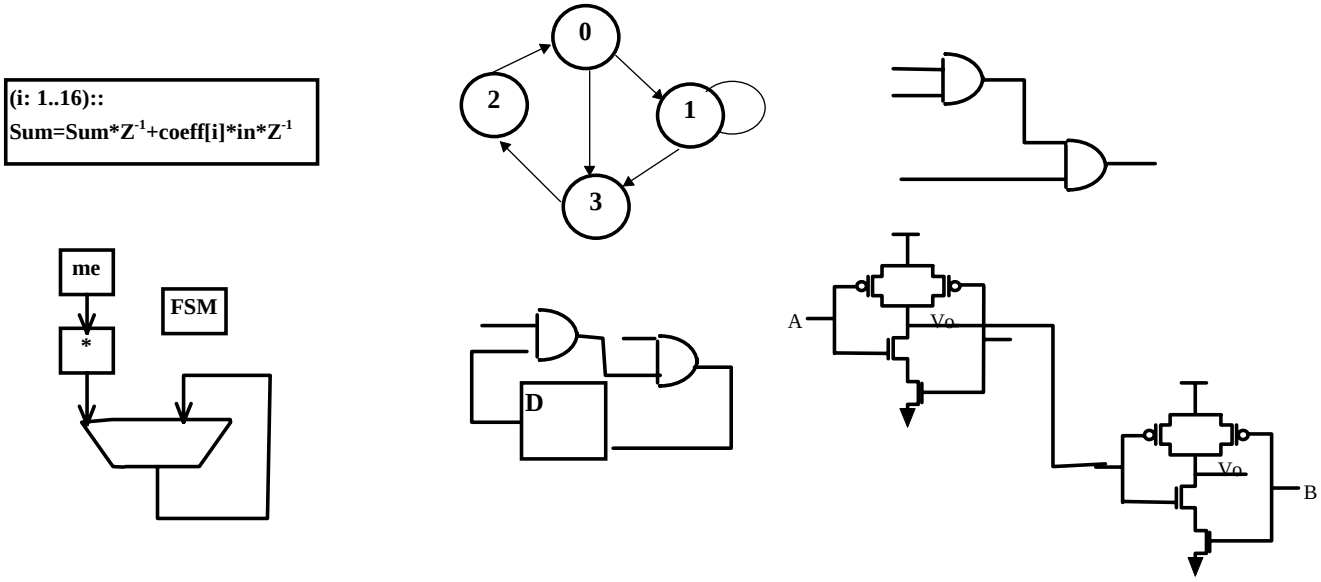
Desgraciadamente este es un problema muy complejo que aún está por resolver.

2.3 SÍNTESIS DE DISEÑOS

En diseño se define como la transformación de una entidad de diseño descrita desde un punto de vista del comportamiento a una descripción estructural.

Según los diferentes niveles de abstracción a los que trabajemos se pueden definir diferentes niveles de síntesis :

- Síntesis de arquitectura
- Síntesis RTL
- Síntesis lógica
- Síntesis a nivel de circuito



2.3.1 SÍNTESIS DE ARQUITECTURA

También llamada de comportamiento o de alto nivel.

Obtiene una descripción estructural de un sistema a partir de una descripción de comportamiento. En definitiva consiste en determinar que recursos se necesitaran para ejecutar dicha tarea :

- Unidades funcionales
- Memorias
- Buses
- Controladores

Emparejando las operaciones a recursos hardware y determinando el orden de ejecución de las operaciones, no perdiendo nunca de vista las ligaduras de área, tiempo o potencia.

Estas herramientas han tenido un amplio desarrollo académico, pero inicialmente su penetración en el mercado no fue todo lo profunda que se podía desear. Las razones fueron:

La falta de una definición concreta de como se definía una arquitectura sobre todo a nivel de microprocesadores

La síntesis de alto nivel presuponía herramientas de síntesis a nivel RTL que se han obtenido solo hace poco tiempo

Durante mucho tiempo la síntesis de alto nivel se concentró solo en unos aspectos dejando otros totalmente olvidados como es el caso de los del impacto de las interconexiones en el diseño total.

2.3.2 SÍNTESIS RTL

Toma una descripción RTL y la convierte a una serie de registros y lógica combinacional.

En esta etapa la arquitectura del diseño ha sido ya capturada.

Habitualmente las descripciones RTL se pueden capturar mediante lenguajes de descripción de HW,

flujo de control mediante sentencias if then else y case

iteraciones

jerarquía

anchos de palabra vectores de bits y campos

operaciones secuenciales y paralelas

especificaciones de registros

2.3.3 SÍNTESIS LÓGICA

Partiendo de una descripciones a nivel lógico se obtiene una netlist de puertas lógicas que suele estar optimada en área , tiempo o consumo de potencia.

La entrada de comportamiento se puede especificar de diferentes formas como puede ser:

- Máquinas de estados finitos
- Diagramas de estados
- Esquemáticos
- Ecuaciones booleanas
- Tablas de verdad
- Descripciones de lenguaje de alto nivel

Las técnicas difieren según el circuito sea combinacional o secuencial o la tecnología final de implementación, como PLAS, celdas estándar, o FPGAS.

La síntesis consiste en una secuencia de pasos de optimización que dependen de los parámetro a optimizar

Generalmente se puede dividir en dos fases

Independiente de la tecnología, donde la lógica se optimiza mediante operaciones algebraicas o booleanas

Correspondencia con la tecnología, en la que se vuelca sobre una tecnología determinada la descripción obtenida en el paso anterior.

Esta síntesis lógica se suele clasificar en síntesis combinacional y síntesis secuencial. La síntesis combinacional a su vez se divide en síntesis de dos niveles y síntesis multinivel.

Las herramientas de síntesis de dos niveles fueron las primeras que estuvieron disponibles

El programa ESPRESSO [Brayton] desarrollado en la universidad de Berkeley es el programa de síntesis de dos niveles más conocido. Estaba pensado para utilizarlo sobre tecnologías muy regulares como por ejemplo las PLAs que se adaptaban perfectamente a la lógica de dos niveles. Al ser de las primeras herramientas automáticas de síntesis que existían forzó a la implementación de lógica aleatoria (estructura multinivel) mediante PLAS (estructura de dos niveles).

Poco después apareció la primera herramienta de síntesis multinivel, MIS (Multilevel, Logic Synthesis) desarrollada también por Brayton en Berkeley. Esto forzó una evolución de la tecnología hacia el uso de celdas estándar, de estructura claramente multinivel.

En cuanto a la síntesis de sistemas secuenciales se orientó principalmente a la reducción del número de estados.

2.3.4 SÍNTESIS DE CIRCUITOS

La tarea de la síntesis de circuitos es trasladar la descripción lógica de un circuito a una red de transistores, generalmente intentando cumplir unas ligaduras de tiempo. Este nivel de síntesis y la correspondencia con la tecnología están muy relacionados. En realidad la correspondencia con la tecnología es el puente entre la síntesis lógica y la síntesis física.

Este proceso se puede dividir en dos etapas.:

La primera consiste en la obtención de un esquemático de transistores a partir de las ecuaciones lógicas. Esta tarea se divide a su vez la definición del estilo del circuito por el diseñador (estática complementaria, de transistores de paso, dinámica, DCVSL…) y en la construcción de una red lógica.

La segunda es la de búsqueda de los tamaños de los transistores que ayudan a alcanzar las ligaduras de rendimiento. La elección de un determinado tamaño, repercute sobre el rendimiento , el área y la disipación de potencia.

Esta herramienta es muy potente pero no se ha desarrollado todo lo que cabría de esperar. La razón es que los resultados de una celda estándar influyen demasiado sobre el resultado final del circuito como para dejarlos a manos de una herramienta .

Dentro de estas fase se incluyen las herramientas de ubicación y rutado automático que han tenido un desarrollo muy importante en los últimos años.

2.4 VALIDACIÓN Y TEST

Una vez que el dispositivo esta fabricado ¿como se sabe si ha sido fabricado correctamente. Una vez que el circuito se integra en el sistema es muy costosa la corrección del fallo.

Por otro lado conviene saber que un diseño correcto no garantiza que el componente funcionará. Existen errores de fabricación debidos a las impurezas del material empleado o a errores de proceso. Además también pueden introducir mal funcionamientos las pruebas posteriores a la fabricación a los que se someten los componentes(resistencia a la temperatura)

Los fallos más típicos son cortocircuitos en los hilos .

Comprobar el comportamiento del circuito bajo todas las posibles condiciones de entrada no es tan sencillo como puede parecer. En la fase de diseño se tiene acceso ilimitado a todos los nodos de la red esto no se puede hacer cuando el circuito está fabricado. En este caso la única información comprobable se reciben través de los pines de entrada salida.

Lo costoso y difícil de la operación de detección de fallos de fabricación hace que se deba tener en cuenta el test del sistema desde las más tempranas fases de diseño, por ejemplo sabiendo que algunas modificaciones en un circuito pueden ayudar a validar la ausencia de fallos. A este estilo de diseño se le denomina diseño para testabilidad (DFT) Esta estrategia contiene dos componentes:

- proporcionar la necesaria circuitería para conseguir un test efectivo y comprensible
- proporcionar los necesarios patrones de test que se deben emplear durante la fase de test.. Por razones de costo es conveniente que la secuencia de test sea lo más corta posible mientras cubre la mayoría de fallos.

2.4.1 TEST DE VALIDACIÓN

Según el objetivo que busquen los test se clasifican en:

Test de diagnostico: se usan durante el depurado del chip. Detectado un fallo del circuito detectar la causa.

Test funcional. Comprobar si cumple las funcionalidad especificada. Este problema es el más simple, puesto que la respuesta esperadas es si o no . Dado que se debe ejecutar sobre cada circuito debe ser un test lo más sencillo posible para que su impacto sobre el costo sea lo menor.

Test paramétrico.- Comprueba parámetros no discretos como los márgenes de ruido, los retardos de propagación, la máxima frecuencia de reloj.

La forma de trabajar es la siguiente se somete el circuito a unos vectores de test. Estos vectores incluyen información sobre las ondas aplicadas, los niveles de voltaje, la frecuencia de reloj y la respuesta esperada. Y se comparan la respuesta esperada con la respuesta obtenida. Si se detectan diferencias la parte correspondiente se marca como fallo

2.4.2 DISEÑO PARA TESTABILIDAD

Vamos a suponer un circuito combinacional, que se quiere validar exhaustivamente mediante todos los vectores de test posibles. Suponiéndole N entradas quiere decir que tenemos que generar 2^N patrones. Suponiendo N=20 entradas esto quiere decir que necesitamos más de un millón de patrones. Suponiendo que la velocidad de test de un patrón es de 1micro segundo, la validación total del módulo es de un segundo. Esto en principio no parece mucho tiempo. Vamos a ver que es lo que ocurre con un sistema secuencial. Estos sistemas no dependen solo de la entrada, también dependen del estado, suponiendo que el sistema tiene N estados el número total de vectores a validar es 2^{N+M} donde M es el número de registros de estado. Para una maquina de estados de tamaño moderado, por ejemplo M=10 esto quiere decir que el número de es próximo al billon lo que tomara 16 minutos, que en principio pare que es una cantidad de tiempo sin importancia.

El problema es que estos dos ejemplos son excesivamente sencillos. Si queremos hacer la prueba para un microprocesador cuya máquina secuencial necesita 50 registros de estado necesitaríamos más de un billon de años para validarlo exhaustivamente. La principal consecuencia es que tenemos que utilizar otro tipo de aproximaciones para tratar el problema. Cualquier solución al problema debe tener en cuenta las siguientes premisas:

una enumeración exhaustiva de todos los patrones contiene gran cantidad de redundancias. Es decir un fallo sencillo en computador puede ser detectado por un conjunto de patrones de entrada.. Es decir la detección del fallo solo necesita de uno de los patrones.

Se puede conseguir una reducción substancial del número de patrones, relajando la condición de que se deben encontrar todos los fallos. Detectar un pequeño porcentaje de fallos puede producir un aumento exagerado del número de patrones necesarias para detectarlas. El porcentaje con el que suelen trabajar es del 95-99% de fallos descubiertos.

Estas premisas solucionan la validación de los módulos combinacionales pero no el de los secuenciales. Para detectar un fallo en un circuito secuencial, no es suficiente el vector de entrada correcto, también el sistema se debe encontrar en el estado correcto. Es decir el sistema secuencial no responde a una única entrada sino a un conjunto de ellas. Esto puede hacer la validación excesivamente costosa.

Una posible solución puede ser reconvertir el circuito secuencial en combinacional. Esto se consigue rompiendo la realimentación. Este es uno de los conceptos clave de la metodología Scan Test. Otra aproximación posible es dejar que el circuito realice un autotest.

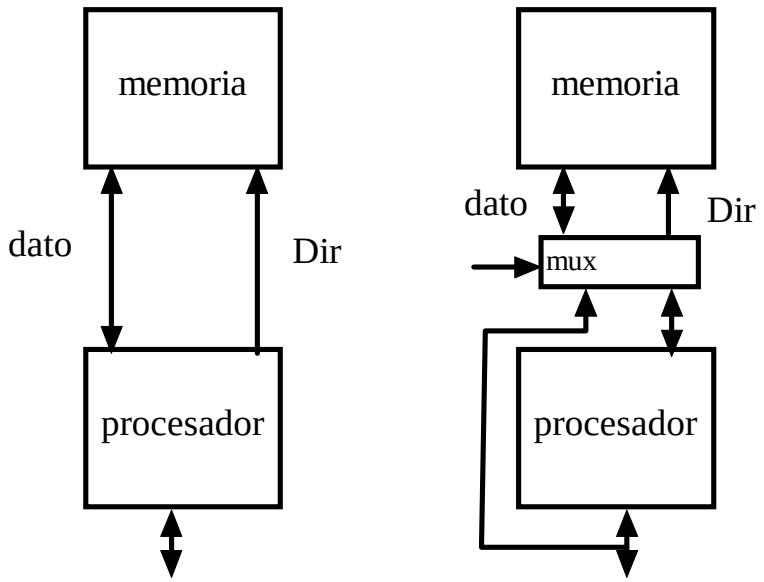
En cualquier caso es importante tener claros los conceptos de controlabilidad y observabilidad.

> **Controlabilidad** mide la facilidad de modificar el valor de un nodo con solo modificar valores de la entrada. Nodos con baja controlabilidad necesitan gran cantidad de vectores para ser testados.
>
> **Observabilidad**.- que mide la facilidad de observar el valor de un nodo a través de la salida.. Un nodo con baja observabilidad necesita de de un gran número de vectores para ser testada.

A continuación se estudian tres soluciones para el problema de los circuitos secuenciales:

- test Ad hoc
- scan based
- selftest

AD HOC TEST



Consiste en combinar una serie de trucos en el diseño que sirven para aumentar la observabilidad y controlabilidad del sistema

Un ejemplo típico consiste en añadir multiplexores que permiten seleccionar caminos alternativos en la fase de test. Por ejemplo ,en un sistema la lectura /escritura de la memoria pasa necesariamente a través del procesador, de manera que se consumen muchos ciclos de reloj. Si se añade un multiplexor entre la memoria y el procesador se mejora notablemente la observabilidad y controlabilidad de la memoria.

Su principal problema es que añade hardware al sistema sin funcionalidad concreta salvo la de test.

Necesidad de mayor número de pines de entrada salida

TEST BASADO EN SCAN

Consiste en convertir todos los registros del sistema en elementos leíbles y escribibles directamente desde el exterior. En definitiva lo que se busca es convertir el circuito secuencial en un circuito combinacional.

Como utilizar un bus de test resulta demasiado caro para el sistema lo que se suele hacer es comunicar los registros de manera secuencial los unos con los otros, de manera que la información de test entra por un extremo y se propaga hasta la salida.

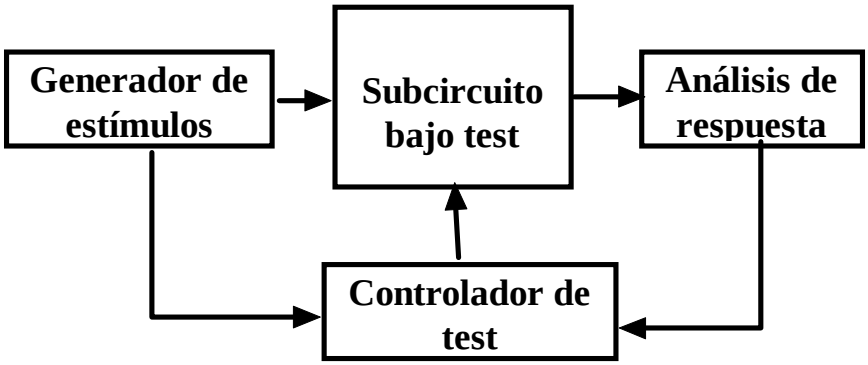
Esta solución tiene el inconveniente de que se deben modificar los registros para que funcionen en dos modos de operación.

BUILT IN SELF TEST (BIST)

El circuito genera sus propios patrones de test en lugar de necesitar una aplicación externa de patrones.

Es una técnica en la que el propio circuito decide si los resultados obtenidos son correctos. Generalmente esta técnica necesita añadir circuitería al diseño para la generación y análisis del diseño

La estructura general es la siguiente:



Existen dos aproximaciones para la generación de patrones. La exhaustiva y la aleatoria. En la exhaustiva se generan todos los patrones de entrada posibles, y la aleatoria que selecciona un subconjunto aleatorio de patrones

En la actualidad los patrones de test se suelen generar mediante herramientas automáticas denominadas ATPG(Automatic Test Patron Generation)

Los fallos de fabricación pueden ser de una gran variedad y se suelen manifestar como cortocircuitos en la líneas de señal o en las de alimentación y nodos flotantes.

3. ESTILOS DE DISEÑO

El diseño físico es un proceso muy complejo, y aún descomponiendo su realización en diferentes pasos, el desarrollo de cada uno de estas subtareas es muy dura, computacionalmente hablando. Sin embargo las necesidades del mercado demandan diseños realizados cada vez en menos tiempo y con un rendimiento de la oblea mayor, (entendiendo por rendimiento de la oblea el % de circuitos de funcionamiento correcto que se pueden obtener de una oblea de silicio).

Para intentar cumplir estas necesidades del diseño, han ido apareciendo a lo largo de los años diferentes estilos de diseño, con ligaduras de mayor o menor importancia que facilitan el diseño físico, generalmente a costa del ahorro de área o de los rendimientos o consumos de potencia.

De manera general los estilos de diseño se pueden clasificar en estilos full-custom o estilos semicustom. La principal característica del diseño full-custom es que los diferentes bloques que forman el diseño se pueden colocar en cualquier parte del area del circuito con la única condición que no se solapen entre sí.

En cuanto a los estilos del semicustom, algunas partes del diseño están prediseñadas y localizadas en partes determinadas del área del silicio.

3.1 EL DISEÑO FULL-CUSTOM

El diseño full custom es aquel en el que no existe ninguna restricción a la hora de realizar la ubicación ni el rutado de los diferentes módulos que lo componen. En este sentido se puede decir que no se utilizan elementos prefabricados y ni prediseñados por el fabricante. Esto significa que el diseñador debe poseer grandes conocimientos de microelectrónica.

En este estilo se debe diseñar todo el circuito de principio a fin, ayudado , como no podía ser de otra forma, por herramientas automáticas que facilitan la tarea. Pero esta ayuda no quita que se requiera un esfuerzo importante para llevar el proyecto a buen término.

El diseñador debe indicar exactamente donde quiere

Se divide el circuito en subcircuitos siguiendo algún tipo de criterio como el funcional. A estos subcircuitos se les llama bloques funcionales. Estos bloques funcionales pueden tener cualquier tamaño

La principal característica de este tipo de diseño es la ausencia total de ligaduras lo que permite diseños muy compactos

como principal inconveniente está la dificultad del proceso de automatización,. Esta es una de las razones de que se utilice cuando el diseño final debe ser de área mínima y además no tiene demasiada importancia el tiempo de diseño.

el espacio no ocupado por bloques se utiliza para el rutado. Inicialmente los bloques se sitúan con el objetivo de minimizar el área pero no se debe olvidar el area necesaria para realizar el rutado. Generalmente se utilizan varias capas de metal,

En un estilo de diseño jerárquico, un bloque a su vez puede estar compuesto de otros bloques, que a su vez pueden usar como estilo de diseño el full-custom u otro cualquiera, como

Como principal ventaja esta la gran flexibilidad a la hora de diseñar lo que permita alcanzar los óptimos de área, rendimiento, o potencia. Como principal desventaja se encuentra el enorme esfuerzo de desarrollo lo que hace que los tiempos de mercado sean elevados. Otra desventaja es que no se puede asegurar que el comportamiento eléctrico, que el diseñador había supuesto a alguno de los módulos sea el correcto.

Solo justificable cuando los costes pueden ser amortizados con un gran volumen de producción los microprocesadores y las memorias semiconductoras.

Cuando los bloques custom pueden ser reutilizados muchas veces por ejemplo mediante librerías de celdas

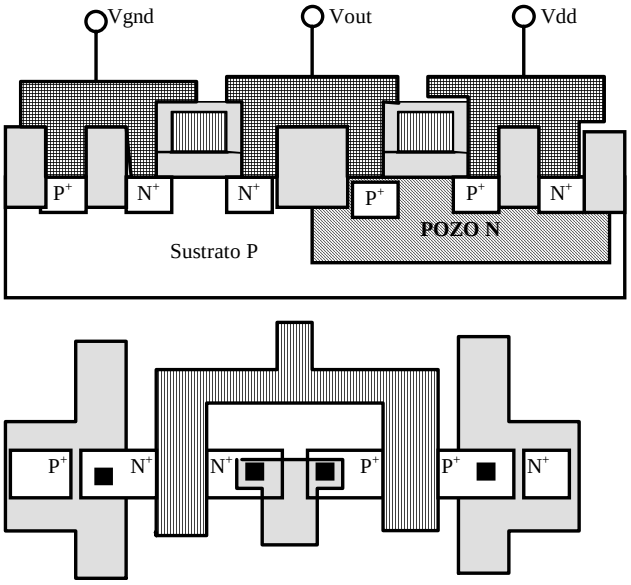
Cuando el coste no es demasiado importante

Debido al crecimiento y desarrollo de las herramientas de diseño automático el rango de diseños custom se reduce de año en año. Incluso algunos procesadores como el Alpha de DEC diseña grandes porciones del mismo mediante estilo semicustom. Solo las unidades criticas como los operadores de coma flotante y de enteros utilizan este estilo.

Aunque las herramientas de diseño para estilo full-custom, no son numerosas si existen algunas de gran utilizadas. como el editor de layout

El editor de layout es la priemra herramienta de trabajo que tuvieron los diseñadores y su objetivo es la generación de la representación física del diseño.

Puesto que el diseño físico ocupa una parte importante del tiempo total de diseño de una celda o nuevo componente este tipo de herramientas está en permanente desarrollo



- Las reglas de diseño son el punto de conexión entre el diseñador de C.I. y el ingeniero de procesos durante la fase de fabricación.
- El principal objetivo de estas reglas de diseño es obtener un circuito con un rendimiento de producción optimo (circuitos validos/circuitos no validos) en un área lo menor posible sin comprometer la fiabilidad del circuito.
- Representan el mejor compromiso entre:

  -Rentabilidad de la fabricación

-Performance del circuito

Las reglas mas conservadoras nos llevan a circuitos que funcionan mejor pero mas lentos y que ocupan mayor área.

Las reglas mas agresivas tienen mayor probabilidad de generar mejoras en el "performance", pero estas mejoras pueden dañar la rentabilidad.

Las reglas de diseño especifican al diseñador ligaduras geométricas y topológicas que deben cumplir los patrones utilizados en el proceso de fabricación.

Estas ligaduras no son leyes rígidas que se deban cumplir inexorablemente para que los circuitos funcionen correctamente, sino mas bien son recomendaciones del fabricante que aseguran una alta probabilidad de una fabricación correcta.

Se pueden encontrar diseños que violan las reglas y viceversa.

Existen dos conjuntos bien diferenciados de reglas de diseño:

Anchura mínima de las líneas

Distancia entre layers

Una anchura demasiado pequeña lleva consigo una discontinuidad en las líneas lo que puede provocar cortocircuitos.

Si los layers están demasiado cercanos se pueden fundir o interactuar el uno con el otro, cortocircuito entre dos nodos de circuitos diferentes

Hay dos aproximaciones para describir las reglas de diseño:

-Reglas "micron"

-Reglas basadas en lambda l.

Las reglas micron dan las anchuras y distancias entre layers en micras m. La forma en que se trabaja en la industria.

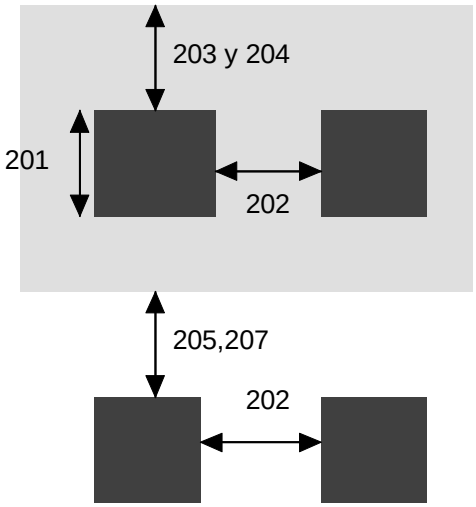
LAMBDA

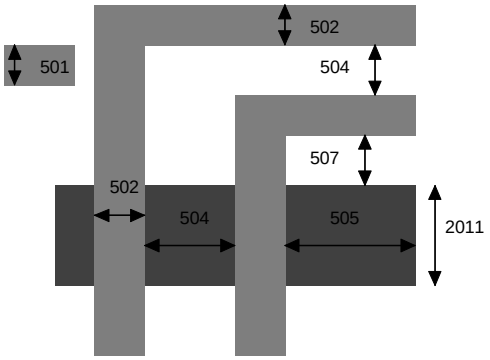
- Es un factor de corrección.
- Fue introducido por Mead-Conway.
- En teoría permite trabajar con diseños independientemente del avance tecnológico.
- Un diseño que utilizase reglas lambda en su descripción serviría para diferentes tecnologías:
- Las reglas lambda se han utilizado con éxito en diseños: 4-1.5 μ.
- No dan buenos resultados para distancias inferiores a las micras.

- En definitiva estas reglas permiten un cierto estado de escalamiento entre procesos diferentes, en este caso seria suficiente reducir el valor de Lambda.
- La experiencia demuestra que las disminuciones no son uniformes.

ENREJILLADO:

- Las herramientas CAD trabajan con enrejillados de dimensiones mínimas en términos de las cuales hay que expresar las reglas de diseño
- para procesos 1.25 m-2 m enrejillados 0.2 m - 0.25 m.
- Por ultimo, alguno de los sistemas de fabricación de mascaras tienen problemas de exactitud digital (de 16 BITS de precisión).





herramienta de extracción eléctrica que obtiene a partir del layout el esquemático del circuito, incluyendo los tamaños de los canales y las interconexiones. El circuito extraido puede utilizarse para comprobar qu eel layout implementa el diseño deseado.

Además el circuito extraído contiene información precisa sobre capacidades parásitas de hilos y difusiones y resistencias, lo que permite una simulación más y un análisis más preciso

3.2 EL DISEÑO SEMICUSTOM

El diseño semicustom es aquel en el que existen ciertas restricciones a la hora de ubicar y rutar los módulos diseñados. Según sean este tipo de restricciones el estilo semicustom se puede subdividir en una serie de subestilos como son:

Basado en celdas estándar

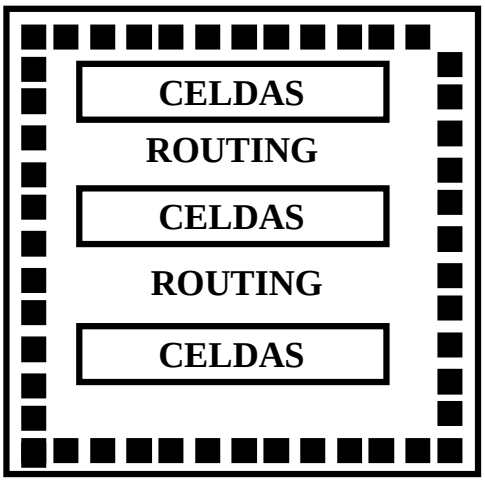
celdas compiladas

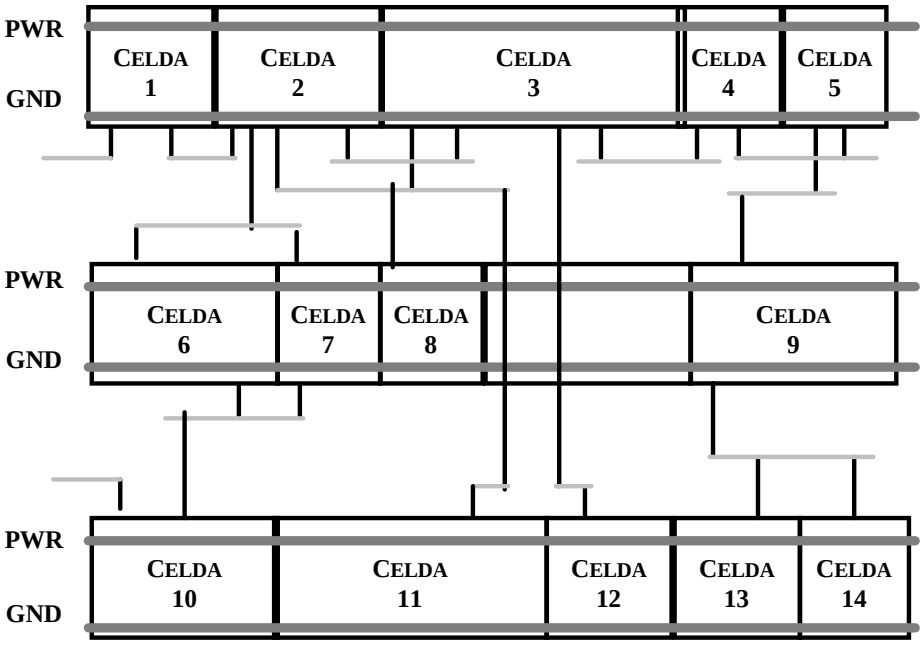
generadores de módulos

FPGAS

3.2.1 BIBLIOTECAS DE CELDAS ESTÁNDARES

- En muchas ocasiones la flexibilidad y grandes prestaciones que se consiguen con el full custom no son necesarias.
- restringe la geometría permitida al circuito y fija una topología de layout específica. Gracias a estas ligaduras se pueden diseñar eficientemente herramientas automáticas, lo acelera enormemente le tiempo de diseño. Como contrapartida los rendimientos y la densidades que se consiguen son menores.
    - Utiliza como elemento básico de diseño un conjunto de celdas que proporciona el fabricante. Cada una de estas celdas implementa una funcionalidad muy básica, como pueden ser puertas OR, AND, biestables, etc. La funcionalidad y características eléctricas de estas celdas están testadas analizadas y probadas por el fabricante para que funcionen correctamente bajo gran número de supuestos.
    - Estas celdas tienen forma rectangular y son todas de la misma altura.
    - todas tiene la toma de alimentación y de tierra en la misma posición que corren horizontalmente a través de las celdas. Y tienen las entradas y salida en las caras superior e inferior
    - Las celdas se colocan en filas y el espacio entre ellas se llama canal. En este estilo se debe fabricar todo el chip.
    - Cada celda puede tener la anchura que necesite para implementar su funcionalidad, por compleja que esta sea.





Como se ve en la figura un chip se compone de filas de celdas estándares intercaladas con zonas de rutado , siendo las filas de celdas es5tándares de la misma anchura, mientras que en las zonas de rutado tienen una anchura que depende de la densidad de rutado en esa zona.

Esta estructura necesita de líneas de rutado verticales que se pueden implementar mediante un layer de metal y añadiendo celdas de trhougput, que son celdas que se dedican exclusivamente a dejar pasar rutados verticales.

Dado que los layout de las celdas estándar están prediseñados por el fabricante, el proceso de diseño se reduce a trasladar la especificaciones a una red de puertas de la biblioteca de celdas estándares. A este paso se le denomina correspondencia con la tecnología (technology mapping). A continuación hay que decidir en que lugar del chip se colocan estas celdas, teniendo como objetivo la minimización del área-intentando que los canales de rutado sean lo más estrechos posibles-.. A esta fase se la denomina placement.

Por último se ejecuta la fase de rutado, que consiste en realizar las conexiones entre las celdas estándar.

Dada la gran regularidad de las celdas estándares, cada uno de esto pasos se puede realizar mediante herramientas de diseño automático.

Dado que el proceso de fabricación de las celdas estándar y del fullcustom es idéntico, los costes y los tiempos de fabricación son los mismos.

El diseñador del full-custom puede realizar importantes optimizaciones de área, consumo o tiempo, que el diseñador de celdas estándares, no puede conseguir. Como contrapartida las celdas estándar pueden ser diseñadas mucho más rápidamente .

No debemos olvidar que el diseño de las celdas estándar, es un diseño Full-custom, es decir es un diseño que consume gran cantidad de tiempo, su ventaja es la enorme reutilización que se realiza de ellas. Esta última característica fuerza a que sean celdas muy robustas, para que funciones dentro de un amplio rango de condiciones.

Dado que el fanin de cada puerta no se conoce hasta que el diseño está realizado, es práctica común asegurar que cada puerta es capaz de trabajar con capacidades de carga elevadas.. Esto simplifica el diseño pero tiene un importante impacto sobre el área y el consumo de potencia.

La información que el fabricante aporta de las celdas estándar, incluye su funcionalidad, su consumo , su fanout, tiempos de subida, tiempos de bajada. Etc.

Las celdas estándar se están utilizando principalmente en los diseños de lógica aleatoria, maquinas de estados finitos, porque se adaptan muy bien a los esquemas multinivel. Además son perfectas para la utilización de herramientas de síntesis lógica.

La síntesis lógica permite tomar como entrada descripciones de sistema mediante lenguajes de alto nivel y obtiene como salida una lista de celdas estándar minimizando el retardo y el área.

En la actualidad el estilo de celdas estándar es el más utilizado en el diseño de Circuitos de aplicación específica e incluso se empieza a utilizar regularmente en el diseño de microprocesadores.

3.2.1.1 Celdas compiladas

Las bibliotecas de celdas estándar tienen la desventaja de ser discretas, es decir de tener el número de opciones limitadas. Cuando se busca como objetivo el rendimiento del sistema son atractivas las celdas con los tamaños optimizados.

Se han generado un conjunto de herramientas para generar layouts según se van necesitando, dando el tamaño del transistor

3.2.1.2 generadores de módulos

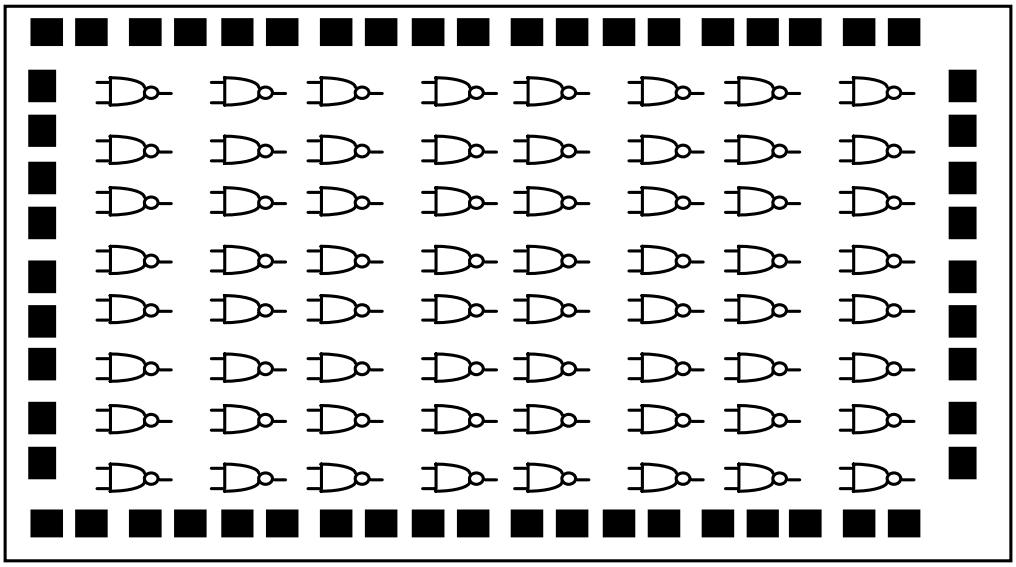
La bibliotecas de celdas son útiles para lógica aleatoria pero son ineficientes para estructuras regulares como desplazadores, sumadores multiplicadores, caminos de datos PLAs o memorias. En todos estos módulos es importante la reducción de las

capacidades internas de los nodos, pero esto es difícil de conseguir utilizando celdas estándar. Además as celdas estándar ignoran la regularidad de estos módulos.

Existen generadores de macroceldas y compiladores de camino de datos

3.2.2 GATE ARRAYS

- Es una simplificación del estilo anterior, sólo que en este caso todas las celdas son iguales.
- En este estilo el chip se prefabrica con un array de puertas idénticas. Estas celdas están separadas por canales verticales y horizontales. El diseño inicial se debe modificar hasta convertirlo en una red de puertas idénticas que se pueda implementar en el chip.
- Como paso final se debe acabar la fabricación realizando el rutado que une las puertas para implementar la funcionalidad deseada



- **FPGAS.** Es un chip ya fabricado formado por arrays de bloque lógicos programables(CLB), en los que se pueden grabar funciones combinacionales, bloques de entrada salida(IOB), que relacionan la lógica con los pines de entrada salida y bloques de interconexión que conectan entre si los bloques lógicos. Con este estilo el diseño no envía a fabricar sino que se implementa programando los CLB, IOB y los bloques de interconexión.

La principal ventaja del estilo semicustom es que utiliza partes ya diseñadas o prefabricadas, con lo que el diseñador puede asegurar con alto porcentaje de probabilidades de acierto su comportamiento eléctrico. Además La utilización de estas partes prediseñadas o prefabricadas ahorra tiempo y esfuerzo de diseño de diseño, y permite diseñadores con menores conocimientos de microelectrónica, pero más versados en temas de diseño lógico o estructura de computadores. Su principal desventaja es que las cotas de optimización son inferiores a las que se alcanzan en el estilo full custom.

3.3 ELECCIÓN DEL ESTILO DE DISEÑO

ν ESTILO FULL-CUSTOM VERSUS ESTILO SEMICUSTOM

La viabilidad de un diseño microelectrónico depende de muchos factores en conflicto como pueden ser el rendimiento en términos de velocidad, el consumo de potencia, el coste y el volumen de producción.

Por ejemplo para que un procesador debe tener un buen rendimiento y un bajo coste para que tenga un rendimiento de mercado adecuado. Conseguir ambos objetivo simultáneamente es sólo posible con volúmenes de producción a gran escala.

Existen otras aplicaciones como el radar o los sistemas espaciales en los que el volumen de producción es pequeño, pero el coste de las partes electrónicas es solo una pequeña parte del total.

Por último, la gran mayoría de los diseños que se realizan solo tiene como objetivo el máximo ahorro de área y el menor tiempo de mercado, para que salgan rentables y competitivos.

Implícitamente hemos podido observar que el coste de un diseño depende de dos factores:

el coste de diseño,

el coste de producción por parte, que depende de la complejidad del proceso, area del diseño y rendimiento del proceso. Vamos a explicar esto con un poco más de profundidad para entender el motivo por el que es tan importante el ahorro de área para los costes de un diseño.

La elección depende del tipo de producto que se esté diseñando. Si el producto es un diseño muy complejo, que se espera producir durante mucho tiempo y en grandes cantidades, la elección serie sin duda el estilo full custom, ya que las ventajas y beneficios finales en optimización de área, rendimiento o consumo superan con creces el esfuerzo de diseño.

En cambio , si el producto que se desea diseñar es un ASIC (Application Specific Integrated Circuit), el estilo de diseño que se elige es el semicustom. Los ASIC son circuitos en los que prima el tiempo de diseño sobre el grado de optimización , es decir diseños que deben estar rápidamente en el mercado, cuya tirada va a ser muy limitada y con un tiempo de vida muy corto porque se rediseñan o modifican muy a menudo.

3.4 ARQUITECTURAS FPGA

3.4.1 ¿QUÉ ES UNA FPGA?

La arquitectura de una FPGA consiste en una matriz o array de bloques lógicos que se pueden programar. Es muy similar a la MPGA (Mask Programmable Gate Array). Las FPGA´s tienen tres componentes principales: bloques lógicos configurables, bloques de entrada - salida y bloques de conexión [Sangiovanni93].

Los bloques lógicos configurables (CLB´S) son los encargados de implementar toda la circuitería lógica del diseño. Están distribuidos en forma de Matriz en el circuito y serán nuestra principal referencia a la hora de hacer el proceso de partición.

Por otro lado están los bloques de entrada y salida (IOB´s) que como su propio nombre índica, son los encargados de conectar la parte del circuito implementada en la FPGA con el mundo exterior. Este "mundo" exterior puede ser directamente la aplicación para la que esté diseñada o como en nuestro caso en el que son necesarias varias FPGA para implementar un circuito, el resto de las FPGA.

Por último están los bloques (switchboxes) y líneas de interconexión que son los elementos de los que dispone el diseñador para hacer el rutado del circuito. En ciertos casos en los que la ocupación de los CLB no es total, estos se pueden utilizar también para llevar a cabo esta tarea.

Los bloques lógicos de una FPGA pueden ir desde algo tan simple como una puerta lógica hasta algo tan complejo como un Microprocesador [Murgai95]. Esto le permite implementar multitud de circuitos tanto combinacionales como secuenciales.

A parte de por la estructura y composición de los bloques lógicos, las FPGA se diferencian también por sus estructuras de rutado y por la tecnología de programación de sus conexiones. Las arquitecturas de rutado de una FPGA puede ser tan simple como una línea de conexión directa entre dos bloques o tan compleja como un multiprocesador (perfect suffle). Por su parte las tecnologías de programación más utilizadas son la SRAM, los antifusibles y las memorias EPROM [Trimberger94].

El proceso de diseño para implementar una FPGA es básicamente el mismo que para un Gate Array. La entrada puede ser tanto un esquemático como una

descripción en un lenguaje de descripción de hardware. El fabricante suministra un software que convierte la descripción del diseño en el programa de la FPGA. El código resultante se puede cargar inmediatamente en el dispositivo y probar el diseño, lo que proporciona una manera muy sencilla de corregir fallos en un diseño.

En la figura 5-1 se muestra la estructura interna de una FPGA. En ella están señalados los bloques lógicos, los IOB´s y las matrices de interconexión [Trimberger93].

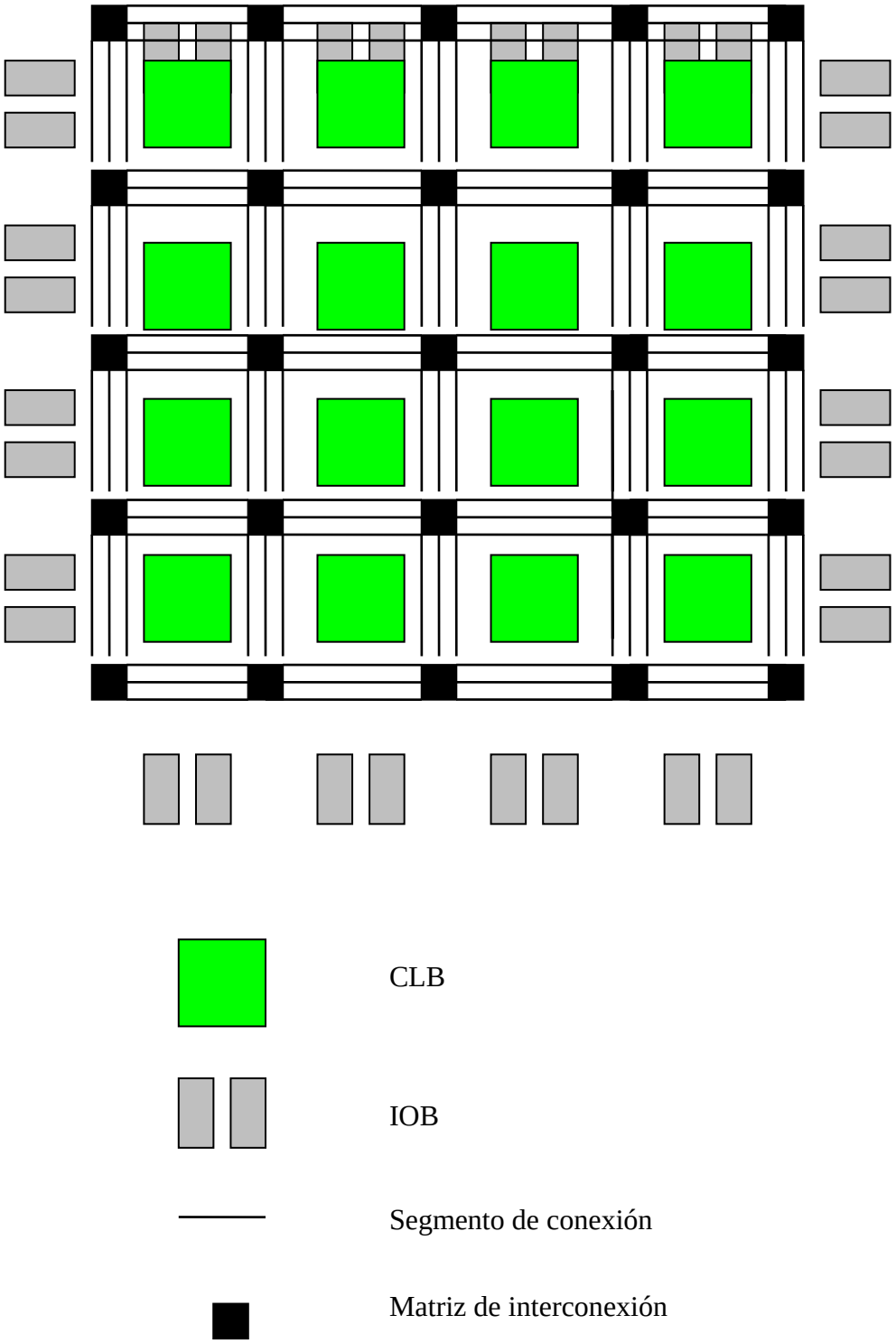


Figura 5-1: Estructura general de una FPGA

3.4.2 TECNOLOGÍAS FPGA

Para programar ala FPGA´s se utilizan diversas tecnologías pero las más importantes son [Rose93] [Rosado95] [Hwang94]:

- SRAM
- Antifusible
- Puerta flotante

3.4.2.1 FPGA´s basadas en SRAM

Las FPGA´s que se programan mediante SRAM, utilizan celdas RAM estáticas para controlar la puerta de paso o los multiplexores. Está tecnología la utilizan los circuito fabricados por Xilinx, Plessey, Algotronic, Concurrent Logic y Toshiba.

En la figuras 5-2 y 5-3 podemos ver dos ejemplos de programación de la FPGA. Si cuando cargamos la SRAM ponemos un uno lógico la puerta de paso estará abierta y se comporta como un interruptor cerrado. Cuando tenemos un cero en la memoria la puerta estará configurada como un interruptor apagado.

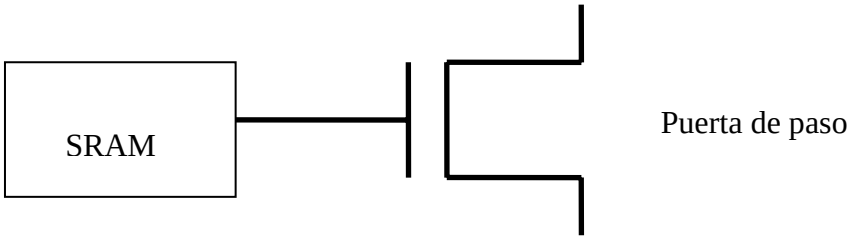


Figura 5-2 Programación con SRAM de una puerta de paso

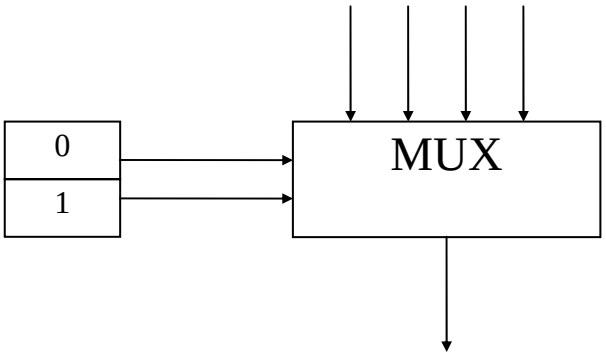


Figura 5-3: Programación con SRAM de un Multiplexor.

3.4.3 VENTAJAS E INCONVENIENTES

Algunas de las ventajas de las FPGA ya se han citado anteriormente. A continuación exponemos algunas de las más importantes [Brown96] [Brown95] [Micro95] .

- El tiempo de programación y puesta en el mercado se reduce considerablemente
- Son programables por el usuario. Esto aparte de dar mayor libertad al diseñador permite una reducción de productos en stock, ya que se pueden utilizar para diferentes aplicaciones
- Algunos tipos son reprogramables, lo que las hace especialmente indicadas para procesos de prototipado en muchos diseños y permite la corrección de errores
- El proceso de diseño es muy simple y asequible.
- Existe una amplia gama de dispositivos que cubren las necesidades de usuarios de todo tipo.
- No necesita procesos de fabricación con máscaras
- Actualmente pueden implementar hasta circuitos de 25000 puertas equivalentes

Los inconvenientes de las FPGA´s son debidos principalmente a su flexibilidad, lo que las hace que en ocasiones sean inapropiadas:

- En primer lugar es más cara que su equivalente programable por máscara, esto es debido a que al tener que dejar los canales de rutado ya delimitados ocupa una mayor área y en una oblea se pueden fabricar menos.
- Es un dispositivo más lento que otros sistemas de propósito específico, debido principalmente a los transistores y matrices de interconexión que utiliza. Para hacernos una idea aproximada los mecanismos de interconexión de una FPGA introducen aproximadamente entre el 30 y el 50 % del retardo total del circuito.
- En ocasiones se desaprovecha parte de la lógica para poder realizar el rutado completo del sistema.

	Full-Custom	**Celdas Estandard**	**Gate Arrays**	**FPGA**
Tamaño celdas	Variable	Altura fija	Fijo	Fijo
Tipo celdas	Variable	Variable	Fijo	Program.
Ubicacion celdas	Variable	Por filas	Fijo	Fijo
Interco-nexion	Variable	Variable	Variable	Program.
Area	Compacto	Compacto a moderado	Moderado	Alta
Performan-ce	Alta	Alta a moderada	Moderada	Baja
Layers fabricados	Todos	Todos	Rutado	Ninguno

4. TEORÍA DEL TRANSISTOR MOS

En este capitulo examinamos las características de los transistores MOS en mayor detalle para predecir su comportamiento real. Nos concentramos en las operaciones estáticas (DC) de los transistores. Este es el primer objetivo de diseño que debe satisfacerse para asegurarse de que las puertas actúan como puertas lógicas. El diseño digital es una abstracción. Todos los diseños son analógicos y la abstracción es solo válida mientras se cumplan los objetivos de diseño.

4.1 TRANSISTOR MOS

La corriente se debe exclusivamente a los portadores mayoritarios y va de la fuente al drenador. Se modula mediante un potencial aplicado en la puerta del transistor. Si los portadores mayoritarios son electrones el transistor es NMOS. Si los portadores mayoritarios son huecos: PMOS

Si el canal por el que se mueven los portadores aparece al aplicar el potencial en la puerta el transistor se llama de **enriquecimiento** (enhancement). Si el canal se fabrica junto con el resto del transistor y al aplicar el potencial en la puerta se elimina el transistor se llama de **empobrecimiento**.

El voltaje de puerta se aplica entre la puerta y el sustrato. Durante todo el estudio, y salvo que se diga lo contrario, la fuente y el sustrato están conectados al mismo potencial y por lo tanto se puede considerar el potencial de puerta entre la puerta y la fuente.

Existen cuatro tipos de transistores MOS:

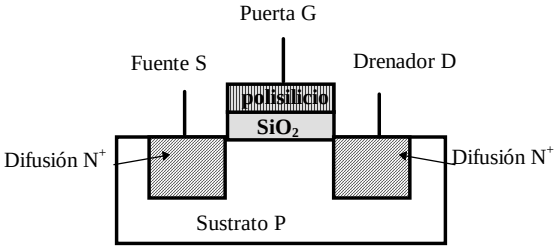
* NMOS de enriquecimiento
* PMOS de enriquecimiento
* NMOS de empobrecimiento o deplexión
* PMOS de empobrecimiento o deplexión

4.2 TRANSISTOR NMOS DE ENRIQUECIMIENTO

4.2.1 DESCRIPCIÓN.

Los portadores mayoritarios son los electrones. Su estructura es la siguiente: tiene un sustrato P ligeramente dopado y dos regiones N+ fuertemente dopadas difundidas dentro del sustrato. Entre estas dos regiones hay una región estrecha del sustrato

llamada canal. Sobre el canal hay una capa (layer) aislante, oxido de silicio SiO2 llamada "Gate Oxide" u óxido de puerta. Sobre el oxido hay un layer de polisilicio llamada puerta (Gate). Como el oxido es aislante la corriente (D.C.) Entre la puerta y el canal es 0. No existe distinción entre fuente y drenador (simetría de la estructura).



4.2.2 MODO DE OPERACIÓN

Suponemos en todo momento Vsustrato=Vfuente

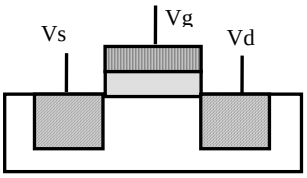
Recordar que:

$V_{GS}= V_G-V_S$

$V_{DS}=V_D-V_S$

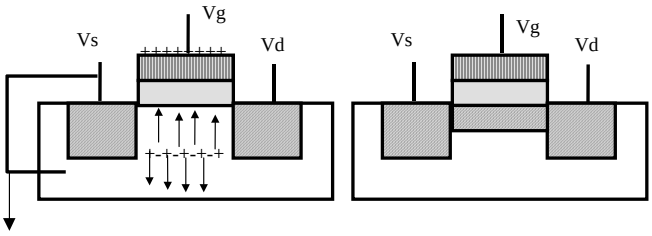
ν **$V_{GS}=0$ Y $V_{DS}>0$**

Suponemos un potencial positivo V_{DS} aplicado entre la fuente y el drenador. Como $V_{GS}=0$ no existe canal → no existe corriente de la fuente al drenador



ν **$V_{GS}>0$ Y $V_{DS}=0$**

Como $V_{GS}>0$ aparece un campo eléctrico E a través del sustrato que atrae a los electrones y repele a los huecos. Si V_{GS} es lo suficientemente grande la región bajo la puerta se convierte de tipo P a tipo N(debido a la acumulación de electrones atraídos) y proporciona un camino entre la fuente y el drenador.



Potencial umbral (Treshold) (Vt): Potencial mínimo necesario para que aparezca el canal N

ν **$V_{GS}>0$ Y $V_{DS}>0$**

Existe canal. Cuando el $V_{DS}>0$ el canal no es uniforme. Esto se debe a la caída de potencial en la componente horizontal del campo debido a la conducción eléctrica. El campo en cada punto del canal tiene dos componentes:

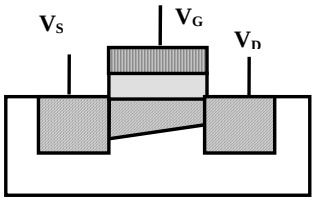
- La vertical que se debe al V_G y que no depende de Y
- La horizontal que se debe a V_{DS} y que varia con X

La intensidad de corriente en el canal depende de la relación que exista entre V_{DS} y $V_{GS} = V_{GB}$

Se llama Voltaje efectivo de puerta: $V_{GS} - V_T$

ν ZONA LINEAL, RESISTIVA, NO SATURADA

Cuando el potencial efectivo de puerta es mayor que el voltaje del drenador el canal es lo suficientemente profundo $V_{GS} - V_T > V_{DS}$



$$Ids = \beta\left[(Vgs - Vt)Vds - \frac{Vds^2}{2}\right]$$

Aunque esta región se llama lineal en realidad solo se cumple esta condición cuando el termino $(V_{DS})^2 / 2$ es muy pequeño, es decir: $V_{DS} << V_{GS} - V_T$

Siendo β el factor de ganancia del transistor MOS. Que viene dado por la siguiente expresión

$$\beta = \frac{\mu\varepsilon ox}{tox}\frac{W}{L}$$

siendo

μ movilidad de canal.

ε_{OX} permisividad del aislante de la puerta [o cte. dieléctrica]

t_{OX}, grosor del aislante

W, anchura del canal.

L, longitud del canal.

KP = parámetro de transconductancia

$$Kp = \frac{\mu\varepsilon ox}{tox}$$

Se puede reexpresar como KP = μ C_{OX}

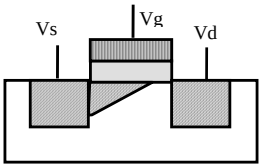
C_{OX} la densidad de capacidad del canal [capacidad / área]

ν ZONA SATURACION

Potencial efectivo de puerta inferior al potencial entre le drenador y la fuente

V_{GS} - V_T < V_{DS} ; V_{GD} < V_T (Pinch - Off).

El canal no alcanza el drenador. En este caso la corriente no se produce a través del canal, sino a través de la zona de deplexión; canal de resistencia muy elevada I_{DS} = cte. El voltaje a través de un canal permanece fijo al valor V_{GS} - V_T



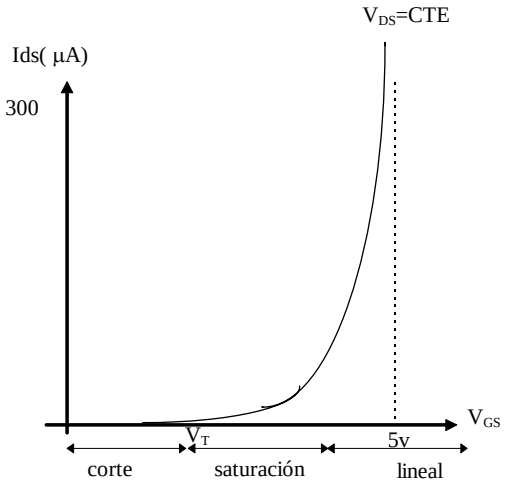
$$Ids = \frac{\beta}{2}(Vgs - Vt)^2$$

Se obtiene de la no saturada sabiendo que la saturación comienza en V_{DS} = V_{GS} - V_T

La corriente de canal es controlada por el potencial de puerta, prácticamente independiente del potencial drenador

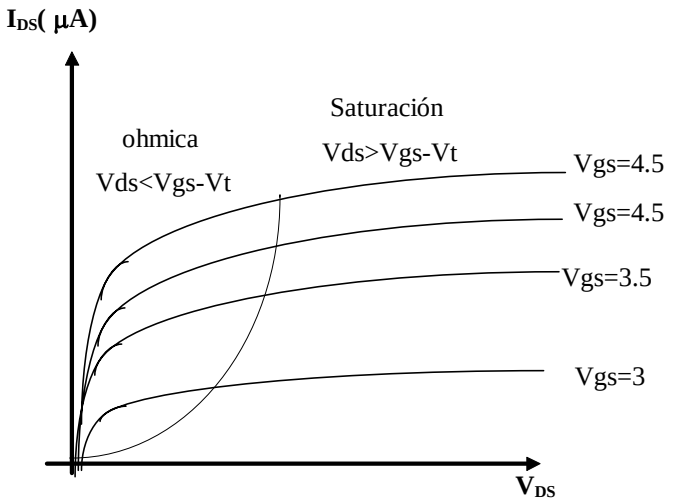
Valores típicos : Para Vdd = 5V y V_T=1V en general V_T=0.2Vdd

ν CARACTERÍSTICAS DE TRANSFERENCIA DE UN NMOS DE ACUMULACION



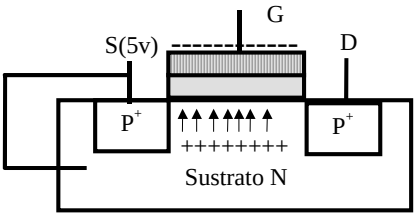
Notar que para un V_{DS} constante al aumentar el V_{GS} el transistor pasa primero por la región de saturación que por la región lineal. Esta es la forma natural de trabajo de la puertas CMOS., Los potenciales Vds no se ven modificados hasta que no se modifican los potenciales de puerta.

ν **CARACTERÍSTICAS DE SALIDA DE UN NMOS DE ACUMULACION.**



Darse cuenta que para un V_{GS} constante las regiones de trabajo por las que pasa el transistor cuando crece el V_{DS} son lineal saturación

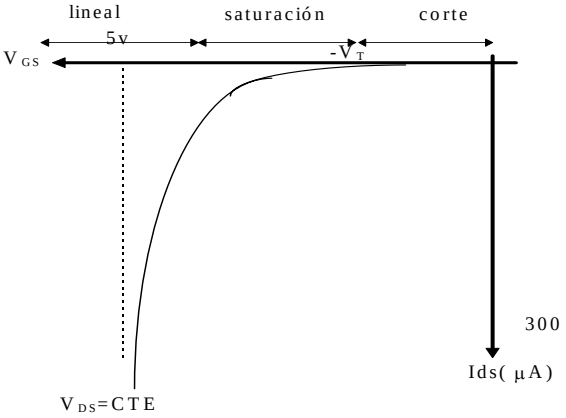
4.3 TRANSISTOR PMOS DE ACUMULACION



El estudio es idéntico al anterior, pero recordado que en este caso los portadores son huecos. El potencial aplicado en D y el potencial aplicado en G deben ser negativos respecto al potencias Vs. Como $V_{GS} < 0$ por inducción se crea un canal P+ entre S y D.Por otro lado como $V_{DS} < 0$ la I_D lleva el sentido de los huecos que irán hacia el potencial negativo D. Por último lógicamente $V_T < 0$.

- **Zona Corte :** $|V_{GS}|<|V_T|$ Es decir, cuando V_T es mas negativo que V_{GS}.
- **Zona lineal:** $|V_{GS} - V_T| > |V_{DS}|$
- **Zona saturación** ; $|V_{GS} - V_T| < |V_{DS}|$

ν **CARACTERÍSTICAS DE TRANSFERENCIA PMOS DE ACUMULACION PARA UN VDS CONSTANTE.**



El estudio de los dispositivos P-MOS suele dar problemas debido al signo de los potenciales:

* $V_{GS} = -|V_{GS}|$ **$V_{GS} < 0$**
* $V_{TP} = -|V_{TP}|$ **$V_T < 0$**
* $V_{DS} = -|V_{DS}|$ **$V_{DS} < 0$**

Independientemente de que las condiciones se expresen en modo absoluto o no, a los potenciales se les debe poner siempre su signo negativo. Esto se hace de una manera o de otra según el potencial de que se trate. Por ejemplo.:

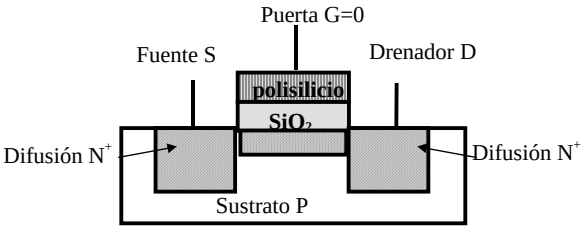
$V_{GS} = V_G - V_S = 0 - Vdd = -Vdd$

De esta manera se sustituye V_{GS} por su valor negativo. Para el potencial umbral siempre se debe hacer : $V_{TP} = -|V_{TP}|$

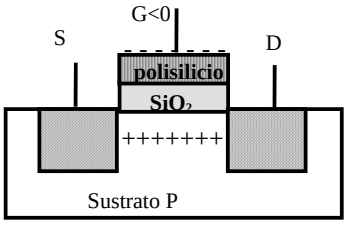
4.4 TRANSISTORES DE EMPOBRECIMIENTO

ν TRANSISTOR NMOS DE EMPOBRECIMIENTO

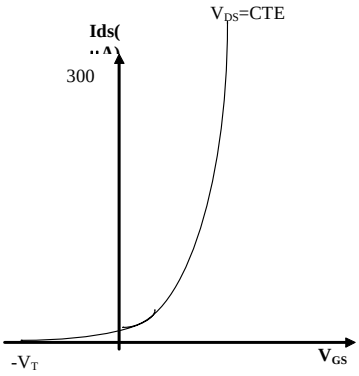
Se fabrica con canal mediante el implante en el canal de los portadores, es decir cuando $V_{GS} = 0$ existe canal.



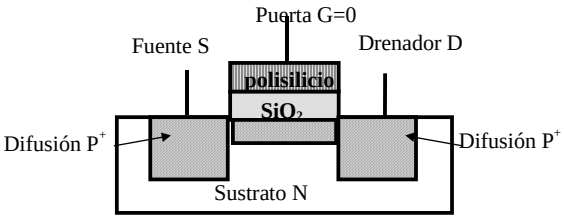
El canal deja de existir para potenciales de puerta negativos [$V_{GS} < 0$]



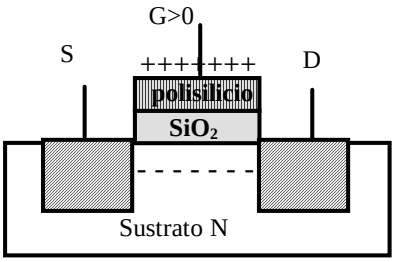
- El potencial umbral es menor que 0, $V_T < 0$.
- Valores típicos de $V_T < -0.8$ Vdd
- Características para NMOS empobrecimiento:



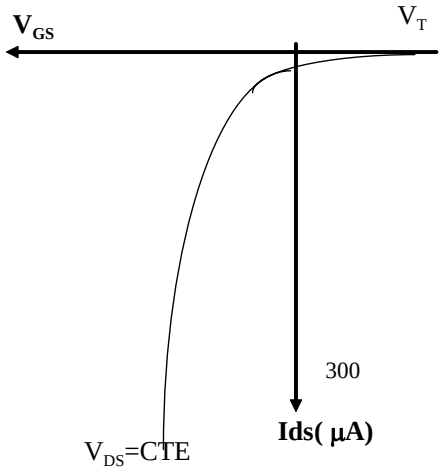
ν **TRANSISTOR DE EMPOBRECIMIENTO TIPO P**



A parte del hecho de los $V_T > 0$, su comportamiento es igual que los transistores PMOS de enriquecimiento



Características para PMOS empobrecimiento:


V_{GS}
V_T
300
Ids(μA)
V_{DS}=CTE

4.5 COMPARACION ENTRE PMOS Y NMOS

Los primeros en utilizarse fueron los PMOS de acumulación por ser de mas fácil fabricación. Las mejoras tecnológicas han dado la primacía a los NMOS, habiendo quedado los PMOS prácticamente obsoletos.

Las razones de la primacía NMOS son las siguientes sabiendo que la movilidad es la velocidad promedio de los portadores(Cm/s) partido del campo eléctrico(Voltio/Cm), la movilidad de huecos: 500 cm^2/Vsg y la de los es electrones: 1300 cm^2/Vsg

Según esto en dispositivos de las mismas dimensiones:

1. La I_{PMOS} menos de la mitad I_{NMOS}
2. Resistencia de PMOS casi tres veces la de un NMOS.

Para que un PMOS y un NMOS alcancen la misma I y R se debe aumentar la relación W/L PMOS, de manera que su superficie debe ser tres veces mayor.

Debido a la mayor densidad NMOS, mayor rapidez en funciones de conmutación. La velocidad esta limitada por la cte. de tiempo RC y C es directamente proporcional a la sección de recta de la unión.

4.6 POTENCIAL UMBRAL

- <u>Def.:</u> Potencial V_{GS} por debajo del cual I_{DS} cae a 0 de manera efectiva.
  (Efectiva porque nunca $I_{DS} = 0$)
- El potencial umbral es función de los siguientes parámetros:
  - Material conductor de la puerta
  - Material aislante de la puerta
  - Grosor del aislante
  - Potencial fuente sustrato
  - Temperatura, su valor disminuye
    - 4 mV / ºC en sustratos muy dopados.
    - 2 mV / ºC en sustratos poco dopados.

4.6.1 ECUACIONES DEL POTENCIAL UMBRAL

$$V_T = V_{TMOS} + V_{FB}$$

- V_{TMOS}: Potencial Umbral del capacitor MOS.
- V_{FB}: Potencial de Flat-Band.

Nota el estudio se realiza para un transistor NMOS . El razonamiento es similar para el PMOS.

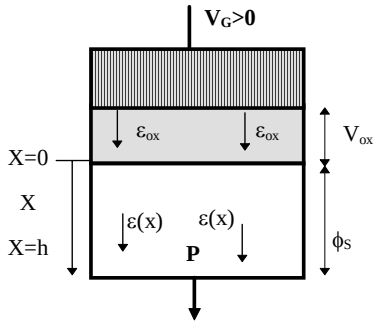
4.6.1.1 Potencial de capacitor MOS, V_{TMOS}

- Se calcula considerando el capacitor ideal que forman la puerta-oxido-sustrato
- **$V_{TMOS} = \phi_B + Vox$**

  Donde

  ϕ_B es el potencial de Bulk

  V_{OX} es la caída de potencial en el oxido



Al aplicar un potencial en V_G se genera un campo eléctrico **εox** en el oxido y un campo eléctrico un ε(x) en el sustrato semiconductor. Este último campo es el que controla la densidad de portadores en el sustrato

A nosotros nos conviene relacionar de alguna manera el V_G aplicado en la puerta con los campos creados εox y ε(x) que genera. El potencia electrostático φ(x) de un semiconductor se relaciona con el campo ε(x) a través de la expresión del gradiente

$$\varepsilon(x)=-d\phi(x)/dx$$

sabiendo que la condición frontera de φ(x=h) es $V_{SUSTRATO}$=0 se puede encontrar que la caída de potencial en el sustrato es φ(X=0) que se le llama potencial de superficie (φs). Por otro lado sabemos que el campo εox provoca la caída de potencial en los extremos del oxido que llamamos Vox. La relación entre los campos creados y el V_G se expresa en forma de suma de potenciales

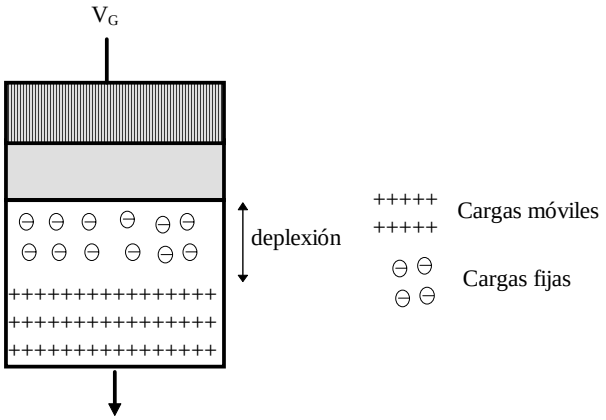
$$V_G= \phi_s +V_{OX}$$

Este valor hay que particularizarlo $V_{T0}=V_G$ mínimo para que aparezca canal.

- **ESTUDIO DEL POTENCIAL DE BULK ϕ_B**

Vamos a encontrar el valor mínimo de φs y le vamos a llamar φb. El campo eléctrico ε(x) es el que empuja las cargas positivas del sustrato hacia abajo generando la zona de deplexión. La profundidad de la zona de deplexión crece con el V_G hasta que alcanza un valor máximo a partir del cual se desplaza hacia abajo.

Def potencial de Bulk potencial de superficie ϕs que hay que aplicar para que se alcance la máxima profundidad de deplexión . Se le representa por $\phi_{B.}$ Cuando se alcanza la máxima profundidad de deplexión está a punto de producirse la inversión dura



El potencial necesario para alcanzar la profundidad de deplexión máxima es $2\phi_F$ siendo ϕ_F el potencial de Fermi

<u>Potencial de Fermi</u>: es la diferencia entre el nivel de energía de Fermi del semiconductor sin dopar y el semiconductor dopado.

$$\phi F = \frac{KT}{q}\ln\left[\frac{Na}{Ni}\right]$$

Na	Densidad de portadores del sustrato dopado
Ni	Densidad de portadores del silicio intrínseco [sin dopar].
K	Cte. de Bolztman.
T	Temperatura absoluta.
q	Carga electrónica.

ν **V_{ox}**

La siguiente cuestión es cuanto debe valer Vox para que ε(X) sea suficiente para alcanzar como mínimo el $\phi_B=2|\phi_F|$. Vox Identifica la caída de potencial a través del oxido de puerta. Hay que tener en cuenta que esta caída se produce en un capacitor y viene dada por la expresión:

$$Vox = Q_B/C_{ox}$$

Siendo Q_B la carga que hay que inducir para generar la zona de deplexión máxima (es decir la carga que hay que inducir para neutralizar las cargas del sustrato P) y que por similitud con la expresión obtenida para un caso similar en los diodos se sabe que es:

$$QB = \sqrt{2q\varepsilon siNa2|\phi F|}$$

Nota :darse cuenta que como estamos hablando de potencial umbral se supone que en el limite casi hay canal pero este todavía no existe, es decir todavía no hay inversión dura bajo la puerta, por eso al estudiar la carga Q_B solo se habla de máxima deplexión que es justo el límite.

es decir la carga inducida es función del potencial de superficie representado en este caso por $\phi_{F.}$ C_{OX} es la capacitancia por unidad de área del oxido de la puerta. Su valor es inversamente proporcional al grosor del oxido Tox.

$$Vtmos = 2|\phi F| + \frac{Q_B}{Cox}$$

$$Vtmos = 2|\phi F| + \frac{\sqrt{2q\varepsilon siNa2|\phi F|}}{Cox}$$

Este potencial es positivo para un transistor NMOS

4.6.1.2 Potencial de Flat-Band

El potencial de Flat band son unas correcciones que se añaden a la expresión anterior para tener en cuenta que el capacitor esta formado por materiales diferentes [polisilicio, silicio monocristal], para compensar la existencia de carga fija [Q_{OX}] y ajustar el Vumbral mediante el implante de impurezas $Q_{I.}$

ν PRIMERA CORRECCIÓN

Se debe a que el capacitor MOS no es perfecto porque el capacitor esta formado por materiales diferentes. Viene dada por la función de trabajo ϕ_{MS} que existe entre el polisilicio de puerta y el silicio. Siendo $\phi_{MS}=\phi_M-\phi_S$ su valor es similar a $\phi_{B.}$ Su magnitud y signo dependen de los materiales de la puerta y el sustrato. Para un transistor de tipo NMOS es un valor negativo aproximadamente igual a -0.9V, para un transistor PMOS su valor es aproximadamente -0.2V

ν SEGUNDA CORRECCIÓN

Se debe al hecho de que las superficie de unión entre el sustrato y el óxido no es perfecta. En esa unión quedan atrapadas cargas positivas que a su vez habrá que compensar induciendo más carga negativa. Como ya vimos en su momento la carga inducida es V=Q/C luego el potencial umbral **aumenta** debido a esto: en una cantidad **$|Q_{OX}|/C_{OX}$**

ν TERCERA CORRECCIÓN

Se debe al hecho de querer modificar el potencial umbral en el proceso de fabricación. Para ello se añade carga extra en la zona de debajo del canal. Esta carga

extra se denomina Q_I . El potencial aplicado en la puerta debe ser capaz de generar la carga inducida que compense esta carga extra

El potencial umbral se ve modificado en una cantidad Qi/C_{OX}.. En este caso el potencial umbral puede aumentar o disminuir en función del signo de la carga inducida. Si esta es de signo positivo el potencial aumenta, si es de signo negativo disminuye. Para el estudio suponemos que es de signo positivo

El potencial total de flatBand es

$$V_{FB} = \phi_{MS} + |Q_{OX}|/C_{OX} + |Q_I|/C_{OX}$$

Luego el potencial umbral total es:

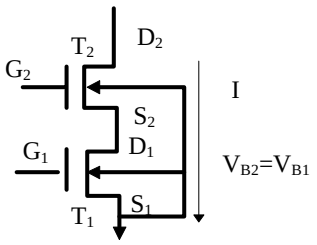
$$V_T = 2|\phi_F| + |Q_B|/Cox + |Q_{OX}|/C_{OX} + |Q_I|/Cox - |\phi_{MS}|$$

La mayoría de los términos de esta expresión son parámetros de la tecnología del material. Q_B es función de V_{SB}.

En ocasiones es necesario modificar el potencial umbral de un dispositivo [p.e. caso del canal-stop para eliminar transistores parásitos]. Existen dos fórmulas sencillas: Modificar la concentración de dopaje en la superficie de unión oxido sustrato o Modificar C_{OX} que a su vez se puede hacer de dos formas diferentes: Utilizando diferentes aislantes, modificar C_{OX}, Modificando el grosor del oxido

4.7 EFECTO SUSTRATO [EFECTO BODY]

Todos los dispositivos MOS de un circuito integrado se fabrican en un sustrato común. Este sustrato se encuentra al mismo potencial para todos los dispositivos [en analógicos puede no ser cierto]. En ocasiones los transistores se deben conectar en serie lo que tiene como efecto un iincremento del potencial puerta-sustrato necesario para que empiecen a conducir.



La condición de conducción que se debe cumplir es $V_{GB} = V_T$, aunque por comodidad se trabaja con V_{GS}. generalmente al polarizar el sustrato del transistor $V_B=V_S$

ν **TRANSISTOR T_1**

El sustrato(B) está polarizado al mismo potencial que la fuente (S) luego $V_{S1} = V_{B1}$. La condición de conducción que hemos utilizado hasta el momento es $V_{GS} = V_T$. Partiendo de estos dos datos vamos a buscar el valor de V_{GB}

$V_G - V_S = V_T$

como $V_S = V_B \rightarrow V_G - V_B = V_T$

$\mathbf{V_{GB} = V_T}$

Que es la misma expresión con la que hemos trabajado hasta el momento. En este caso $V_{SB} = 0$

ν **TRANSISTOR T_2**

Como los dos transistores están conduciendo debe haber una caída de potencial

$V_{S2} > V_{S1} \rightarrow V_{S2} = V_{S1} + K$ (1)

Ademas $V_{S1} = V_B$

Condición de conducción $V_{GS} = V_T$

$V_G - V_{S2} = V_T$

$V_G = V_T + V_{S2}$

de (1)

$V_G = V_T + V_{S1} + K \rightarrow V_T + V_B + K$

$\Rightarrow \mathbf{V_{GB} = V_T + K}$

Efecto sustrato o efecto Body es: incremento del potencial umbral debido a la existencia de un potencial entre el sustrato y la fuente (Vsb)

4.8 EFECTOS DE SEGUNDO ORDEN

Hasta el momento se ha discutido el comportamiento de un dispositivo MOS casi ideal. Pero no se ha tenido en cuenta que el comportamiento del transistor se puede desviar bastante del ideal. Esto es especialmente cierto cuando las dimensiones del dispositivo se acercan al orden de micras. En estos casos el canal pasa a tener longitudes similares a otros parámetros, y se debe tener en cuenta la profundidad de deplexión en las uniones y la profundidad de las difusiones. A estos dispositivos se les llama de canal corto. En el modelo de canal largo [el habitual] el modelo de una dimensión (la vertical) es válido pero en el de canal corto esta descripción se aparta de la realidad.

4.8.1 VARIACIONES DEL POTENCIAL UMBRAL

Se van a tratar tres casos:

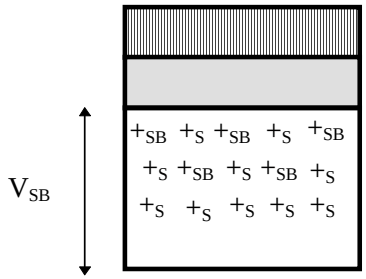
- Variación del potencial umbral debido al efecto cuerpo
- Variación debido a la existencia de las zonas de deplexión en las difusiones
- Variación debido al V_{DS}

ν **EFECTO DEL POTENCIAL SUSTRATO BIAS**

En el punto del potencial umbral NMOS supusimos que la carga que debía inducirse en el capacitor venía dada por la siguiente expresión

$$Q_{B0} = \sqrt{2 \cdot q \cdot \varepsilon_{si} \cdot Na \cdot 2|\phi_F|}$$

Esta expresión solo tiene en cuenta el potencial de superficie aplicado y la densidad de portadores del sustrato. Cuando existe un potencial entre la fuente y el sustrato (efecto cuerpo) hay que introducir una variación en la expresión que tenga en cuenta la variación de portadores. A la anterior expresión vamos a llamarla a partir de ahora $Q_{B0.}$ Vamos a estudiar las modificaciones que hay que introducir para tener en cuenta el efecto cuerpo $V_{SB}>0$:



Siendo $+_S$ el exceso de carga positiva del semiconductor P y $+_{SB}$ la carga en la superficie, debida a que como V_{SB} es positivo y no existe intensidad almacena estas cargas en la superficie.

Como el V_{SB} en definitiva es una modificación del potencial de superficie ϕs, se debe incluir en Q_B

$$\mathbf{Q_B = \sqrt{2 \cdot q \cdot \varepsilon si \cdot Na \cdot (2 \cdot |\phi_F| + |V_{SB}|)}}$$

NOTA: si esta carga no fuera motivada por un potencial de superficie, sino que fuera debida a un implante o a irregularidades se incluiría en la expresión del potencial umbral como Q/C

La tensión umbral con esta modificación queda

$$\mathbf{V_T = 2|\phi_F| + |Q_B|/Cox + |Q_{OX}|/COX + |Q_I|/Cox - |\phi_{MS}|}$$

Esta expresión se puede reorganizar de la siguiente manera:

$$Vt = Vto + \gamma\left[\sqrt{2|\Phi_F| + |Vsb|} + \sqrt{|2\Phi_F|}\right]$$

donde

$$V_{T0} = 2|\phi_F| + |Q_B|/Cox + |Q_{OX}|/COX + |Q_I|/Cox - |\phi_{MS}|$$

V_{TO}: Potencial umbral para $V_{SB} = 0$.

$\gamma = [\, 2q\varepsilon_{si} N_A \,]^{1/2} / Cox$ coeficiente de efecto cuerpo

$C_{OX} = \varepsilon_{OX} / T_{OX}$. La capacitancia por unidad de área

ε_{OX} : Permisividad del oxido [Cte. dieléctrica]

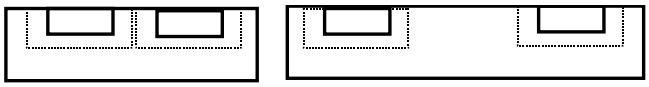
T_{OX}: Grosor del oxido.

[en SPICE V_{TO} = VTO ; N_A = NSUB; $2\Phi f = \Phi s$]

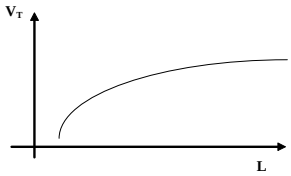
Conclusión: **El efecto cuerpo produce un aumento del potencial umbral de los dispositivos.**

Ejemplo: sea un transistor NMOS con un potencial umbral VT=0.75 . Se sabe que el coeficiente de efecto cuerpo es γ=0.54. Calcular el potencial umbral del transistor sabiendo que V_{SB}=5V y que $2\phi_F$=-0.6 V (notar el signo negativo de $2\phi_F$, por eso se pone entre valores absolutos con un signo menos delante). Sustituyendo en la expresión se obtiene VT(5 V)=0.75 + 0.86=1.6 V,Que es más del doble del potencial cuando no existe el efecto cuerpo

ν **EFECTO DEPLEXION DE LAS DIFUSIONES**



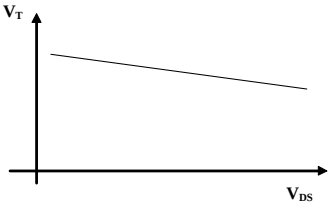
Para tamaños grandes de canal con la descripción del modelo conseguida es suficiente (V_T constante). Pero, ¿que ocurre cuando el transistor tiene un canal pequeño?. En el estudio anterior se habló de las zonas de deplexión bajo puerta pero no de la deplexión de las difusiones. Cuando L es grande la deplexión de las difusiones se puede despreciar, Pero cuando L pequeño es muy importante. Disminuye el V_{BULK} necesario puesto que esa zona de deplexión es más fácil de crear



ν **VARIACIÓN DEL POTENCIAL UMBRAL CON V_{DS}**

Cuando V_{DS} es pequeño, un aumento produce un aumento en la zona de deplexión y el drenador.

- Disminución V_{TO}
- Efecto DIBL (DRAIN INDUCED BARRIER LOWERING.)



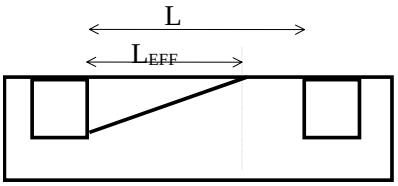
4.8.2 MODULACIÓN DE LA LONGITUD DE CANAL

Las ecuaciones que hemos estudiado hasta el momento suponen que en saturación la longitud del canal L es constate

$$I = (K \cdot W / 2L)\,[V_{GS} - V_T]^2$$

y por lo tanto

$I \ddagger f(V_{DS})$



En una segunda aproximación se puede ver que la longitud real del canal no es L, sino $L_{EFF} = L - L_{SHORT}$

donde

$$L_{SHORT} = \{ 2\, \varepsilon_{si} / qN_A [V_{DS} - [V_{GS}-V_T]] \}^{1/2}$$

Es decir

$L = f(V_{DS})$

$I = f(V_{DS})$

Estas variaciones se pueden incluir en la I_{DS} con la siguiente expresión:

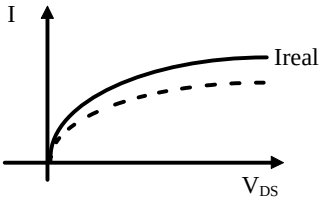
$$I_{DS} = (KW/2L)\, [V_{GS}-V_T]^2\, [1+\lambda V_{DS}]$$

donde

K es la ganancia del proceso $\mu\varepsilon/t_{OX}$

λ: Modulación empírica de la longitud de canal.

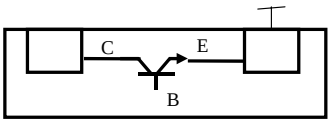
SPICE: LAMBDA -0.02 V^{-1} - 0.005 V^{-1}



Conclusión: El decremento efectivo del canal aumenta la Intensidad calculada inicialmente

4.8.3 CONDUCCIÓN SUBUMBRAL

Existe intensidad de corriente entre la fuente y el drenador aunque el $V_{GS}<V_T$



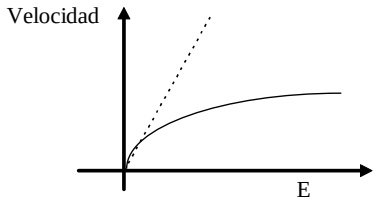
Hasta el momento habíamos visto que si $V_{GS} < V_T$ entonces $I_{DS} = 0$

- Los difusores N y el sustrato P forman un transistor bipolar parásito que puede conducir en inversa
- $V_{BE}=V_B-V_E= 0-V_{DD}$

- La I_{DS} incrementa exponencialmente con V_{DS} y V_{GS}.
- **Ventaja**: diseño de baja potencia.
- **Inconveniente:** mal funcionamiento en dispositivos dinámicos , en estos dispositivos la capacidad es la memoria del mismo y debido al efecto se producen pérdidas de información.

4.8.4 SATURACIÓN DE LA VELOCIDAD DE LOS PORTADORES

Hasta el momento habíamos supuesto que la velocidad de los portadores es proporcional al campo eléctrico, independientemente de su valor, es decir μ=constante. La realidad es que la velocidad de los electrones se satura, por lo tanto por mas que crezca el campo la velocidad no crece más.



- En el silicio de tipo P:
  - Campo de saturación Es = $1.5 \cdot 10^4$ Vol/cm
  - Velocidad de saturación $V_{SAT} = 10^7$ cm/seg.
  - Para un canal NMOS de L= 1μ señal suficiente V_{DS} = 2 volt. para alcanzar esta condición.(E=V/L)
- Los huecos (silicio de tipo N) saturan a la misma velocidad aunque necesitan un campo mayor.
- La Ids en saturación cuando la velocidad se satura viene dada por la expresión:

$$I_{DSAT} = V_{SAT}\, C_{OX}\, W\, (V_{GS} - V_{DSAT} - V_T)$$

- Consecuencias:
  - La Id tiene una dependencia lineal de V_{GS} (no cuadrática como se había visto). Por lo tanto reducir los potenciales de trabajo no tiene efectos sobre los dispositivos de canal pequeño
  - Las variaciones de V_{GS} no afectan tanto a I_{DS}.
  - I_{DS} no depende de la longitud, por lo tanto el dispositivo no se puede mejorar reduciendo el canal.

4.8.5 TUNEL FOWLER-NORHEIM

Cuando el layer del oxido es muy fino puede aparecer un flujo de corriente entre la puerta y la fuente o el drenador; es decir aparece una conexión a través del oxido.

- $I_{FN} = C_1 wL\ C_{OX}^2\ e^{[-Eo/Eox]}$
- E_{OX}; campo eléctrico a través del oxido.
- $E_{OX}=V_{GS}/T_{OX}$.
- Eo y C_1 ctes.
- **Efecto**: limita la anchura mínima del oxido
- Se utiliza en dispositivos lógicos programables alterables electrónicamente.

4.8.6 PERFORACION DE CANAL (DRAIN PUNCHTROUGH)

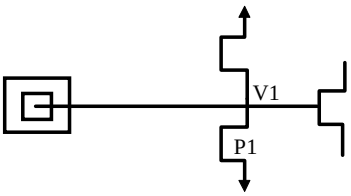
Cuando V_{DS} es lo suficientemente elevado aparece una I_{DS} independiente de V_G Se rompe la barrera de potencial entre S y D. Debido a la naturaleza bidimensional del problema es difícil obtener ecuaciones exactas que lo describan. La I_{DS} aumenta linealmente con la densidad de dopaje y cuadráticamente con L

Este efecto se usa para la protección de los circuitos, de manera que se limitan los voltajes a través de los nodos del circuito. Las señales de entrada de una puerta lógica atacan casi siempre al contacto puerta G del transistor. El máximo potencial que se puede colocar en G sin temor a perforar el oxido es 40-100V. El voltaje que aparece o crece en la puerta viene dado por la siguiente expresión:

$$V=I\ (T_1-T_2)/C_G$$

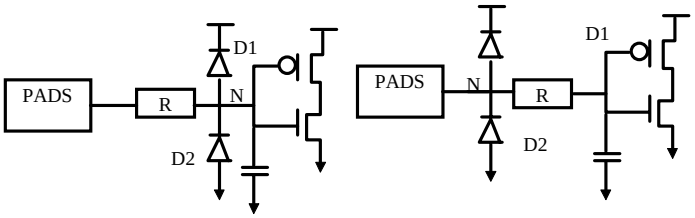
por ejemplo para los valores $I=10\mu A, C_G=00{,}03pF, t_1-t_2=1\mu sg$, el potencial producido es de V=330V

Se colocan transistores de Punch-Trough sin puerta que rompen al conducir con potenciales próximos a los 50V.



Si V_1>50V P1 rompe a conducir cuando el potencial en V1

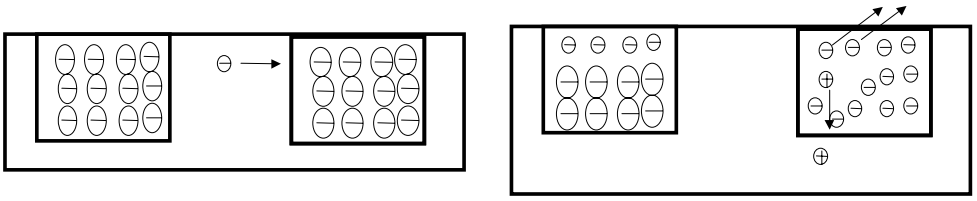
NOTA: generalmente se puede utilizar una combinación de diodos y resistencias para limitar el voltaje potencialmente destructivo. Otros circuitos típicos son los siguientes:



Los diodos conducen si el potencial en N sube por encima de V_{DD} o baja por debajo de gnd, las resistencias se utilizan para limitar la corriente que fluye en los diodos, R=200-3kohmios,D1 conduce en directa cuando $V_N>V_{DD,}$ D2 conduce en directa cuando $V_N<$gnd

4.8.7 ELECTRONES CALIENTES. IONIZACIÓN DE IMPACTO

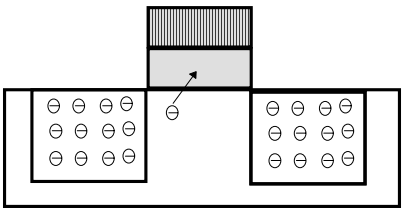
Cuando la longitud del canal es muy reducida, el campo eléctrico del drenador en el dispositivo aumenta (para un V_{DS} = cte.) Cuando la longitud es menor que micras, los electrones toman suficiente energía como para chocar contra otros electrones del drenador generando huecos.



Los huecos que aparecen por impacto son atraídos por el sustrato que tiene el potencial negativo y son repelidos por el drenador de potencial positivo. Esto genera una corriente de huecos a través del sustrato. A este proceso se le conoce como **ionización por impacto.**

Los electrones con suficiente energía se llaman **electrones calientes**, fenómeno similar a la conducción por avalancha. En algunas ocasiones estos electrones tienen suficiente energía para penetrar en el oxido, produciendo una corriente de puerta. Esto puede llevar a la degradación de los dispositivos MOS

- Potencial umbral
- Corriente subumbral
- Transconductancia



En ocasiones la corriente de sustrato se puede utilizar positivamente para evaluar la importancia del efecto electrón caliente. En el aspecto negativo: Dan lugar a pobres tiempos de refresco en las memorias dinámicas. Ruidos en los sistemas de señales mixtas y Posible generación de Latchup.

Los “Hot Holes” no presentan problemas normalmente, debido a su movilidad. La existencia de los electrones calientes ha guiado el desarrollo de la tecnología CMOS en los últimos años.

5 LOS INVERSORES MOS

5.1 DEFINICIONES Y PROPIEDADES

El inversor es el núcleo de todos los diseños digitales. Una vez que sus propiedades y modos de operación están claros, el estudio del resto de diseños más complejos se simplifica. El estudio de estos circuitos complejos se puede realizar extrapolando los datos obtenidos en el inversor

5.1.1 RUIDO

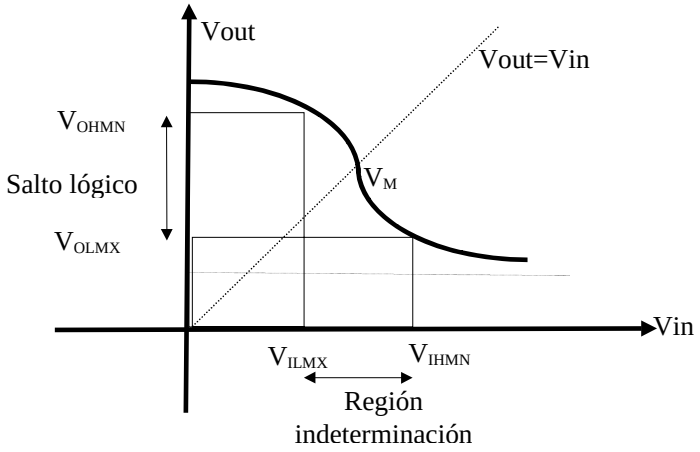
Lo mínimo que se le puede pedir a una puerta lógica es que ejecute la función para la que fue diseñada. El problema es que aparecen desviaciones del comportamiento debido factores de fabricación como son variaciones en el potencial umbral o en las dimensiones del transistor, o bien desviaciones debidas a las fuentes de ruido internas o externas al chip.

Definición ruido: Variaciones de intensidad o potencial no deseadas en los nodos lógicos. Las causas más habituales son:

* Inductancias acopladas.- Las variaciones de I en un provocan variaciones de intensidad en hilos cercanos
* Capacitancias acopladas
* Ruidos en los potenciales de alimentación y tierra

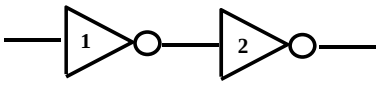
5.1.2 MARGENES DE RUIDO

Es la máxima fluctuación que se permite a los potenciales de salida para que sigan representando el mismo valor lógico de entrada

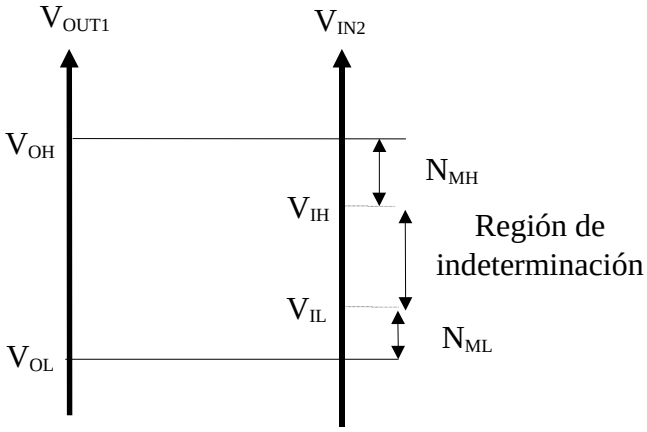


- $\mathbf{V_{IL}}$- máximo potencial que representa un 0 de entrada
- $\mathbf{V_{IH}}$- mínimo potencial que representa un 1 de entrada
- $\mathbf{V_{OL}}$.- máximo potencial que representa un 0 en la salida
- $\mathbf{V_{OH}}$.- mínimo potencial que representa un 1 lógico a la salida

$\mathbf{V_M}$ potencial umbral de interruptor o de puerta. Es el punto en el que se realiza la transición lógica y la condición que debe cumplir es $V_M=f(V_M)$ o lo que es lo mismo $V_{IN} = V_{OUT}$. No se debe confundir con el potencial umbral de transitar V_T. Para realizar el estudio se supone que la salida de un inversor es la entrada de otro, $V_{OUT1}=V_{IN2}$



A continuación se dibuja el salto lógico del 1 y la región de transición del dos



Es importante tener en cuenta que la salida de uno es la entrada del otro, el problema es que el valor lógico representado por la salida puede no ser el potencial lógico representado por la entrada.

¿Como se interpreta la figura anterior?. Para que el funcionamiento del inversor sea correcto se debe cumplir

- $V_{OH} > V_{IH}$ el 1 de salida siempre es un 1 de entrada, incluso cuando la salida fluctúa al rededor de V_{OH}
- $V_{OL} < V_{IL}$ el cero de salida siempre es un cero de entrada incluso cuando la salida fluctúa alrededor de V_{OL}

- $\mathbf{N_{MH}=|V_{OH}-V_{IH}|}$ Margen de ruido de alta
- $\mathbf{N_{ML}=|V_{IL}-V_{OL}|}$ Margen de ruido de baja

Por ejemplo: si el máximo 0 lógico de salida es 1,5 y el máximo 0 lógico de entrada es 2,5, el cero de salida siempre es un cero lógico de entrada.

¿Como mejorar los márgenes de ruido

- Mayor salto lógico
- Región de transición más pequeña

* Lo deseable $V_{IL}=V_{IH}=V_{DD}/2$

5.2 EL INVESOR CMOS DE CARGA DINÁMICA

Es la puerta de diseño básica de la tecnología FULLY CMOS estática.

Atención: No confundir el tipo de carga y el tipo de lógica

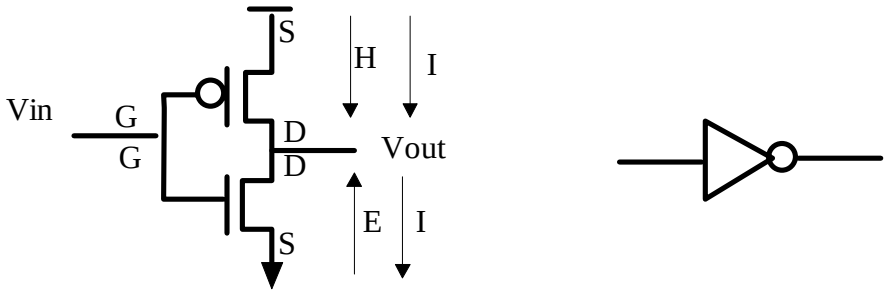
Tipos de lógica

- **Lógica estática** es cuando en cualquier instante de tiempo la salida está conectada al V_{DD} o al Gnd o a ambos
- **Lógica dinámica** utiliza capacidades parásitas como elemento de memoria. Puede ocurrir que exista un instante de tiempo en el que la salida no este conectada a V_{DD} o gnd.

Tipo de carga (pull-up)

- **Dinámica:** Cuando el transistor de carga conduce y esta cortado alternativamente
- **Estática**: Siempre conduce aunque en diferentes regiones

5.2.1 REGIONES DE TRABAJO DEL INVERSOR



Para encontrar las diferentes regiones del inversor vamos a estudiar las regiones de trabajo de cada uno de los transistores que lo forman. Para que se puedan observar en la característica de entrada/salida se van a poner las regiones de trabajo en función de V_{IN} y $V_{OUT.}$ En general

* $V_{GS}=F(V_{IN})$
* $V_{DS}=F(V_{OUT})$

Particularizando para cada transistor

- **Para el PMOS**

  $S=V_{DD}$

  $D=V_{OUT}$

  $G=V_{IN}$

  $B=V_{DD}$

  $V_{GS} = V_G - V_S = V_{IN}-V_{DD}$

$V_{DS} = V_D - V_S = V_{OUT} - V_{DD}$

- **Para el NMOS**

  V_S=GND=0

  $V_D=V_{OUT}$

  $V_G=V_{IN}$

  $V_{GS}=V_G-V_S=V_{IN}$

  $V_{DS}=V_D-V_S=V_{OUT}$

ν **Regiones de trabajo del PMOS**

- **Corte**

  $V_{GS} > V_{TP} \rightarrow \mathbf{V_{IN} > V_{TP} + V_{DD}}$

- **Saturación**

  $V_{GSP}<V_{TP} \rightarrow V_{IN}<V_{TP}+V_{DD}$

  $V_{DSP}<V_{GSP}-VTP \rightarrow \mathbf{V_{OUT}<V_{IN}-V_{TP}}$

- **Lineal**

  $V_{GS}<V_{TP} \rightarrow V_{IN}<V_{TP}+V_{DD}$

  $V_{DSP}>V_{GSP}-V_{TP} \rightarrow \mathbf{V_{OUT}>V_{IN}-V_{TP}}$

ν **Regiones de trabajo del NMOS**

- **Corte**

  $V_{GSN}<V_{TN} \rightarrow \mathbf{V_{IN}<V_{TN}}$

- **Saturación**

  $V_{GSN}>V_{TN} \rightarrow V_{IN}>V_{TN}$

  $V_{DSN}>V_{GS}-V_{TN} \rightarrow \mathbf{V_{OUT}>V_{IN}-V_{TN}}$

- **Lineal**

  $V_{GSN}>V_{TN} \rightarrow V_{IN}>V_{TN}$

  $V_{DSN}<V_{GSN}-V_{TN} \rightarrow \mathbf{V_{OUT}<V_{IN}-V_{TN}}$

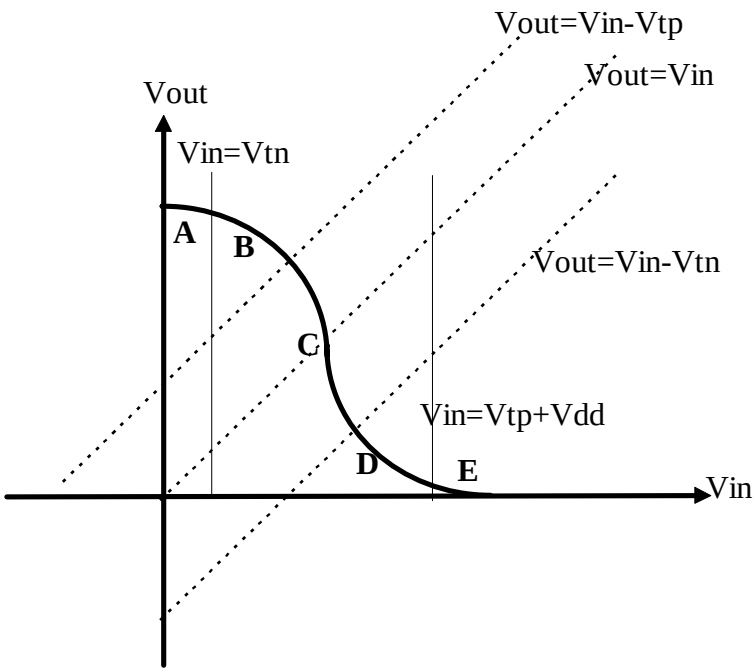
ν **Regiones de trabajo del inversor**

Vistos los valores de potenciales que delimitan las zonas de trabajo de cada uno de los transistores pasamos a estudiar las regiones de trabajo del inversor.

- Para el transistor NMOS hay dos fronteras:
  - ∗ Paso de corte a saturación $V_{IN}=V_{TN}$
  - ∗ Paso de saturación a lineal $V_{OUT} = V_{IN}-V_{TN}$
- Para el transistor PMOS hay dos fronteras:
  - ∗ Paso de corte a saturación $V_{IN}=V_{DD}+V_{TP}$

- ∗ Paso de saturación a lineal $V_{OUT}=V_{IN}-V_{TP}$

- Suposiciones
  - ∗ Recordar que Vin es el mismo para ambos transistores
  - ∗ $I_{DSN}=-I_{DSP}$
  - ∗ Su potencial umbral de puerta $V_M=V_{DD}/2=V_{IN}=V_{OUT.}$ Esta es un ligadura para determinar la relación βn/βp. En realidad lo que estamos diciendo es que se quiere que los tiempos de subida y de bajada sean iguales.
  - ∗ Durante la transición están conduciendo ambos transistores lo que provoca un pulso de corriente de corte desde Vdd a gnd
  - ∗ La entrada es de tipo escalón



En esta figura se tienen determinadas las cinco regiones de trabajo del inversor que son:

A .-N cortado y P lineal

B .-N saturación y P lineal

C .-N Saturación y P saturación

D .-N lineal y P saturación

E .-N lineal y P cortado

Estas regiones son fáciles de determinar con solo estudiar cuales son las fronteras de trabajo de cada transistor . A continuación buscamos el valor del V_{OUT} en cada una de las regiones. Para encontrar el V_{OUT} en un instante determinado

1. Lo primero es comprobar si cumple las condiciones de conducción.
2. En caso que conduzcan $|I_N|=|I_P|$ para hallar Vout
3. En caso que no cumpla las condiciones de conducción se estudia el potencial que debe alcanzar Vout para que el transistor deje de conducir.

ν **REGIÓN A**

NMOS cortado y PMOS lineal, y se debe cumplir $I_N = - I_P$ como $I_N=0 \rightarrow I_P=0 \rightarrow V_D=V_S$, Luego $\mathbf{V_{OUT} = V_{DD}}$. Lo msmo se obtiene despejando de la ecuación de la instensiada lineal de PMOS

¡¡¡ Atención !!! un transistor puede tener sus potenciales polarizados para conducir pero no tener ninguna intensidad a su través.

ν **REGIÓN B. NMOS en saturación y PMOS en lineal.**

Para calcular el potencial igualamos las Intensidades

* $I_{DSN}=-\beta_N(V_{GS}-V_{TN})^2/2=-\beta_N(V_{IN}-V_{TN})^2/2$ (1)
* $I_{DSP}=\beta_P[(V_{GSP}-V_{TP})V_{DS}-V_{DS}^2/2]$

Modificamos esta última expresión para que quede $V_{GS}=F(V_{IN})$ y $V_{DS}=F(V_{OUT})$

* $V_{GSP}=V_{IN}-V_{DD}=-|V_{GSP}|$
* $V_{DSP}=V_{OUT}-V_{DD}=-|V_{DSP}|$
* $V_{TP}=-|V_{TP}|$

luego la intensidad queda:

* $I_{DSP}=\beta_P[(V_{IN}-V_{DD}-V_{TP})(V_{OUT}-V_{DD})-(V_{OUT}-V_{DD})^2/2$ (2)

Igualando (1) y (2) y despejando Vout

$$V_{OUT}=(V_{IN}-V_{TP})+[(V_{IN}-V_{TP})^2-2(V_{IN}-V_{DD}/2-V_{TP})V_{DD}-\beta_N/\beta_P(V_{IN}-V_{TN})^2]^{1/2}$$

ν **REGIÓN C. NMOS en saturación y PMOS en saturación**

Vamos a ver como encontrar la relación entre las dos β's para que el potencial umbral de puerta V_M cumpla unas condiciones determinadas.

$$Idsp = \frac{\beta p}{2}(Vgs - Vtp)^2 = \frac{\beta p}{2}(Vin - Vdd - Vtp)^2$$

$$Idsn = \frac{\beta n}{2}(Vin - Vtn)^2$$

Igualando las intensidades y despejando V_{IN} (darse cuenta que V_{OUT} no aparece)

$$Vin = \frac{Vdd + Vtp + Vtn\sqrt{\frac{\beta n}{\beta p}}}{1+\sqrt{\frac{\beta n}{\beta p}}}$$

Con esta expresión podemos elegir el valor que deben tener βn y βp para que el potencial umbral de interruptor tenga un valor u otro. (recordar que en C $V_M=V_{IN}=V_{OUT}$)

Por ejemplo: suponiendo que $|V_{TP}|=|V_{TN}|$ ¿que valor tendríaβn/βp para que $V_{IN}=V_{DD}/2$ sustituyendo en la expresión queda:

$$\frac{Vdd}{2}\left(1+\sqrt{\frac{\beta n}{\beta p}}\right)=Vdd+Vtp+Vtn\sqrt{\frac{\beta n}{\beta p}}$$

y esta expresión solo se cumple cuando βn=βp

ν **REGIÓN D. PMOS saturación y NMOS lineal**

Para calcular el potencial de salida se deben igualar las intensidades:

$$Idsp=\frac{-\beta p}{2}(Vin-Vdd-Vtp)^2$$

$$Idsn=\beta n\left[(Vin-Vtn)Vout-\frac{Vout^2}{2}\right]$$

$$V_{OUT}=(V_{IN}-V_{OUT})-[(V_{IN}-V_{TN})^2-\beta_P/\beta_N(V_{IN}-V_{DD}-V_{TP})^2]^{1/2}$$

ν **Región E. PMOS está cortado y NMOS lineal**

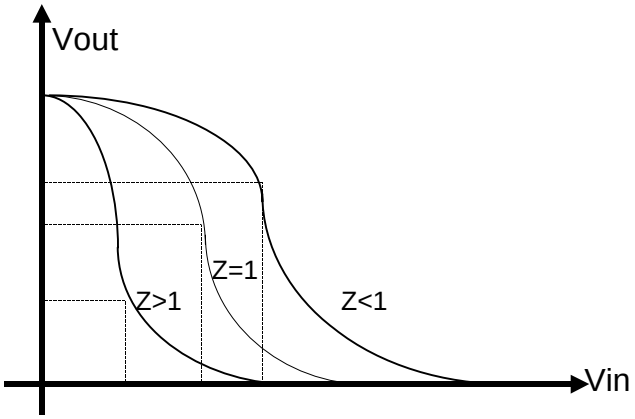
$V_{OUT}=0$

5.2.2 INFLUENCIA DE LA RELACIÓN βN/βP EN LA CARACTERÍSTICA DE ENTRADA SALIDA

$$\beta=\frac{\mu\varepsilon_{ox}}{t_{ox}}\left(\frac{W}{L}\right)$$

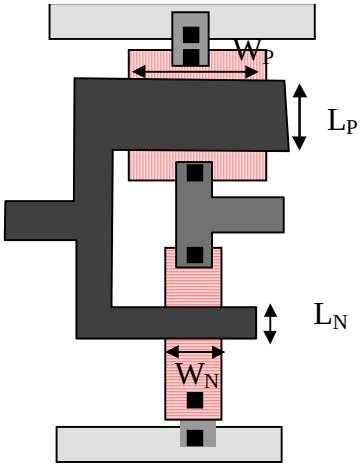
Según crece Z=βn/βp se desplaza a la izquierda el potencial umbral de puerta V_M.

- El salto lógico permanece prácticamente constante
- Z fija una relación geométrica entre los canales de los recursos.



Para el caso en que se desee que $V_M=V_{DD}/2$ ya se vio en un apartado anterior que $\beta n/\beta p=1$ luego $W_P/L_P=2.5W_N/L_N$

En las lógicas en que la relación $\beta n/\beta p$ no tenga importancia se deben coger los tamaños mínimos de los transistores



5.2.3 INVERSORES MOS DE CARGA ESTATICA

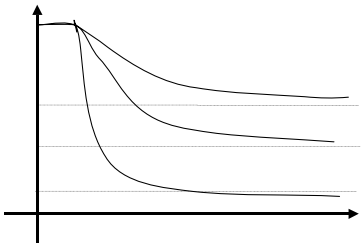
A parte de los inversores CMOS existen otras formas de implementar inversores MOS. En la carga dinámica el dispositivo conduce - no conduce. En la carga estática el dispositivo de carga conduce siempre

5.2.3.1 Inversor NMOS Genérico

Utiliza una resistencia o una fuente de intensidad como carga

La diferencia entre ambos

- Cuando la carga se hace a través de una resistencia la intensidad varia en función del potencial de salida Vout ya que $I=(V_{DD}-V_{OUT})/R$
- Cuando la intensidad la proporciona un generador ésta es independiente del V_{OUT}
- Vamos a ver la gráfica el comportamiento del inversor de carga resistiva



Cuando la resistencia decrece:

- Mayor intensidad → mayor consumo
- Menor salto lógico y mayor zona de indeterminación → peores márgenes de ruido
- Aumento de la velocidad de carga (pull-up)
- Peor cero lógico

Conclusión: La elección del valor de la resistencia de carga debe ser un compromiso entre:

⇒ La tolerancia al ruido (V_{OL})

⇒ El consumo de potencia (I)

⇒ La velocidad de carga Pull-up

Aunque la resistencia de carga (o la fuente de intensidad) se pueden fabricar utilizando polisilicio no dopado de alta resistencia, o mediante algún circuito complejo, lo habitual es que se haga mediante transistores:

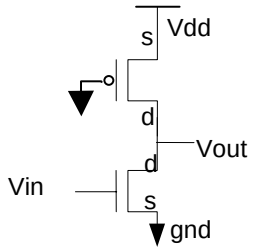
- **Inversor de carga saturada.-** si el transistor se usa en saturación es equivalente a usar una fuente de intensidad cte, (esta es la mejor opción como $I_S<I_L$ se consiguen mejores 0's)
- **Inversores de carga no saturada**.- la carga a través de un transistor que trabaja en la región lineal. Equivalente a usar una resistencia.

Las razones de usar carga estática es reducir el número de transistores ya que la carga se hace a través de un solo transistor esto da lugar menos área y a una mayor velocidad esto último debido la reducción de la capacidad total del circuito y a la mayor facilidad en cargar el 1 lógico

- **Inconveniente** Aunque disminuye el consumo dinámico de potencia porque tiene menos capacidades que cargar y descargar, aumenta el consumo estático

  Peores márgenes de ruido

5.2.4 EL INVERSOR PSEUDONMOS



Se caracteriza por utilizar un transistor P, permanentemente conectado a tierra, para realizar la carga. El transistor N, controlado por Vin, realiza el pull down. A este transistor se le suele llamar driver. Se le llama así porque su comportamiento es parecido al del transistor NMOS que realiza la carga a través de un NMOS de deplexión.

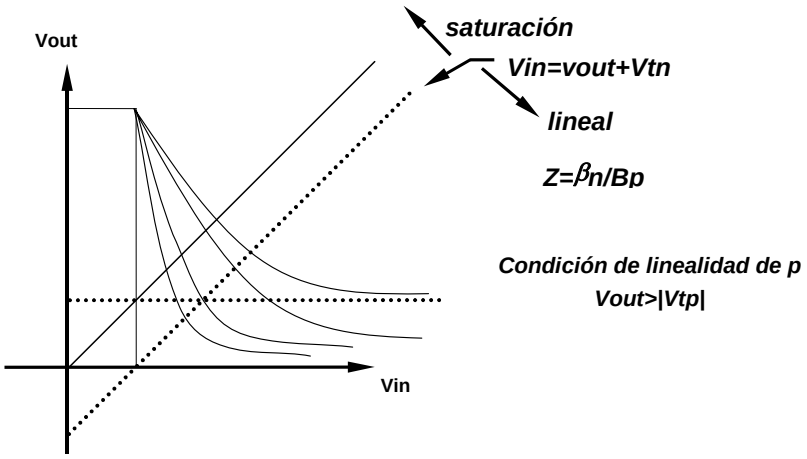
A esta lógica se la llama proporcional porque hay que buscar una proporción entre el transistor de load (pmos) y el driver (nmos) para que funcione correctamente. En la lógica proporcional es necesario un estudio cuidadoso del tamaño de los canales para que las puertas funcionen correctamente. En este tipo de puertas hay un mayor consumo de potencia que en la lógica CMOS complementaria. Se suele usar en circuitos de alta velocidad debido a que la carga del 1 es muy rápida y tiene menos capacidades.

ν **REGIONES DE TRABAJO**

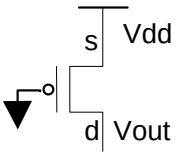
Primero vamos a ver cuales son las regiones de trabajo de cada uno de los transistores. El dispositivo N pasa por tres regiones:

- Corte $V_{IN}<V_{TN}$
- Saturación $V_{IN}-V_{TN}<V_{OUT}$
- Lineal $V_{IN}-V_{TN}>V_{OUT}$

La frontera entre la zona lineal y la saturación es la línea punteada de la figura y cumple $V_{IN}=V_{OUT}+V_{TN}$



El dispositivo P puede comportarse de dos maneras diferentes según sea la relación Z=βn/βp Permanecer siempre en lineal o bien empezar en lineal y acabar en saturación. Vamos a estudiar el motivo de que ocurra esto:



La condición de linealidad para el transistor p es $V_{GS}-V_{TP}<V_{DS}$ sustituyendo lo svalores se obtiene:

$$-V_{DD} + |V_{TP}| < V_{OUT} - V_{DD}$$

$$V_{OUT} > |V_{TP}|$$

Luego la condición de linealidad es función del Vout, pero Vout es función de βn/βp

- Cuanto menor sea Z=βn/βp mas elevado será el V_{OUT}→ P permanece lineal
- Cuanto mayor sea βn/βp menor será V_{OUT} → P pasará a saturación

Visto lo anterior las regiones de trabajo son diferentes según sea la relación βn/βp. ¿Que le interesa al diseñador? Que el comportamiento sea lo más parecido a un inversor CMOS completo es decir $V_M=V_{DD}/2$ y que el estado final del PMOS sea saturación puesto que esto permite alcanzar más rápidamente el 0 (o alcanzar un 0 mejor, lo que reduce la influencia de los ruidos al aumentar el salto lógico). Añadiendo además la condición $|V_{TP}|=|V_{TN}|$ vamos a estudiar los tres puntos que determinan el comportamiento del inversor

1. $V_{OUT}=F(V_{IN}=0)$
2. $V_{OUT}=F(V_{IN}=V_{OUT})$
3. $V_{OUT}=F(V_{IN}=V_{DD})$

ν SI $V_{IN}=0$ esta N está cortado. Vamos a demostrar que P se encuentra en la zona lineal cuya condicion es $V_{GS}-V_{TP}<V_{DS}$

$$V_{GS}=-V_{DD}$$

$$V_{DS}=0$$

$$V_{GS}-V_T<V_{DS} \rightarrow -V_{DD}+V_{TP}<0$$

Luego P se encuentra en la zona lineal. Como el dispositivo P esta en condiciones de conducir y en N está cortado I=0 el V_{OUT} =5V.

■ $Vm=V_{IN}=V_{OUT}$,

El transistor N esta en saturación demostración: la condición de saturación es $V_{IN}-V_{TN}<V_{OUT}$ como $V_{IN}>V_{TN}$ está conduciendo, como $V_{IN} = V_{OUT}$ se cumple la condición

El transistor P cumple la condición de linealidad $V_{OUT} > |V_{TP}|$. Demostración por reducción al absurdo. Suponiendo que $V_{OUT}<|V_{TP}|$ Como $V_{OUT}=V_{IN}$ y $|V_{TP}|=V_{TN}$ → $V_{IN}<V_{TN}$ y esto es imposible porque en este caso N estaría cortado. Luego P se

encuentra en la zona lineal. Vamos a ver a continuación cual es el valor de $V_{M.}$ Para encontrarlo igualamos las intensidades

$I_{DSN}=\beta_N(V_{IN}-V_{TN})^2$

$I_{DSP}=\beta_P[(-V_{DD}-V_{TP})(V_{OUT}-V_{DD})-(V_{OU}T-V_{DD})^2/2]$

despejando Vout se obtiene

$V_{OUT}=-V_{TP}+[(V_{DD}+V_{TP})^2-C]^{1/2}$ siendo $C=K(V_{IN}-V_{TN})^2$ y $K=\beta n/\beta p$ [1]

Como $V_{IN}=V_{OUT}$ se tiene que $V_M=V_T+(V_{DD}-V_T)[\beta_P/(\beta_P+\beta_N)]^{1/2}$

La expresión se puede reordenar de la siguiente manera:

$$\frac{\beta n}{\beta p}=\frac{(Vdd+Vtp)^2-(Vout+Vtp)^2}{(Vin-Vtn)^2}$$

Dando valores $V_{IN}=V_{OUT}=V_{DD}/2$, con $|V_{TP}|=|V_{TN}|=0.2V_{DD}$ y se obtiene $\beta n/\beta p=6$

ν $V_{IN}=V_{DD}$

N esta en lineal y P en saturación esto lo fijamos nosotros para que su comportamiento se aproxime al ideal que buscamos

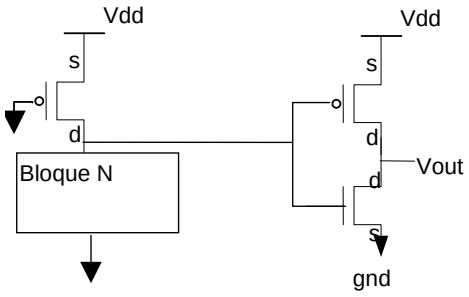
$$In=\frac{\beta n}{2}\left[(Vin-Vtn)Vout-\frac{Vout^2}{2}\right]$$

$$Ip=\frac{\beta p}{2}(-Vdd-Vtp)^2$$

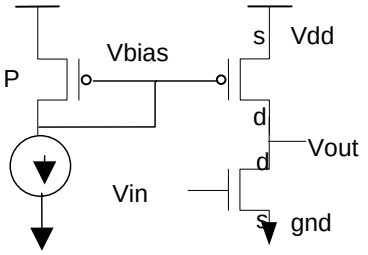
Igualando y despejando Vout (suponemos $|V_{TP}|=|V_{TN}|$)

$$Vout=(Vdd-Vt)\left(1-\sqrt{1-\frac{\beta n}{\beta p}}\right)$$

De lo anterior se puede deducir lo mucho que influye Bn/Bp en los márgenes de ruido y en la rapidez de respuesta. Una de las características que se debe tener en cuenta es la diferencia de márgenes de ruido de alta y de baja. . por ejemplo para una relación de ganancias $\beta n/\beta p=8$→ NML=-8.6 y NMH=2.6. Esta descompensación entre los márgenes de ruido se puede mejorar colocando un inversor CMOS completo a continuación del pseudonMOS



5.2.5 INVERSOR PSEUDONMOS DE CARGA SATURADA



La principal ventaja de que la carga sea saturada es que se alcanza el cero más rápidamente (se obtiene un cero de mayor dureza) debido a que $I_S<I_L$. El dispositivo P tiene en serie con él un generador de I constante lo que hace que el V_{BIAS} = Vconstante. Visto el sentido de la I del generador V_{BIAS}>Gnd=0 y $V_{DD}>V_{BIAS}$ el potencial de bias se elige para que el transistor de carga del inversor funcione exclusivamente en la región de saturación, es decir:

$$V_{GS}-V_{TP}=V_{BIAS}-V_{DD}-V_{TP}>V_{OUT}-V_{DD}$$

$$V_{BIAS}-V_{TP}>V_{OUT}$$

En el limite superior cuando Vout = 5v =Vdd

$$V_{BIAS}-V_{TP}>V_{OUT}\rightarrow V_{BIAS}-V_{TP}>V_{DD}$$

$$\rightarrow V_{BIAS}-(-|V_{TP}|)>V_{DD}\rightarrow V_{BIAS}+|V_{TP}|>V_{DD}$$

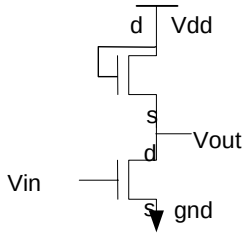
¿Es esto posible sabiendo que $V_{BIAS}<V_{DD}$?→ Se debe diseñar un generador de intensidad constate que haga cumplir a Vbias esta condición. Por ejemplo suponiendo $|V_{TP}|$=1v y un V_{BIAS} = 4,5<V_{DD} se cumple la condición 4,5+1>5

En el resto de los casos la Vout va disminuyendo luego la condición de saturación se cumple siempre. Como en saturación I=f(V_{GS},V_T) y en este caso V_{GS}=Cte → es equivalente a realizar la carga a través de un generador de I constante. Con este inversor se mejora bastante el comportamiento respecto al anterior. La variación βn/βp prácticamente no se refleja en las gráficas.

5.2.6 INVERSOR CON CARGA NMOS DE ENRIQUECIMEINTO

Los inversores que estudiamos a continuación se han utilizado en el pasado pero es poco probable que se utilicen en la actualidad. Se pueden considerar de carga dinámica puesto que el transistor de carga conduce y se corta y siempre conduce en saturación.

La diferencia con los vistos hasta ahora es que el load y el driver no se cortan alternativamente, sino que se cortan los dos al tiempo cuando la entrada es un 0 y por la tanto la salida es un 1 esto tiene como principal efecto malos 1 y 0



Se utilizó en la tecnología NMOS antes que aparecieran los transistores de deplexión. El transistor de carga se encuentra siempre en saturación. Demo:

$V_{GS}=V_{DD}-V_{OUT}$

$V_{DS} = V_{DD}-V_{OUT}$

luego $V_{GS}=V_{DS}$→ $V_{GS}-V_T<V_{DS}$ expresión que siempre se cumple.

El driver siempre pasa por las tres regiones:

- Corte $V_{IN}<V_{TN}$
- Saturación $V_{IN}-V_{TN}<V_{OUT}$
- Lineal $V_{IN}-V_{TN}>V_{OUT}$

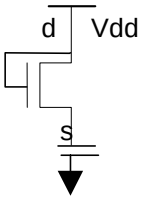
Hay que darse cuenta que debido al efecto body $V_{TL}>V_{TD.}$ El efecto aparece por que hay dos transistores en serie en la el mismo sustrato.

$V_{GS}>V_T$→ $V_G-V_S>V_T$→$V_G>V_T+V_S$ (siendo Vs=0)$V_G>V_T$

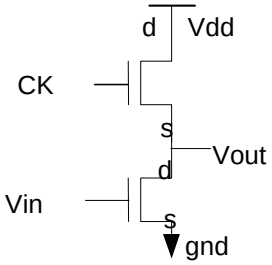
Cuando $V_S\neq 0$ →$V'_T=V_T+V_S$

Vamos a estudiar a continuación los puntos característicos de la $V_{OUT}=F(V_{IN})$

ν V_{IN}=0.Suponemos el circuito equivalente:



De manera que suponiendo que V_{IN} pase de 1 a 0 de manera instantáneas la capacidad de salida está descargada, es decir Vout = 0 . En ese instante $V_{GS}=V_{DD}>V_{TN}$ luego se cumplen las condiciones de conducción luego Vout empieza a incrementarse. Según se incrementa el V_{OUT} el $V_{GS}=V_{DD}-V_{OUT}$ empieza a disminuir. ¿Cual es el Vout final. El transistor de carga deja de conducir cuando $V_{DD}-V_{OUT} = V_{TN}$ luego $V_{OUT}=V_{DD}-V_{TN.}$ Este inversor se podría modificar de la siguiente manera:

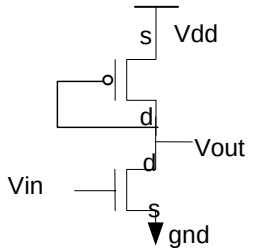


De esta manera el transistor de carga sólo conduce en algunos instantes . Los otros dos puntos característicos serán

b)$V_{IN}=V_{OUT}=V_M$ que hay que fijar para que V_M tome el valor que se desee

c)$V_{IN}=5V$

5.2.7 INVERSOR CON CARGA PMOS DE ENRIQUECIMIENTO



El dispositivo P conduce siempre en saturación. Demostración:

$V_{GS}-V_{TP}>V_{DSP}$

$V_{GSP}=V_{OUT}-V_{DD}$

$V_{DS}=V_{OUT}-V_{DD}$

luego $-V_{TP}>0 \rightarrow |V_{TP}|>0$

El dispositivo N conduce en diferentes regiones dependiendo de la relación con Vin y Vtn

* Corte $V_{IN}<V_{TN}$
* Saturación $V_{IN}-V_{TN}<V_{OUT}$
* Lineal $V_{IN}-V_{TN}>V_{OUT}$

ν **Vamos a estudiar los tres puntos que definen la curva característica**

$V_{OUT}=F(V_{IN}=0)$

$V_{OUT}=F(V_{IN}=V_{OUT})$

$V_{OUT}=F(V_{IN}=V_{DD})$

- Para $V_{IN}=0$ el transistor N esta cortado. Para su estudio suponemos que V_{IN} pasa de 1 a 0 de manera instantánea y que la capacidad está cargada a 0. En esta situación si $V_{IN}=0$ de manera constante el V_{OUT} tiende a 1. Cuando $V_{OUT}=V_{DD}-|V_{TP}|$ el transistor P se corta que es a lo que queríamos llegar. Vamos a demostrarlo:

  $V_{GS}=V_G-V_S=V_{OUT}-V_S=V_{DD}-|V_{TP}|-V_{DD} \rightarrow V_{GS}=V_{TP}$ que es la condición de corte

  Luego en esta fase N y P cortados Y el potencial Vout máximo es $V_{DD}-|V_{TP}|$

- Para $V_M=V_{IN}=V_{OUT}$ los dos transistores están en saturación
  * $V_{IN}-V_{TN}<V_{OUT}$
  * Para encontrara V_M se igualan las intensidades y se despeja ,$I_{NSAT}=I_{PSAT}$

  $I_{DSN}=\beta_N/2(V_{IN}-V_{TN})^2$

$I_{DSP}=\beta/2(V_{OUT}-V_{DD}-V_{TP})^2$

Despejando $V_{OUT}=V_{DD}+V_{TP}+K^{1/2}(V_{IN}-V_{TN})$ donde $K=\beta n/\beta p$

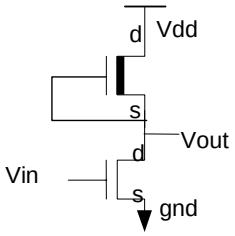
- Para $V_{IN}=V_{DD,}$ N esta en lineal porque $V_{IN}-V_{TN}>V_{OUT}$

  In lineal, Ip saturación

  ∗ $V_{OUT}=F(\beta n/\beta p)$

  ∗ Márgenes de ruido de baja muy malos

5.2.8 INVERSOR DE CARGA NMOS DE EMPOBRECIMIENTO



- El transistor de carga (Load)

  $V_{GS}=V_G-V_S=V_{OUT}-V_{OUT}=0$

  $V_{DS}=V_{DD}-V_{OUT}$

  $V_{GS}>V_{TN}$ porque recordar que en los transistores NMOS de empobrecimiento el potencial umbral es negativo

  ∗ Este transistor conduce siempre

  ⇒ Cuando $|V_{TN}|<V_{DD}-V_{OUT}$ en saturación

  ⇒ Cuando $|V_{TN}|>V_{DD}-V_{OUT}$ en lineal

- El driver Funciona en tres regiones diferentes

  ∗ Corte $V_{IN}<V_T$

  ∗ Saturación $V_{IN}-V_T<V_{OUT}$

  ∗ Lineal $V_{IN}-V_T>V_{OUT}$

ν **Estudio de los tres puntos característicos:**

- Para $V_{IN}=0$ Nd cortado y Nl conduce: vamos a comprobarlo . Para hallar el Vout

  $V_{GS}=V_{OUT}-V_{OUT}=0$

  $V_{GS}>V_{TN}=-|V_{TN}|$

  $I_D=0=I_L \rightarrow V_D=V_S=V_{OUT}=V_{DD-}$

- Para $V_{IN}=V_{OUT}$

  ∗ Nd saturación $V_{IN}-V_{TN}<V_{OUT}$

  ∗ Nl depende del valor V_{OUT} que a su vez depende de $\beta d/\beta l$

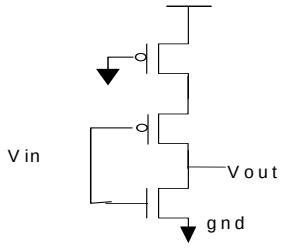
  Fijando $V_{IN}=V_{OUT}=V_{DD}/2$

  $|V_{TN}|<V_{DD}/2$

  ⇒ luego load en saturación

- $V_{IN}=V_{DD}$
  - ∗ Nd lineal
  - ∗ Nl saturación

5.2.9 INVERSOR DE CONEXIÓN A LOGICA TTL



Estamos modificando los umbrales de entrada para que las salidas se ajusten a la lógica TTL. El primer transistor P alimenta un inversor CMOS convencional con una carga reducida con ello modificamos el potencial umbral de entrada para que se adapte a una salida TTL

- TTL V_{IL}=0.8 V_{IH}=2V
- CMOS VIK=2,3 VIH=3,3

Para ello lo que hago es que la curva característica del CMOS se aproxime a la de TTL, es decir que la caída sea más rápida, o lo que es lo mismo que cueste más la subida. ¿Como consigo empeorar la subida? desplazando el V_{TP} a la izquierda, es decir incrementando el valor absoluto de $V_{TP.}$ ¿Y como consigo esto? Utilizando para nuestro provecho el efecto sustrato

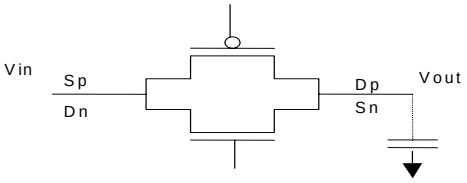
Potencial $V_{TP2}>V_{TO1}$

→le cuesta más empezar a conducir

→ le cuesta más subir

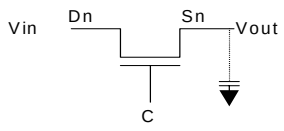
→ la curva se desplaza a la IZQUIERDA.

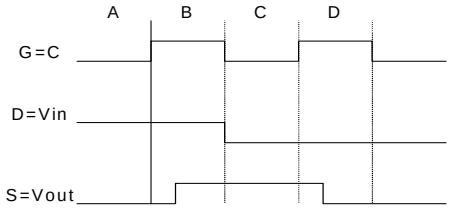
5.3 LA PUERTA DE TRANSMISIÓN



Para estudiar la puerta tendremos que en cuenta la carga y descarga de una capacidad a través de la puerta de transmisión. Se hace un estudio de los dispositivos por separado y posteriormente unidos.

5.3.1 TRANSISTOR DE PASO NMOS





Suponer la capacidad inicialmente descargada $V_{OUT}=V_{GND}$

A) C=0→V_{GS}=0→I_{DS}=0 V_{OUT} =V_{GND}

B) C=1 y Vin =1

- El transistor de paso conduce
- Se carga la capacidad
- Como $V_{IN}>V_{OUT}$ la I fluye por el transistor
- la I fluye de derecha a izquierda
- Cuando la salida ≈V_{DD}-V_T el dispositivo se corta

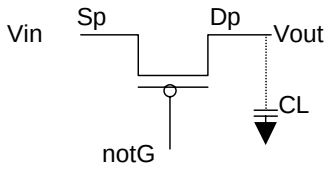
C) C=0

- La capacidad permanece cargada y aislada→ $V_{OUT}=V_{DD}-V_{TN}(V_{DD})$
- Siendo $V_{TN}(V_{DD})$ el potencial umbral afectado por el efecto cuerpo→ la señal se degrada al pasar por la puerta

D) C=1 y Vin=0

- Inicialmente V_{OUT} =$V_{DD}-V_{TN}(V_{DD})$
- El transistor conduce
- Se descarga la capacidad $V_{OUT}=V_{GND}$
- La intensidad va de izquierda a derecha
- El 0 no se degrada

5.3.2 TRANSISTOR DE PASO PMOS



- NotG=0→$V_{IN}=V_{DD}$, $V_{OUT}=V_{GND}$ →la capacidad está descargada
- NotG=0 y $V_{IN}=V_{DD}$ I fluye y se carga la capacidad a V_{DD}
- NotG=1 la capacidad permanece cargada y asilada
- NotG=0 y Vin=0 inicialmente $V_{OUT}=V_{DD}$.

  El transistor conduce

La capacidad se descarga

$V_{OUT} = V_{TP}(gnd) \rightarrow$ se degrada la transmisión del 0 lógico.

Recordar que los NMOS conducen bien el 0 y los P conducen bien el 1, luego combinando los dos transistores N y P en paralelo se construye una puerta de paso que transmite bien el 0 y el 1

- G=0 N off, P off

  V_{IN}=gnd$\rightarrow V_{OUT}$ =Z

  Capacidades de salda aislada$\rightarrow$ se conserva el valor anterior

- G=1 N on, P on

  V_{IN} =gnd$\rightarrow V_{OUT}$=gnd

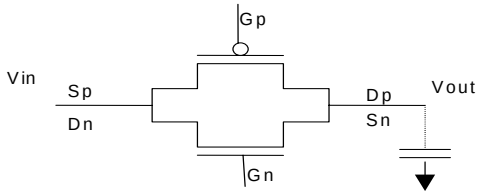
  $V_{IN}=V_{DD} \rightarrow V_{OUT}=V_{DD}$

Las puertas de transmisión son importantes y aparecen en muchos diseños MOS:

- Como multiplexores
- Como elementos de carga (en biestables)
- Interruptores analógicos
- Implementando estructuras lógicas.

Podemos aproximar el comportamiento de las puertas lógicas al de una resistencia que conecta una entrada y una salida estando la resistencia controlada por un potencial.

Vamos a estudiar que ocurre en una puerta de paso cuando la señal de entrada está constante y la señal de control varía



$V_{IN}=V_{DD}$=cte

Gn es un pulso 0$\rightarrow$1

Gp es un pulso negativo 1$\rightarrow$0

Estado de los transistores en el instante inicial

Condiciones iniciales: $V_{IN}=V_{DD}$ y C_L descargada$\rightarrow V_{OUT}$=gdn

En general

$V_{GSN}=V_{GN}-V_{OUT}$

$V_{DSN}=V_{DD}-V_{OUT}$

$V_{GSP}=V_{GP}-V_{IN}$

$V_{DSP}=V_{OUT}-V_{DD}$

Particularizando para el problema

V_{GSN}=0-0=0

V_{GSP}=5-5=0

Luego ambos dispositivos están cortados

ν **Estado de los transistores en el transitorio**

Suponemos V_{GN} es un pulso positivo 0→1 instantáneo

Suponemos que V_{GP} es un pulso negativo 1→0 instantáneo

Vamos a ver que ocurre en el **transistor N**

$V_{GS}=V_{DD}-V_{OUT}$

$V_{DS}=V_{DD}-V_{OUT}$

Luego está en saturación porque se cumple

$V_{GS}-V_T<V_{DS}$

y seguirá en saturación mientras se cumpla

$V_{GS}>V_{TN}$

es decir mientras se cumpla

$V_{DD}-V_{TN}>V_{OUT}$

- **Transistor P**

$V_{GSP}=0-V_{DD}$

$V_{DS}=V_{OUT}-V_{DD}$ de tal manera que si se incrementa el potencial de salida se decrementa el V_{DS}

El transistor P esta en saturación mientras cumpla la condición

$V_{GS}-V_T>V_{DS}$

$-V_{DD}-V_T>V_{OUT}-V_{DD}$

$|V_{TP}|>V_{OUT}$-> si V_{OUT} crece P pasa a la zona lineal

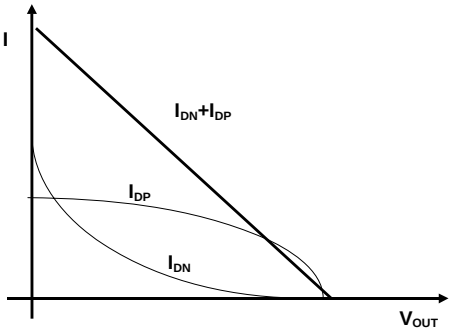
Aparecen tres regiones

N saturado P saturado $V_{OUT}<|V_{TP}|$

N saturado P lineal $|V_{TP}|<V_{OUT}>V_{DD}-V_{TN}$

N cortado P lineal $V_{DD}-V_{TN}<V_{OUT}$

- Si reflejamos en una gráfica I_{DP}, I_{DN}, $I_{DP}+I_{ND}$ en función del potencial de salida



Se puede observar que el comportamiento de la I de la puerta es lineal respecto al Vout

Señal de entrada variable, señal de control constante: Suponemos que la señal de salida sigue de cerca la de entrada→V_{IN}-V_{OUT}=-0'1

- Suponiendo que Vin es pulso positivo puro vamos a ver las regiones de trabajo de N

  $V_{GS}=V_G-V_S= V_{DD}-V_{OUT}$

  $V_{DS}=V_{IN}-V_{OUT}=0,1=Cte$

La condición de conducción $V_{GS}>V_{TN}$

$V_{DD}-V_{OUT}>V_{TN}$

$V_{DD}-V_{TN}>V_{OUT}$

$5-1>V_{OUT}$

Mientras 4 >Vout el dispositivo N conduce. ¿En que región?

⇒ $V_{GS}-V_{TN}>V_{DS}$ en lineal

⇒ $V_{DD}-V_{OUT}-V_{TN}>0,1$

$5-1-0,1>V_{OUT}$

3,9>Vout mientras $V_{OUT}<3,9$ conduce en lineal

⇒ Corte $V>V_{OUT}$

⇒ Saturación $3,9<V_{OUT}>4$

⇒ Lineal $V_{OUT}<3,9$

- Regiones de trabajo de P

  $V_{GS}=-V_{IN}$

  $V_{DS}=0,1$

  - Conducción $V_{GS}<V_{TP}$

    $-V_{IN}<-|V_{TP}|$

    $V_{IN}>|V_{TP}|$→ conduce p

  - ¿En que región conduce?

    $V_{GS}-V_{TP}<V_{DS}$ condición de lineal

    $-V_{IN}+|V_{TP}|<-0,1$

    $V_{IN}-|V_{TP}|>0,1$

    $V_{IN}>|V_{TP}|+0,1$

    ⇒ Corte $V_{IN}<|V_{TP}|$

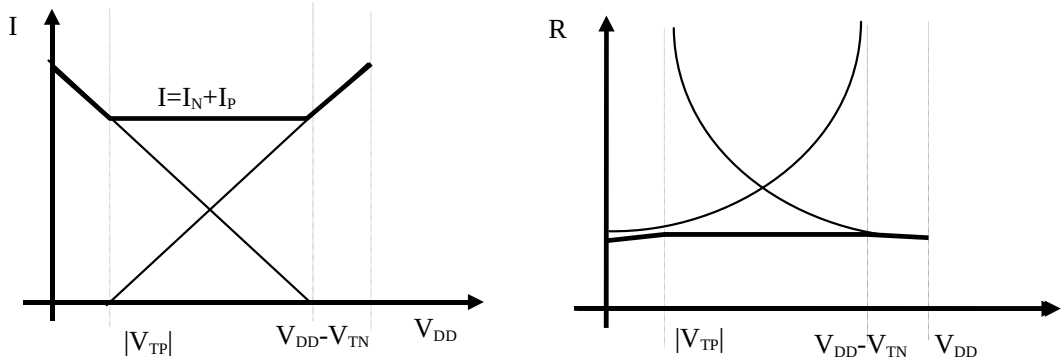
    ⇒ Saturación $|V_P|<V_{IN}<|V_{TP}|+0,1$

    ⇒ Lineal $V_{IN}>|V_{TP}|+0,1$

Se puede observar, tanto para el NMOS como para el PMOS, que se pasa muy brevemente por los estados de saturación, estos estados los despreciamos a la hora de estudiar las zonas de trabajo. Existen tres regiones de trabajo:

- N lineal P cortado $V_I<|V_{TP}|$
- N lineal P lineal $|V_{TP}|<V_{IN}<V_{DD}-V_{TN}$
- N cortado P lineal $V_{IN}>V_{DD}-V_{IN}$

En este modo de funcionamiento la puerta de paso tiene una resistencia que se puede considerar constante.

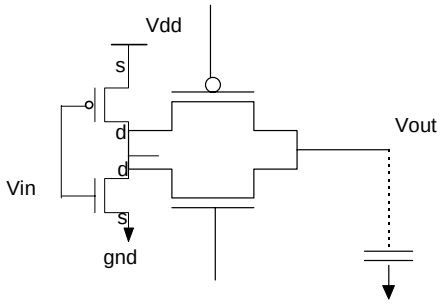


Notar que en el caso de las resistencias, la suma no equivale a resistencias en serie sino resistencias en paralelo y por lo tanto:

$$1/R_T = 1/R_1 + 1/R_2$$

5.4 EL INVERSOR TRIESTATE

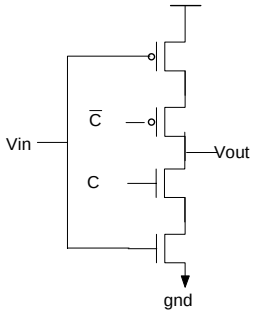
Se puede construir colocando en cascada un inversor con una puerta de transmisión



Cuando C=0 la salida se encuentra en la condición triestate luego la entrada no se comunica con la salida luego la salida conserva el ultimo valor

Cuando C=1 la puerta de transmisión transmite la señal de entrada

Existen otros diseño de este inversor, por ejemplo eliminando la conexión entre el transistor N y el P del inversor. El modo de operación es el mismo:

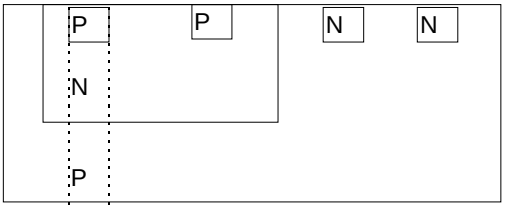


Para el mismo tamaño de transistores este inversor trabaja a la mitad de velocidad que el inversor normal, porque para llegar a la salida debe atravesar el doble de inversores. Este inversor se discute más adelante por ser fundamental para:

- Ciertas lógicas síncronas
- Cargas
- Bus
- Drivers
- Estructuras de entrada salida.

5.5 3.6 TECNOLOGÍA BICMOS

Es la formada por recursos bipolares y CMOS. Los recursos bipolares pueden ser parásitos o deliberados. Los recursos parásitos son dispositivos bipolares indeseados que pueden dar lugar a la aparición del fenómeno del latch up, es decir transistores realimentados que producen un importante aumento de la intensidad a través del circuito integrado. En estos casos en el peor de los casos se destruye el circuito y en el mejor de los casos se producen mal funcionamientos. En un dispositivo CMOS de pozo N aparecen transistores bipolares PNP:



Siendo el sustrato el colector, el pozo la base, el emisor la fuente o el drenador

ν DELIBERADOS

Los inversores CMOS tienen unas características prácticamente perfectas:

- Buenos 0 y1
- Márgenes de ruido simétricos
- Alta ganancia el transitorio
- Bajo consumo de potencia

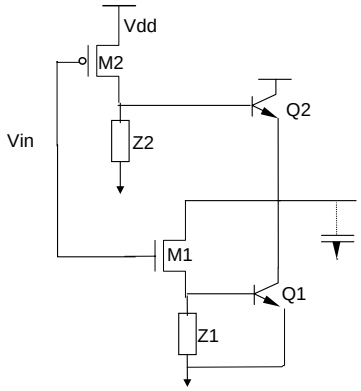
- Alta densidad de empaquetamiento

Su principal inconveniente es la velocidad sobre todo cuando tiene que cargar capacidades grandes.

Los transistores bipolares ECL se caracterizan por su gran rapidez. Para los mismos fanout y tecnologías comparables los retardos son 5 veces menores que los CMOS. Su principal inconveniente es el elevado consumo de potencia lo que hace difícil su integración VLSI. Y otro problema son los pequeños márgenes de ruido que tienen. La tecnología BicMOS combina ambos tipos de circuitos a un coste razonable.

5.5.1 INVERSOR BICMOS

ν **PUERTA GENERICA**



El transistor bipolar Q1 proporciona el 0 y el transistor bipolar Q2 proporciona el 1. Vamos a ver el funcionamiento:

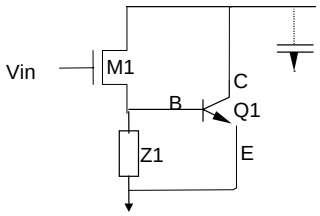
- **V_{IN}=1**

  ⇒ M_1 conduce → Q_1 conduce

  ⇒ M_2 cortado → Q_2 cortado

Si se considera la caída de potencial en el transistor M_1 despreciable se puede asegurar que V_{OUT} =V_{BE} por lo tanto cuando V_{OUT} <V_{BE}(ON) el transistor Q_1 se corta.

$$V_{OUT}=V_{BE}(on)$$



- **V_{IN}=0**

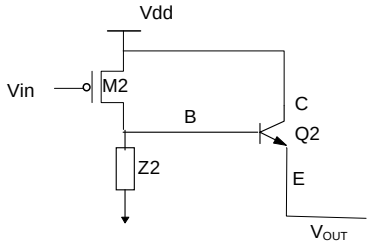
  ⇒ M_2 conduce → Q_2 conduce

  ⇒ M_1 cortado → Q_1 cortados

Suponiendo que en el transistor M_2 prácticamente no se produce caída de potencial podemos asegurar que $V_B=V_{DD}$=CTE, por otro lado el $V_E=V_{OUT}$ se incrementa, luego en el límite de conducción del transistor Q_2

$$V_{BE}=V_{DD}-V_{OUT} = V_{BE}(ON) \rightarrow$$

$$V_{OUT}=V_{DD}-V_{BE}(ON)$$



para estudios DC Q_1 y Q_2 nunca conducen simultáneamente y esto implica bajo consumo. Las resistencias Z_1 y Z_2 se utilizan para eliminar rápidamente la carga almacenada en la base del transistor bipolar cuando se quieren cortar estos. El principal efecto es:

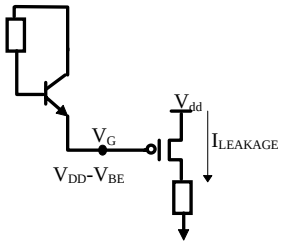
- Reducir los tiempos de transición,
- Reduce el consumo puesto que Q_1 y Q_2 permanecen menos tiempo conduciendo simultáneamente

En el transitorio hay un pequeño periodo de tiempo en que Q1 y Q2 conducen simultáneamentees decir se abre un camino entre Vdd y gnd y por lo tanto se produce un consumo potencia dinámica

La característica de transferencia. El salto lógico es menor que el suministrado por los voltajes

- para $V_{IN} = 0$ $V_{OUT} = V_{DD} - V_{BE}(on)$
- para $V_{IN}=1$ $V_{OUT} = V_{BE}(ON)$

Luego el salto lógico es $V_{DD}-2V_{BE}(on)$. Esto tiene una **desventaja**: aparecen problemas de corrientes de leakage al conectar el siguiente inversor en cascada. Vamos a suponer que alimentamos con un 1 logico es decir con un potencial $V_G= V_{DD}-V_{BE}(on)$ al siguiente inversor. Entonces el potencial Vgs del transistor PMOS de la siguiente puerta es $V_{GS}=V_{DD}-V_{BE}on-V_{DD} = -V_{BE}on<V_{TP}$, Luego el transistor PMOS conduce cuando debería estar cortado.



- **Región de trabajo de los transistores Q1 y Q2**

Para V_{IN}=0 M_2 actúa como un resistor asegurando que el potencial del colector sea siempre mayor que el potencial de la base, como las condiciones de saturación son:

$$V_{BE}>V_{BESAT}$$

$$V_{BE}>V_{BC}$$

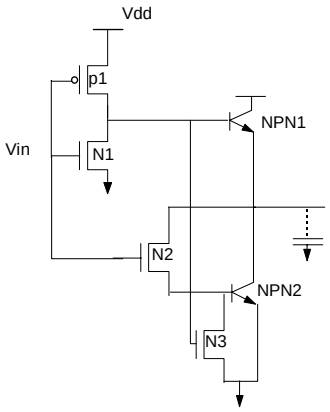
Pero como $V_{BE}<V_{BC}$ siempre , **Q2 nunca entran en saturación**

Lo mismo ocurre en el caso contrario cuando V_{IN} = 1

Esto da lugar a una ventaja **:**el principal retardo que se produce en un transistor bipolar es la carga y descarga de las capacidades de base cuando el transistor conduce en saturación. si evitamos que los transistores Q_1 y Q_2 entren en saturación el inversor trabaja más rápidamente.

ν **INVERSOR BICMOS**

Se sustituyen las resistencias del circuito genérico estudiado por transistores MOS



Introduce mejoras en la velocidad de las salidas debido a las elevadas ganancia de corriente del dispositivo NPN

- V_{IN}=0: P1 conduce

  en NPN_1 →V_B=Vdd→conduce

  en N_3→V_G=Vdd→N_3 conduce

  si N_3 conduce →NPN_2 cortado

  N_2 cortado porque Vin=0

no existe camino a tierra

$V_{OUT}=V_{DD}-V_{BE}$→se carga la capacidad de salida

- V_{IN}=1

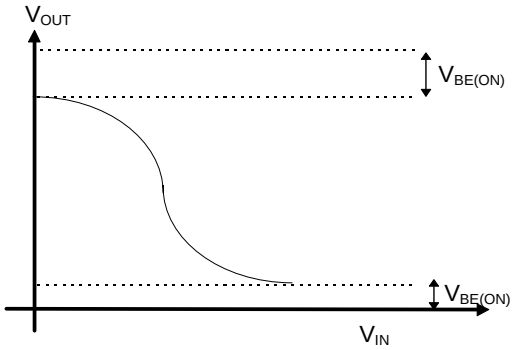
  NPN_1 cortada porque N_1 conduce

  N_2 conduce si suponemos que la capacidad esta cargada a 1→NPN_2 conduce

  N_3 cortado

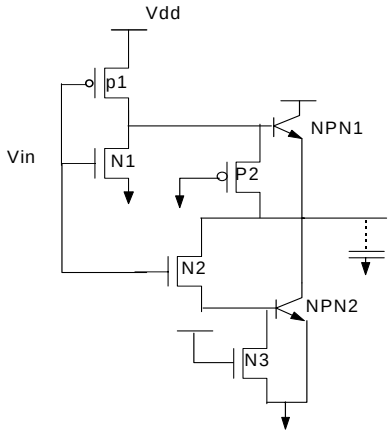
  La capacidad se vacía a través de N_2 y NPN_2 luego se vacía muy rápidamente

  $V_{OUT}=VBE_{ON}$



Como se puede ver en la figura el principal problema de este inversor es que degenera el valor de las señales

Vamos a ver un segundo inversor que regenera totalmente los valores eléctricos de las señales



Suponemos que V_{OUT}=0→ capacidad descargada

- VIN=0

  P1 conduce→alimenta NPN_1→NPN_1 proporciona rapidez para cargar la capacidad

  P2→ conduce un 1 puro→ se encarga de evitar la degeneración del la señal.

  La dimensión del P2 debe ser lo suficientemente grande como para acelerar la carga del 1máximo, pero no tan grande que se trague la intensidad de base de NPN_1

- VIN=1 NPN_1 cortado porque N_1 conduce

  N_2 conduce→suponiendo V_{OUT} =V_{DD}(capacidad cargada) NPN_2 conduce

  N_3 conduce → existe un camino a tierra→V_{OUT}=V_{GND}

- La principal ventaja de este inversor es que no se degenera ni el 1 ni el 0.
- Como los dispositivos NPN tienen mayor ganancia y mejores respuestas a la alta frecuencia que los PNP son aquellos los que se utilizan siempre en la tecnología BICMOS
- Esta tecnología bicMOS se utiliza para:

* Mejorar la velocidad de los CMOS sobretodo donde sea necesario una elevada capacidad de carga
* Drivers BUS
* Drivers IO
* Sensores amplificadores
* Chip mixtos

6 TECNOLOGIA DE PROCESOS CMOS

6.1 FABRICACION BASICA DE DISPOSITIVOS

- El silicio puro o intrínseco es un semiconductor. Su resistencia se encuentra entre la de los conductores y los aislantes
- La conductividad se puede variar añadiendo impurezas en el cristal de silicio.
- las impurezas pueden ser electrones y huecos
    - Silicio dopado con e^- :Tipo N
    - Silicio dopado con h+: Tipo P
- Fuertemente dopado se representa N+; P+
- Unión es la región en la que el silicio cambia de tipo P a tipo N.

6.1.1 FABRICACIÓN DE LA OBLEA {WAFER PROCESING}

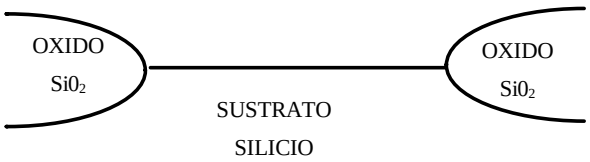
- El material puro que se usa actualmente es una oblea o disco de silicio
- Tamaño: 75-230 mm de diámetro y 1mm de grosor
- Los discos se cortan de un lingote de cristal sencillo de silicio que se obtiene a partir de silicio policristalino puro fundido
- Se añaden cantidades controladas de impurezas al polisilicio fundido para alcanzar las características eléctricas deseadas

ν **MÉTODO DE CZOCHRALSKI**

- La orientación de cristal se consigue mediante un cristal semilla que se introduce en el silicio fundido para iniciar el crecimiento del cristal
- **La fusión**
    - ∗ Se realiza en un crisol de cuarzo rodeado de un radiador de grafito
    - ∗ El grafito se calienta mediante inducción de radio frecuencia
    - ∗ La temperatura se mantiene unos grados por encima del punto de fusión (1425ºC)
    - ∗ Se trabaja en atmósferas de helio o argón (gases nobles) para que no se produzcan reacciones con el o_2
- Después de introducir la semilla se extrae muy lentamente al tiempo que se rota.
- El diámetro del lingote es función de:
    - ∗ Velocidad a la que se saca la semilla
    - ∗ Velocidad de rotación.
    - ∗ Crecimiento a razón de 30-180mm/h. Las obleas que se obtienen se pulen por una cara para evitar errores finales.

6.1.2 OXIDACIÓN

- **Misión**: funcionar como aislante
- La oxidación del silicio para obtener dióxido de silicio (SiO_2) se consigue calentando la oblea en una atmósfera rica en oxigeno.
- Existen dos aproximaciones al problema:
    - **Oxidación húmeda**:
        - La atmósfera de vapor de agua
        - Temperatura : 900-1000ºC
        - Proceso es rápido
    - **Oxidación seca**:
        - La atmósfera contiene oxigeno puro
        - Temperatura: 1200ºc
        - El proceso es mas lento
- El proceso de oxidación consume silicio.
    - Como SiO2 tiene el doble de volumen del silicio el layer crece en ambas direcciones verticales→ **pico de pájaro**



6.1.3 EPITAXIS, DEPOSICIÓN, IMPLANTACIÓN IONICA Y DIFUSIÓN

Objetivo: variar la concentración de portadores en el silicio

Como: introducir impurezas mediante:

- Epitaxis
- Deposición
- Implantación
- Difusión.

ν **EPITAXIS**

- Crecimiento de una película de cristal sencillo sobre la superficie de silicio (que también es un cristal sencillo)
- Se obtiene sometiendo la superficie a temperaturas elevadas y a una fuente de material dopante

ν **DEPOSICIÓN**

- Material dopante evaporado sobre la superficie de silicio
- Todo ello seguido de un ciclo térmico utilizando para conducir las impurezas de la fase anterior.

ν **IMPLANTACIÓN IONICA**

- Se somete la superficie de silicio al impacto de impurezas con energías elevadas
- Cuando las impurezas golpean la superficie de silicio viajan a través suya

ν **DIFUSIÓN**

- Se produce a elevadas temperaturas (800C) entre silicios con densidades de impurezas diferentes
- Las impurezas se trasladan de zonas de alta concentración a zonas de baja concentración.

- Como casi todos los procesos se producen a elevadas temperaturas, conviene que para mantener las concentraciones constantes se mantenga la temperatura tan baja como sea posible.
- La **capacidad** de construir transistores y otras estructuras de interés depende de la capacidad de controlar **donde, cuanto y que tipo** de impurezas se introducen en la superficie de silicio
- El **tipo** se consigue mediante la fuente de impurezas utilizada
  - ∗ Boro: huecos
  - ∗ Arsénico y Fósforo: electrones
- Para controlar la **cantidad** de dopantes:
  - ∗ Implantación:
    - ⇒ Energía de los portadores
    - ⇒ Tiempo de actuación
  - ∗ Deposición y difusión
    - ⇒ Temperatura
    - ⇒ Tiempo
- El **donde**
  - ∗ Se determina utilizando mascaras de materiales especiales
  - ∗ En las zonas cubiertas por la mascara no se produce la implantación iónica, o los dopantes no entran en contacto con la superficie de silicio.
  - ∗ Los materiales mas usados en las mascaras son:
    - ⇒ Fotoresistentes
    - ⇒ Polisilicio (silicio policristalino)
    - ⇒ Dióxido de silicio (SiO2)
    - ⇒ SiNi Silicon Nitride.
  - La capacidad de estos materiales de actuar como barrera es un factor muy útil en el proceso de difusión selectiva.

ν DIFUSIÓN SELECTIVA

- Colocar las mascaras con sus ventanas sobre la superficie de silicio.
- Someter las áreas expuestas a la acción dopante.
- Eliminar las mascaras que no se necesitan.

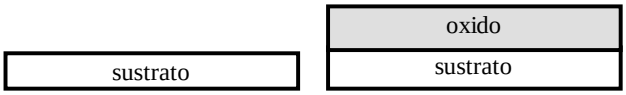
A continuación vamos a ver como se utilizan algunas de las máscaras que hemos nombrado con anterioridad.

6.1.4 FABRICACIÓN DEL AREA ACTIVA

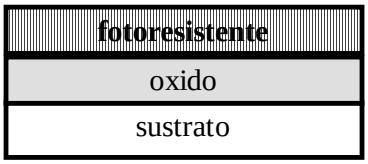
- El área activa es la región del sustrato en la que se fabrica un transistor
- en este caso la zona activa se caracteriza por quedar rodeada de óxido aislante
- En este apartado vamos a ver como generara ese óxido
- Vamos a necesitar una capa auxiliar que al final del proceso se elimina

ν **PASOS:**

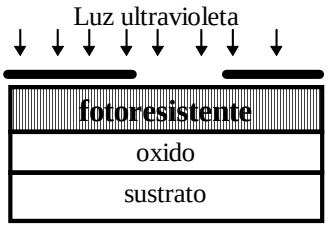
- Se cubre el silicio con una capa de oxido



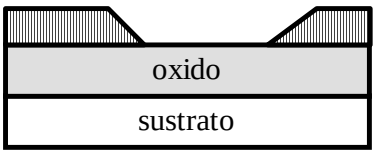
- Se cubre la capa de oxido con una capa de material fotoresistente e inerte al ácido.



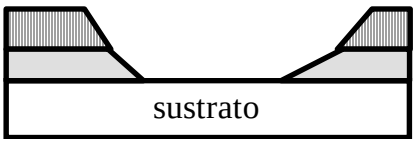
- La capa antiácido se puede polimerizar mediante luz ultravioleta.



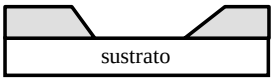
- Las áreas polimerizadas de la capa antiácido se eliminan mediante disolventes orgánicos.



- se elimina el oxido de las zonas libres de capa antiácido mediante aguafuerte.



- Se elimina el resto de la capa antiácido.



- Los efectos de la difracción en los bordes de la mascara limita las anchuras de las líneas a 0.8nm.

6.1.5 POLISILICIO

- Un silicio compuesto de mas de una forma cristalina
- Se usa para fabricar las puertas en los circuitos integrados
- Es el layer que lleva la señal de control.

ν **SELF-ALIGNED**

- Se utilizan como mascara para fabricar la fuente y el drenador (zonas de difusión
- definición muy precisa.
- Mínimo solapamiento entre la puerta y la fuente del drenador
- mejora mucho los rendimientos del circuito.
- En el caso MOS se deposita polisilicio sin dopar sobre el aislante de puerta. A continuación se dopa tanto el polisilicio como la fuente y el drenador. El polisilicio sin dopar tiene una resistencia elevada.

6.1.6 PASOS PARA UN PROCESO TIPICO DE FABRICACION DE UNA PUERTA (TRANSISTOR MOS)

- Se realiza mediante fotomasking y eliminación del oxido (aguafuerte, ácido). Ambos vistos con anterioridad.
- Partiendo de la ultima figura. Nueva capa fina de oxido, luego tenemos oxido de dos tamaños.

  * **Fino**

    ⇒ Muy controlado,

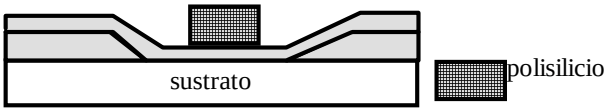
    ⇒ Define las zonas activas del transistor

    ⇒ se le llama oxido de puerta: thinox

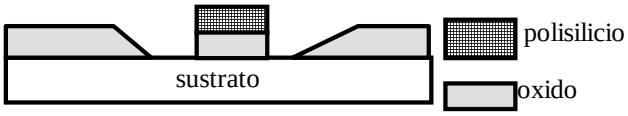
  * **Grueso**,

    ⇒ Aísla los transistores unos de los otros.

- Se deposita la capa de polisilicio, mediante el proceso de aguafuerte se da forma a las interconexiones y las puertas de los transistores.



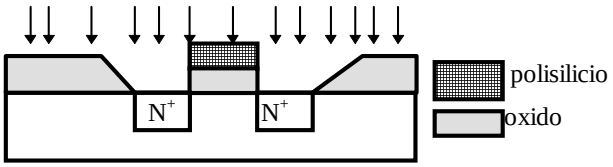
- Se elimina el oxido que cubre las futuras zonas de difusión mediante el proceso de aguafuerte.



- Se somete a la zona de fuente - puerta - drenador a un proceso de dopaje;

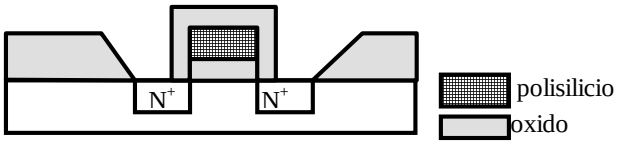
- Dopaje del polisilicio.

- Generación de las zonas de difusión.

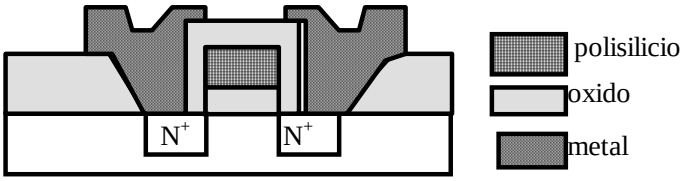


  * Proceso "Self-Aligned" no se produce solapamiento entre la fuente-drenador y la puerta.

- Se vuelve a cubrir todo con una nueva capa de oxido, mediante aguafuerte se generan los agujeros de contacto bajo los layers.

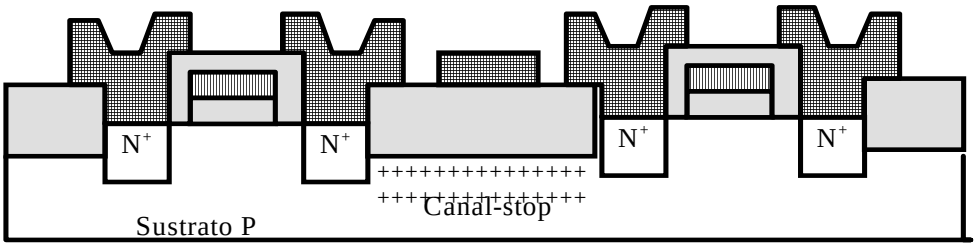


- Se añade aluminio (u otro metal) a los contactos mediante evaporacion. Por ultimo se añaden nuevas capas de oxido y nuevos contactos para permitir la utilización de mas de un metal.



6.1.7 TRANSISTORES MOS PARASITOS

- Su potencial umbral es mucho mas elevado que el del transistor regular.
- Este potencial umbral se asegura mediante
  - una capa de oxido suficientemente gruesa,
  - introduciendo mediante difusión impurezas en las zonas del sustrato donde no se necesitan transistores = "**Canal-Stop**", en este caso del tipo R.
- Estos dispositivos pueden tener alguna utilidad utilizándose como protección de otros circuitos.



6.2 TECNOLOGIA CMOS BASICA

- La más importante de los sistemas VLSI
  - ∗ Bajo consumo de potencia estática; producto
  - ∗ Retardo x potencia pequeña. Similar al de bipolares, nMOS, AG
- Existen 4 tecnologías de fabricación :

  -Pozo N

  -Pozo P

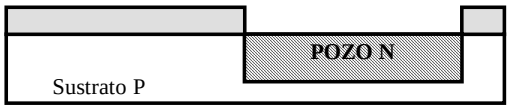
  -Twin-Tub.

  -Silicio sobre aislante.

6.2.1 CMOS DE POZO N

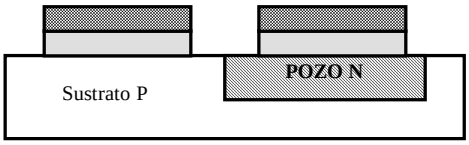
1. FABRICACIÓN DEL POZO N.

- En este pozo se fabrica el transistor P.
- La mascara se llama de N-WELL.
- Mediante:
  - ∗ Implantación iónica:
    - ⇒ Pozos poco profundos,
    - ⇒ Bien definidos,
    - ⇒ Compatible con dimensiones mas finas.
  - ∗ Deposición y difusión:
    - ⇒ Pozos mas profundos
    - ⇒ Propagación lateral
    - ⇒ Afecta a lo cerca que se pueden poner unos dispositivos de otros.



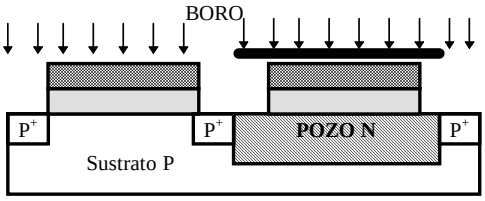
2. GENERACIÓN DE LAS ZONAS ACTIVAS

- Crecimiento del thinox que define las zonas activas.
- Active Mask (Thinoxide, Island, Mesa).
- Se crea una capa fina de SiO2
- Se cubre con una capa de SiN que se usa como mascara para los siguientes pasos



3. GENERACIÓN CANAL-STOP

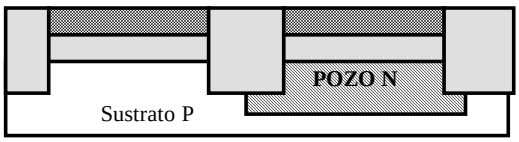
- Mascara Canal-Stop solo cubre la zona de pozo N.
- Se aplica una fuente de boro (huecos) de tal manera que entre las zonas activas aparece regiones con mayores impurezas P.
- Recordar que la mascara SiN protege la zona activa del transistor N.
- Para llevarlo a cabo se utilizan mascaras fotoresistentes (ya visto su funcionamiento).



4. CRECIMIENTO DEL OXIDO GRUESO

- Aísla transistores entre si.
- Mediante el proceso de oxidación ya estudiado.
- Crece en las zonas libres de SiN(Nitruro).
- El oxido crece en todas las direcciones, incluso por debajo de las zonas en las que se encuentra el SiN (nitruro). →Forma de pico de pájaro.
- La invasión del oxido reduce el arrea activa quedando mas pequeña que el patrón que se fijo.

- Step Coverage
  - ∗ El excesivo crecimiento del Field-Oxide (oxido grueso).
  - ∗ La diferencia de alturas entre Field Oxide y el Thinox provoca que el conductor de la siguiente cara disminuya hasta romperse



5. VARIACIÓN DEL POTENCIAL UMBRAL DEL DISPOSITIVO P

- Se añade una carga adicional de N en el pozo n entre el oxido y el silicio
- Con las cantidades de impurezas normales:
  - ∗ V_{TN}= 0.5-0.7 V
  - ∗ V_{TP}= -1.5-2.0 V.
- Están muy desequilibradas y conviene igualarlas
- El dispositivo NMOS tiene en valores absolutos un potencial menor que el dispositivo PMOS
- Convendría que el dispositivo PMOS tuviera un potencial umbral menor próximo en valor absoluto al del potencial NMOS. Con esto conseguimos características de transferencia simétricas.

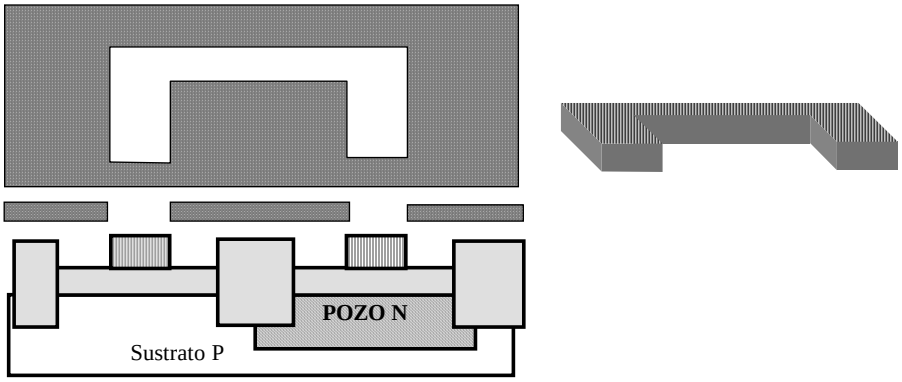
$V_T = V_{TMOS} + V_{FB}$

$V_{FB} = -|\phi_{MS}| + |Q_{FC}|/C_{OX}$

Q_{FC} Carga fija debido a las imperfecciones de la superficie y al dopado. Y este valor es lo estamos aumentando

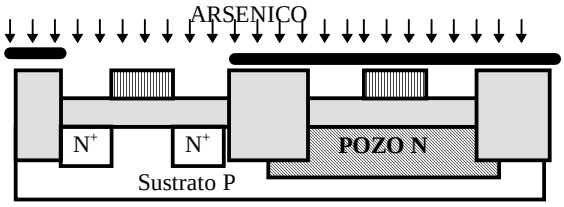
6. DEFINICIÓN DE LAS ZONAS DE POLISILICIO

- Se crea el layer de polisilicio mediante el proceso de aguafuerte
- la mascara “Polisilicio Mask” (en forma U).
- La utilización del polisilicio como mascara provoca fuentes-drenadores autoalineadas.



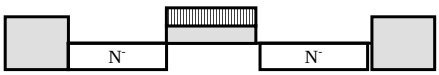
7. DIFUSIÓN N^+

- Mascara N^+.
- Si descubre el sustrato P indica las zonas en las que implantan la fuente
- El drenador del transistor
- Si descubre el pozo N la implantación sirve para generar un contacto ohmico, contactos con los sustratos, a esa mascara se la conoce como select porque selecciona las regiones N.

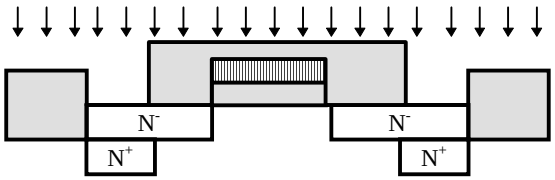


- Estructuras **LDD Lightly Doped Drain:**
  - ∗ El la actualidad en los procesos de pequeña dimensión uno de los principales esfuerzos se orienta a la eliminación del Hot Electron.
  - ∗ proceso

    A. Implantaciones poco profundas de N^-

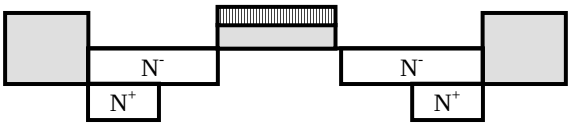


A. Se genera un oxido sobre el polisilicio.



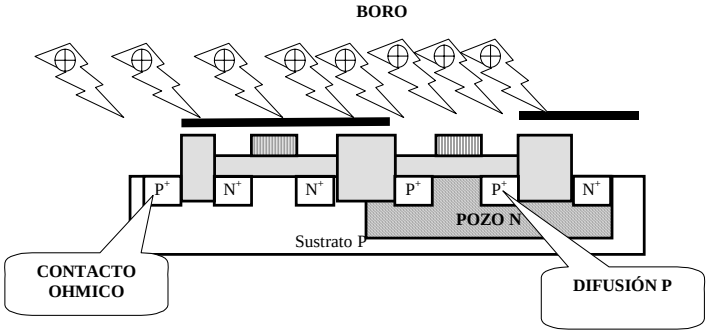
A. Se realiza una implantación N+.

B. Se elimina la zona de oxido resultando una estructura mas resistente al efecto de los electrones calientes.



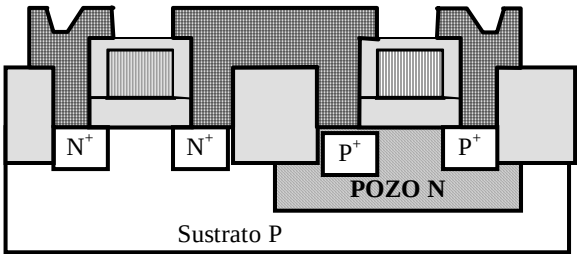
8. DIFUSIÓN P+

- Mascara complementaria a la N+ (mascara P+).
- Una difusión P+ sobre la región de pozo N genera un transistor P.
- Una difusión P+ sobre una zona del sustrato P genera un contacto ohmico con el sustrato.
- El paso de LDD no es necesario en los transistores P porque n tiene importancia el efecto de hueco caliente.



9. **METALIZACIÓN**

- Se cubre la superficie de SiO2
- se definen los “Contact Cuts” para que los metales entren en contacto con la difusión del plisilicio.

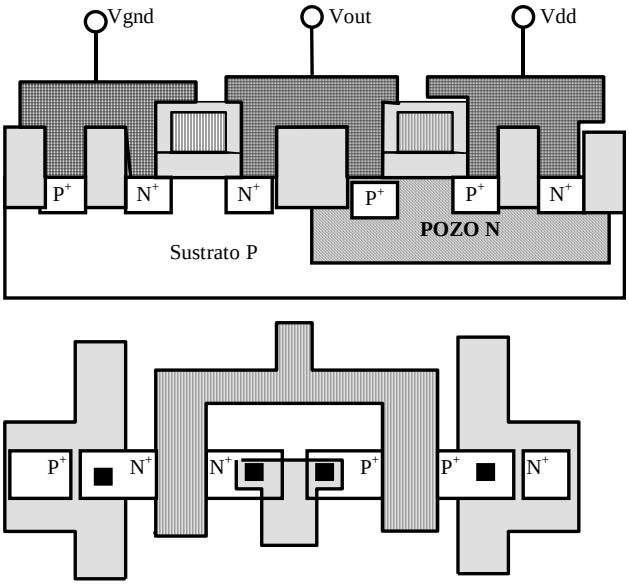


10. **PASIVACIÓN**

- Se recubre el chip de un material protector,
- Se deben dejar descubiertos los pads, los puntos internos de test.

6.2.1.1 Polarización de los substratos

- Vamos a estudiarlo en el proceso de pozo N
- El sustrato P se conecta a la fuente negativa VSS.
- El pozo N se conecta a la fuente positiva V_{DD}.
- El contacto con el sustrato P se realiza mediante la implantación P+
- El contacto en el pozo N mediante la implantación N+
- A estos contactos también se les llama
  - Well contacts
  - Body ties
  - Tub ties



6.2.2 CMOS DE POZO P

- Proceso similar al del pozo N.
- Se utiliza cuando se necesita un equilibrio entre las características de los transistores N y P.
- Como el transistor fabricado sobre el sustrato tiene mejores características que el fabricado en el pozo los dispositivos P son mejores que los N. es decir:
    - Pozo N; Acentúa las diferencias entre los dispositivos N y los P.
    - Pozo P; Reduce las diferencias entre los dispositivos N y P.

6.2.3 PROCESO TWIN TUB

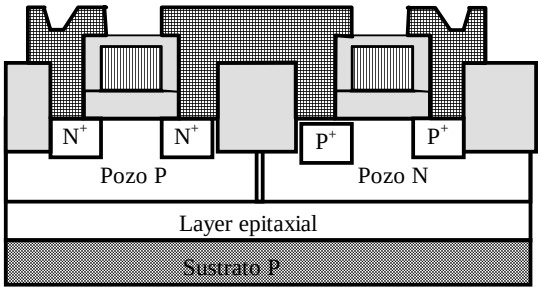
- Bañeras gemelas
- Optimizar por separado los transistores N y P.
    - ∗ Potencial umbral
    - ∗ Efecto cuerpo
    - ∗ Ganancia.
- El material inicial es un sustrato dopado con cargas positivas o negativas
- Se hace crecer un layer de epitaxial,
    - ∗ Conocido como EPI,
    - ∗ Protección Latch-Up.

- Con la epitaxis se consigue:
    - ∗ Crecimiento controlado de layer de silicio puro
    - ∗ Concentración de dopantes exacta distribuida homogéneamente por toda la superficie

- El proceso de creación es similar al del pozo N con la diferencia de la formación de las dos bañeras en las que se implementan los transistores:
    1. Formación de las bañeras
    2. Construcción del thinox
    3. Implantaciones de fuentes y drenador
    4. Definiciones de los contactos
    5. Metalización

Ventaja: equilibrio en la optimización de los dispositivos P y N

∗ se incluye un paso de ajuste del potencial umbral

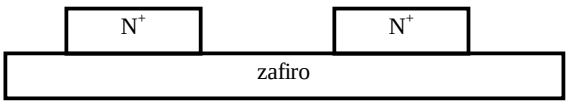

N+
N+
P+
P+
Pozo P
Pozo N
Layer epitaxial
Sustrato P

6.2.4 SILICIO SOBRE AISLANTE (SOI).

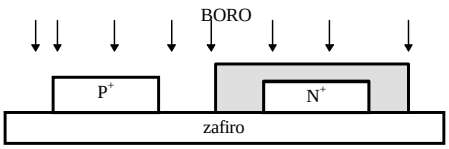
- Mejora problemas tradicionales:
  - ∗ Velocidad.
  - ∗ Latch-up.
  - ∗ Mejor empaquetado
  - ∗ Disminuye las capacidades parásitas del sustrato
- Consiste en conseguir un crecimiento epitaxial sobre un aislante, generalmente zafiro.
- En ocasiones se hace crecer sobre SiO2.
- Es compatible con los métodos tradicionales de fabricación de CMOS.
- Desaparecen las corrientes parásitas entre el sustrato y los transistores y los dispositivos adyacentes.

ν PASOS

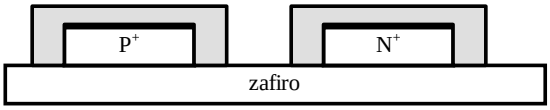
- Crecimiento de un layer tipo N ligeramente dopado (grosor entre 7 y 8 μ)
- Mediante técnicas habituales se elimina todo el layer salvo en las zonas en las que se fabrican los transistores a parecen dos islas N



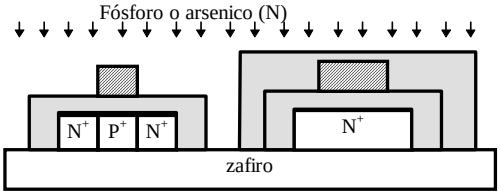
- Se genera una isla P mediante implantación de boro



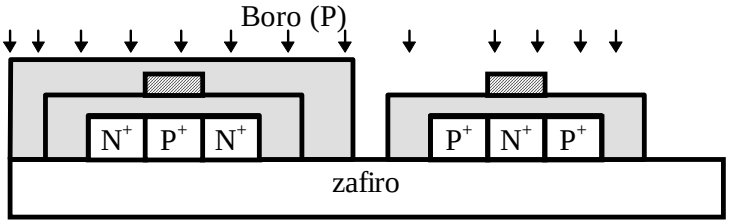
- Crecimiento del thinox sobre las islas-crecimiento muy controlado→para fabricar las capacidades



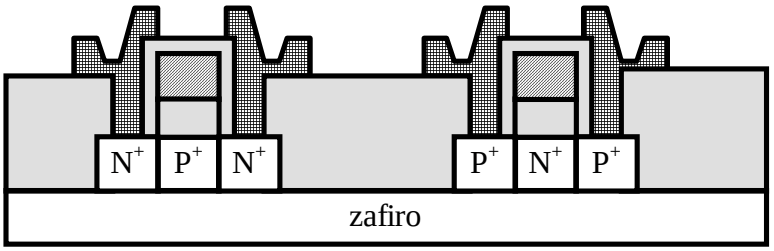
- Se deposita polisilicio sobre las islas y se le da forma mediante máscaras



- Se dopa la isla P mediante fósforo (portador tipo N+ para generar el dispositivo de canal N
- Se forma el dispositivo tipo P bombardeando la isla N+ con boro(P+)



- Se deposita un layer aislante sobre toda la estructura
- Se generan los contactos
- Metalización



ν **VENTAJAS:**

- Debido a la ausencia de pozos las estructuras son muy densas
- Bajas capacidades del sustrato, circuitos rápidos
- No existe Latch-up, se aíslan los transistores del sustrato
- No existe efecto cuerpo porque no hay corrientes en el sustrato
- Aumento de la tolerancia a fallos

ν **DESVENTAJAS:**

- La ausencia de contactos ohmicos hace el circuito de difícil protección
- Aunque se reducen las capacidades parásitas del sustrato, las capacidades entre hilos todavía existen luego la reducción de la capacidad total es menor de lo que cabría esperar
- La densidad no es particularmente importante, debido a que en la actualidad la densidad depende sobre todo de los layers metálicos de conexión
- El sustrato de zafiro encarece el producto

6.3 MEJORAS EN LOS PROCESOS CMOS

Objetivos:

- Incremento de la rutabilidad
- Capacitores de alta calidad para las memorias
- Resistencias de características variables

6.3.1 MEJORA DE LA RUTABILIDAD

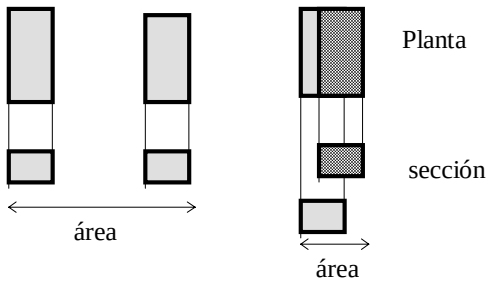
Existen dos posibilidades:

Mas de un nivel de metal

Mejoras en el layer de polisilicio

6.3.1.1 Mas de un nivel de metal

- Layers adicionales de distribución de señales y potencias.
- Facilita el rutado automático
- Mejora la distribución de Vdd, gnd y ck que son las tres líneas mas difíciles de trazar
- La utilización de mas de un layer de metal tiene como efecto la utilización de nuevas conexiones entre estos layers
- En la actualidad todos los diseño utilizan como poco dos layers de metal
- Mas de dos layers de metal para circuitos de alta velocidad y densidad
- Diferentes metales implica diferentes niveles
- Diferentes niveles implican mas hilos en el mismo área

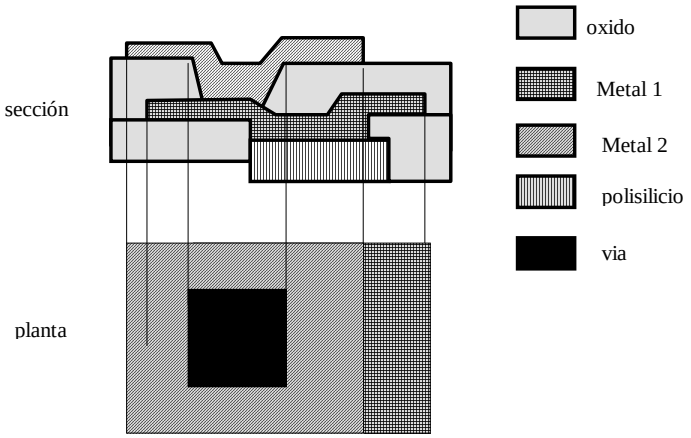


- **El metal 1**
  - ∗ Se utiliza siempre para contacto con las capas inferiores.
    - ⇒ polisilicio
    - ⇒ Difusiones.
- **El metal2**
  - ∗ se conecta con polisilicio y difusiones a través del metal1
  - ∗ En este caso debe existir distancia mínima entre ambos contactos
  - ∗ Hago llegar el Vdd a la difusión a través del metal
  - ∗ Efectos:
    - ⇒ Mayor densidad de empaquetado
    - ⇒ Caminos mas cortos
    - ⇒ Menor retardo

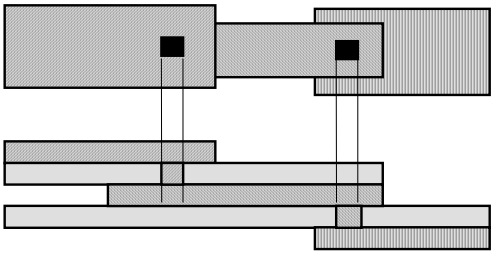
- El metal más usado es el aluminio
- **Problema**
  - ∗ Es la aparición del step coverage si el metal 1 no es lo suficientemente plano y delgado
  - ∗ Esto provoca que no se puedan montar exactamente una línea de metal sobre la otra
  - ∗ Si se emplea algún método para aplanar la primera capa de metal se puede conseguir que el segundo metal tenga la misma orientación del primero
  - ∗ Efecto mayor densidad de empaquetamiento
  - ∗ Esto se puede evitar colocando una segunda capa de metal mas gruesa o metieno mayor separación entre el metal 1 y metal2

ν **VIAS**

- Las conexiones entre metales se realizan mediante vías.
- generalmente al rededor de la VIA debe haber una zona de metal



- Conexión de la señal que lleva el metal 2 con el polisilicio o la difusión, Mediante un puente de metal 1.- esto produce un elevado consumo de área pero es un mal menor



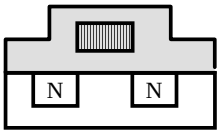
6.3.2 MEJORA DEL LAYER DE POLISILICIO

- El layer de polisilicio se utiliza para crear las puertas, pero debido a su alta resistencia, no se utiliza para líneas largas pues produce elevados retardos.
- Recordar que se utiliza sobre todo por su capacidad de usarse como mascara para producir fuente-drenadores autolineados.
- La solución sin niveles extra (sin utilizar un metal) es reducir la resistencia del polisilicio mezclándolo con un metal refractario.
- Existen tres soluciones al problema según se combine el polisilicio y el metal:

  * **Silicide**

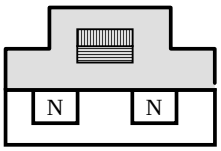
    ⇒ Silicio + tantalio

    ⇒ Las resistencias que se obtienen son 1-5 ohmio/square



  * **Polycide**

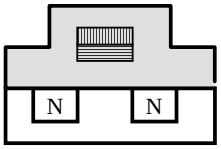
    ⇒ bocadillo: silice-polisilicio.



  * **Salicide.**

    ⇒ Formación de los drenadores y la fuente también con silicide

    ⇒ (Self Aligned Silicide)



- El efecto principal de todos estos procesos es reducir la resistencia de conexión del segundo layer. Permite al material de puerta usarse para trasladar la señal a distancias de tipo medio.

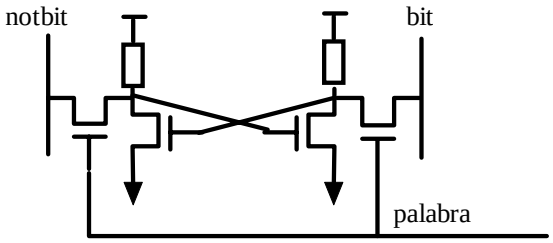
6.3.3 ELEMENTOS PASIVOS DEL CIRCUITO.

Resistencias

Capacitores

6.3.3.1 Resistencias

- El polisilicio sin dopar tiene una resistencia elevada.
- Se puede usar para construir resistencias.
- Se suele utilizar en el diseño de celdas de memoria RAM estáticas

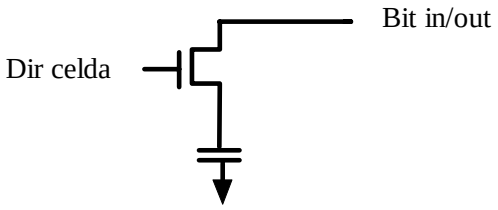


(Los dispositivos P se sustituyen por resistencias).

- La resistencia debe ser lo suficientemente grande como para evitar que en uno de los inversores salga 1 cuando debería salir 0
- Por ejemplo:
  - Si esa celda tiene cargado un 0→ en el inversor 2 se carga un 0 y en el 1 se carga un 1.
  - Si la resistencia del inversor no es lo suficientemente grande puede ocurrir que la Intensidad proporcionada por la resistencia sea tan grande que el valor 0 de salida se aproxime peligrosamente a la zona de indeterminación.
- La forma de prevenir el dopaje del polisilicio (se produce en la fase de generación de las difusiones) es cubrir las zonas deseadas con una mascara fotoresistente.

6.3.3.2 Capacitores

- Orientados al diseño de memorias dinámicas
- Alta capacidad en el mínimo espacio
- Diseño en tres dimensiones, se explota la dimensión vertical para aumentar la relación capacidad/área
- La estructura mas habitual es el **capacitor de trinchera** gracias al cual se ha conseguido aumentar notablemente la densidad de las memorias.

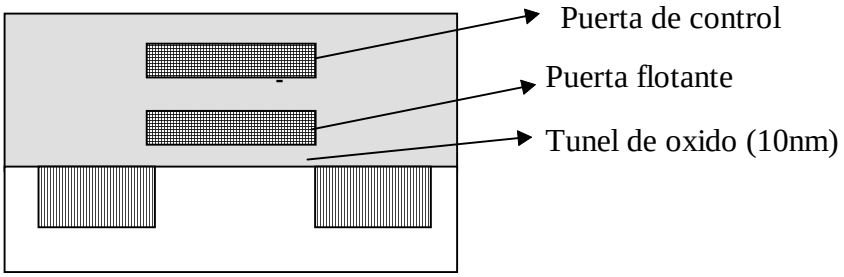


Funcionamiento

Carga y descarga de la capacidad que es donde se almacena el dato

6.4 ROM'S ALTERABLES ELECTRICAMENTE

- ROM: Read Only Memory.
- EAROM/EEROM (memorias ROM alterables eléctricamente.
- Se suele añadir en los procesos CMOS memorias ROM reprogramables, pero permanentes.
- Esto se consigue mediante un nuevo layer de polisilicio.
- Los dos layers de polisilicio forman una capacidad aislada del resto de los dispositivos.



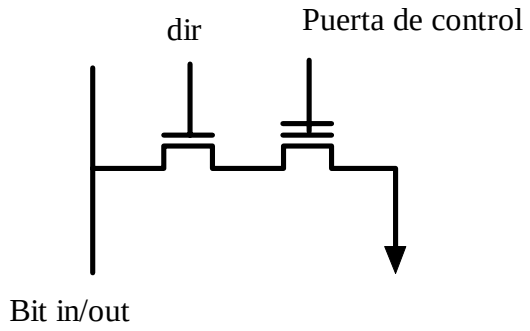
- El problema es como programar (cargar o descargar) esta capacidad, para ello se utilizan dos efectos ya estudiados:

  - Tunel Fowler-Nordheim
  - Electron caliente

- Aplicando los potenciales necesarios en la puerta de control, la fuente y el drenador se consigue que se produzca una corriente de portadores (e-) entre la puerta y la fuente o el drenador según se desee cargar o descargar la capacidad.
- Método alternativo: electrón caliente, los electrones con exceso de energía se desvían chocando contra el oxido fino y atravesándolo.

ν **MODO DE FUNCIONAMIENTO**

- Esquema de una memoria ROM reprogramable:



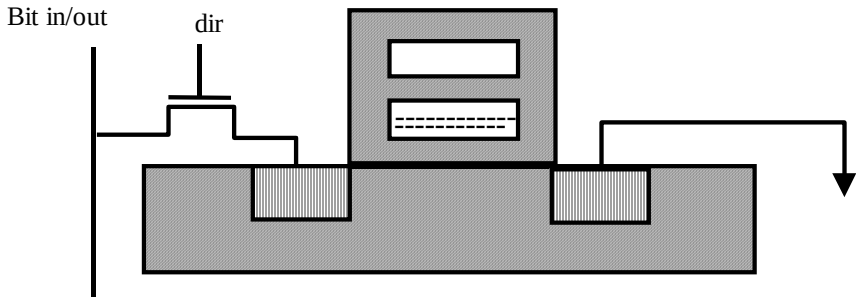
- Si se quiere almacenar un cero se modifica el potencial umbral de la puerta de control para que exista un camino a tierra siempre
- Si se quiere almacenar un uno se modifica el potencial umbral de la puerta de control para que no conduzca nunca
- **Precarga:**
  - ∗ Entre operaciones de lectura se carga la línea de bit de in/out con un 1
  - ∗ Si el contenido de la celda es un cero se descarga la línea a través del camino atierra
  - ∗ Si el contenido de la celda es un 1, no existe camino a tierra y el 1 sigue almacenado en la línea

- **Modificación de los potenciales umbral**
  - ∗ Se somete al transistor a las condiciones de Efecto túnel

    ⇒ V=14V en la puerta control

    ⇒ V=12V en el drenador

    ⇒ V=0 en la fuente



- Se carga la puerta flotante con los e-.
- Aumenta el potencial umbral hasta un valor de 7V.
- Para valores de trabajo de potencial normales, en la puerta control:

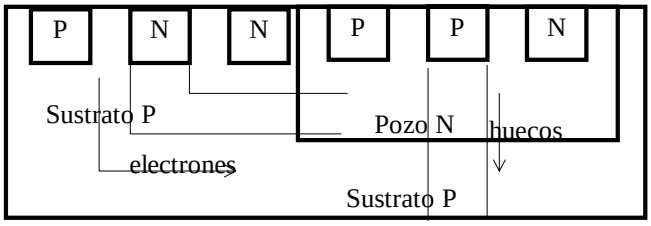
  V_{GS}= 5, V_T = 7 luego V_{GS}<Vt siempre

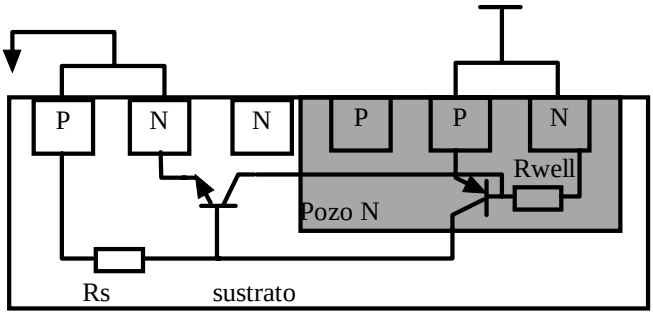
- En definitiva si se puede considerar la puerta de control como la puerta flotante, forman una capacidad aislada que conserva el valor de carga.
  - Indefinidamente?
    - 10 años a 125º
    - A mayor temperatura menos tiempo

6.5 LATCH-UP

- Cortocircuito entre V_{DD} Y V_{SS}
- El efecto no se produce siempre, se tienen que dar ciertas circunstancias y ,si no se quema el circuito, no es permanente
- Debido a la aparición de
    - Transistores bipolares parásitos
    - Resistencias de pozo y sustrato
- Efecto
    - En el peor de los casos se produce la destrucción del circuito integrado
    - En el mejor de los casos mal funcionamiento y dificultades para alcanzar el 0

6.5.1 ESTUDIO DEL CIRCUITO REALIMENTADO

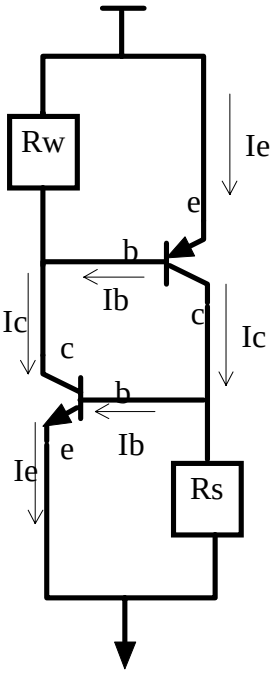




- Aparecen dos corrientes que unen V_{SS} Y V_{DD} para producir el cortocircuito:

  La que atraviesa PNP mayoritariamente de huecos

  La que atraviesa PNP de electrones



ν **TRANSISTOR VERTICAL PNP**

- Emisor formado por fuente drenador del dispositivo P
- La base es el pozo N
- El colector es el sustrato P
- Tanto la fuente como el drenador pueden actuar de emisor pero solo la fuente puede mantener la condición de Latch-up debido a su conexión con Vss

ν **TRANSISTOR HORIZONTAL NPN**

- También llamado lateral
- Emisor formado por fuente-drenador del dispositivo N.
- La base es el sustrato P.
- El colector es el pozo N

ν **REALIMENTACIÓN:**

- Colector NPN= Base PNP
- Base NPN =Colector PNP

- La resistencia de sustrato y de pozo se deben a la resistividad de los semiconductores
- <u>Recordar:</u>

NPN	V_{BE} <0.65; corte
	V_{BE} >0.65; activa
	V_{BE}>= 0.75; saturación y V_{BE}>V_{CE}
PNP	V_{BE} > -0.65; corte
	V_{BE} <-0.65; activa

ν **MODO DE OPERACIÓN**

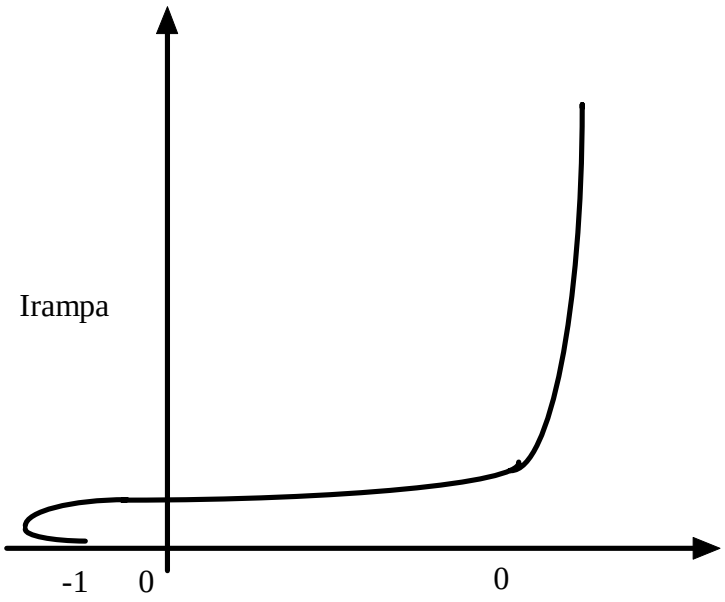
- Supongamos que por alguna causa el dispositivo NPN empieza a conducir:

I_{C1} sube→ V_{C1} baja→ V_{B2} baja→ PNP comienza a conducir→ I_{C2} sube→I_{B2} sube→ V_{B1} sube→ I_{C1} sube.

- Aparece una realimentación positiva que hace que la intensidad que atraviesa todo el dispositivo se incremente bruscamente de manera notable.

ν **CARACTERÍSTICA VI**

- describe el comportamiento del dispositivo de Latch-up.



- Igual que antes no vamos a ver porque comienza el fenómeno. Solo estudiamos que es lo que ocurre.
- Suponemos que inicialmente no conduce NPN y V_B=Vss=0 luego para que conduzca:

$$V_{BE} > 0.65 \rightarrow V_{BE} - V_E = -V_E > 0.65 \rightarrow V_E = -0.65.$$

- Si por algún motivo Vne=VeNPN =-0.65 (disminuye) comienza a conducir el NPN.
- Llega un momento llamado punto de disparo "trigger point" en que se invierte el sentido de Vne y empieza la conducción de PNP y la realimentación, alcanzando el Vne un valor de mas o menos 4V. y la I un valor muy elevado.
- Fija el valor de la fuente del dispositivo a 4V en lugar de 0; no carga bien el 0. A este valor potencial se le llama potencial de Holding. Este estado persiste mientras el potencial entre los dos transistores sea mayor que el potencial de Holding.

- El repentino aumento de consumo de intensidad generado por el Latch-up puede provocar le destrucción de las líneas de metal que trasladan las señales Vdd y Vss.
- Cuando se estudia la realimentación se dice:
  - I_{C1} sube; V_{C1} baja, el razonamiento es el siguiente;
  - $I = V_1 - V_2 / Rw = V_{DD} - V_{C1} / Rw$

    Rw cte.

    Vdd.cte
  - Si V_{B2} sube; V_{C1} baja.
  - Si V_{C1} baja; V_{B2} baja.
- Para PNP $V_{BE} = V_B - V_E = V_B - V_{DD}$. Como VB2 baja habrá un momento en que $V_{BE} < -0.65$ y empiece a conducir.
- Si I_{C2} sube; $V_{C2} - V_G / R_S$ sube como Rs cte. y Vgnd cte; V_{C2} sube; V_{B1} sube.

6.5.2 DISPARO DEL LATCH-UP

- Para que ocurra se tienen que dar dos condiciones:

  Circuitos bipolares parásitos cumplan las condiciones de conducción

  El potencial de Holding se mantenga

- Los circuitos bipolares pueden empezar a conducir debido a dos factores:
  - Internos
    - Voltajes en Transitorios
    - Corrientes en Transitorios
  - Externos
    - Voltajes que sobrepasan las condiciones normales de trabajo
    - Corrientes que sobrepasan las condiciones normales de trabajo
- Además los pulsos de radiación (partículas α) también pueden producir el disparo.

6.5.2.1 Tipos de disparo

ν **DISPARO LATERAL**

- Fluye I a través del emisor NPN;
- modifica el potencial de base del PNP y hace que comience a conducir.
- El punto de disparo de este transistor viene dado por la expresión:

$$I_{NPN} = V_{PNPON} \ / \ \alpha_{NPN} R_{WELL}$$

donde:

V_{PNP} = 0.65V. Voltaje umbral del dispositivo.

α_{NPN}= Ganancia en base común del transistor NPN.

R_{WELL} = Resistencia de pozo.

Como R_{WELL} es constante V=IR por lo tanto cuanto mayor sea la resistencia de pozo menor debe ser la intensidad npn para que se alcance el punto de disparo

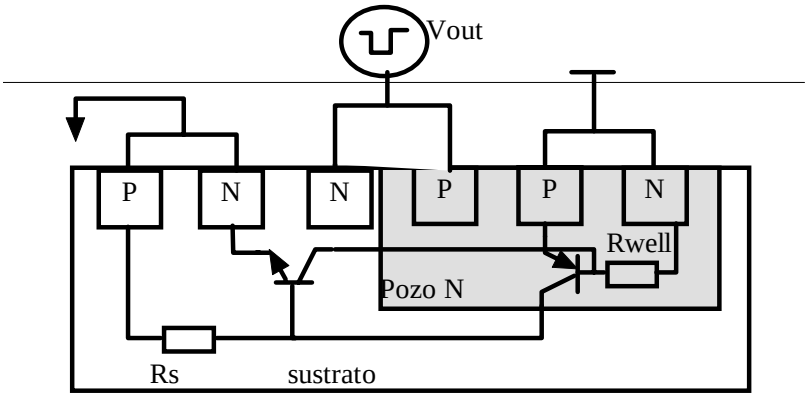
ν **DISPARO VERTICAL**

- Cuando una corriente lo suficientemente elevada se inyecta por el emisor del dispositivo PNP.
- Igual que antes esta intensidad se multiplicaba por la ganancia de corriente en base común del dispositivo PNP.
    - ∗ Disparo vertical fluye I a través del PNP: modifica el potencial de base del NPN y empieza a conducir el NPN.
    - ∗ En IR=V para I pequeñas y R grandes se alcanza V con facilidad.
    - ∗ Notar que la I es del dispositivo A y V del dispositivo B.

6.5.2.2 Ejemplos de carga vertical y horizontal

- Hemos visto que la forma de producir el disparo Latch-up es inyectar una intensidad por el emisor en cualquiera de los dos emisores.
- Esta intensidad se suele producir por motivos internos como aparición de potenciales transitorios (es poco probable).
- En los circuitos de I/O del circuito donde los circuitos internos se alimentan y alimentan al mundo exterior. Mucho cuidado con el diseño de los circuitos periféricos.

ν **CARGA HORIZONTAL**



- Si por cualquier motivo en la salida Vout se produce una fluctuación por debajo del Vss, por causas ajenas al circuito propiamente dicho:

$$Vbe_{NPN} > 0.65 \rightarrow \text{NPN conduce} \rightarrow \text{Conduce PNP}$$

ν **VERTICAL**

Por causas ajenas al circuito se produce una elevación entre el circuito de salida.

$$Vbepnp = Vb - Ve = Vdd - Ve$$

$$\text{Supongamos } Ve = Vdd + 0.2$$

$$Vbe = -0.7$$

Comienza a conducir, se dispara el Latch-up.

Que en estos dos ejemplos se produzca el fenómeno de Latch-up depende de la anchura de los pulsos y de la rapidez de los transistores parásitos.

6.5.3 PREVENCIÓN DEL LATCH-UP

- De un análisis del circuito parásito de Latch-up se extrae la inecuacion:

  $\beta npn \beta pnp > 1 + [(\beta npn+1)(Irsustrato + Irwell\beta pnp)]/Idd-Irsustrato$

  Donde:

  $Irsustrato= Vbenpn / \beta Rsustrato$

  $Irwell= Vbepnp / \beta Rwell$

  Idd = Intensidad total proporcionada.

De esta inecuacion se deduce que hay dos formas de prevenir el Latch-up:

Disminuyendo los valores de las resistencias.

Disminuyendo las ganancias de los transistores parásitos.

- El Latch-up se puede evitar actuando desde dos frentes bien diferentes:

  -Modificación y mejora de los procesos de fabricación cMOS.

  -Nuevas formas de diseño que de alguna manera mejoren el comportamiento.

6.5.3.1 Mejoras en los procesos

ν **CAPA EPITAXIAL**

- Reduce la ganancia de los transistores.
- Si el sustrato sobre el que se asienta esta fuertemente dopado reduce las resistencias.
- Si además se polariza a cero se convierte en un sumidero para la corriente de transistor, con lo que se evita la realimentación.

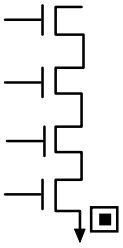
ν **POZOS RETRÓGRADOS**

- Fuertemente dopados en el fondo para reducir las resistencias.
- Ligeramente dopados en la superficie lo que ayuda al buen funcionamiento de los dispositivos.

6.5.3.2 Mejoras en el diseño de Layouts

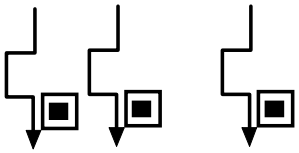
- La mejor forma de prevenir el Latch-up es mediante la utilización de los contactos ohmicos
- R_S
    - Resistencia de sustrato P
    - Conduce los huecos
    - Si se añade el contacto ohmico (contacto de sustrato) se esta introduciendo una cuña de material fuertemente dopado se reduce la resistencia.
- R_W
    - Resistencia de pozo N.
    - Conduccion de e-.
    - Contacto de pozo: cuña de material fuertemente dopado que reduce la resistencia (cuanto mas dopado menor resistencia)
- La utilización de contactos de pozo reduce mucho la posibilidad de que se produzca el Latch-up por motivos internos.
- El uso de estos contactos se puede reglamentar en parte mediante unas pocas condiciones:

    1. Cada pozo su contacto (un pozo puede tener varios contacto a la alimentación).
    2. Cada contacto de sustrato unido a metal directamente (ni en difusión ni en polisilicio).
    3. Colocar los contactos de sustrato tan cerca como sea posible de la fuente de los transistores que estén conectados directamente a la línea de alimentación.

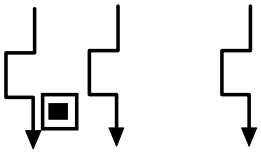


3a.Regla conservativa .Localizar un contacto de sustrato en cada conexión de alimentación (Vdd o Vss).

3b.Regla agresiva. Localizar un contacto de sustrato cada 5-10 transistores, 25-100n.



Regla 3a



Regla 3b

4. Intentar en el empaquetado y diseño que los dispositivos N queden orientados al Vss y los P queden orientados a Vdd.

6.5.3.3 Prevención del Latch-up en los dispositivos de E/S

- La forma de evitar el Latch-up en los dispositivos de entrada salida es rodearlos de guard rings.
    - Son contactos circulares alrededor del transistor.
    - Efecto; Reducir la resistencia. Reducen la ganancia de los bipolares parásitos.
    - Difusiones P+ en el sustrato P. Difusiones N+ en el pozo. Generalmente, incluso en los algunos diseños se utilizan celdas de PADS.

6.6 REGLAS DE DISEÑO

- Las reglas de diseño son el punto de conexión entre el diseñador de C.I. y el ingeniero de procesos durante la fase de fabricación.
- El principal objetivo de estas reglas de diseño es obtener un circuito con un rendimiento de producción optimo (circuitos validos/circuitos no validos) en un área lo menor posible sin comprometer la fiabilidad del circuito.
- Representan el mejor compromiso entre:

  -Rentabilidad de la fabricación

  -Performance del circuito

Las reglas mas conservadoras nos llevan a circuitos que funcionan mejor pero mas lentos y que ocupan mayor área.

Las reglas mas agresivas tienen mayor probabilidad de generar mejoras en el "performance", pero estas mejoras pueden dañar la rentabilidad.

Las reglas de diseño especifican al diseñador ligaduras geométricas y topológicas que deben cumplir los patrones utilizados en el proceso de fabricación.

Estas ligaduras no son leyes rígidas que se deban cumplir inexorablemente para que los circuitos funcionen correctamente, sino mas bien son recomendaciones del fabricante que aseguran una alta probabilidad de una fabricación correcta.

Se pueden encontrar diseños que violan las reglas y viceversa.

Existen dos conjuntos bien diferenciados de reglas de diseño:

Anchura mínima de las líneas

Distancia entre layers

Una anchura demasiado pequeña lleva consigo una discontinuidad en las líneas lo que puede provocar cortocircuitos.

Si los layers están demasiado cercanos se pueden fundir o interactuar el uno con el otro, cortocircuito entre dos nodos de circuitos diferentes

Hay dos aproximaciones para describir las reglas de diseño:

-Reglas "micron"

-Reglas basadas en lambda l.

Las reglas micron dan las anchuras y distancias entre layers en micras m. La forma en que se trabaja en la industria.

LAMBDA

- Es un factor de corrección.
- Fue introducido por Mead-Conway.
- En teoría permite trabajar con diseños independientemente del avance tecnológico.
- Un diseño que utilizase reglas lambda en su descripción serviría para diferentes tecnologías:
- Las reglas lambda se han utilizado con éxito en diseños: 4-1.5 μ.
- No dan buenos resultados para distancias inferiores a las micras.
- En definitiva estas reglas permiten un cierto estado de escalamiento entre procesos diferentes, en este caso seria suficiente reducir el valor de Lambda.
- La experiencia demuestra que las disminuciones no son uniformes.

ENREJILLADO:

- Las herramientas CAD trabajan con enrejillados de dimensiones mínimas en términos de las cuales hay que expresar las reglas de diseño
- para procesos 1.25 m-2 m enrejillados 0.2 m - 0.25 m.
- Por ultimo, alguno de los sistemas de fabricación de mascaras tienen problemas de exactitud digital (de 16 BITS de precisión).

6.6.1 REPRESENTACIÓN DE LAYERS

El proceso de fabricación de un dispositivo CMOS es muy complejo y necesita muchos niveles de mascaras.

Si el diseño tuviera que describir todas las mascaras este se haría engorroso y pesado.

Generalmente se hace una abstracción del diseño y solo se utilizan las mascaras estrictamente necesarias.

Los elementos que se utilizan para describir un diseño CMOS son:

Dos sustratos (P Y N) (no es un pozo).

Regiones difusión N y P.

La puerta del transistor.

Los Paths de conexión.

Los contactos entre layers.

La forma de representar estos layers:

Diferentes colores.

Diferentes tipos de bordes.

Diferentes rellenos o fondos.

Representación de layers para un proceso CMOS de pozo N:

LAYER	COLOR	SIMBOLIC
Pozo-n	marron	
Thinox	green	transistor n
Poli???	rojo	
P+	amarillo	transistor p
Metal 1	azul	
Metal 2	tostado(tan)	
via	negro	
Metal 3	gris	

6.6.2 ENUMERACIÓN DE ALGUNAS REGLAS

(basadas en lambda)

ν Pozo N

1. Tamaño minimo 10 l - 5 m.
2. Espacio entre dos pozos de mismo potencial 6l-3???
3. Espacio minimo entre dos pozos a diferente potencial 8l-4??

ν AREAS ACTIVAS

1. Tamaño mínimo 3
2. Espacio mínimo 3 l.
3. Solapamiento pozoN, a una difusión P+ 5 l.
4. Solapamiento pozoN, difusión N+ 3 l.
5. Distancia de un pozo N a un N+ 5 l.
6. Distancia de un pozo N a un P+ 3 l.

(Recordar que las zonas activas son las difusiones que implementan los transistores).

ν Polisilicio.

1. Tamaño mínimo 2 l.
2. Distancia mínima 2 l.
3. Distancia a la zona activa 1 l.
4. Extensión de la puerta 2 l.

ν MASCARAS P+ Y N+

(Según estemos en un proceso de pozoN o pozoP).

1. Mínimo solapamiento en la zona activa 2 l.

2. Tamaño mínimo 7 l.

Para determinar el área activa donde irán los dispositivos se utilizan dos mascaras;

-La de thinox, oxido fino. Verde si va a ser N. Amarilla si va a ser P.

-La mascara P+ que se utiliza en la implantación o difusión de los drenadores y puertas.

Como la mascara N+ = P+ con definir una sola mascara es suficiente.

ν CONTACTOS.

Metal 1: Poli. Difusion

1.Tamaño minimo 2 l.

2. Minima distancia en polisilicio 2 l.

3. Minima distancia en zona activa 2 l.

4. Minimo solapamiento con zona activa 2 l.

5. Minimo solapamiento con poli l

6. Minimo solapamiento con metal 1 1 l.

7. minimo espacio a una puerta 2 l.

ν METAL 1

Distribucion local de Vdd y Vgnd y de las lines de señales.

1. Tamaño minimo 3 l.

2. Distancia minima 3 l.

ν REGLAS DE POZO

Vamos a ver a continuación el motivo de la existencia de alguna de las reglas que se han listado con anterioridad.

1. Debido a que los pozos son mas profundos que las difusiones, cuando se fabrican tienen una mayor difusion alteral, esto puede provocar un cortocircuito en las difusiones externas. B5 y B6.

2. Para que sirve la zona de solapamiento interno entre las zonas Thinox (difusiones) y el pozo?. Para evitar que la zona de oxido grues en su crecimiento irregular se coma la zona que corrsponde a las difusiones.

3. Debido a las elevadas resistencias sheet (Rs = s/t) del pozo y para prevenir los elevados saltos de potencial que las corrientes de los sustratos provocan se deben colocar contacto ohmicos (contactos de sustratos) siempre que se pueda y que no vaya en contra de otras reglas. Esto esta muy relacionado con el Catch-Up. (Contacto sustrato en pozoN, difusion N+).

ν Reglas de transistor.

1. El polisilicio se usa de mascara para generar las zonas de difusion (fuente, drenador) autoalineadas.Si el poli se queda corto puede provocar un cortocircuito entre la fuente y el drenador.

   El polisilicio debe sobrepasar la zona activa, a esto se le llama “extension de puerta”.

2. El thinoxid (la zona activa) debe existir mas alla del polisilicio para que las zonas de difusion puedan meter y sacar portadores en el canal.

3. Las zonas de poli y las zonas activas que no forman transistores deben estar suficientemente separados para evitar posibles transistores indeseados.

4. Los dos tipos de transistores se fabrican cruzando:

-Zonas activas: Difusion

Implante

-Polisilicio.

Los dispositivos P se fabrican dentro de un pozo N.

Un dispositivo N se fabrica dentro de una zona de mascara N+ (N-Plus). Esta region que se dibuja dentro del layer no significa que este dentro de una difusion N+, sino que es la mascara utilizada para generar la zona de difusion.

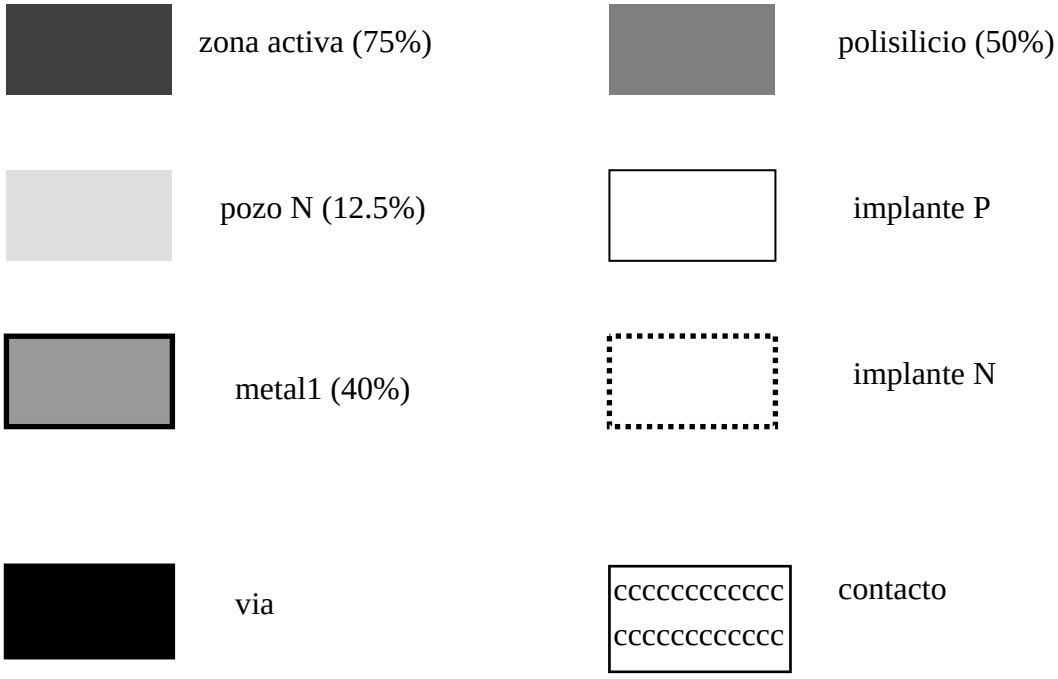
Para simplificar, en ocasiones la mascara P+ o N+ se coloca automaticamente al final del diseño.

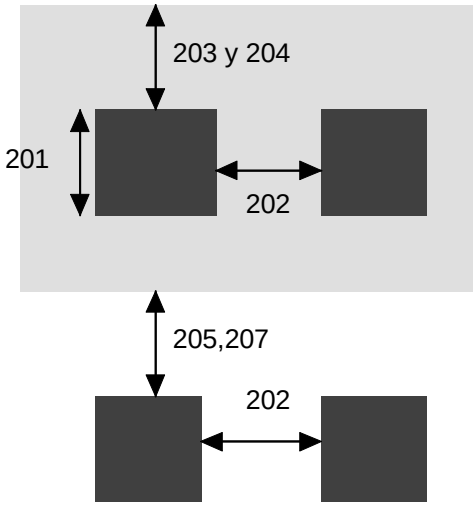
Resumen:

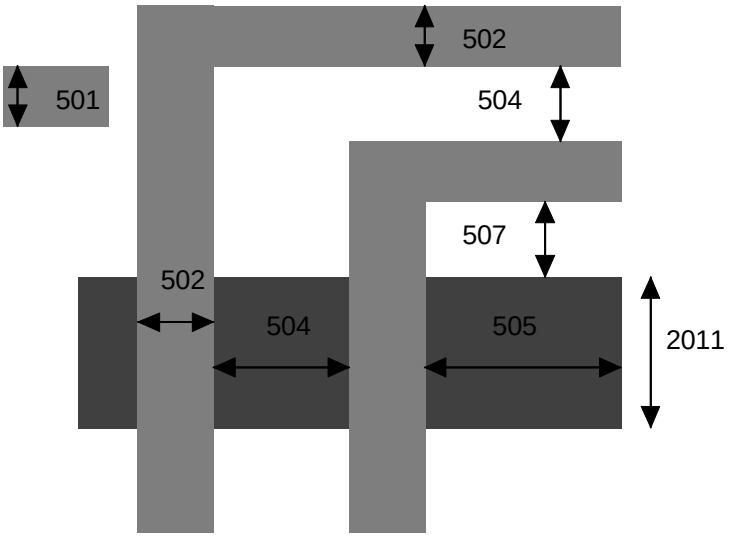
- Dispositivo P: rodeados pozo N

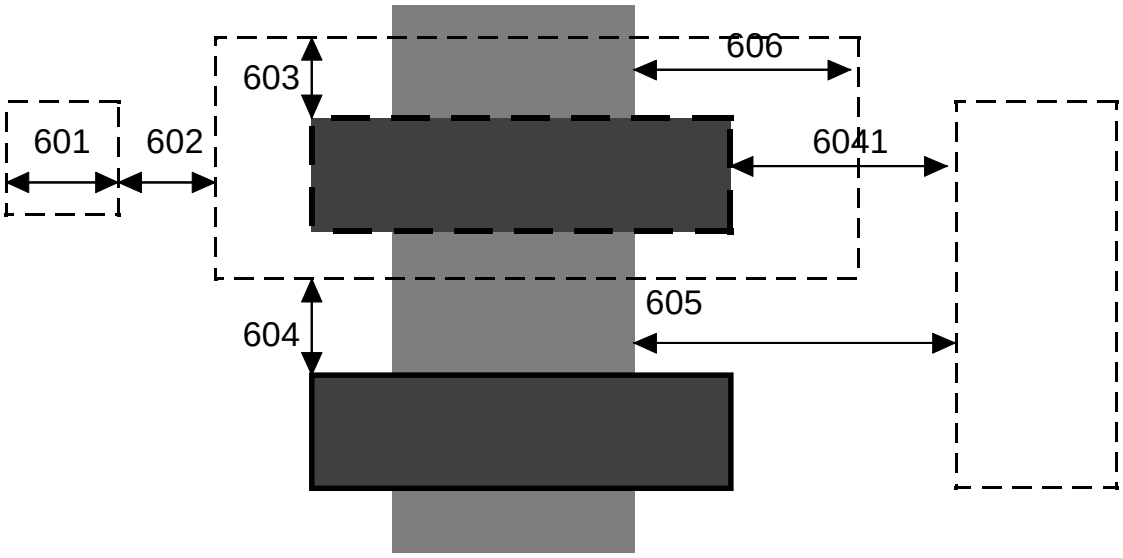
- Dispositivo N: rodeados Pmask+?????.

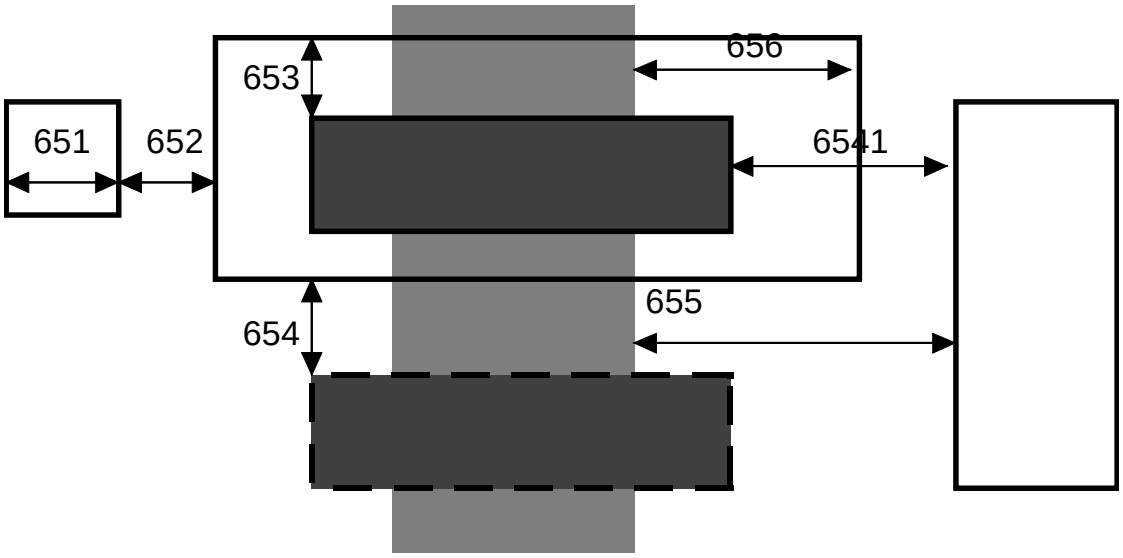
6.7 REGLAS DE DISEÑO DE ES2 PARA LA TECNOLOGÍA DE 0.7 MICRAS ECPD7

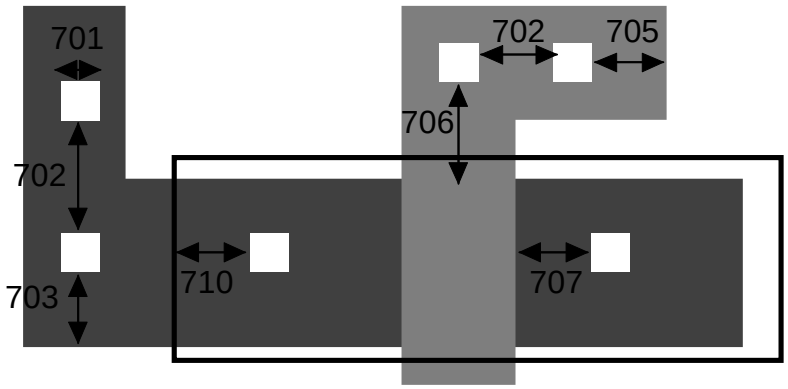



203 y 204
201
202
205,207
202


502
501
504
507
502
504
505
2011


606
603
601
602
6041
605
604


656
653
651
652
6541
655
654


701
702
705
706
702
710
707
703

7 CARACTERIZACIÓN DE CIRCUITOS

Un circuito integrado se compone de diferentes capas o layers de materiales como son

- El sustrato de silicio
- Las líneas de metal (Generalmente Aluminio)
- El oxido (SIO_2)
- El polisilicio de las puertas

Cada layer tiene resistencias y capacidades, cuyo conocimiento es fundamental para el calculo de las estimaciones de comportamiento de los dispositivos.

También existen inductancias que condicionan el comportamiento de los sistemas, estas inductancias suelen ser insignificantes y solo son importantes en los estudios de comportamiento de las entradas y salidas

7.1 ESTIMACIÓN DE RESISTENCIAS

7.1.1 RESISTENCIA DE UNA PLANCHA DE MATERIAL

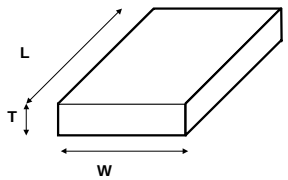
$$R = \left(\frac{\rho}{t}\right)\left(\frac{l}{w}\right)$$

ρ- resistividad del material

T grosor

L longitud

W anchura



Puesto que T es constante para una tecnología dada, esta expresión se puede reordenar de la siguiente manera:

$$R = Rs\left(\frac{l}{w}\right)$$

donde R_S es la resistencia sheet cuyas unidades son Ω/square

Las planchas de materiales con la misma R_S y la misma relación L/W tienen la misma resistencia.Si la longitud crece la resistencia aumenta , si la anchura crece la resistencia disminuye. A continuación se particulariza para los siguientes casos**:** Plancha de metal, Polisilicios y difusiones, canal.

ν **PLANCHA DE METAL:**

En los layers de metal la resistencia depende casi exclusivamente de la relaciones entre L,W y T porque ρ es constante. Según sea la aplicación que se desea dar al metal así será el tamaño que deberá tener la plancha:

Los layers superiores de metal suelen ser más gruesos para que tengan menos resistencias. En las memorias suelen ser un poco más delgados .Con esto se empeora la conductividad pero se evitan problemas de step coverage lo que permite usar varias lineas de metal prácticamente superpuestas y esto a su vez permite el ahorro de área. Con esto se aumenta la productividad de la oblea y se consigue una mayor concentración de celdas por unidad de área

En los ASIC las líneas de Vdd gnd o ck suelen ser líneas gruesas para poder distribuir de manera efectiva estas señales.

Los valores típicos que podemos encontrar para tecnologías de fabricación de entre 1-0.5 micras son para metal1 y metal 2 de 0.07 ohmios/square y para el metal3 →0.04 ohmios/square.

ν **RESISTENCIAS DE DIFUSIONES Y POLISILICIOS**

A diferencia de los metales que para un Rs se conoce perfectamente la resistencia en estos casos las resistividades están muy influidas por la densidad de concentración de impurezas y por los cambios que se producen al aplicar mezclas de materiales (como el silicide que es una mezcla de silicio y tantalio).Valores típicos para tecnologias de entre 1-0.5 micras son:

- Polisilicio 20 Ohmios/square
- Silicide 3 ohmios/square
- Difusión 25 ohmios/square
- Pozo N 2Kohmios/square

ν **RESISTENCIA DE UN CANAL MOS**

Aunque en los dispositivos NMOS generalmente se trabaja en la región no lineal (saturación) en la que se cumple $V_{DS}>V_{GS}-V_T$, se hace una descripción lineal con $V_{DS}<V_{GS}-V_T$ para simplificar la descripción y poder describir el comportamiento del transistor como la resistencia del canal. La Resistencia de canal viene dada por la siguiente expresión:

$$Rc = K\frac{L}{W}$$

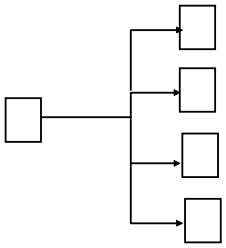
$$K = \frac{1}{\mu Cox(Vgs - Vt)}$$

Donde K es un valor similar a la resistencia de hoja ya vista. Tanto para canales N como para canales P, la resistencia de hoja puede tener valores 1-30 kΩ/square (para tecnologías de entre 1.0,5 micras). Como La resistencia es función de la movilidad μ y del potencial umbral y ambos dependen de la temperatura la resistencia también será función de la temperatura

7.2 ESTIMACIÓN DE CAPACIDADES

La respuesta dinámica de un sistema MOS depende en gran medida de las capacidades parásitas asociadas tanto a los transistores MOS como a las líneas que las unen. La capacidad de carga total de salida de una puerta de lógica CMOS es la suma de un conjunto de capacidades:

Capacidad de puerta.- Depende del número de entradas conectadas a una misma salida (fan-out)



Capacidad de la difusión.- Drenadores conectado a la salida

Capacidad de los hilos.- Capacidad de rutado

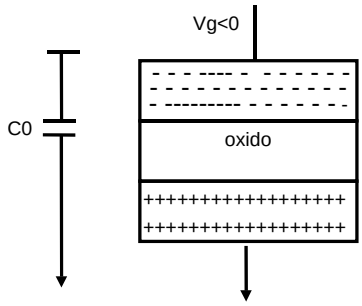
7.2.1 CARACTERÍSTICAS DEL CAPACITOR MOS

Antes de empezar el estudio se verá el comportamiento de un capacitor MOS puro. El capacitor MOS es un transistor MOS sin fuente ni drenador. Según el voltaje de puerta se puede suponer que el capacitor se encuentra en una de estas tres fases:

- Acumulación $V_G<0$
- Deplexión $V_G\approx0$
- Inversión $V_G>0$

Para el estudio suponemos un dispositivo de sustrato p

ν **ACUMULACIÓN $V_G<0$**



La carga negativa de la puerta atrae los agujeros a la superficie de silicio. Se puede aproximar la capacidad como la de un capacitor plano paralelo cuya expresión es:

$$C0 = \frac{\varepsilon sio2 \varepsilon 0 A}{Tox}$$

Siendo

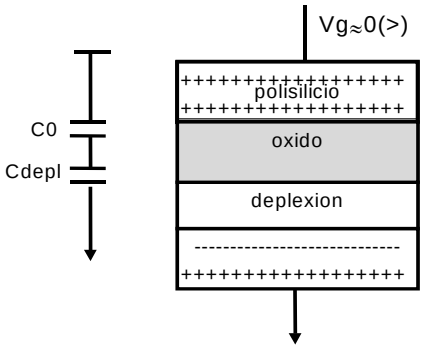
A el área del capacitor

ε_{sio2} constante dieléctrica del oxido (permisividad relativa)

ε_0 permisividad del vacío

Deplexión VG≈0 (>)

Cuando se aplica un pequeño potencial en la puerta se forma una zona de deplexión en el sustrato p situado directamente bajo la puerta



El voltaje positivo de la puerta repele los huecos del sustrato p dejando la zona vacía de huecos. Se puede aproximar a una capacidad plano-paralela con dos dieléctricos diferentes en su interior: el oxido y el semiconductor neutro

La densidad de carga en la superficie de la región de deplexión es función de:

Densidad de dopaje n

Carga electrónica (carga de electrón)

Profundidad de la zona de deplexión que es función del potencial aplicado en la puerta

la capacidad que aparece debido a la región de deplexión

$$Cdep = \frac{\varepsilon 0 \varepsilon siA}{D}$$

siendo

D profundidad de la región de deplexión

ε_{si} es la constante dieléctrica del silicio

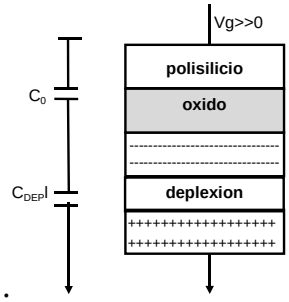
La capacidad total de la puerta viene dada por la suma de las capacidades C_0 y C_{DEP} en serie

$$\frac{1}{Cgb} = \frac{1}{C0} + \frac{1}{Cdep}$$

$$Cgb = \frac{C0Cdep}{C0 + Cdep}$$

ν **INVERSIÓN V_G>0**

los portadores minoritarios del sustrato p (los e^-) son atraídos hacia la superficie de silicio lo que provoca la inversión generando un canal N bajo la puerta



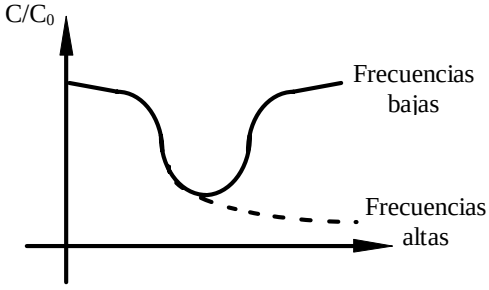
El valor de la nueva capacidad depende de la frecuencia a la que opere la señal V_G. Si la **frecuencia es baja** la situación a la que se llega es similar a la de la región de acumulación, pero con los valores de carga invertidos en este caso

$$\mathbf{C_{GB}=\varepsilon_{SIO2}\varepsilon_0 A/T_{OX}}$$

Si la **frecuencia es muy elevada** (mas100hz) los portadores minoritarios no pueden seguir el movimiento impuesto por los cambios de potencial

$$Cgb = \frac{C0Cdep}{C0 + Cdep} = Cmin$$

Esta capacidad mínima depende de la profundidad de la región de deplexión que a su vez es función del potencial aplicado



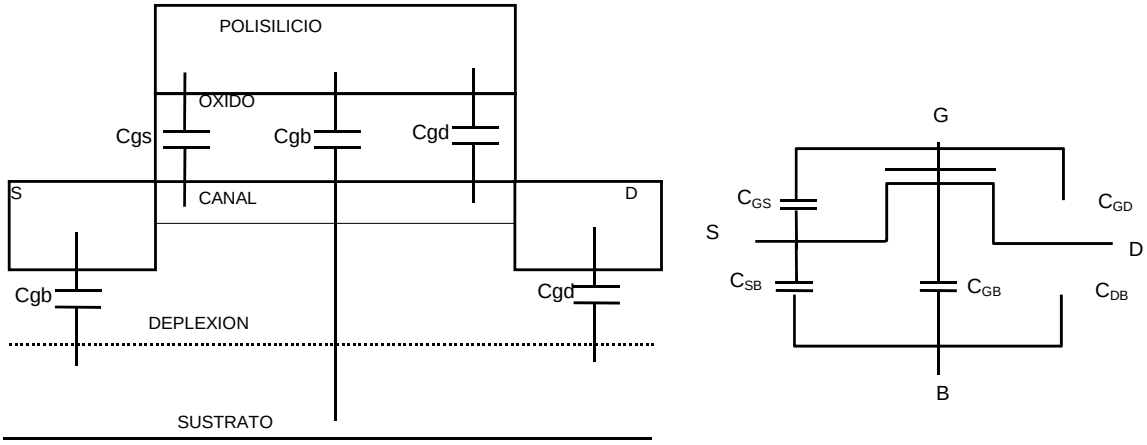
7.2.2 CAPACIDAD DE LOS DISPOSITIVOS MOS

Vamos a suponer que no existe solapamiento entre la puerta el drenador y la fuente.

En un transistor MOS se pueden encontrar las siguientes capacidades:

- **C_{GS},C_{GD}** capacidades puertas- canal que se agrupan en las regiones de drenador y fuente
- **C_{SB},C_{DB}** capacidades entre la fuente drenador con el sustrato(de difusión)
- **C_{GB}** capacitor MOS(puerta sustrato)

7.2.2.1 Capacidad de puerta



La capacidad total de la puerta es la suma en paralelo de cada una de sus capacidades:

$$C_G=C_{GS}+C_{GD}+C_{GB}$$

Vamos a estudiar el comportamiento de la capacidad total de puerta C_G para cada región en la que puede trabajar el transistor

ν **CORTADO $V_{GS}<V_T$**

No existe canal luego $C_{GS}=C_{GD}=0$, por lo tanto $C_G=C_{GB}$ que en este caso es la capacidad del capacitor mos puro en la zona de acumulación y por lo tanto su valor es:

$$C0=\frac{\varepsilon sio2\varepsilon 0A}{Tox}$$

ν **NO SATURADO $V_{GS}-V_T>V_{DS}$**

C_{GS} y C_{GD} =F(V_{GS}) cuyos valores estimados conservativamente son:

$$C_{GD}=C_{GS}=\frac{\varepsilon_0\varepsilon_{SIO2}A}{2T_{OX}}$$

Cuando se trabaja a altas frecuencias se vio que la capacidad $C_{GB}\approx 0$, luego la capacidad total de puerta es:

$$C_G=C_{GD}+C_{GS}=\frac{\varepsilon_0\varepsilon_{SIO2}A}{T_{OX}}$$

La capacidad no desaparece solo se modifican sus componentes

ν **Saturación V_{GS}-V_T>V_{DS}**

El Canal está fuertemente invertido, y en la región del drenador se produce estrangulamiento de canal luego el valor de cada una de las componentes de la capacidad es:

- $C_{GD}=0$
- se incrementa
- $C_{GB}=0$ (para altas frecuencias)

El valor de la capacidad C_{GS} realizando una estimación conservativa es

$$C_G=C_{GS}=\frac{2\varepsilon_0\varepsilon_{SIO2}A}{3T_{OX}}$$

Generalizando para los tres casos estudiados y para propósitos de retardos se puede suponer que la capacidad de puerta es

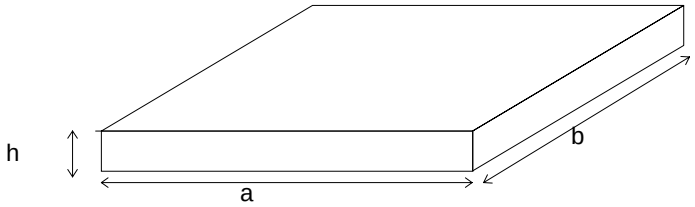
$$C_G=\frac{\varepsilon_0\cdot\varepsilon_{SIO2}A}{T_{OX}}$$

7.2.3 CAPACIDAD DE DIFUSIÓN

Aparece debido al contacto entre las zonas de difusión y el sustrato (o pozo) . Siempre que se une una zona N y una P aparece entre ambas una región de deplexión. Esta capacidad se puede descomponer en dos:

- La capacidad de base Cja
- La capacidad de longitud Cjp

Para ver el valor de cada una de ellas vamos a suponer que la difusión tiene una base de dimensiones a x b y una altura h.



- Cja es la capacidad debida a la unión de la base de la difusión con el sustrato y viene expresada por unidades de área. Su valor es

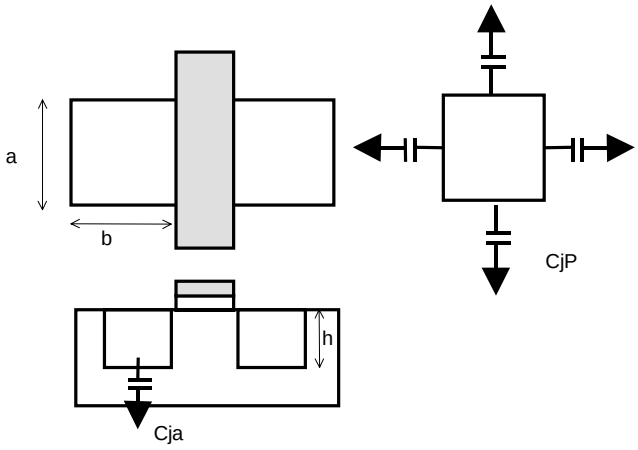
  $C_{ja} \cdot ab$

- **Cjp** es la capacidad debida a la unión de los laterales de la difusión con el sustrato perímetro de la difusión. Se mide en capacidad por unidad de longitud. suponiendo una h despreciable frente a y b su valor es

  * $\mathbf{C_{JP} \cdot (2^a + 2b)}$

La capacidad total de difusión será la suma de ambas

$$C_{DIFUSION} = C_{JA} \cdot (ab) + C_{JP} \cdot (2a+2b)$$



ν **MODIFICACIÓN DE $C_{DIFUSIÓN}$**

Conviene hacer notar que C_{JP} se ve modificada por el canal y la zona de deplexión. En los cálculos anteriores hemos supuesto un potencial a través de la unión $V_J=0$. Puesto que la profundidad de la zona de deplexión depende del voltaje a través de la unión esto implica que este potencial modifica C_{JA} y $C_{JP.}$ La expresión general que describe esta variación es :

$$C_{Ji} = C_{JOi} \left(1 - \frac{V_J}{V_B} \right)^{-m}$$

$$C_{JOi} = C_{Ji} \Big|_{V_J=0}$$

Donde

V_J es el potencial a través de la unión

Vb es el potencial de built in en la unión.

Potencial a través de la región de deplexión debido al contacto entre los semiconductores de tipo N y tipo P todo ello en ausencia de potenciales externos[pg 27 del Tsividis (aproximadamente 0,6v)]

M es una constante que depende de la distribución de impurezas cerca de la unión . Su valor entre 0,3 y 0,5

7.3 CAPACIDADES DE CONEXIONADO

Mientras que los hilos ideales no afectan al rendimiento de los sistemas, la realidad es bien distinta ya que los hilos reales introducen capacitancias, resistencias e inductancias parásitas que pueden tener una influencia perjudicial en el comportamiento del circuito. Estos efectos negativos se incrementan con los avances tecnológicos por dos causas: Cuanto menor es la tecnología menores son los retardos internos de las puertas y por lo tanto mayor efecto tienen los retardos de hilos. Según avanzan las tecnologías se hace mayor el tamaño de los circuitos y por lo tanto se hacen mayores los hilos introduciendo nuevos problemas de retardo. En cualquier caso los hilos largos tienen dos efectos diferentes sobre el rendimiento de los circuito: Introducen ruidos y Producen retardos

En la actualidad para las velocidades a que trabajan los dispositivos CMOS el elemento parásito más importante es el de las capacidades de interconexión. De hecho es casi el único problema del que se tienen que encargar un diseñador de circuitos estándares sin ligaduras demasiado fuertes. En una primera aproximación la capacidad de un hilo puede modelarse como la de un capacitor plano paralelo.

$$C=\varepsilon_{OX}WL/T_{OX}$$

Donde

W·L= A es el área

T es el grosor del aislante

ε permisividad del material aislante

- Valores típicos: $fF/\mu m^2$
  - metal1- sustrato 0.031
  - metal2-sustrato 0.015
  - metal3 -sustrato 0.010

7.3.1 EFECTO DE LAS MEJORAS TECNOLÓGICAS EN LA CAPACIDAD CONEXIONADO

pg440 Rabaey

Como mejora tecnológica en este caso definimos el escalado que se produce de una tecnología a otra

- El escalado por S de una tecnología reduce alguno de los parámetros que definen la capacidad de un hilo. Entre ellos destacan: W anchura y Tox

El comportamiento total de la capacidad del hilo cuando se escala la tecnología viene dado por

$$S_{C,hilo} = S_L \cdot S_W / S_T$$

- Suponiendo $S_W = S_T = S$ siendo S el factor de escalado de la tecnología la expresión resultante queda $S_{C,hilo} = S_{L,}$ en la que S_L depende de que la conexión sea local o intermódulos

El escalado de la longitud de hilo L depende de el tipo de localidad del hilo . Existen dos tipos de localidad de hilo:

- Conexiones locales entre puertas lógicas próximas
- Conexiones intermódulo como por ejemplo CK, V_{DD} o GND

La longitud promedio de un hilo global viene dado por la siguiente expresión

$$L_{av} = \frac{\sqrt{Ad}}{3}$$

Siendo Ad el área del chip

Otra longitud que se suele tomar como referencia es la longitud diagonal

$$L_D \approx (Ad)^{1/2}$$

las **conexiones locales** se reducen en la misma escala que se reducen el resto de los parámetros $S_L = S$

- luego las capacidades también se reducen en la misma escala

En las **conexiones globales** la longitud promedio es proporcional al área de circuito

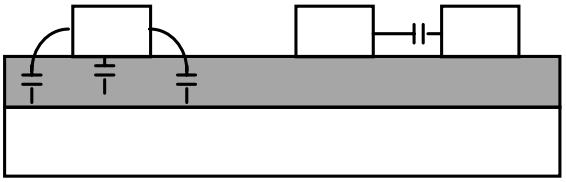
- el área crece con cada nueva tecnología
- la capacidad debida a los conexionados globales aumenta con la disminución de la tecnología. $S_L = 1/S$

En definitiva se puede observar que con la disminución de la tecnología se produce una separación entre los retardos del hilo y los retardos de las puertas.

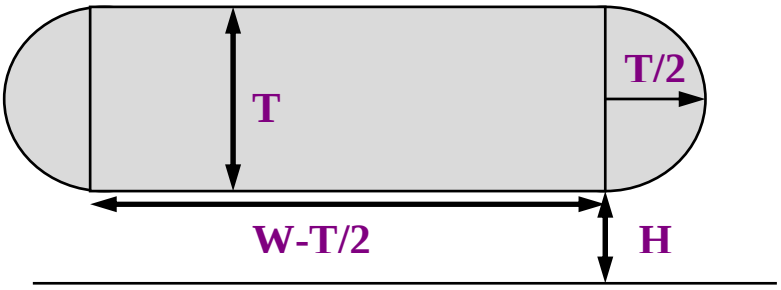
Problema El escalado aplicado a la longitud y anchura del hilo no es aplicable a la altura h del hilo. La reducción de h provoca un aumento de la resistencia como se verá más adelante, **solución** aplicar a h un escalado diferente del resto de los parámetros, esto produce una reducción de la relación W/H → $S_W/S_H \neq 1$

7.3.2 CAPACIDAD DE LOS CAMPOS LATERALES

El modelo plano paralelo es incorrecto porque ésta aproximación ignora: los campos laterales que aparecen en los flancos del conductor, debido a que su grosor es finito y las capacidades que aparecen entre los conductores del mismo layer



El efecto de los campos laterales es aumentar el área efectiva del condensador. Se puede hacer un estudio aproximado de la capacidad total suponiendo que el conductor está formado por una sección rectangular en el medio y dos secciones hemisféricas en sus extremos



La capacidad total es la suma de dos capacidades la del capacitor plano - paralelo de anchura W-T/2 y la del capacitor cilíndrico de radio T/2

Donde:

W es la anchura del conductor

H grosor del aislante

T grosor del conductor

ε permisividad

$$C = \varepsilon \left[\frac{W - \frac{T}{2}}{H} + \frac{2\pi}{LN\left\{1 + \frac{2H}{T} + \sqrt{\frac{2H}{T}\left[\frac{2H}{T} + 2\right]}\right\}} \right]$$

Esta fórmula solo tiene validez para los valores W≥t/2, t≈h

Existe una fórmula empírica bastante exacta

$$C=\varepsilon L[(W/H)+0,77+1.06(W/H)^{0.25}+1.06(T/H)^{0.5}]$$

ν **CONCLUSIONES:**

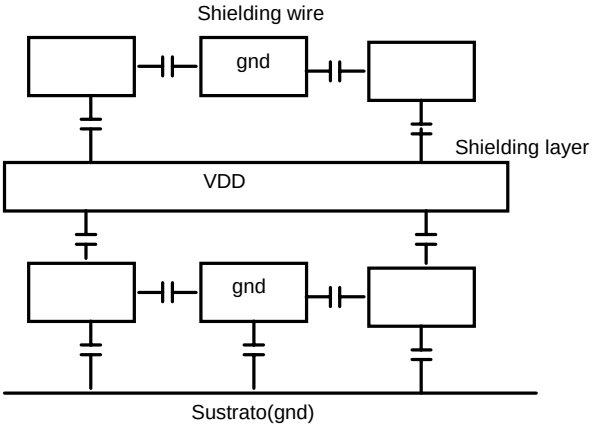
- La capacidad es mayor que la que se obtiene considerando exclusivamente el modelo plano paralelo
- La capacidad no es lineal con la anchura del metal
- La capacidad de los campos laterales pueden incrementar la capacidad total un factor entre 1.5 y 3 para líneas de pequeña anchura (W/H <1.5)

ν **CAPACITANCIAS ENTRE HILOS**

Un último factor a tener en cuenta es el de las capacitancias entre hilos. Al reducir el tamaño de la tecnología se reduce mucho la distancia entre hilos del mismo layer o de diferente layer. Cuando W llega a ser mas pequeña que 1.75h esta capacidad puede empezar a dominar y aparece el fenómeno de cross-talk (que se estudia en el siguiente apartado)

7.3.3 CAPACIDADES ENTRE MÚLTIPLES CONDUCTORES CROSS-TALK

En la actualidad los Circuitos Integrados tiene varios layers de routing → las interacciones entre los layers son muy complejas



El acoplamiento indeseado entre hilos de señales vecinas introduce interferencias que se llaman cross-talk. Estas interferencias actúan como fuentes de ruido que pueden conducir a errores intermitentes, puesto que los ruidos dependen de los valores transitorios de las señales vecinas. Este acoplamiento puede ser capacitivo e inductivo. $Cgb = \frac{C0Cdep}{C0+Cdep}$ El efecto de cross-talk capacitivo es un efecto dominante en las velocidades de los circuitos. Para reducir sus efectos se deben evitar hilos largos paralelos a largas distancias. Esto suele ocurrir por ejemplo

cuando se rutan las dos señales de reloj para sincronizar sistemas dinámicos de dos fases o en los buses

Los efectos del cross talk entre señales del mismo layer se pueden reducir guardando una distancia lo suficientemente grande entre los hilos e insertando hilos de shielding (hilos protectores) que suelen ser hilos de tierra o alimentación entre dos señales.

7.4 TIEMPO DE RETARDO RC EN LA PROPAGACIÓN DE LA SEÑAL

También llamado retardo RC distribuido, es el retardo de propagación de una señal a lo largo de un hilo depende de muchos factores incluyendo:

- Las resistencias distribuidas
- La capacitancia distribuida del hilo
- la capacidad de entrada del receiver
- La resistencia de salida del driver

El retardo RC es dominante en casos como Layers de elevada resistencia como puede ser el polisilicio, y Layers de menor resistencia pero muy largos como pueden ser los metales o los silicides

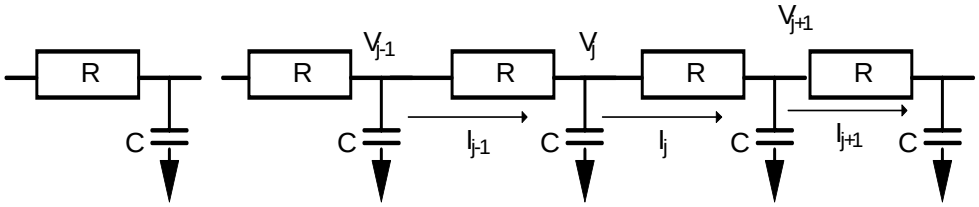
Existen diversos modelos de retardo RC nosotros estudiaremos el Modelo Lumped y el Modelo distribuido

- **MODELO LUMPED**

Es una primera aproximación que reúne en una resistencia toda la resistencia distribuida del hilo y en una sola capacidad toda la capacidad distribuida en el hilo. Es una aproximación pesimista e inexacta para hilos largos. El valor de su constante de tiempo es $\tau = R_T C_T$ siendo $R_T = r \cdot L$ y $C_T = c \cdot L$ luego $RC = \tau = rcL^2$ siendo r la resistencia por unidad de longitud y c la capacidad por unidad de longitud.

- **MODELO DISTRIBUIDO**

Un hilo largo se representa como un conjunto de secciones RC



Cuando una señal se propaga en un hilo el cambio de potencial no se produce simultáneamente en todo el hilo sino que en cada punto el V_I se va modificando poco a poco siguiendo la siguiente expresión:

$$C \cdot dVj/dt = I_{J-1} - I_J$$

La variación de potencial en el punto j es directamente proporcional a la intensidad efectiva en ese punto (la dedicada a llenar la capacidad) e Inversamente proporcional a la capacidad. (símil de la piscina Ij-1 agua que llega, Ij agua que sale dVj/dt variación del nivel del agua)

La intensidad en el punto j viene dada por la expresión

$$Ij = \frac{Vj - Vj + 1}{R} = \frac{dV}{R}$$

Si se supone que el número de secciones en que se divide el hilo aumenta esto implica que el tamaño de cada sección disminuye luego el estudio se puede hacer diferencial

$$\frac{CdVj}{dt} = dI = d\left(\frac{dV}{R}\right)$$

Como C=c·L y R=r·L y en forma diferencial se pude expresar como C=c·dx y R=r·dx sustituyendo se obtiene:

$$cdx\left(\frac{dVj}{dt}\right) = \left(\frac{dVj}{rdx}\right) \text{ y operando } cr\frac{dVj}{dt} = \left(\frac{d^2vj}{dx^2}\right)$$

Donde:

c es la capacidad por unidad de longitud

r resistencia por unidad de longitud

x distancia desde la entrada a la sección

V el potencial en un punto del hilo

Esta ecuación no tiene una solución exacta y por lo tanto hay que utilizar aproximaciones. Una aproximación puede ser la siguiente:

$$t = rc(\Delta L)^2\left(\frac{N(N+1)}{2}\right)$$

Donde N es el número de secciones N=L/ΔL. Cuando N→∞ la ecuación se reduce

$$t = \frac{rcL^2}{2}$$

la constante de tiempo crece cuadráticamente con la longitud del hilo. Por eso el retardado de la señal suele ser importante en hilos largos en los que domina el efecto RC

Notar que el modelo lumped no es exacto ya que en este caso predice un retardo rcL^2 que es dos veces mayor que el valor distribuido. Genéricamente RC= $krcL^2$ donde K

es una constante que se halla empíricamente. Los resultados de los estudios de los retardos en una línea son diferentes según sea el modelo de retardo elegido.

Rango de voltaje	RC lumped	RC distribuido
0→50% (tp)	0.69RC	0.38RC
0→63%(ι)	RC	0.5RC
10%→90%(tr)	2.2RC	0.9RC

Siendo

Tp lo que se conoce como retardo

Tr el tiempo de subida

ι la constante de tiempo RC

7.4.1 REDUCCIÓN DE LOS RETARDOS RC

Una primera opción para reducir los retardos es usar materiales de interconexión mejores. Por ejemplo un metal 1 (aluminio) de longitud 5mm tiene un retardo de 0.08ns que se puede ignorar

- **Problema** En ocasiones no se puede elegir el tipo de material. Por ejemplo las líneas de dirección de las memorias son hilos largos de polisilicio. Implementar estas líneas en polisilicio contribuye de manera importante a la densidad de circuito
- **Solución** Introducir buffers intermedios llamados repetidores. Esto tiene el efecto de hacer una línea de interconexión más corta→ reduce su retardo de propagación cuadráticamente, lo que es suficiente para no tener en cuenta los retardos extra que introducen los buffers
  - La expresión general del retardo rc de una línea de longitud l viene dada por la siguiente expresión obtenida analíticamente (mirar la tabla anterior) :

$$Tp=0.38\cdot r\cdot c\cdot L^2$$

  - Al introducir M-1 repetidores en la línea la expresión del retardo queda

$$0.38rc[L/M]^2M+(M-1)t_{buffe}$$

Donde

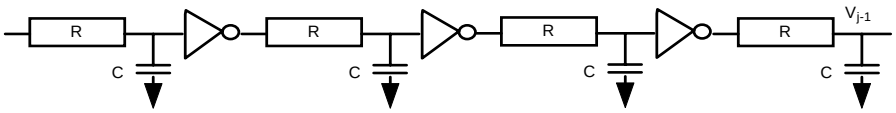
t_{buffe} es el retardo de propagación del buffer

M el número de secciones en que se divide la línea

  - El número óptimo de buffers a utiliza se consigue haciendo

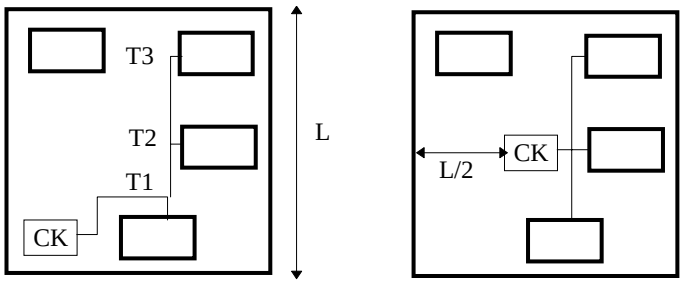
$$\partial Tp/\partial M=0$$

y es $\mathbf{M=L(0.38rc/t_{buffer})^{1/2}}$


R
R
R
R
C
C
C
C
V_{j-1}

7.4.2 REDUCCIÓN DEL RETARDO DE LA LÍNEA DE RELOJ

- Es típico el estudio del retardo de las líneas de reloj que deben sincronizar todo el circuito
- Estas líneas suelen ser largas por lo que tienen dos efectos:
  - ∗ Importantes retardos en las líneas
  - ∗ Skew o desviaciones de unos módulos a otros



Solución errónea

- **solución**
  - ∗ consiste en colocar el buffer del reloj en medio de circuito
  - ∗ Ventajas:
    - ⇒ La máxima longitud se reduce a L/2
    - ⇒ Las desviaciones que se producen son muy parecidas

  Aunque se aumenta la capacidad de entrada del buffer debido a que el hilo de entrada es más largo, se disminuye la resistencia de salida

7.4.3 CUANDO SE DEBEN TENER EN CUENTA LA LONGITUD DEL HILO

- Para hilos suficientemente cortos podemos ignorar el retardo RC
- En este caso los hilos se pueden tratar simplemente como capacidades de carga
- Es útil definir reglas sencillas que sirvan de guía para determinar la máxima longitud de un hilo que funcione como un nodo.
- La condición para despreciar el retardo RC de un hilo es que el retardo del hilo sea inferior al de la puerta lógica:

$$T_{hilo}<T_{puerta}$$

Como $T_{hilo}=r\cdot c\cdot L^2/2$

Se obtiene que

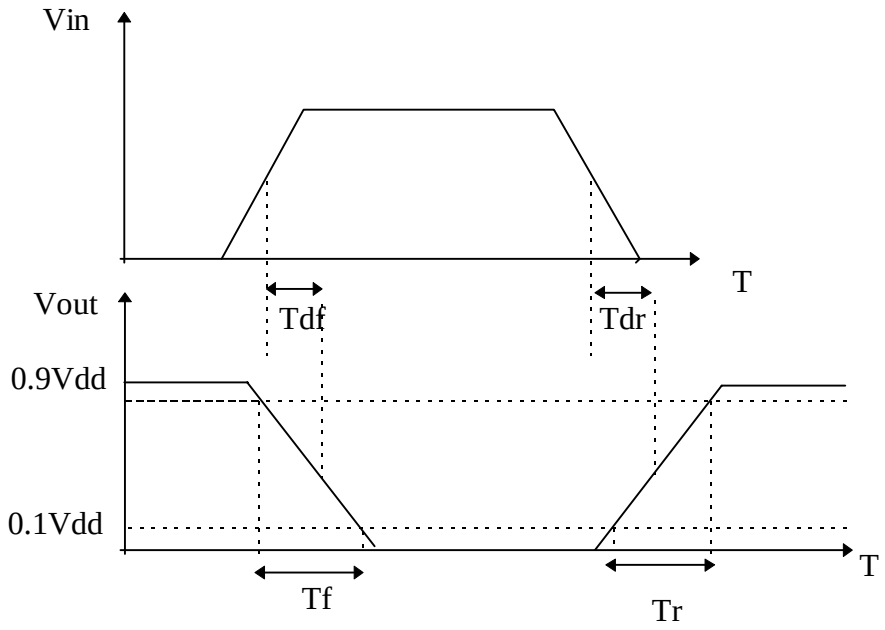
$$\mathbf{L<(2T_{puerta}/rc)^{1/2}}$$

7.5 CARACTERÍSTICAS DE COMUTACIÓN

- El estudio de modelos analíticos y empíricos de los inversores MOS se realiza porque se extraen conclusiones generalizables al resto de las puertas lógicas
- El tiempo de comutación de una puerta lógica está limitado por el tiempo necesario para cargar o descargar la capacidad de carga de salida
    - C_L se carga a Vdd
    - C_L se descarga a Vss
- Definiciones
    - ∗ **Tiempo de subida** T_R= tiempo que tarda la onda de salida en pasar del 10% al 90% del Vdd
    - ∗ **Tiempo de bajada** T_F tiempo que tarda la onda de salida en pasar del 90% al 10 de Vdd
    - ∗ **Tiempo de retraso**
        - ⇒ T_D la diferencia entre el tiempo que se tarda en alcanzar el 50% del Vin y el 50% del Vout.
        - ⇒ Existen dos tiempos de retardo diferentes :

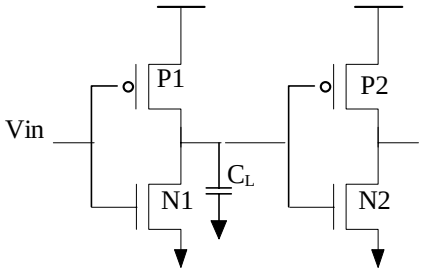
          Tdr tiempo de retardo de subida

          Tdf tiempo de retardo de bajada:



7.6 MODELOS ANALÍTICOS DEL RETARDO

7.6.1 TIEMPO DE BAJADA



C_L capacidad de carga de salida del inversor

7.6.1.1 Estudio cualitativo

- Suponer
  - ∗ Vin(t) es un pulso cuadrado es decir Vin=Vdd de manera instantánea
  - ∗ La capacidad esta inicialmente cargada a Vdd
  - ∗ El dispositivo n cortado
- En el momento que V_G=1 el dispositivo N empieza a conducir → se descarga la capacidad a través del transistor N
- El dispositivo n pasa por dos fases
  - ∗ Saturación
  - ∗ No saturación
- **Saturación**

  Vin=V_G=5v y Vout=Vdd porque C_L cargada

  Luego V_{GS}-V_T<V_{DS}→ Vdd-V_T<Vdd

  En esta fase flujo de intensidad máximo
- **No saturación**

  Según se va descargando la capacidad C_L el Vout va disminuyendo hasta que se cumple Vdd-V_T>Vout

7.6.1.2 Estudio cuantitativo

- El estudio del tiempo de bajada se debe dividir en dos $T_F = T_{FS} + T_{FL}$
  - ∗ **T_{FS}** Tiempo de bajada en saturación
  - ∗ **T_{FL}** Tiempo en bajada en lineal

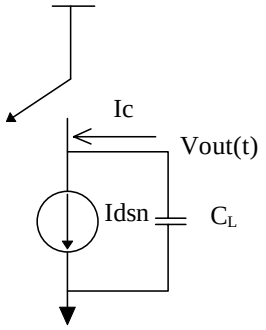
la condición que les separa es $Vdd-V_{TN}=Vout$

T_{FS} tiempo que el potencial de salida tarda en pasar de Vout= 90%Vdd a $Vout=Vdd-V_{TN}$

T_{FL} tiempo que el potencial de salida tarda en pasar de $Vdd- V_{TN}$ a 10%Vdd

ν **CALCULO DE T_{FS}**

- El circuito equivalente en saturación es el siguiente:



La variación del potencial es directamente proporcional a la intensidad e inversamente proporciona a la capacidad y viene dada por la expresión:

$$C_L \frac{dVout}{dt} + Idssat=0 \quad [1]$$

La intensidad del transistor N en saturación viene dada por la expresión:

$$Idssat=\beta n/2(Vdd-V_{TN})^2 \quad [2]$$

sustituyendo [2] en [1] se obtiene

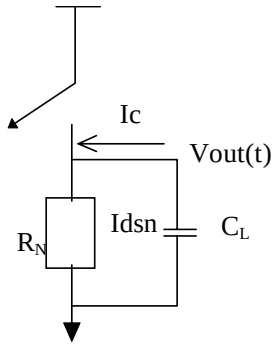
$$C_L \frac{dVout}{dt} + \beta_N/2(Vdd-V_{TN})^2=0$$

despejando dt e integrando

$$T_{FS}= \frac{2C_L}{\beta_N(Vdd-V_{TN})^2}\int_{Vdd-V_{TN}}^{0.9Vdd} dVout = \frac{2C_L(V_{TN}-0.1Vdd)}{\beta_N(Vdd-V_{TN})^2}$$

ν **CÁLCULO DE T_{FL}**

En este caso el dispositivo N trabaja en la zona lineal
el circuito equivalente es:



La variación del potencial con respecto del tiempo viene dado por la expresión:

$$C_L \frac{dVout}{dt} + Idslin=0 \qquad [3]$$

La intensidad del transistor N en la zona lineal viene dado por

$$Idslin=\beta n[(Vdd-V_{TN})Vout-Vout^2/2] \quad [4]$$

Igual que en el caso anterior sustituyendo [4] en [3] , despejando dt e integrando se obtiene:

$$T_{FL}= \frac{C_L}{\beta_N(Vdd-V_{TN})} \int_{0.1Vdd}^{Vdd-V_{TN}} \frac{dVout}{\frac{Vout^2}{2(Vdd-V_{TN})}-Vout} = \frac{C_L}{\beta_N(Vdd-V_{TN})} Ln\left(\frac{19Vdd-20V_{TN}}{Vdd}\right)$$

siendo $T_F = T_{FS}+T_{FL}$

El resultado se puede aproximar a la siguiente expresión

$$T_F \approx K \frac{C_L}{\beta_N Vdd}$$

Donde K puede valer entre 3 y 4 para

Valores del Vdd entre 3 y 5 V y

V_{TN} entre 0.5 y 1V

- **Conclusiones:**
  - ∗ Retardo directamente proporcional a C_L → a menor C_L menor retardo
  - ∗ Retardo inversamente proporcional a Vdd→ a menor Vdd mayor retardo
  - ∗ Retardo inversamente proporcional β → a mayor anchura o menor L→ menor retardo

ν **TIEMPOS DE SUBIDA T_R**

El estudio es idéntico al de los tiempos de bajada y la expresión final que se obtiene es:

$$T_R \approx K \frac{C_L}{\beta_P Vdd}$$

- **Conclusiones:**
  - ∗ Para transistores N y P de idéntico tamaño tenemos que $\beta_N=2\beta_P$ debido a las diferentes movilidades de los portadores → $T_F=T_R/2$
  - ∗ se puede asegurar que el inversor es más rápido descargando la capacidad que cargándola
  - ∗ Si se quiere $T_F=T_R$ → $\beta_N=\beta_P$→ $W_P=2\text{-}3W_N$

ν **TIEMPO DE RETARDO**

- Recordamos que el tiempo de retardo es el tiempo que se tarda en alcanzar el 50% del valor final
- Existen varias aproximaciones

$$T_{DR} \approx \frac{C_L V_{DD}}{\beta_P (V_{DD} - V_{TP})^2}$$

$$T_{DR} \approx \frac{C_L}{\beta_P V_{DD}} \qquad V_{TP} << V_{DD}$$

$$T_{DF} \approx \frac{C_L V_{DD}}{\beta_N (V_{DD} - V_{TN})^2}$$

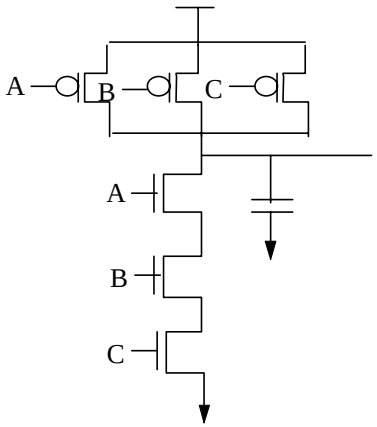
$$T_{DR} \approx \frac{C_L}{\beta_N V_{DD}} \qquad V_{TN} << V_{DD}$$

- Se puede aproximar como $T_{DR} = T_R/2$ y $T_{DF}=T_F/2$
- El tiempo de retardo promedio = $(T_{DR}+ T_{DF})/2$

$$T_D \approx \frac{C_L}{2Vdd}\left(\frac{1}{\beta p} + \frac{1}{\beta_N}\right)$$

7.6.2 INVERSOR EQUIVALENTE

- El retardo correspondiente a una puerta lógica se puede aproximar construyendo el inversor equivalente
- inversor equivalente es aquel cuyo transistor de pulldown y su transistor de pullup tienen un tamaño que refleja el comportamiento real de pulldown y de pullup de la puerta lógica.
- Vamos a calcular el inversor equivalente de una puerta NAND de tres entradas,



- **Transistor N equivalente**
  - Cuando se lleva a cabo la descarga de la capacidad todos los transistores N deben estar conduciendo, para que exista un camino a tierra
  - La ganancia del transistor N equivalente viene dado por la expresión

$$1/\beta_{Neq} = 1/\beta_{N1} + 1/\beta_{N2} + 1/\beta_{N3}$$

Suponiendo

$$\beta_{N1}=\beta_{N2}=\beta_{N3}=\beta_N$$

la expresión queda

$$\beta_{Neq}=\beta_N/3$$

La ganancia total de una rama con transistores en serie viene dada por la expresión

$$\mathbf{1/\beta_{Neq} = 1/\beta_{N1} + \ldots + 1/\beta_{N3}}$$

- **Transistor P equivalente**

Se pueden dar tres casos:

$\Rightarrow$ conduzca un transistor P

$\Rightarrow$ conduzcan 2 transistores P

$\Rightarrow$ conduzcan los 3 transistores P

- La diferencia está en la cantidad de intensidad que llega al punto Vout y por lo tanto lo rápido que será la carga del uno.
- **Siempre que existan ramas en paralelo se debe escoger la rama que proporciona la menor intensidad o la menor ganancia o la mayor resistencia, es decir el peor de los casos**
- ¿Porque?- Es preferible utilizar estimaciones más pesimistas para asegurarnos el buen funcionamiento del circuito en cualquier situación
- En el ejemplo caso debemos tomar el peor de los tres casos es decir cuando solo conduce 1 de los transistores.$\rightarrow \beta peq=\beta_P$
- **Si en uno de los bloque de lógica aparecen ramas serie y paralela mezcladas, se debe ir buscando la ganancia equivalente de cada una de las ramas hasta reducirla a un único y transistor con la ganancia equivalente que sea**

ν **TIEMPOS DE SUBIDA Y BAJADA DEL INVERSOR EQUIVALENTE**

$$T_{Req} \approx K \frac{C_L}{\beta_{Peq} Vdd}$$

$$T_{Feq} \approx K \frac{C_L}{\beta_{Neq} Vdd}$$

$T_{Requ}/T_{Fequ} = \beta_{Neq}/\beta_{Peq}=\beta_N/3\beta_P$

¿que relaciones de ganancia debían tener los transistores de la puerta NAND para que los tiempos de subida y bajada fueran idénticos?

$T_{Req}/T_{Feq}=1 \rightarrow \beta_P \approx \beta_N.0,3$

7.6.3 CIRCUITOS DE FAN-OUT ELEVADO

El incremento de las capacidades de interconexión, especialmente la de los hilos globales, enfatiza la importancia de tener circuitos drivers que puedan cargar y descargar las capacitancias con suficiente rapidez.

Esta necesidad se ve incrementada por el hecho de que los circuitos son cada vez más complejos lo que tiene como consecuencia que puertas sencillas tengan fan-outs cada vez más elevados y por lo tanto una carga también superior

- En un circuito integrado se pueden encontrar :

  Cargas on chip

  Cargas out chip
- Ejemplos de cargas elevadas on chip son:
  - ∗ Buses
  - ∗ Redes de reloj
  - ∗ Hilos de control
  - ∗ Memorias
- El peor caso sin lugar a dudas es el de las cargas off chip que en este caso se deben a:
  - ∗ Pines de empaquetamiento
  - ∗ Al hilado de la tarjeta impresa
- El rango típico de las capacitancias off chip se encuentra entre 20 y 50 **pico faradio,** tres ordenes de magnitud por encima de la carga normal de una puerta (**femtofaradios**)
- Situaciones de fan-out elevado son aquellas en que un pequeño transistor, con una capacidad de puerta C_G alimenta una capacidad C_L de salida muy grande:

$$C_G <<< C_L$$

- **Problema**
  - Conseguir cargar estos nodos con suficiente velocidad
  - Causa del problema

    ⇒ El retardo es proporcional a la capacidad de carga C_L

    ⇒ Por lo tanto la propagación del retardo off chip podría ser miles de veces mayor que el de nodo on chip normal, si el circuito driver que se utiliza es el mismo.
  - ∗ Falsa solución

    ⇒ Aumentando el tamaño de los canales del drive (W)

⇒ Efecto: el aumento del tamaño aumenta la carga de las etapas lógicas anteriores

⇒ el problema no se soluciona

- **Solución -** utilizar circuitos buffers entre la etapa lógica y la capacidad de carga C_L
- Antes de empezar vamos a ver que cuando se colocan varios inversores en cascada la razón entre los tamaños de los transistores N y P del inversor varía del caso en que solo hay un inversor.

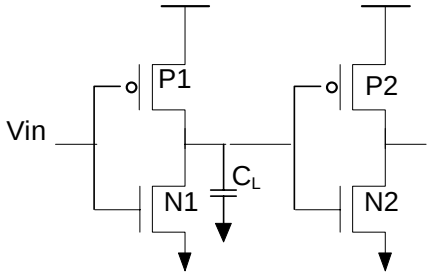
7.6.3.1 Etapas de transistores en cascada

- Ya se vio en su momento que para que los tiempos de subida y de bajada de un inversor CMOS aislado fueran iguales los tamaños de los transistores N y P que lo forman debían tener cierta relación:

$$(W/L)_P=\varepsilon(W/L)_N$$

Siendo $\boldsymbol{\varepsilon=\mu n/\mu p}$

- **Esto no es valido** cuando en lugar de un transistor hay varios en cascada, la relación es diferente y se va a estudiar a continuación



- La capacidad de carga del primer inversor es la suma de:
  - Sus propias capacidades de difusión
  - Las capacidades de fanout
  - Y la capacidad del hilo

$$C_L=(C_{DP1}+C_{DN1})+(C_{GP2}+C_{GN2})+CW$$

Donde

- C_{DP1} y C_{DN1} son las capacidades de difusión del transistor P y del N del primer inversor
- C_{GP2} y C_{GN2} las capacidades de puerta del segundo inversor
- C_W es la capacidad del hilo

Suponiendo que P es diseñado α veces mayor que N

$$\boldsymbol{\alpha=(W/L)_P/(W/L)_N}$$

Sabemos que las capacidades se escalan en la misma proporción que los canales luego

$$C_{DP}=\alpha C_{DN}$$

$$C_{GP}=\alpha C_{GN}$$

Luego

$$C_L = (1+\alpha)(C_{DN}+C_{GN})+C_W = (1+\alpha)C_N+C_W$$

La propagación del retardo se puede expresar como función de α sabiendo que:

$$T_D \approx \frac{C_L}{2Vdd}\left(\frac{1}{\beta p} + \frac{1}{\beta_N}\right)$$

$\beta_N/\beta_P=\varepsilon(W/L)_N/(W/L)_P$

$\alpha=(W/L)_P/(W/L)_N$

$$Tp= \frac{(1+\alpha)Cn+Cw}{(2VddKn)}\ (1+\varepsilon/\alpha)$$

El valor óptimo de α se determina haciendo $\partial tp/\partial\alpha=0$

$$\alpha_{optimo}=\sqrt{\varepsilon(1+ Cw/Cn)}$$

Nota : para un solo inversor el $\alpha=\varepsilon$

Para $C_W <<< C_N \rightarrow \alpha_{optimo} \approx \varepsilon^{1/2}$ (este es el caso de una cadena de inversores.)

Para $\varepsilon=2.5$ α toma el valor de 1.6

conclusión cuando se utilizan cadenas de inversores es mas ventajoso hacer el transistor PMOS más pequeño que la relación ε determinada para un inversor único.

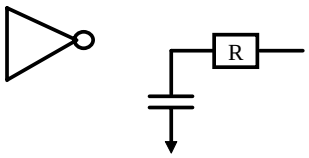
Esto da lugar a una respuesta de tiempos asimétrica pero optimiza la respuesta total del sistema., al tiempo que reduce el área.

7.6.4 SOLUCIÓN AL PROBLEMA DEL FAN-OUT ELEVADO

- Solución es colocar buffers en cascada cada uno un poco mayor que el anterior
- Vamos a calcular la disminución del retardo total al introducir los buffers y la razón de crecimiento del buffer entre dos etapas
- def.- Td retardo promedio de un inversor de tamaño mínimo que conduce a otro inversor de tamaño mínimo viene dado por

$$Td=(Tr+Tf)/2$$

- sustituyendo un inversor por su modelo de interruptor RC tenemos lo siguiente



- el retardo en el inversor viene dado por la expresión

$$Td=RmCm \ [1]$$

donde

Rm resistencia de mínimo canal

Cm capacidad de mínimo canal

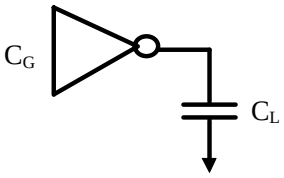
- suponiendo aproximaciones lineales para Rm y Cm tenemos

$$R \propto L/W$$

$$C \propto LW$$

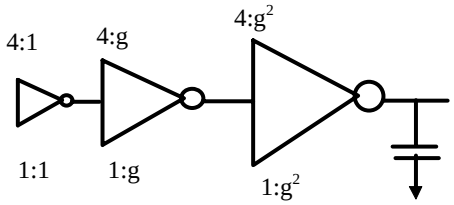
y sustituyendo en [1] se tiene que $T_D=L^2$

- T_D es una constante para todos los inversores que tengan la misma longitud de canal independientemente de la anchura
- Se va a estudiar el siguiente caso



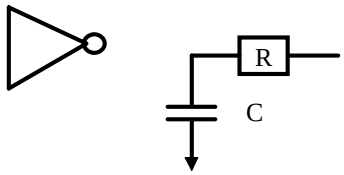
donde $C_G <<< C_L$

- añadimos M=2 etapas de inversores cada uno de ellos g veces mas ancho que el anterior hasta que el último tiene un canal lo suficientemente grande como para cargar C_L en un tiempo mínimo.



donde N:M N longitud del canal y M es la anchura

Sustituimos cada inversor por su modelo de interruptor RC

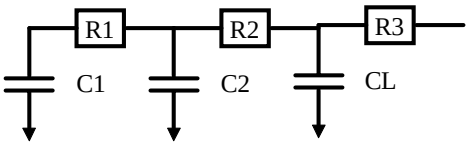


donde

C es equivalente a la capacidad de entrada

R la resistencia que opone el inversor al conducir

quedando:



Se puede ver el retardo de la transmisión como

$$T=R_1C_2+R_2C_3+R_3C_L$$

suponemos que $C1=C_g$ no interviene porque empezamos a contar a partir de que el inversor 1 ha empezado a conducir, es decir a partir de que C_1 ya se ha llenado

$$T=\sum_{I=1}^{M+1=3} Ri\cdot C_{i+1} \qquad [2]$$

sabemos que

$RiCi = T_D$ = constante para inversores de la misma longitud de canal

luego

$Ri=Td/Ci$

y sustituyendo en [2] se obtiene:

$$T=T_D\sum_{I=1}^{M+1=3}\frac{C_{i+1}}{C_i}$$

además hemos supuesto que

$$C_{i+1}=g\cdot C_i \quad [3]$$

donde g se llama razón de etapa

$C_2=g\cdot C_1$

$C_3=g\cdot C_2=g^2\cdot C1$

$C_4=g\cdot C_3=g^3\cdot C1$

como 3=M+1

$C_1=C_G$

$C_4=C_L$

$C_L=g^{M+1}C_G$ [4]

y de [2] y [3] se obtiene

$$\mathbf{T=Td(M+1)g}$$

de [4] se obtiene

$$\mathbf{g=(C_L/C_G)^{1/(M+1)}}$$

para una relación C_L/C_G dada, al aumentar el número de etapas M disminuye la razón de etapa g

ν **OPTIMIZACIÓN DEL RETARDO T**

- Según se aumenta el Nº de etapas disminuye el retardo hasta llegar a un mínimo a partir del cual empieza a crecer
- De las dos siguientes expresiones:

  $T=Td(M+1)g$

  $M+1=Ln(C_L/C_G)$

se obtiene el valor del retardo

$$T=Td\cdot\frac{Ln\left(\frac{C_L}{C_G}\right)}{Lng}\cdot g$$

para hallar el mínimo se deriva con respecto a g y se iguala a cero:

$$T=F(g)\rightarrow dT/dg=0\rightarrow \mathbf{g=e}$$

$$\mathbf{T=e\cdot T_d\cdot Ln\left(C_L/C_G\right)}$$

para un C_L/C_G dado el retardo es mínimo si se toma un **g=e**

fijando g fijamos M

se ve que el retardo con etapas depende de C_L/C_G

atención C_L/C_G nunca podrá ser 1 porque

$$g=(C_L/C_G)^{1/(M+1)} = e=(1)^{1/(M+1)}\neq e$$

si la expresión (1)

$$T=T_D\sum_{I=1}^{M+1}\frac{C_{i+1}}{C_i}$$

se particulariza para un inversor de tamaño mínimo, sabiendo que $C_{i+1}=C_L$ y $Ci=C_G$ y se obtiene:

$$T_1=T_dC_L/C_G$$

que es el tiempo que tarda en cargar C_L el inversor sin etapas adicionales

Por otro lado la expresión

$$T=eT_dLn\left(C_L/C_G\right)$$

es el tiempo que tarda en cargar la capacidad CL M etapas con una razón de etapa g=e

Cuando C_L/C_G pequeño $T\approx T_1$ pero el área es tan grande que no merece la pena la utilización de etapas en este caso

C_L/C_G grande $T<<<T_1$ en este caso consumo de área grande pero delay pequeño

7.7 DISIPACIÓN DE POTENCIA

- El consumo de potencia se descompone en
  - ∗ Potencia estática
  - ∗ Potencia dinámica

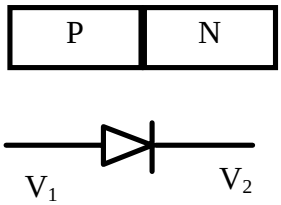
7.7.1 POTENCIA ESTÁTICA

- Consumo de potencia debido a las perdidas de intensidad desde Vdd en las fases estables de las señales
- En los dispositivos CMOS solo existe camino Vdd→ Vss en los transitorios
- cuando las señales son estables

  Vin=1 P cortado y N conduce

  Vin=0 P conduce y N cortado
- En este caso se puede suponer el consumo estático =0
- Sin embargo no se ha tenido en cuenta
  - ∗ La intensidad inversa en los diodos parásitos
  - ∗ La conducción subumbral
- Ambas intensidades son pequeñas pero existen
- Aunque su magnitud es despreciable frente al consumo de potencia dinámica vamos a estudiar el consumo debido a las corrientes inversas de los diodos

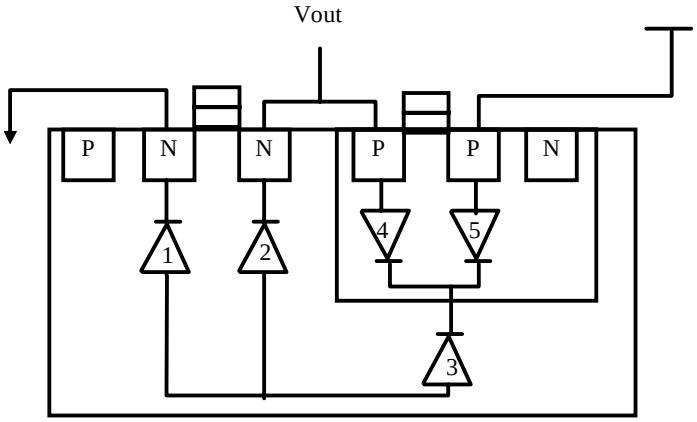
Recordando que:



cuando $V_1>V_2$ el diodo conduce en directa

cuando $V_1<V_2$ el diodo conduce en inversa

Cuando $V_1=V_2$ el diodo no conduce



ν **ESTUDIO DE LOS DIODOS PARÁSITOS CUANDO VIN=0**

D_1=Vss-Vss cortado

D_2=0-5 inversa

D_3=0-5 (el 5 debido a la polarización del pozo) inversa

D_4=5-5 cortado

D_6=5-5 cortado

podemos suponer D_1 y D_3 conduciendo en paralelo

ν **ESTUDIO DE LOS DIODOS PARÁSITOS CUANDO VIN=5**

D_1=0-0 cortado

D_2=0-0 cortado

D_3=0-5 inversa

D_4=0-5 inversa

D_5=5-5 cortado

Podemos suponer D_3 y D_4 conduciendo en paralelo

La intensidad inversa de un diodo viene dada por la expresión

$$\mathbf{I_{diodo}=I_S(^{eqv/kt}-1)}$$

para V<0:

$$(1/e^{qv/kt})-1\approx-1$$

luego $\mathbf{I_D=-I_S}$

siendo

q carga del electrón 1.6º2.10-19

K constante de boltzman 1.38.10-23 J/K

T temperatura

I_S corriente inversa de saturación

para un inversor el consumo de potencia estática viene dado por:

$$\mathbf{P_S = Is \cdot Vdd}$$

Suponiendo que hay N inversores funcionando la potencia estática total viene dada por

$$\mathbf{PS = N \cdot I_S \cdot Vdd}$$

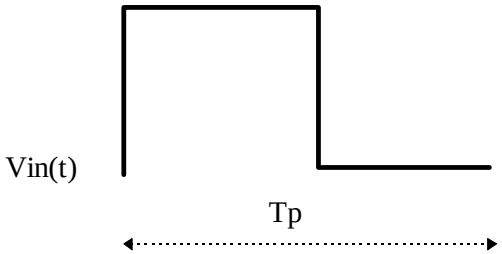
7.7.2 POTENCIA DINÁMICA

Se descompone en

- **P_{SC}**
  - consumo de potencia debida a la existencia de un camino entre Vdd-Vss (short current)
  - 10% del total
- **Pd**
  - consumo de potencia debido a la carga y descarga de las capacidades
- **Pd>>Psc**

7.7.2.1 Potencia debido a las capacidades

Suponemos además los tiempos de subida y bajada despreciables frente a tiempo de ciclo Tp



- La potencia tendrá en cuenta
  - en los primeros $T_P/2$ la descarga de la Capacidad
  - en los últimos $T_P/2$ la carga de la capacidad

Le expresión del consumo de potencia debido a la carga y descarga de las capacidades es:

$$Pd= \frac{1}{Tp}\int_{0}^{Tp/2} In(t)\cdot Vout\cdot dt + \frac{1}{Tp}\int_{Tp/2}^{Tp} Ip(t)\cdot(Vdd-Vout)\cdot dt$$

El primer término se debe a la in de descarga de la capacidad

el segundo término a la carga de la capacidad

Como:

$i_N(t)= C_L\cdot dVout/dt$

$i_P(t)=C_L\cdot d(Vdd-Vout)$

Sustituyendo modificando los límites de integración0→Vdd y Vdd→0 se obtiene:

$$Pd= \frac{C_L}{Tp}\int_{0}^{Vdd} Vout\cdot dVout + \frac{C_L}{Tp}\int_{Vdd}^{0} (Vdd-Vout)\cdot d(Vdd-Vout)$$

Integrando se obtiene

$$Pd= \frac{C_L\cdot Vdd^2}{Tp} = F_P \cdot C_L\cdot Vdd^2$$

Donde F_P es la frecuencia de trabajo

La potencia dinámica es proporcional

La frecuencia de trabajo

Proporcional a cuadrado de la fuente de alimentación

Es independiente de los parámetros de dispositivo

7.7.2.2 Disipación de corto circuito

$Psc=I \cdot Vdd$

- Recordar que por estar en serie P y N cumplen $|I_N|=|I_P|$
- Mientras estén los dos conduciendo siempre hay uno, como poco, que está en saturación→ se puede utilizar la expresión de la I en saturación

$$I=\beta/2(Vin(t)-Vt)^2$$

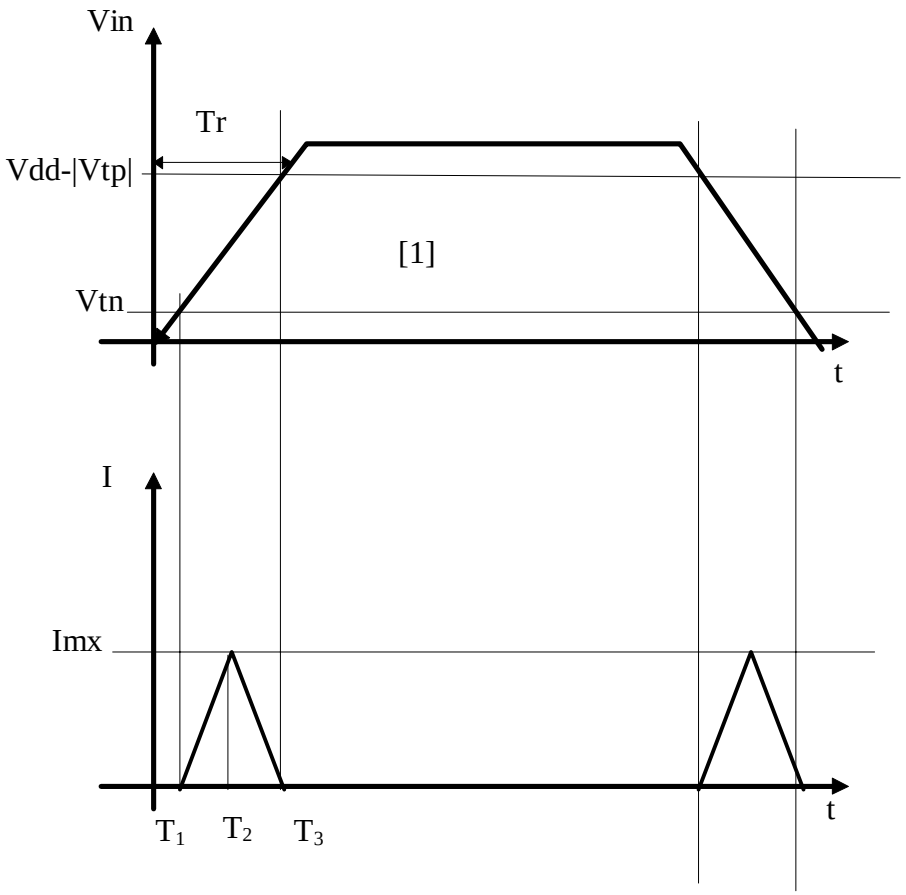
Para el estudio suponemos

$Vtn=-Vtp$

$\beta_N=\beta_P$

Por lo tanto la intensidad consumida en el transitorio de subida es igual que la consumida en el transitorio de bajada

Vamos a ver en la siguiente figura como se debe realizar el estudio:



$$I = 2\left[\frac{1}{Tp}\int_{T_1}^{T_2} I(t)\cdot dt + \frac{1}{Tp}\int_{T_2}^{T_3} I(t)d\right]$$

Ademas se puede suponer

T_2-T_1=T_3-T_2

Luego la expresión queda

$$I = \frac{4}{Tp}\int_{T_1}^{T_2} I(t)\cdot dt$$

Como hemos supuesto que La menor intensidad que atraviesa ambos transitores en todo momento es la de saturación tenemos que

$$I = \frac{4}{Tp}\int_{T_1}^{T_2} \beta/2(Vin(t)-Vt)^2\cdot dt$$

De la figura [1] se puede obtener

Vin(t)=Vdd·t/Tr [2]

para encontrar el valor de T1 se sabe que Vin = V_{TN} , sustituyendo en [2] y despejando t se obtiene

$$T_1 = V_T\ Tr/Vdd$$

Para encontrar el valor de T2 se sabe que Vin=Vdd/2 sustituyendo en[2] y despejando t se obtiene

$$T_2 = Tr/2$$

el resultado final del consumo de potencia dinámica debido a las corrientes de corto circuito es

$$P_{SC} = \frac{\beta}{12} (Vdd-2Vt)^3 \frac{T_R}{T_P}$$

Donde

$T_R=T_F$

T_P= periodo

este cálculo se lleva a cabo suponiendo que no hay que cargar ni descargar capacidades

se puede observar que el consumo de potencia es proporcional a Tr cuanto más lentos son los circuitos mas potencia consumen.

7.7.3 RELACIÓN DEL CONSUMO DE POTENCIA CON LA TEMPERATURA

- Aunque el consumo de potencia de los CMOS es mucho menor que la de los circuitos integrados bipolares que implementan la misma función es un factor que limita bastante el tamaño de los circuitos integrados
- La disipación de potencia se convierte en calor que debe eliminar a través del empaquetamiento
- El empaquetamiento ofrece una resistencia a la eliminación del calor → diferencia de temperatura en el empaquetamiento similar a la diferencia de potencial eléctrica
- Dado
    - Un tipo de empaquetamiento,
    - La mx temperatura que soporta el circuito integrado
    - una temperatura ambiente
- Se puede especificar el mx consumo de potencia permitido que puede tener el circuito

$$P = \frac{T_J - T_A}{\theta_{JA}}$$

donde

P potencia

T_J mx temperatura en el CI

T_A temperatura ambiente

θ_{JA} resistencia térmica

8. LÓGICA COMBINACIONAL ESTÁTICA

ν **LÓGICA COMBINACIONAL**

- no realimentada
- En cualquier instante de tiempo las salidas están relacionadas con las señales de entrada mediante expresiones booleanas
- Esta definición ignora los retardos de propagación de las puertas
- La lógica combinacional forma el corazón de la mayoría de los circuitos integrados como son l
    - Las unidades aritmético lógicas
    - Los controladores

ν **LÓGICA SECUENCIAL**

- La salida es función de la entrada y de la historia del sistema
- Se caracteriza por la realimentación

ν **LIGADURAS DE UN CIRCUITO**

- Pueden variar enormemente, pero están limitadas por:
    - El área
    - El rendimiento
    - Consumo de potencia
- El área es la ligadura de mayor importancia, ya que es la que determina el coste final del Circuito Integrado
- En la actualidad el rendimiento también es muy importante
- El consumo de potencia es crucial en el diseño de aplicaciones transportables.

- La existencia de diferentes ligaduras en los circuitos tiene como efecto la utilización de diferentes estilos de diseño y tecnologías de fabricación para satisfacerlas

8.1 DISEÑO CMOS ESTATICO

ν **DISEÑO ESTÁTICO**

- Es en cualquier instante de tiempo la salida esta conectada a Vdd o gnd. A través de un camino de baja resistencia.
- La salida toma en cualquier instante de tiempo el valor de la función implementada por el circuito

ν **DISEÑO DINÁMICO**

- Utiliza las capacidades parásitas para almacenar momentáneamente información.
- Puertas mas simples y rápidas.
- Su diseño y operación son mas complejos. Incremento de la sensibilidad al fluido

ν **INVERSOR CMOS ESTÁTICO**

- Tiene un excelente comportamiento:
  - ∗ Baja sensibilidad al ruido.
  - ∗ Buena velocidad
  - ∗ Bajo consumo

Este comportamiento es extrapolable al resto de las puertas lógicas,

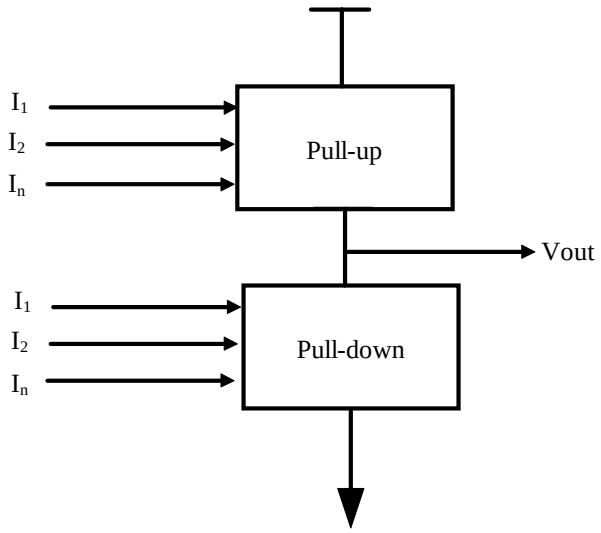
INCONVENIENTE

- Puertas complejas son grandes y lentas.
- Este es el motivo por el que aparecen otros estilos de diseño

A continuación se estudia en detalle algunas lógicas estáticas

8.2 LÓGICA CMOS COMPLEMENTARIA

- Lógica estática
- Combinación de dos redes:
- Pull-Up
  - ∗ Carga el 1 de máxima dureza
  - ∗ PMOS
- Pull-Down
  - ∗ Carga el 0 de máxima dureza
- Estas dos redes deben cumplir la condición de que en las fases estáticas una y solo una de ellas conecta la salida con la fuente.



- Al diseñar una puerta lógica CMOS complementaria se deben tener en cuenta las siguientes consideraciones:
  - ∗ Conexiónes serie de interruptores equivale a puertas AND
  - ∗ Conexiónes paralelas de interruptores equivale a puertas OR
  - ∗ Las puertas de Pull-Up y Pull-Down son duales
    - ⇒ La conexión de interruptores en paralelo en el árbol de Pull-Up se corresponde con interruptores serie en el Pull-Down

- La lógica CMOS es inversa
  - ⇒ Implementa siempre funciones negadas NAND, NOR, NOT.
  - ⇒ Para diseñar puertas no invertidas hay que utilizar una etapa de inversores.

ν PROPIEDADES

- Márgenes de ruidos altos:
  - Voh=Vdd
  - Vol=Gnd
- No existe consumo de potencia estática
- No existe camino entre Vdd y Gnd en continua
- Tiempos de subida y bajada comparables, con las condiciones de escalado apropiadas

ν INCONVENIENTES

- Aparecen sobre todo en las puertas de **Fan-in** elevado
  - Una puerta de N entradas necesita 2 N transistores
    - ⇒ Aunque para puertas pequeñas puede no ser significativo, para puertas complejas si lo es
  - La propagación del retardo se deteriora rápidamente al aumentar el Fan-iN
    - ⇒ El gran numero de transistores aumenta la capacidad total de puerta
    - ⇒ Los transistores en serie aumentan la resistencia
- El **Fan-Out** también tiene un elevado impacto en los retardos
  - Cada salida conecta un NMOS y un PMOS
- Los efectos del Fan-In y Fan-Out en el retardo de propagación se pueden expresar:

$$T_P = a_1 FI + a_2 FI^2 + a_3 FO$$

donde

FI = Fan-IN

FO= Fan-Out

a1,a2,a3 factores de peso que dependen de la tecnología

8.2.1 TÉCNICAS DE DISEÑO PARA EVITAR LOS PROBLEMAS DE FAN-IN ELEVADO

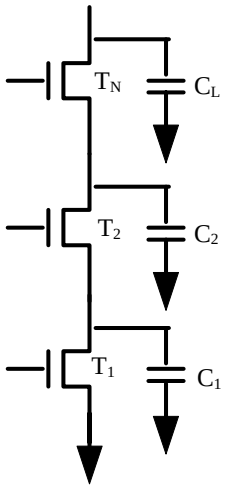
ν **AUMENTAR EL TAMAÑO DE LOS TRANSISTORES.**

Solo es valido en algunas ocasiones.

Se puede producir un efecto contrario puesto que se aumenta la capacidad

ν **TAMAÑO PROGRESIVO DE LOS TRANSISTORES.**

- Hasta el momento se ha hablado de transistores serie se ha simplificado el estudio suponiendo que solo existe la capacidad C_L de salida de la puerta lógica.
- El esquema representado a continuación se aproxima más a la realidad



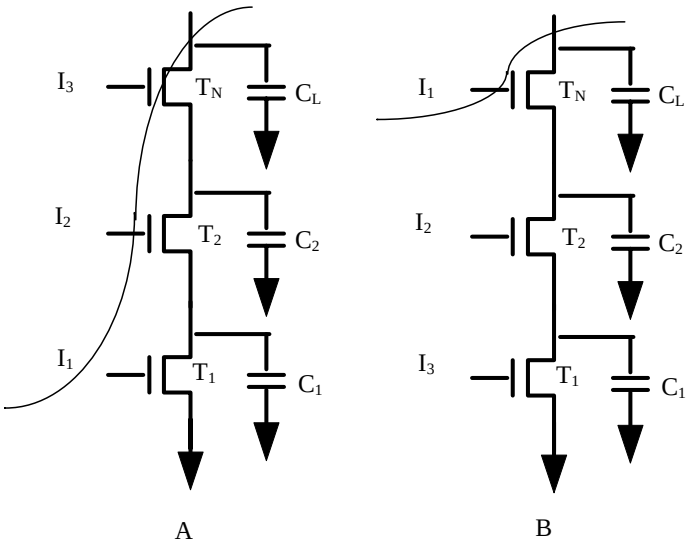
- Suponiendo que se esta descargando la capacidad de salida
- El transistor N-esimo debe conducir la I que proviene de C_L,
- El transistor T_1 debe conducir la I que viene de C_l, ... , C_3, C_2 , C_1
  - ∗ Es decir, debe conducir mucha mas intensidad.
- Una manera de tener esto en cuenta:

$$T_1 > T_2 > ... > T_N$$

- Esta técnica se utiliza Habitualmente en los descodificadores de memoria que utilizan puertas de Fan-In elevado

ν **REORDENAMIENTO DE LOS TRANSISTORES DE LA PUERTA.**

- No todas las señales llegan simultáneamente a las entradas de una puerta lógica
- Una señal se llama critica si es la ultima señal de salida que alcanza un valor estable
- Camino critico es el camino a través de la lógica que determina la velocidad de la estructura.
- Las señales criticas deben situarse cerca de las salidas
- Ejemplo



Vamos a suponer que I_3=I_2=1 cuando I_1 todavía no ha empezado a cambiar

En el caso A

la señal critica es la mas alejada de la salida.

Aunque I_3 e I_2 están en condiciones de conducir no lo hacen porque I_1 les cierra el camino a tierra.

Cuando por fin I_1=1 se debe descargar las capacidades C_l, C_1 y C_2

En el caso B

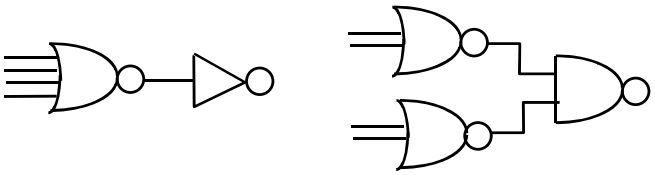
I_2 yI_3 están en condición de conducción

C_2 y C_3 se descargan. Cuando por fin I_1 se pone a 1 solo debe descargar C_L

Mas rápido

ν **MEJORAS EN EL DISEÑO LÓGICO**

Manipular las ecuaciones lógicas para reducir las ligaduras de Fan-In y por tanto reducir el retardo



ν **USAR OTRO ESTILO DE DISEÑO**

esta opción es la que estudiaremos en el resto del tema a continuación.

Las soluciones vistas solucionan parte de los problemas generados por el fan-in

En cuanto al el Fan-Out es el factor que domina la velocidad de las puertas en muchas ocasiones

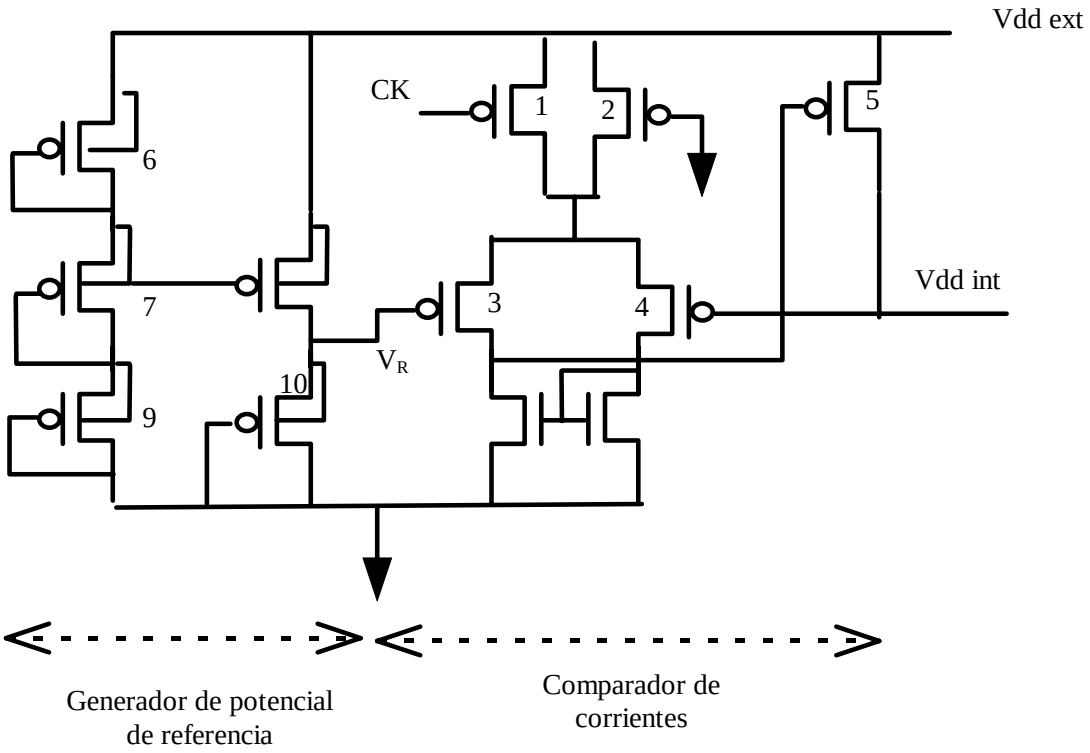
Cargar grandes capacidades mediante puertas complejas es caro, debido a que todos los transistores deben ser escalados hacia arriba para poder proporcionar suficiente intensidad.

Es mejor introducir los Buffers como ya vimos en su día. Así se puede minimizar el tamaño de la puerta compleja.

8.3 REGULADORES DE POTENCIAL INTERNOS

Suele ser habitual utilizar potencias inferiores a la de alimentación en algunas partes del diseño.

Esto da lugar a reguladores de potencia en el interior del Circuito integrado



ν **Circuito generador**

genera un V_R < Vddex debido a que todos los dispositivos están conduciendo VR = Cte.

en el no se ven reflejados las modificaciones que puedan ocurrir en Vddex.

ν **Circuito comparador**

- Intenta mantener un Vddint = constante independientemente del Vddext utilizando V_R como potencial de referencia.
- Esto lo consigue mediante una realimentación
  - ∗ La salida Vddint. de P5 es la entrada de P4.

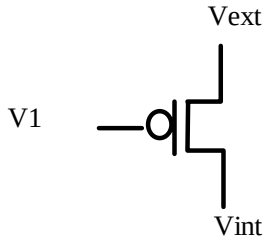
- La salida de P3 es la entrada de P5.

Vamos a ver el funciónamiento del comparador:

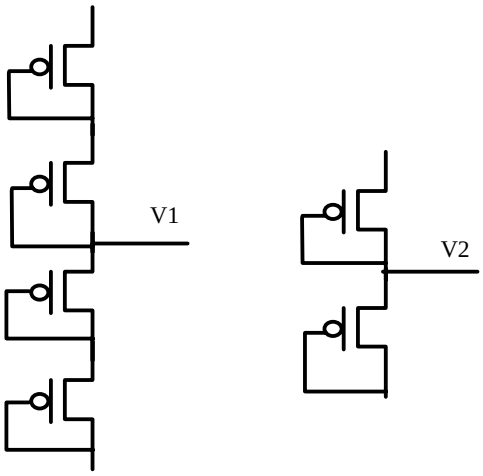
Si disminuye V_{DD} int→ aumenta I_4→disminuye I_3→ aumenta V_{OUT3} → disminuye I_5→aumenta V_{DD} int

V EL VOLTAJE DE REFERENCIA

- Los circuitos que generan un potencial independiente del Vddex o de los cambios de temperatura se llaman generadores de potencial de referencia. pp 354 McGraw-Hill.
- En nuestro caso, si se colocara solo el circuito y regulando convenientemente V_1.

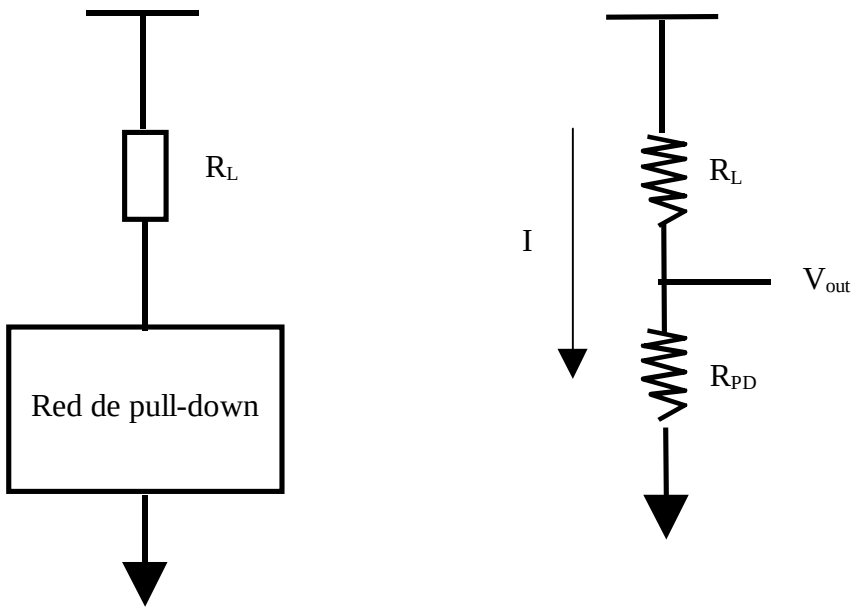


- Cualquier variación de Vext produce cambios en Vin introduciendo errores
- Para evitar esto se coloca el circuito generador de potencial y el circuito comparador vistos.
- En el circuito potencial de referencia hay que tener en cuenta que el generador de la izquierda es más robusto que el de la derecha pero V1<V2:



8.4 LÓGICA PROPORCIONAL

- Se componen de
  - ∗ Red Pull-Down NMOS que realiza la función lógica
  - ∗ un dispositivo sencillo de carga
- La carga puede ser :
  - ∗ un dispositivo pasivo como una resistencia o un
  - ∗ elemento activo como un transistor
- Vamos a suponer que tanto la carga como el bloque lógico se representan por resistores linealizados cuando los potenciales de entrada = 1.
- Tanto el inversor de carga como el bloque lógico conducen:



Vdd/Rl+Rpd = Vout/Rpd

$$\mathbf{Vout = Vol = \frac{Rpd\ Vdd}{Rl+Rpd}}$$

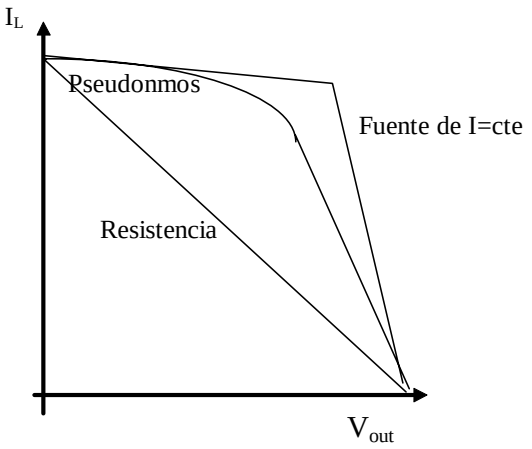
- Para que los márgenes de ruido sean aceptablemente bajos Rpd <<< R_L
- A esta lógica se le llama proporcional porque hay que tener cuidado con la elección del tamaño de los transistores (resistencias) para que las puertas funcionen correctamente.
- Por otro lado la carga R_L debe ser tal que permita el suficiente paso de corriente para cargar el "1" lo mas rápido posible. Vamos a estudiar brevemente tres casos diferentes:

8.4.1 CARGA A TRAVES DE UNA RESISTENCIA

- La I_L en función del potencial de salida

$$\mathbf{I_L = (Vdd - Vout\ )/\ R_L}$$

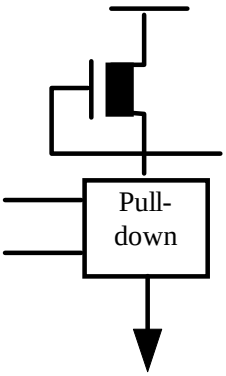
La carga del "1" se hace lenta porque la I disminuye linealmente con el aumento del Vout



- Esta figura
  - ∗ Refleja la variación de la intensidad de carga frente al Vout.
  - ∗ Cuanto mayor sea la disminución del I del circuito mas lento es.
  - ∗ El comportamiento ideal es el de una fuente de intensidad constante
  - ∗ Utilizando una resistencia el tiempo de subida es un 25% mas lento
    - ⇒ en la figura se ve que la disminución de la intensidad con el Vout es muy grande
- el área necesaria para fabricar una resistencia es muy grande.
- ejemplo
  - ∗ para generar una Resistencia de 40 Kohmios se necesitan 3,200 nM^2
  - ∗ Material utilizado utilizando difusión N+
  - ∗ Espacio en el que entran mas de un centenar de transistores de tamaño mínimo

8.4.2 CARGA MEDIANTE UN TRANSISTOR DE DEPLEXION

- Fue la primera configuración que se utilizo para implementar la carga.
- Tuvo gran popularidad en la era NMOS
- El transistor conduce siempre, incluso cuando Vgs = Vout =0, debido a que su potencial umbral es negativo



Inicialmente la I_L se comporta como la de una fuente de intensidad y se puede aproximar a su corriente en saturación:

$$I = \frac{\beta_N |V_{TN}|^2}{2}$$

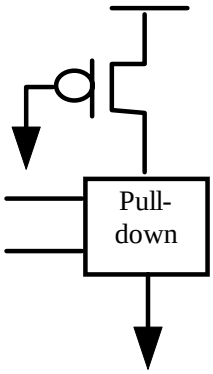
que se obtiene aproximando Vgs y Vds a 0 en la expresión

$$\mathbf{Is = (K \cdot W/2 \cdot L)\ (Vgs - V_{TN})^2\ (1+\lambda \cdot Vds)}$$

- Se desvía el comportamiento de la fuente de I por dos razones:
  - ∗ La modulación de la longitud del canal λ que modula la I disponible en modo saturación
  - ∗ **Efecto Cuerpo**
    - ⇒ Disminuye el valor de $|V_{TN}|$ (ya que $V'_{TN}=-|V_{TN}|+K$)→ disminuye la I_L para valores crecientes de Vout.
    - ⇒ El efecto cuerpo produce una importante degeneración de la característica de carga.
    - ⇒ Aun así su funcionamiento es mucho mejor que la carga mediante una resistencia

8.4.3 CARGA A TRAVÉS PSEUDO-NMOS

- Es la mejor de las 3 aproximaciones
- Se llama así porque se parece al NMOS estudiado



- No experimenta efecto cuerpo→ Vsb = 0
- Vgs = - Vdd
- Proporciona la corriente de carga mayor de todos los dispositivos de carga de tamaño similar que se pueden utilizar
- La corriente viene dada por:

$$I_L = \beta p/2\ (V_{DD} - |V_{TP}|)^2$$

- La gráfica correspondiente a esta intensidad aparece en la figura I=F(Vout)
  - Debemos recordar que en la gráfica las I están normalizadas.
  - En la gráfica absoluta el valor se acerca mucho más al valor ideal, aunque el pseudo NMOS abandona el modo saturación antes que el NMOS de empobrecimiento. Esto se puede deber a que el valor de Vgs es mayor.
  - Cuando se comparan las curvas del pseudoNMOS y del NMOS de empobrecimiento (ambas con un tamaño similar) se ve que el pseudoNMOS proporciona mas corriente aunque la movilidad de los portadores sea más pequeña esto tiene como consecuencias un comportamiento AC y DC mejor

- La característica DC del pseudo NMOS es similar a la del CMOS:
  - ∗ En ambos casos $V_{OH} = V_{DD}$
  - ∗ La diferencia aparece para el Vol que en los pseudo NMOS ≠Vgnd debido a la naturaleza proporcional
- **Búsqueda del Vol**
  - ∗ se puede obtener igualando las corrientes cuando V_{IN}=1
  - ∗ En este punto los transistores se encuentran en las siguientes regiones de trabajo

    ⇒ N lineal

    ⇒ P saturación
  - ∗ Igualando las intensidades se obtiene

    $\beta n [\{Vdd - Vtn\} Vol - (Vol^2/2)] = \beta p/2 (Vdd - lVtpl)^2$

Suponiendo lVtnl=lVtpl=Vt

$$Vol = (Vdd - Vt) \left(1- \sqrt{1- \frac{\beta_P}{\beta_N}}\right)$$

- **Búsqueda del potencial umbral de puerta**
  - ∗ La condición que debe cumplir es Vin = Vout
  - ∗ Se igualan las intensidades
  - ∗ En ese punto los transistores se encuentran en las siguientes regiones de trabajo:

    ⇒ N saturación

    ⇒ P lineal

$$V_M=V_T + (Vdd - Vt)\sqrt{\frac{K_P}{K_P+K_N}}$$

ν VENTAJAS FRENTE A CMOS COMPLEMENTARIA

- Menor área
  - ∗ N+1 transistores por cada entrada,
- Menor capacidad parásita.
  - ∗ Las capacidades de carga de entrada se reducen a la mitad

ν PROBLEMAS

- Consumo de potencia estática cuando la salida esta a 0

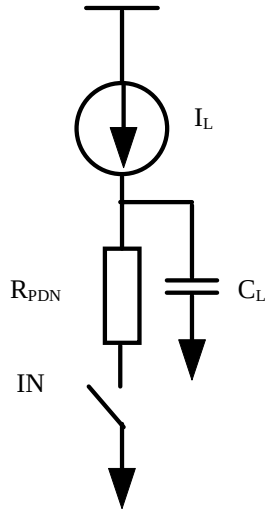
  debido a que un camino directo entre Vdd y gnd.
- El consumo promedio de potencia viene dado por

$$\mathbf{Pbaja = Vdd\ I_{LOW} = \beta p/2Vdd\ (\ Vdd - Vt\ )^2}$$

- La característica de transferencia es asimétrica,
  - ∗ El valor Vm no se encuentra en el medio del valor del voltaje
  - ∗ Los tiempos de subida y bajada no son iguales
  - ∗ Los caminos de carga y descarga no son iguales:
    - ⇒ La carga a través del dispositivo PMOS
    - ⇒ La descarga a través de la red de transistores N

8.4.3.1 Ajuste del comportamiento

Vamos a estudiar sobre el circuito equivalente de la figura como ajustar el comportamiento de los pseudo NMOS



1. Para reducir el consumo de potencia, I_L debería ser pequeña
2. Para obtener un margen de ruido NM_L razonable el producto Vol = I_L Rpdn debería ser bajo
3. Para reducir el tiempo de subida $T_{LH} = C_L Vdd/2I_L$ I_L debería ser alta
4. Para reducir el tiempo de bajada T_{HL} = 0.69 Rpdn Cl→ Rpdn debería ser pequeña

- Para que el margen de ruido sea razonable debe existir una relación entre los tamaños de los transistores:

$$\gamma = (W/L)n / (W/L)p$$

  - Esta relación depende de la tecnología y del margen de ruido deseado
  - Ejemplo Para Vol = 0.2 V en una tecnología de 1.2nm CMOS con Vdd = 5V se necesita γ=3.

- Para reducir el tiempo de bajada R_{PDN} debe ser pequeña → los NMOS deben ser anchos.
- las condiciones 1 y 3 son opuestas. Una puerta mas rápida implica mayor consumo y menor margen al ruido

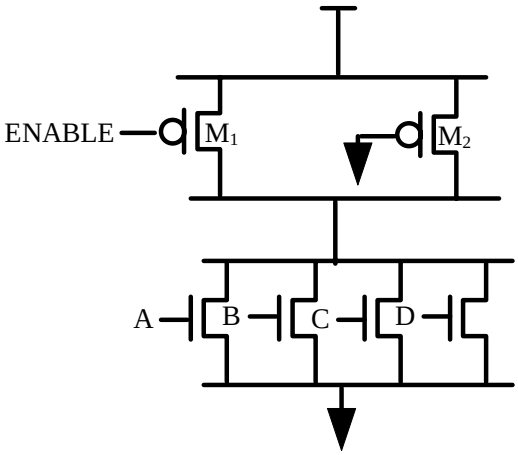
8.4.3.2 Consideraciones de diseño

- Atractivo para puertas complejas de elevado Fan-In.
- Su elevado consumo de potencia elimina la aplicabilidad de los pseudo NMOS a circuitos grandes.
- Las puertas de tamaño mínimo consumen 1mw
    - ∗ 100000 puertas consumen 50w (suponiendo que solo la mitad de las puertas estén a baja
    - ∗ Todo sin tener en cuenta el consumo de potencia dinámica
- Son útiles para
    - ∗ Pequeños circuitos en los que la velocidad es importante
    - ∗ En circuitos en los que se sabe que la mayoría de las salidas permanecen en alta y por lo tanto no existe consumo de potencia.

        ⇒ Este ultimo caso descodificadores de direcciones de las memorias.

8.4.3.3 Modificaciones del dispositivo de carga PMOS

- Buscan:
  - ∗ Mejora de la velocidad
  - ∗ Mejora del consumo de potencia.
- Consideremos el caso de una puerta lógica que se activa solamente durante determinados intervalos de tiempo
  - ∗ descodificadores de direcciones de memoria
  - ∗ Funciona solo después de un cambio de dirección.
  - ∗ En modo espera deberían tener bajo consumo de potencia y grandes márgenes de ruido en
  - ∗ debería funcionar rápidamente al detectar un cambio de dirección, aunque fuera a expensas del consumo de potencia
  - ∗ Todo esto se consigue en el circuito de la figura



  - ∗ Se añade un transistor M1 a la carga controlado por una señal de capacitación
  - ∗ Los dos transistores no tienen el mismo tamaño

    M_2 es pequeño

    M_1 es grande
  - ∗ Similar a la lógica dinámica que se trata mas adelante
  - ∗ En el modo de espera el dispositivo M_1 esta cortado

* El dispositivo M_2 tiene

  $\Rightarrow$ una gran resistencia

  $\Rightarrow$ bajo consumo de potencia

  $\Rightarrow$ un valor reducido de Vol.

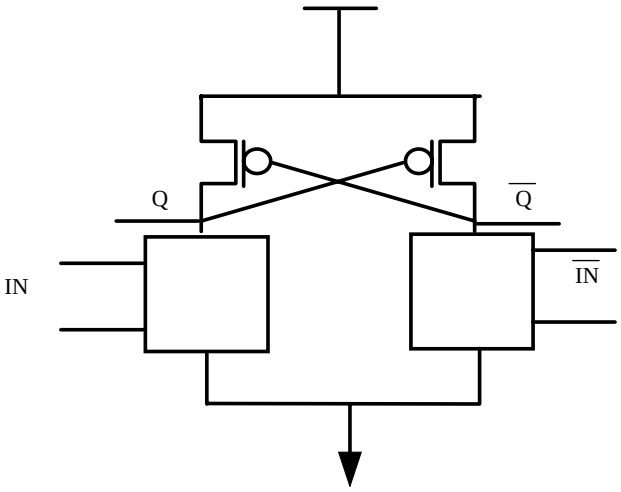
* Después de detectar un cambio de dirección el M_1 pasa a conducir provocando

  $\Rightarrow$ una gran cantidad de corriente

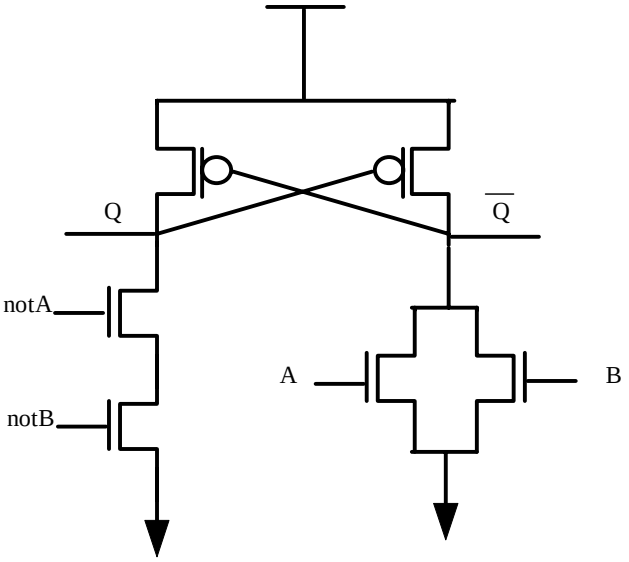
  $\Rightarrow$ un transitorio a 0 muy rapido.pg72.

8.4.4 DIFERENTIAL CASCADE VOLTAGE SWITCH (DCVS)

- Carga a través de un PMOS
- Esta lógica permite eliminar completamente la corriente estática
- Cada puerta genera las dos polaridades de la señal de salida



- Las redes de descarga implementan una función y la inversa.
- Para un conjunto de entradas un nodo conduce y el otro no
- Vamos a ver un ejemplo



Q = not(A + B) = notA notB

notQ = not(notA not B) = A + B

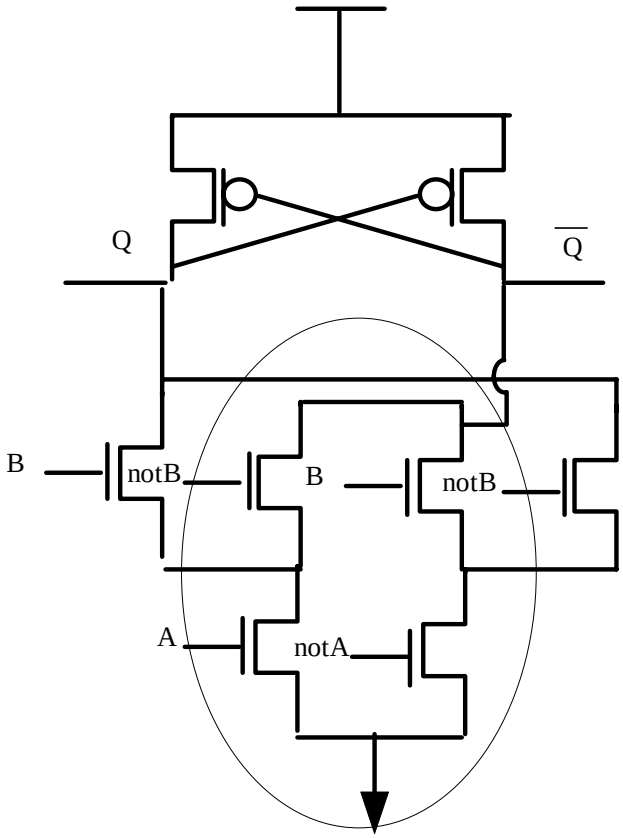
ν **VENTAJAS**:

- Velocidad similar a los pseudo NMOS (debido a que la reducción de las capacidades parásitas de los nodos de salida produce respuestas mas rápidas.
- elimina el consumo de potencia estática aunque la corriente durante el transitorio se incrementa.

ν **DESVENTAJA**

- Aumento del área

- Esto no es tan malo como puede parecer ya que la puerta proporciona dos señales diferentes lo que permite eliminar las etapas extras de inversores.
- En ocasiones haya lógica común a las dos redes que se pueda compartir.
- Un ejemplo es la implementación de puertas XOR, que es importante para la síntesis automática

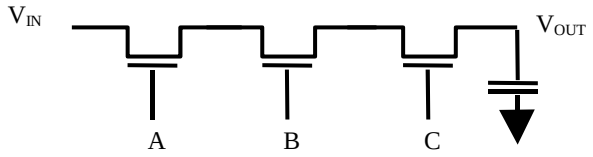


8.5 LÓGICA DE INTERRUPTORES

Se basa en los transistores de paso o en las puertas de transmisión.

Existen diferentes configuraciones que estudiamos a continuación

8.5.1 TRANSISTORES DE PASO



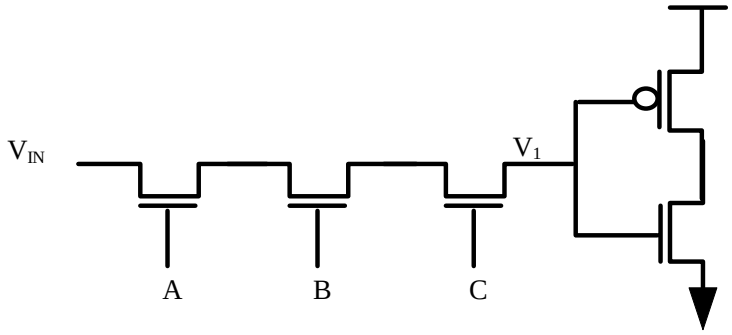
Vout = A·B·C·Vin – Vtn(efecto body)

- **Problema.**

  Degenera la señal de salida el Vtn afectado por el efecto cuerpo.

- **Solución**.

  Añadir al final un inversor que de fuerza a la señal

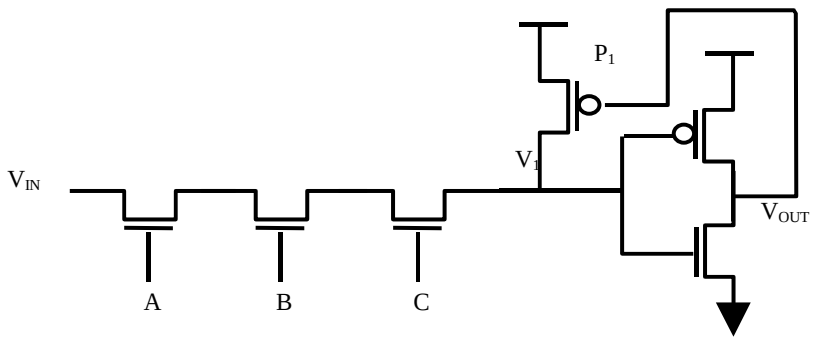


- **Problema**

  * En V_1 se produce una degeneración de la señal. Por ejemplo si $V_1 = 3.5$ V (perdida aproximadamente igual a la del Vt). $V_1 = 5 - Vtn$. Esto afecta tanto a los márgenes de ruido como al consumo de potencia. Al margen de ruido porque el máximo 1 que entra al inversor no es de máxima dureza, y al consumo de potencia porque debido al retardo añadido en las señales de subida y de bajada las corrientes de corto circuito son mayores.

  * El problema puede se más grave todavía si el potencial de umbral de NMOS es tan grande que el transistor PMOS no se corta. En este caso se degenera la señal de salida del inversor y se agrava el consumo de potencia.

- **Solución**



  * Este dispositivo da fuerza al "1"

  * Cuando V1 es aproximadamente 1 → Vout es aproximadamente 0 → empieza a conducir P1 aparece una realimentación

  * Con la transmisión del 0 no hay problema porque no se degenera la señal y a V_1 llega un 0 puro.

- **Problema**

  * ¿Que ocurre cuando en la figura A =0?

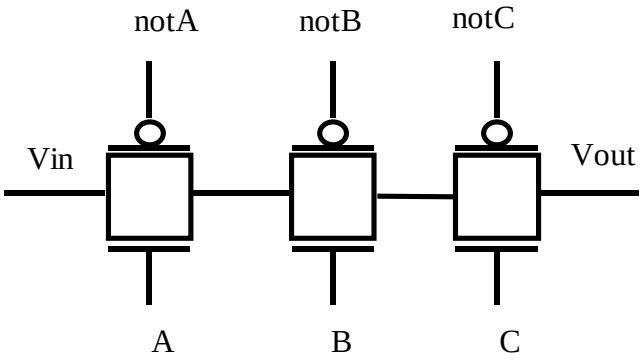
  * Como teóricamente implementa una función **AND** si A=→ Vout = 0.

  * Pero como el transistor controlado por A se corta la capacidad de salida queda aislada y cargada con el con el último valor

* Por esto hay que tener **mucho cuidado** al implementar esta lógica para procurar que siempre estén la función lógica y su complemento

8.5.2 PUERTAS DE TRANSMISION

- Eliminan el problema de la degeneración de la señal
- Continua sin existir camino a tierra



∗ **Inconvenientes**

∗ Se necesitan dos tipos de transistor que están localizados en diferentes pozos

⇒ Reducción de eficacia del LAYOUT

∗ Las señales de control deben presentar ambas polaridades

⇒ Influencia negativa del LAYOUT.

∗ La conexión paralela de pMOS y nMOS

⇒ Incrementa las capacidades

⇒ Reduce el rendimiento.

∗ El retardo es proporcional a N2, siendo N el numero de puertas en serie

- **Solución**

∗ Para N grandes se recomienda romper la cadena mediante bufares.

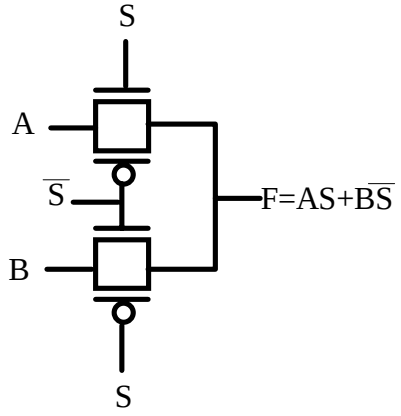
∗ En este caso el retardo solo tiene un retardo lineal con el numero de puertas

- **Tamaño de las puertas**
  - ∗ Son convenientes diseños con tamaños mínimos.
  - ∗ Las características no se ven afectadas por los tamaños
  - ∗ Para cadenas largas puede ser beneficioso usar un escalado progresivo

    ⇒ El ultimo tiene que descargar todas las capacidades

8.6 LÓGICA DE MULTIPLEXORES

La siguiente figura nos muestra un circuito multiplexor típico:



- Se usa principalmente en el diseño de biestables
- Existe una forma de implementar lógica con multiplexores utilizando la expansión de Shanon cuya expresión es la siguiente:

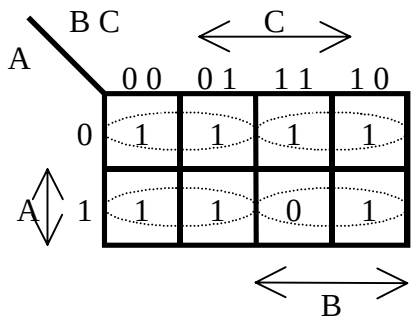
$$E(X)=\Sigma m_i(x_i)\cdot F(x,i)$$

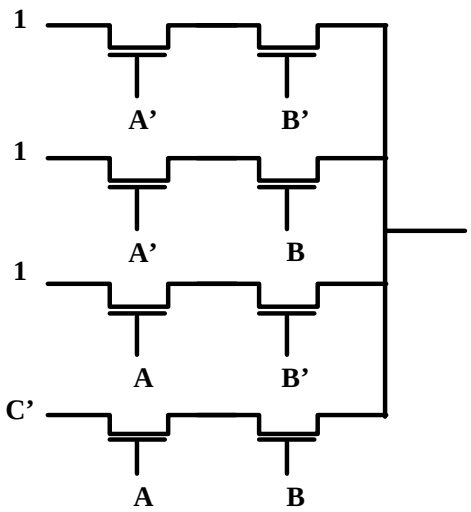
- A continuación vemos un ejemplo de como utilizar la expansión de Shanon. Sea la función NAND de tres entradas:

A	B	C	NAND
0	0	0	1
0	0	1	1
0	1	0	1
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	0

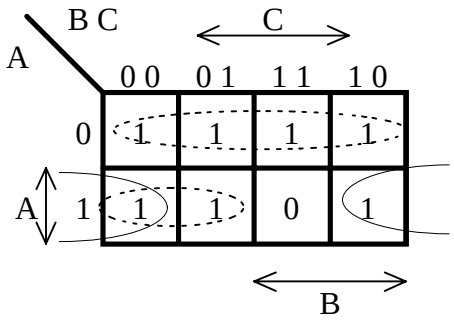
De las tres variables de entrada hay que escoger las que queremos que sean de control. Por ejemplo A y B. Aplicando la descomposición de Shanon

$F=\Sigma_{i=0}{}^{3}\ mi(A,B)\cdot F(mi(A,B),C) = A'\cdot B'\cdot 1+A'\cdot B\cdot 1+A\cdot B'\cdot 1+A\cdot B\cdot C'$

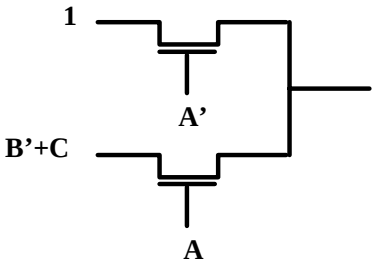




Ejemplo 2:Vamos a suponer ahora que A es la única señal de control.



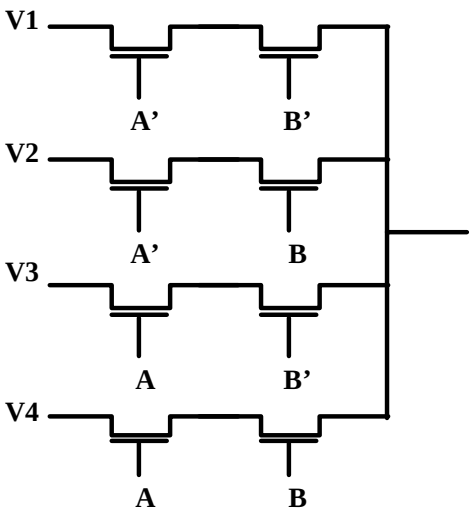
F=A'+A·(B'+C')



Con esta lógica no existe el problema de que no exista un camino a tierra, puesto que existe un dato de entrada por cada posible configuración de las señales de control.

8.6.1 IMPLEMENTACIÓN DE UNA UNIDAD GENERADORA DE FUNCIONES DE DOS ENTRADAS

mintérminos	AB	AND	OR	XOR	entrada unidad
A'B'	00	0	0	0	V1
A'B	01	0	1	1	V2
AB'	10	0	1	1	V3
AB	11	1	1	0	V4



8.7 CPL COMPLEMENTARY PASS-TRANSISTOR LOGIC

- La lógica CPL es una solución que se utiliza en ocasiones para evitar la degeneración de la señal que aparece en las lógicas que utilizan puertas de transmisión debido al efecto del potencial umbral.
    - Recordemos que cuando un 1 lógico atraviesa un transistor NMOS la señal se degenera en una cantidad equivalente al VT y que cuando una señal 0 lógico atraviesa un transistor PMOS, la señal se degenera en una cantidad equivalente al potencial umbral
    - Conclusión, a menor potencial umbral menor degeneración de la señal

- La lógica CPL consiste en introducir los cambios necesarios en el proceso de fabricación para aproximar el potencial umbral a 0.
- Solo es posible si lo permite el fabricante.
- Características:
    - Es lógica diferencial; es decir necesita las dos polaridades de las entradas y genera las dos polaridades de la salida:
        - Esto es una ventaja para puertas complejas como la XOR.
    - Entran dentro de la clasificación de estáticas;
        - la salida siempre conectada a Vdd o Gnd;

            Efecto → reduce el ruido

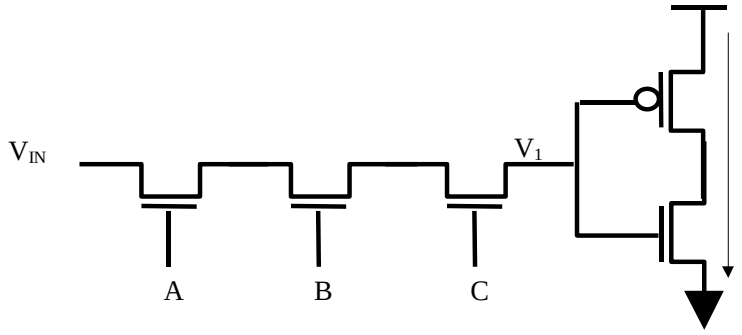
Ventaja:

Reducción o eliminación consumo estático.

Demostración,

Dado el circuito de la figura, se ve que el potencial de entrada al inversor llega muy degenerado cuando el Vin=1. Exactamente el V1=Vin - 3VTN esto implica que Vin no es lo suficientemente elevado como para cortar el transistor PMOS y por lo tanto aparece una corriente estática entre VDD y Gnd

En cambio si Vt≈0 entonces V1≈Vdd por lo tanto se corta o eliminan el consumo estático



Efecto secundario

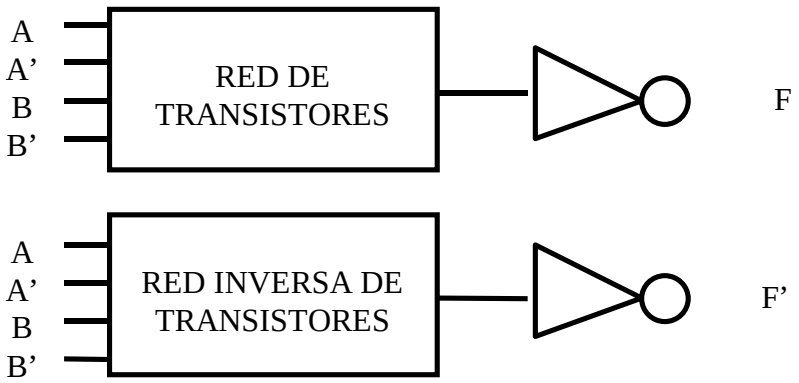
aumento de la velocidad de conexión debido a que la I que atravisa el transistor es proporcional $(Vgs - Vt)^2$ cuando el dispositivo esta saturado.

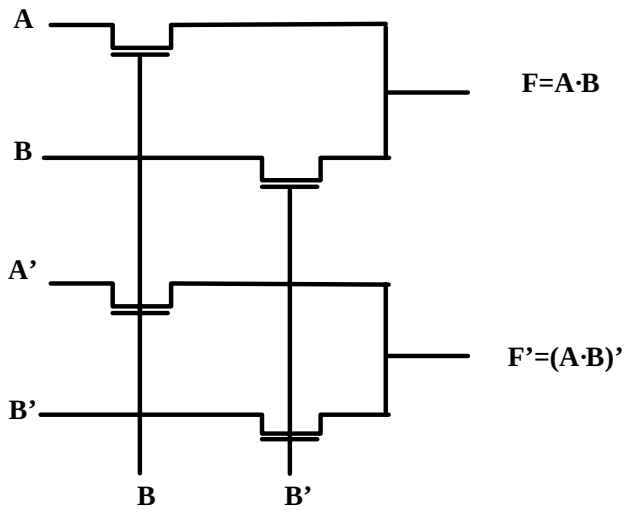
El diseño es muy modular,

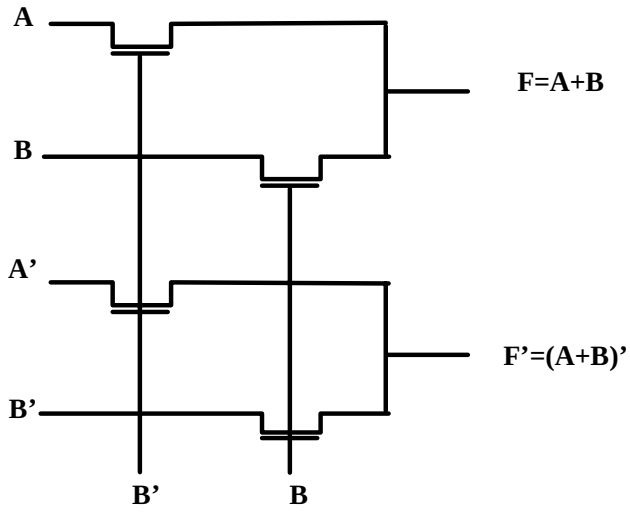
todas las puertas usan exactamente la misma topología, solo cambian las entradas.

El diseño de librerías es muy simple,

las puertas complejas se diseñan poniendo en cascada puertas sencillas.







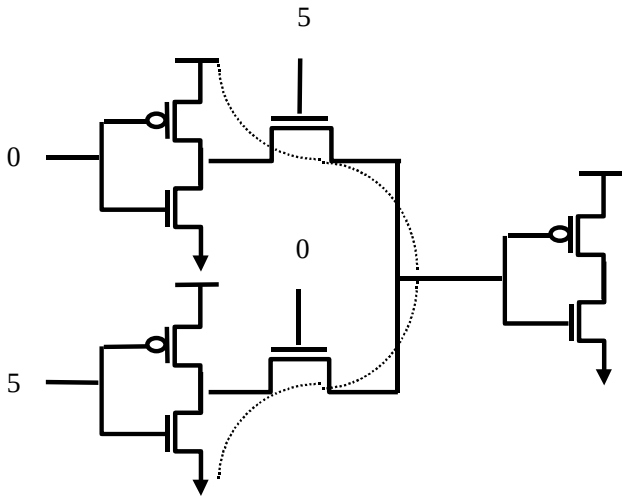
DESVENTAJAS:

El uso de transistores con potencial umbral ≈0 puede ser peligroso debido a la reducción de los márgenes de ruido.

Cortar un dispositivo con Vt=0 es muy difícil

Las corrientes subumbral fluyen con fuerza a través de la puerta de paso incluso si Vgs < Vt, aparece un camino de corriente continua

Exceso de rutado debido a la necesidad de implementar una función lógica y su complementaria



VENTAJA:

La disponibilidad de una puerta XOR sencilla de implementar hace esta lógica atractiva para sumadores y multiplicadores.

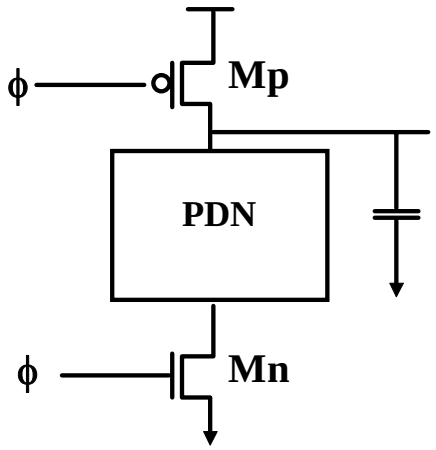
I.

9 LÓGICA COMBINACIONAL DINÁMICA

9.1 PRINCIPIOS

La lógica dinámica usa un numero de transistores similar al de la lógica proporcional N+1, pero evita el consumo de potencia estática. Tiene dos fases:

- **Precarga**
- **Evaluación.**



El bloque PDN (pull-Down-In) es similar a los vistos hasta ahora. La fase de carga y evaluación determina la señal de reloj 0.

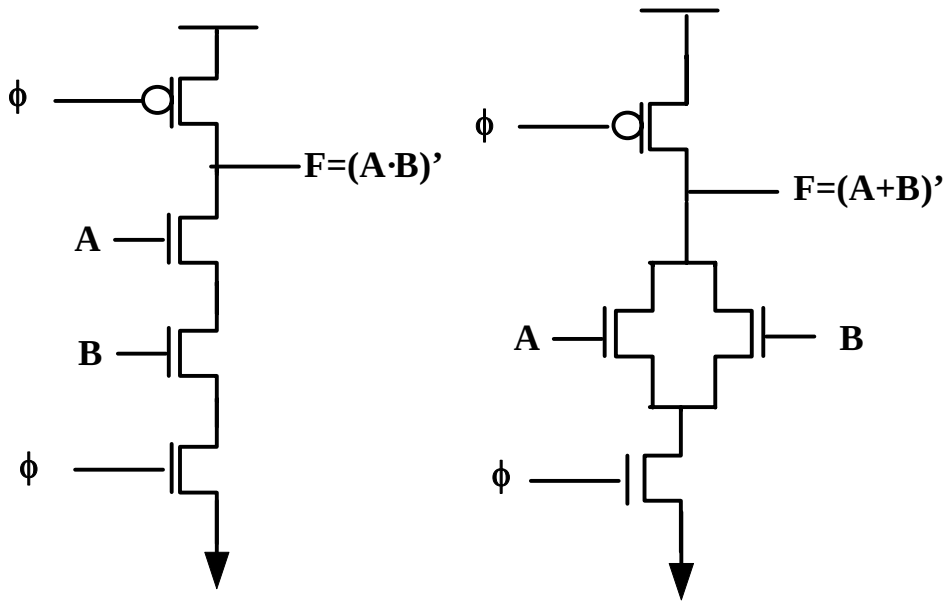
ν **PRECARGA**

- $\phi = 0$
- La capacidad de salida CL se carga a Vdd a través del transistor. Es decir la salida permanece a 1 independientemente de las entradas
- Durante este tiempo, independientemente de estado del PDN, no existe camino a tierra, ya que esta cortado por el transistor Mn

ν **EVALUACIÓN**

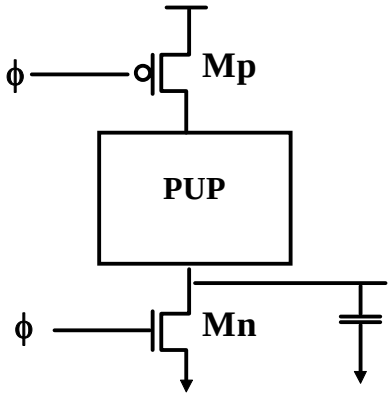
- ϕ= 1
- Mp se corta y Mn se abre.
- En función de las señales de entrada al bloque PDN, **existe un camino a tierra** o no.
    - Si existe camino Cl se descarga.
    - si no existe camino el valor precargado en Cl permanece.
- Una vez descargada la capacidad de salida Cl no se vuelve a cargar hasta la siguiente precarga, esto obliga a que las señales de entrada al bloque PDN puedan hacer más de una transición durante la evaluación.

Ejemplos:



ν **LÓGICA DINÁMICA DE BLOQUE P**

- También se puede diseñar la lógica dinámica utilizando bloque P en lugar de bloques N
- A este bloque se le llama de Pull-UP



- La precarga la realiza el transistor N conectado a la tierra y es a 0

ν **PRECARGA**

- ϕ=1
- transistor Mp cortado
- el transistor Mn conduce→ aparece un camino a tierra→ se carga la capacidad CL a 0

ν **EVALUACIÓN.**

- ϕ=0
- Transistor Mn cortado
- transistor Mp conduciendo se abre un camino a la alimentación VDD según sea la función implementada
- debido a la baja movilidad de los huecos frente a la de los electrones son más lentos
- En caso que se decidiera usarlos se pueden escalar para intentar mejorar la información
    - escalar es modificar los tamaños de los transistores

9.2 CARACTERÍSTICAS

- El numero de transistores es N+2, siendo N el numero de entradas.
- No proporcional, es decir los márgenes de ruido no dependen de las proporciones entre transistores.
- Solo consume potencia dinámica.
- **la velocidad de conexión es rápida**
    - Debido al reducido numero de transistores por puerta, y a que solo se tiene un transistor por cada señal de entrada las capacidades son pequeñas;
- Los márgenes de ruido y el umbral de puerta de la lógica dinámica son función de la duración del periodo de evaluación

  si T demasiado largo la señal uno se degrada debido a la descarga de la capacidad de salida CL

9.3 ANÁLISIS DE LOS TIEMPOS DE SUBIDA Y BAJADA

TIEMPOS DE EVALUACIÓN

- Después de la precarga Vout = "1".
- En la fase de evaluación se pueden dar dos casos:
  - que no exista un camino entre Vout y tierra
    - no existe descarga de la capacidad de salida→ no existe transición→ el tiempo en obtener un 1 es nulo (tlh=0)
  - que exista un camino entre el Vout y tierra
    - se descarga Cl a través de la red Pull-Down.
    - el tiempo que se tarda en pasar de 1 a 0 (tpHL) es proporcional a CL y a la I que transmite el PDN.
    - Además la presencia del transistor evaluador relentiza algo las puertas.

.

TIEMPOS DE PRECARGA

Vienen determinado por el tiempo necesario para cargar Cl a través del transistor de precarga P.

Durante este tiempo la puerta, lógicamente, no puede utilizarse, aunque esto no es un problema si se sabe aprovechar.

A menudo los sistemas digitales se puede diseñar de manera que los tiempos de precarga de un modulo coincidida con la evaluación de otros módulos. Ej.:

La precarga de la unidad aritmética puede coincidir con la descodificación de la Instrucción

El diseñador puede elegir el tamaño del transistor de precarga que desee, en contraste con el pseudo NMOS.

Haciendo este transistor mas grande disminuyen los tiempos de precarga.

si este transistor se hace demasiado grande se produce un incremento de tpHL, debido al aumento Cl.

9.4 CORRIENTES DE PERDIDA

Uno de los problemas mas importantes es ser poco robusto frente al ruido.

La lógica dinámica se basa en el almacenamiento dinámico del valor de una salida en una capacidad.

Debido a las perdidas de corriente este valor se devalúa y esto provoca un mal funcionamiento de la puerta.

Existe dos fuentes de intensidad:

Diodos polarizados en inversa

Corrientes subumbral.

ν **LOS DIODOS**

La capacidad Cl esta formada entre otras , por la capacidad de la difusión del drenador de los transistores N de Pull-Down.

Esta difusión forma un diodo que se encuentra inversamente polarizado

Esta densidad es función de la temperatura

A temperatura ambiente (T=300K) la densidad de la corriente de los diodos es aproximadamente entre 0.1 y 1μA/cm^2.

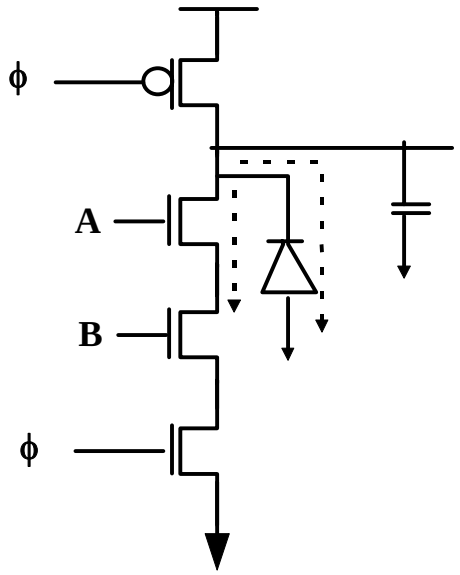
y se dobla cada 10K.

Para una difusión con un área de 1μm^2 la corriente inversa de este diodo se encuentra entre 10^{-15} y 10^{-14} A.

ν **LAS INTENSIDADES SUBUMBRAL**

fluyen del drenador a la fuente. Este efecto es mas importante cuando A no es un 0 exacto ya que se aproxima al Vtn en presencia de ruidos, por eso es importante que el potencial umbral sea suficientemente alto, con valores superiores a 0.5 V.

Su principal efecto es la degradación del “1” en las fases de evaluación, necesidad de una frecuencia mínima (250 Hz - 1 kHz).



ν **Consecuencia.:**

- Las técnicas dinámicas son poco atractivas para ciertos productos tales como relojes o productos de bajo consumo, debido a que funcionan a frecuencias menores para conservar la potencia.
- Difíciles de testar, provocar la permanencia del Ck.

- **Tiempo de descarga de la capacidad CL**

La expresión que calcula el tiempo que tardea en descargarse una capacidad es:

$$\mathbf{T = \frac{C_L \cdot V_{DD}}{i}}$$

Siendo

i la I de perdida (Leackage.)

V_{DD} la alimentación

C_L la capacidad de salida

Ejemplo:

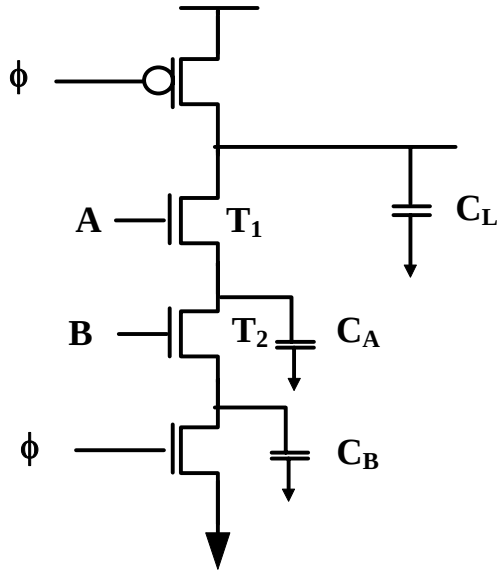
i = 1nA.

C_L = 0.02 pF t = 100 ns.

$V_{DD} = 5$ V

9.4.1 DISTRIBUCIÓN DE CARGA

Sea el circuito dela figura



Fase de precarga:

el nodo de salida se precarga a Vdd

Suponemos que todas las entradas al bloque lógico son = 0 y que C_A esta descargada.

Evaluación.

Si suponemos que B = O T2 está cortado no existe camino a tierra

Suponemos que A sube de 0 a 1 T1 empieza a conducir

La carga almacenada originalmente en CL se distribuye entre CL y C_A, esto provoca una caída de potencial en Vout que no se puede recuperar debido a la característica dinámica de esta lógica.

Se puede aproximar cual será el potencial final de salida;

$$V' = \frac{C_L \cdot V}{C_A + C_L}$$

Vout se degenera:

- Si las C son muy grandes.
- Si las C son muchas.
- Por ambos motivos.

Si V' es demasiado pequeña, al conectar la salida a un inversor puede provocar consumos estáticos debido a que no se corta correctamente el transistor PMOS

Problemas :

Tanto las corrientes de leackaje, como la distribución de carga tienen como principal efecto degenerar la señal Vout que por las características dinámicas del sistema no se puede regenerar

Solución 1

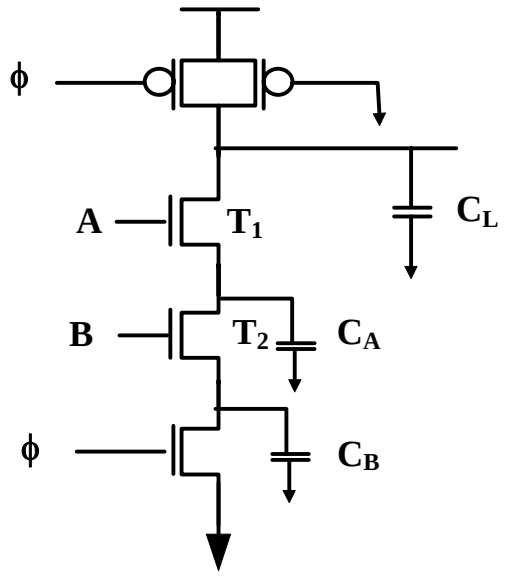
hacer bloques pseudoestaticos.

Esto se consigue añadiendo un **pequeño PMOS de alta resistencia,** Con su puerta conectada a gnd en paralelo con el transistor de precarga: efecto pseudo NMOS.

A este transistor se le llama Bleeder:

reduce la impedancia del nodo de salida (mayor paso de I).

Introduce consumo de potencia estática, lo que se minimiza haciendo este transistor largo y estrecho.



solución 2

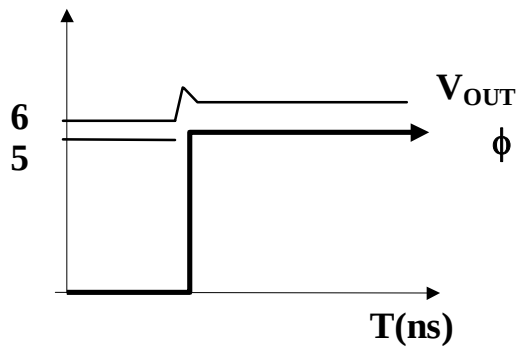
precargar los nodos internos usando un transistor.

Un problema en ambas soluciones es la repercusión en el área.

9.4.2 ACOPLAMIENTO DE RELOJ (CLOCK-FEEDTROUGH).

NOTA: es un problema del condensador, la diferencia de potencial en la capacidad no puede variar instantáneamente, de manear que si la diferencia de potencial entre los extremos de una capacidad es de 1V (1-0) y en el extremo del 0 se produce un incremento instantáneo de 0 a 5, en el extremo del 1 se debe producir un incremento instantáneo a 6 para que la diferencia de potencial entre ambos puntos se siga conservando.

El acoplo quiere decir que la variación en ϕ afecta a Vout.



Otro causa de ruido en esta lógica:

La señal de reloj se acopla con la capacidad de salida a través de las capacidades del dispositivo de precarga.

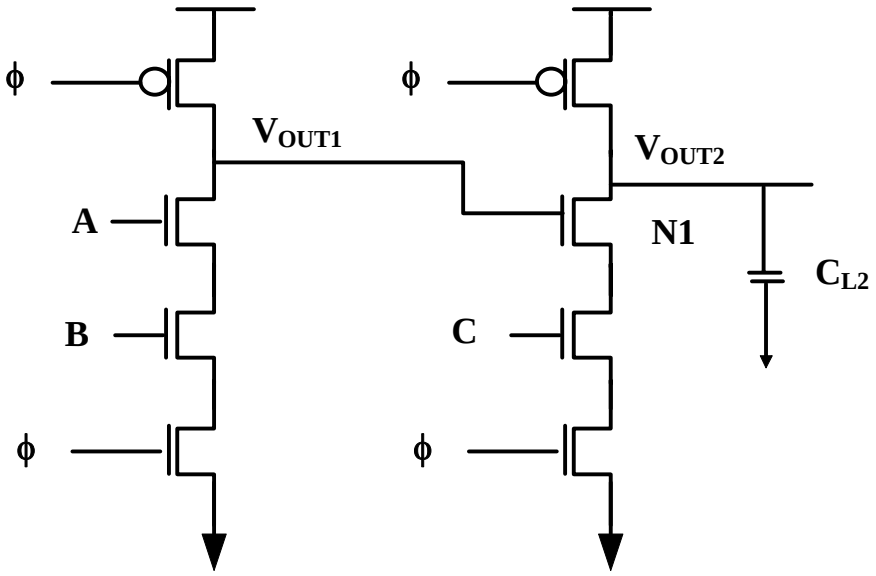
El problema de este acoplamiento es que el potencial en un instante dado sube por encima del potencial de alimentación Vdd.

Esto puede hacer que los diodos parásitos pasen de estar cortados o en inversa, a estar en directa. Esto provoca una inyección de electrones en el sustrato que pueden ser recogidos por un Vout cercano que se encuentre en alta impedancia (fase de evaluación) en el estado “1. La inyección de electrones en el nodo aislado provoca la disminución del potencial positivo que almacena. Y por lo tanto en su caso más extremo convertir un 1 en un valor indeterminado o en el peor de los casos en un cero.

Otro efecto pernicioso es el disparo del Latch-Up. Este fenómeno se puede evitar colocando un numero de contactos de pozo cerca de los dispositivos de precarga para recoger los electrones inyectados.

9.5 PUERTAS DINÁMICAS EN CASCADA

Sea el circuito de la figura:



ν **PROBLEMA**

Todas las salidas de la primera puerta se encuentran a la vez en fase de precarga.

Debido a los retardos de las puertas esto perjudica al funcionamiento de la red.

ν **DEMOSTRACIÓN:**

Precarga

$V_{OUT1} = 1$,

$V_{OUT2} = 1$, y

las entradas A y B de la primera puertas toman el valor 1

La entrada C de la segunda puerta toma el valor 1

evaluación

el V_{OUT1} no cambia instantáneamente al valor 0 que le corresponde.

Durante el tiempo que tarda en cambiar V_{OUT1} se está abriendo un camino desde V_{OUT2} a tierra, luego esa capacidad se esta descargando.

Cuando finalmente V_{OUT1} alcanza el valor cero el que el transistor N1 se corta, pero ya se ha producido la descarga de $C_{L2.}$

Como la lógica es dinámica la salida V_{OUT2} no puede alcanzar el valor eléctrico que le corresponde.

Se produce una descarga parcial o total de la C_{L2}, lo que se traduce en el mejor de los casos en una reducción de los márgenes de ruido y en el peor en funcionamientos incorrectos

Esta claro que el problema reside en que la precarga se produce a 1, y por lo tanto existe la posibilidad de que al inicio de la fase de evaluación exista un camino a tierra por el que descargar la capacidad.

ν **SOLUCIÓN**

Precarga a 0

De esta manera se cierran todos los caminos a tierra hasta que se haya realizado correctamente la evaluación de la entrada

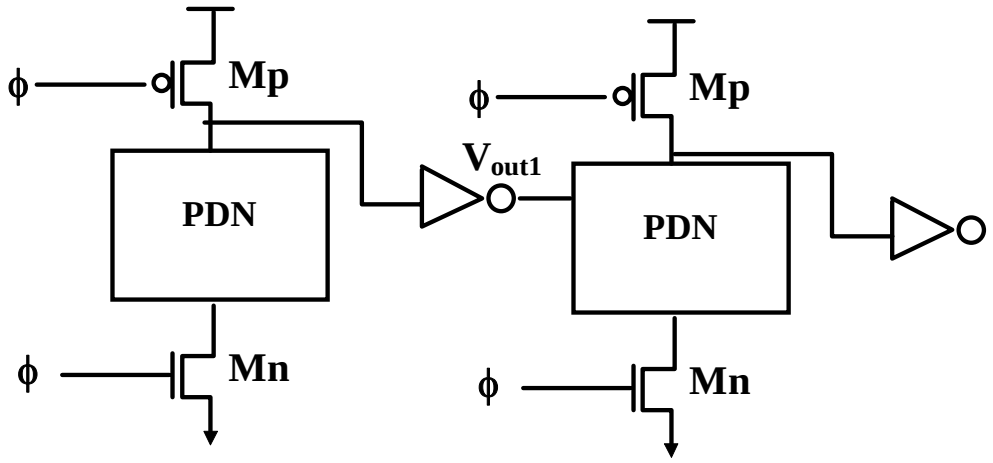
Con la precarga a cero se soluciona el problema de la descarga en el caso de nodos internos, pero no queda resuelto el caso de las entradas exteriores, entendiéndose como tal las que no provienen de otro bloque de lógica dinámica anterior. También en este caso si la señal varia durante la fase de evaluación de 1 a 0 se produce una descarga de la CL por existir un camino a tierra .

Solución

las entradas externas a una puerta lógica estén estables en la fase de evaluación; es decir solo pueden cambiar en la precarga.

Una regla similar se puede aplicar a la lógica de la dinámica de bloque P; se fija la precarga a 1; las únicas transiciones permitidas durante la fase de evaluación son de 1 a 0.

9.6 LÓGICA DOMINO



- Se incluye un inversor a la salida del bloque lógico

  Se asegura que todas las entradas que provienen de otro nodo estén a 0 durante la precarga.

  La única transición posible durante la evaluación es de 0 a 1.

  la ventaja de aumentar la inmunidad al ruido dado que el fan -out de la puerta se alimenta a través del inversor estático que tiene una muy baja resistencia de salida

  el inversor se puede diseñar para que trabaje de manera optima a altas velocidades.

- Modo de operación de una cadena de puertas

  Precarga.

  $\phi = 0$.

  Se precargan todas las puertas a la vez

  $V_{OUT1} = 0$; impide que se abran caminos hasta tierra en la siguiente puerta.

  Evaluación.

  $\phi = 1$

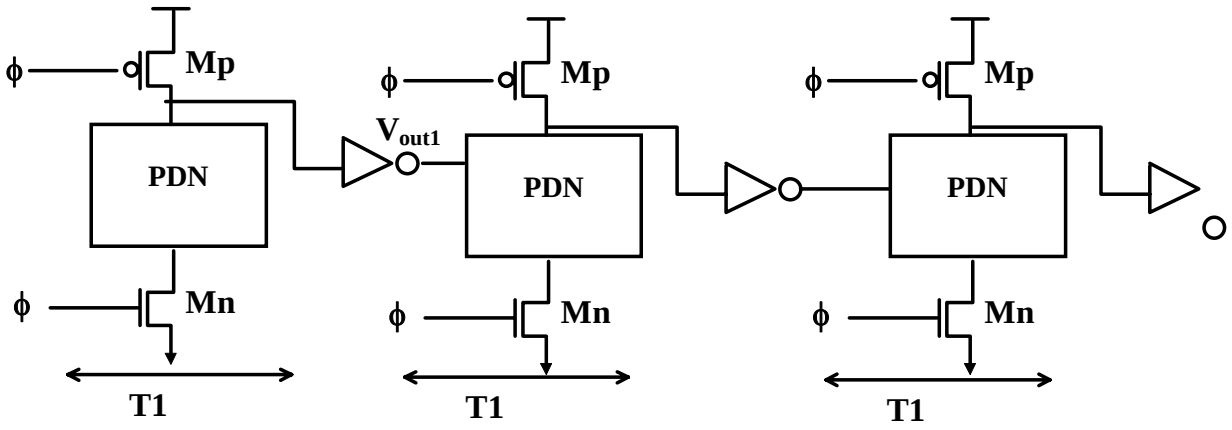
  Cada etapa se evalúa y causa la evaluación de la siguiente (como la caída de juego de domino)

- Máximo número de etapas en cascada

  Se pueden colocar en cascada tantas etapas como se desee con solo tener en cuenta que se deben evaluar con medio ciclo de ϕ, que es lo que dura la fase de evaluación.

  Ejemplo

La fase de evaluación debe durar como poco:



Siendo

T1 el tiempo de evaluación del de la primera puerta ,

T2 el tiempo de evaluación del de la segunda puerta ,

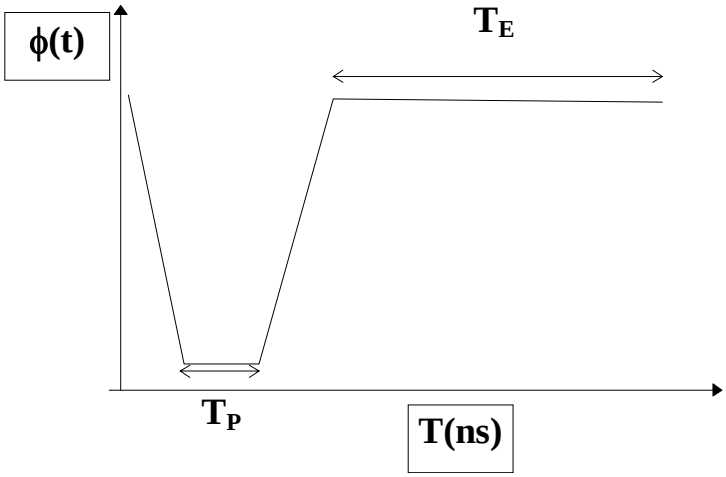
T3 el tiempo de evaluación del de la tercera puerta ,

La fase de evaluación debe durar como poco

T_E=T1+T2+T3

La fase de precarga debe ser lo suficientemente larga como para cargar las capacidades CL a “Vdd” . A este tiempo le llamamos T_P

El ciclo de reloj debe ser como mínimo igual a T_E+T_P



Recordar que la ϕ es común para todas las etapas.

9.6.1 EJEMPLO DE IMPLEMENTACIÓN CON LÓGICA DOMINÓ

En está lógica se debe tener cuidado con los diseños porque al ser puertas que implementan funciones lógicas positivas no es una lógica completa

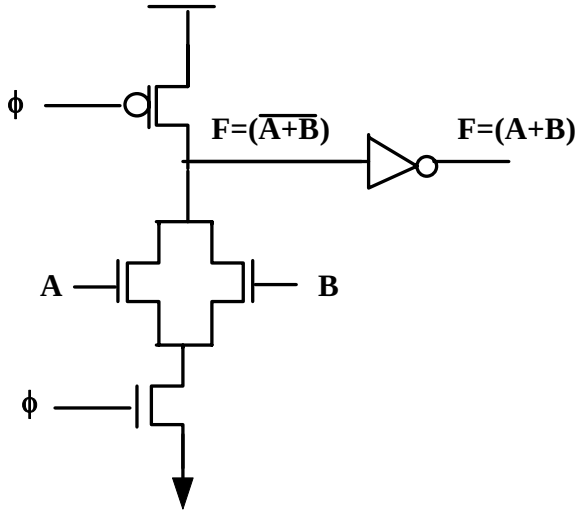
Vamos a ver un ejemplo de implementación:

F=(a+b)'

G=c·d

T=F·G

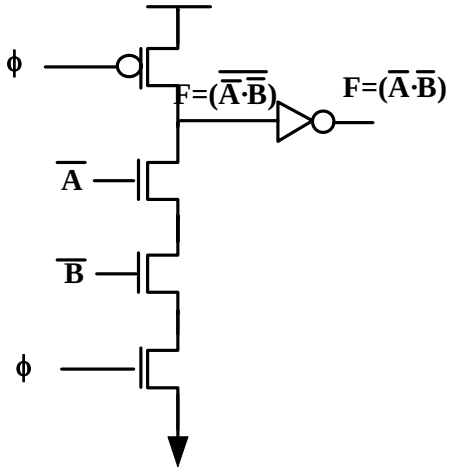
La función F a dar problemas porque no se puede implementar directamente.



Una falsa solución sería añadir un inversor a continuación del existente, pero con esto se incumple la condición de precarga a 0 y cambios en evaluación 0 a 1.

La solución consiste en aplicar la lógica booleana:

F=(a+b)'=a'·b'



9.6.2 PROPIEDADES DE LA LÓGICA DOMINÓ

Solo implementa lógica no invertida. Como hemos visto esto puede ser un problema a la hora de implementar ciertas puertas lógicas.

Su Fan out es menor que el de la lógica CMOS complementaria porque una puerta tiene la mitad de capacidades de entrada que en el caso del CMOS complementario lo que la hace apropiada para circuitos complejos con elevado Fan-Out: como por ejemplo

Unidades Aritmético Lógicas

Unidades de Control.

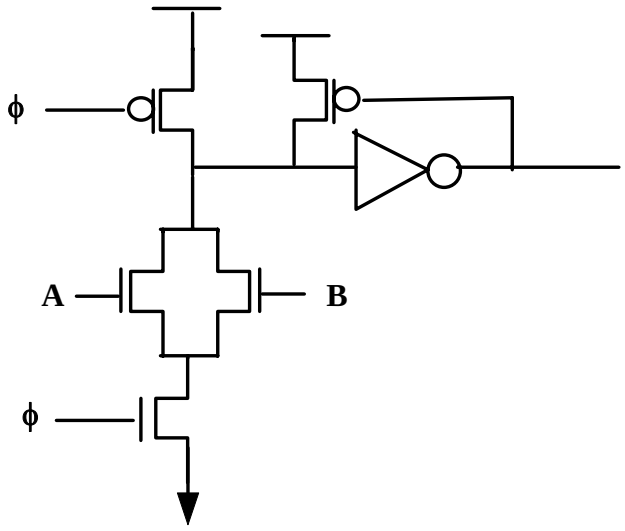
Velocidades elevadas:

solo existen los tiempos de subida

precarga a cero

El inversor pude diseñarse para que optimice la carga del Fan-out,

Se puede disminuir su vulnerabilidad a la distribución y perdidas de carga añadiendo un inversor que restaure el nivel.



Se han usado en una gran cantidad de circuitos de alta velocidad.

P.e.: primer microprocesador de 32 bits. Bellmac 32 de ATT.

Esta cayendo en desuso debido a la dificultad de no implementar lógica negada.

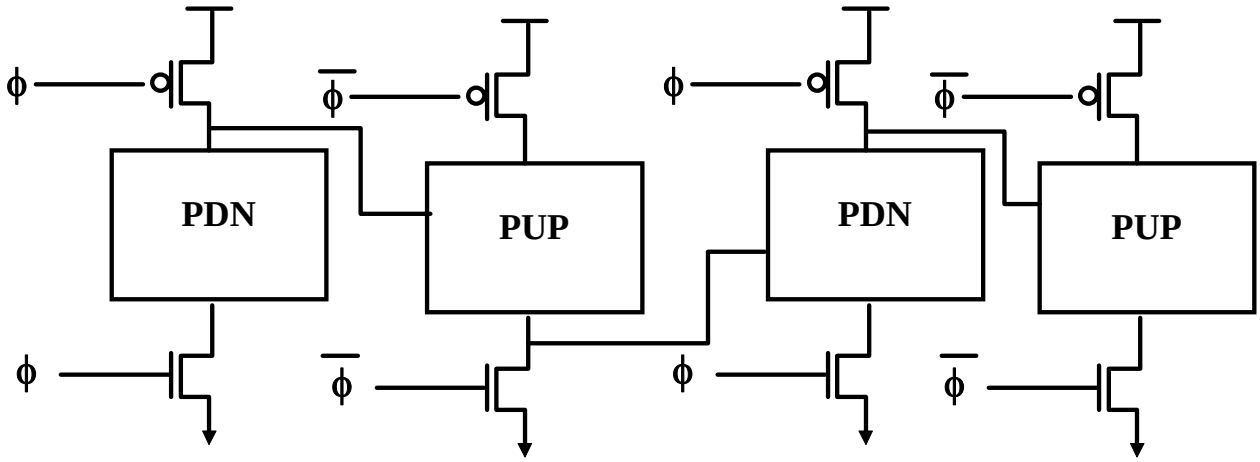
9.7 LÓGICA DOMINO NP O ZIPPER

Se elimina el inversor estático que se colocaba en la lógica dominó entre las etapas

Se intercalan bloques de lógica N y bloques de lógica P.

Lógica N: Precarga a 1. Cortan los dispositivos P que le siguen.

Lógica P: Precarga a 0. Corta los dispositivos N.



FASE DE PRECARGA:

$\phi = 0$ y $\phi' = 1$

Todos los bloques precargan **a la vez tanto los N como lo P.**

Los N precargan porque $\phi=0$ → se corta el camino a tierra y se abre a VDD

Lo P precargan porque $\phi'=1$ → se corta el camino a VDD y se abre un camino a tierra

FASE DE EVALUACIÓN

$\phi=1$ y $\phi'=0$

los bloques N cierran el camino a Vdd y se evalúan

Los Bloque P cierran su camino a tierra y se evalúan

Luego todos los bloques se evalúan en una fase común de evaluación y se sigue produciendo el efecto dominó.

Igual que ocurría en la lógica domino el ciclo de evaluación y precarga debe tener una duración mínima

el tiempo de evaluación es la suma de cada uno de los tiempo sde evaluación

Te = Te1 + Te2 + ... + Ten

el tiempo de precarga debe ser el mayor de todos los tiempos

Tp = MAX(Tp1,Tp2,Tp3,..)

el tiempo total del ciclo debe ser Tt = Te +Tp

Este estilo lógico es la base del estilo NORA, muy popular en el diseño secuencial, que se vera mas adelante.

Desventaja:

Los bloques P son mas lentos.

los retardos de propagación se pueden igualar modificando los canales, pero esto supone más área.

Bajas tolerancias al ruido

Ventajas:

Alta densidad de empaquetamiento: eliminar el inversos estático.

Se pueden conseguir velocidades muy altas

20% mas rápida que el domino debido a

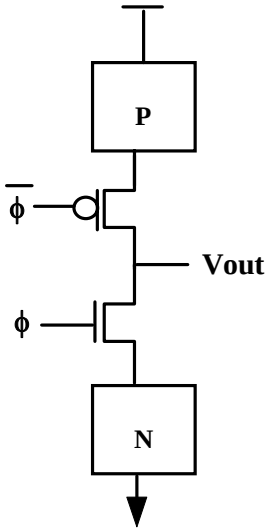
eliminación inversor,

menor capacidad.

Ejemplo:

El procesador Alpha L DEE (primer microprocesador cMOS a 250 Mhz..

9.8 C^{2}MOS CLOCKED CMOS



El bloque P y el N son una función y su complemento.

Misma capacidad de entrada que la lógica CMOS. Tiempos de subida y bajada mayores debido a los transistores P y N conectados en serie y a la salida Vout.

Atención: no tienen fase de precarga y evaluación

:

ϕ= 1: Se comporta como un inversor CMOS normal pero algo mas lento.

ϕ = 0: Cl queda cargada con su ultimo valor y aislada.

A diferencia del resto de lógicas estudiadas se dispone del dato en todo el ciclo de CK pero solo cuando ϕ = 0 se puede considerar estable.

Su principal uso es el diseño

- De caches
- De interfaces con otras estructuras cerrojo.
- De registros de desplazamiento, que se verán en profundidad mas adelante.

9.9 CONSUMO DE POTENCIA EN PUERTAS CMOS

En un tema anterior ya se estudió el consumo de potencia de inversores CMOS complementarios.

Los resultados obtenidos son por lo general extrapolables a puertas mas complejas, aunque se deben tener en cuenta algunas consideraciones.

En este apartado se hablara de:

- La actividad de conexionado (de interrupción)
- Glitching
- La corriente de corto circuito.

9.9.1 ACTIVIDAD DE CONEXIÓNADO (INTERRUPCION) EN LÓGICA ESTÁTICA

En los CMOS la potencia se consume principalmente durante los transitorios, es decir es un consumo dinámico. Este consumo dinámico se debe a:

La carga de Cl.

La aparición de un camino directo entre Vdd la tierra

De esto dos el más importante se debe a la carga de las capacidades y viene dado por la expresión hallada para los inversores CMOS complementarios:

$$W = Cl\ Vdd^2\ f$$

Siendo:

Cl: Capacidad

Vdd: Alimentación.

f: frecuencia con que el circuito pasa de 0 a 1

ACTIVIDAD DE INTERRUPCIÓN

A la frecuencia con que una puerta lógica cambia su salida de 0 a 1 se le llama **actividad de interrupción**.

Este factor es fácil de calcular para un inversor, pero la dificultad de calculo crece con la complejidad de la puerta.

En redes complejas la actividad es función de :

La naturaleza de las entradas

La estadística de las entradas

Estilo (dinámico, estático)

La función a implementar

Estos factores se pueden incorporar a la expresión inicial modificada:

$$W = Cl \cdot Vdd^2 \cdot P \cdot f$$

donde

f: frecuencia de eventos de entrada.

P: probabilidad de que se produzca una transición en la entrada que provoque en la salida un cambio de 0 a 1

Vamos a calcular $P_{0\rightarrow 1}$ para lógica estática de entradas independiente

A) Entradas con la misma probabilidad:

$P_{0\rightarrow 1} = P_0 \cdot P_1$

donde:

P0 probabilidad de que la salida sea 0.

P1 probabilidad de que la salida sea 1.

como $P_0 = 1-P_1$

$$\mathbf{P_{0\rightarrow 1} = (1-P1)\cdot P1}$$

B) Entradas no distribuidas uniformemente

suele ocurrir en circuitos lógicos con múltiples puertas en cascada.

En este caso se debe hallar la probabilidad de salida sea 1 en función de las probabilidades de entrada

P1 = f(Pa,Pb, ... ,Pm)

siendo Pa, Pb, ... ,Pm las probabilidades de que las entradas a, b y m sean 1.

ν **Ejemplos**

Vamos a ver a continuación un ejemplo para la función NOR de dos entradas implementada en lógica CMOS complementaria.

- Las entradas distribuidas uniformemente

A	B	NOR
0	0	1
0	1	0
1	0	0
1	1	0

La probabilidad de que la salida sea 0 es 3/4.

La probabilidad de que la salida sea 1 es 1/4.

$$P_{0\rightarrow 1} = P0 \cdot P1 = 3/4 \cdot 1/4 = 3/16$$

- Entradas no uniformemente distribuidas

P1 = f(Pa,Pb)

siendo:

Pa probabilidad de que la entrada a sea 1.

Pb probabilidad de que la entrada b sea 1.

En el caso de la puerta NOR, para que la salida sea 1, ambas entradas deben ser 0:

$$P1 = P_{A0} \cdot P_{B0} = (1-P_{A1})\cdot(1-P_{B1}), \text{ Luego}$$

$$P_{0\rightarrow 1} = P_0 \cdot P_1(1-P_1)\cdot(P_1) = [1-(1-P_{A1})\cdot(1-P_{B1})]\cdot (1-P_{A1})\cdot(1-P_{B1})$$

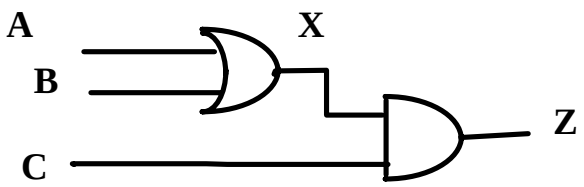
$$= [1 - (1-Pa)(1-Pb)]\ (1-Pa)(1-Pb)$$

ν **Puertas en cascada:**

El calculo de $P_{0\rightarrow 1}$ se complica un poco cuando se colocan varias puertas lógicas en cascada.

La técnica consiste en evaluar las señales y probabilidades de transición avanzando desde las entradas a las salidas.

Ejemplo:



Teniendo A,B y C igual Posibilidad de ser 0 a 1:

Pa = Pa1 = 1/2

Pb =Pb1 =1/2

Pc = Pc1 =1/2

Como A y B están uniformemente distribuidas es fácil ver P_{X1}:

A	B	OR
0	0	0
0	1	1
1	0	1
1	1	1

$P_{X1} = 3/4$

El caso de la AND es diferente porque las entradas no están uniformemente distribuidas.

Px = 3/4

Pe = 1/2

X	C	AND
0	0	0

0	1	0
1	0	0
1	1	1

$Pz1 = Px \cdot Pc = 3/4 \cdot 1/2$ $P_{Z0 \rightarrow 1} = P_{Z0} \cdot P_{Z1} == (1-P_{Z1}) \cdot (P_{Z1}) = = (1 - ¾ \cdot 1/2) \cdot (3/4 \cdot 1/2)$

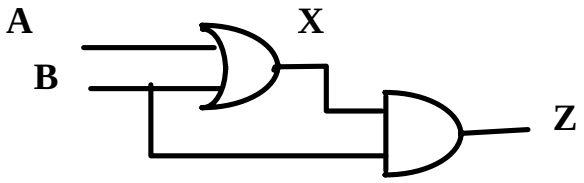
ν **DESVENTAJAS**

Esta forma de estudiar los circuito tiene importantes fallos:

Redes con realimentación (c. secuenciales).

Fan-Out reconvergente. Cuando la entrada ataca a dos puertas diferentes.

Uno de las suposiciones en que se basa el estudio es que las señales de entrada son independientes y en este caso esto no se cumple. En la actualidad esta independencia de entradas es prácticamente inexistente.



El estudio de este segundo caso se hace mas complejo todavía:

$Pz = P(z=1) = P(B=1, x=1)$

En el caso de independencia de entradas como ya vimos;

$Pz = Pb \cdot Px$

En el caso de dependencia de entradas hay que utilizar probabilidades condicionadas:

$Pz = P(x=1|B=1) \cdot P(B=1|x=1)$

siendo:

$P(x=1|B=1)$; Probabilidad de que x=1 cuando B=1Como x se obtiene a través de una OR

$P(x=1|B=1) = 1.$

$P(B=1|x=1)$ probabilidad de que B=1 cuando x=1 como la puerta es una AND es igual a $P(B=1)=P_B$ luego :

$Pz = Pb = 1/2$

(En realidad se ha aplicado la ley de absorción (a+b)·b = b = z)

$$P_{0\rightarrow 1} = (1-Pz)Pz = 1/2 \cdot 1/2 = 1/4$$

En cualquier caso debe quedar constancia de que las herramientas CAD son imprescindibles para el estudio de estas probabilidades.

9.9.2 ACTIVIDAD DE INTERRUPCIÓN EN LA LÓGICA DINÁMICA

En la lógica dinámica la señal se precarga en cada ciclo de CK. El consumo solo se produce cuando en la fase de evaluación ha descargado la capacidad.

Se consume independientemente de los valores anteriores o posteriores.

En las puertas dinámicas el consumo esta determinado únicamente por la probabilidad de que la salida sea 0 y no por la probabilidad de transición de 0 a 1.

Las probabilidades de señal son siempre mayores que las probabilidades de transición (que es producto de dos probabilidades de señales).

En un CMOS estático si la señal de entrada no cambia no se consume potencia. En la dinámica esto no es cierto.

Ejemplo de calculo de probabilidad de una puerta.dinámica NOR; suponiendo probabilidad de entrada idénticas.

A	B	NOR
0	0	0
0	1	1
1	0	1
1	1	1

$P_0 = ¼$

La actividad que se obtiene para una puerta dinámica (1/4) y para un puerta estática (3/16) es prácticamente idéntica pero existen dos factores que se deben tener en cuenta:

a. La Cl dinámica es menor.

b. En la dinámica las líneas de CK también consumen potencia.

Para el caso de que las entradas no estén uniformemente distribuidas el tratamiento es similar al ya visto.

Para una puerta OR de dos entradas A,B para que la salida sea 0

$P_{A0} \cdot P_{B0} = (1-P_A)\cdot(1-P_B) = P_{0\rightarrow 1.}$

9.9.3 GLITCHING (FALLOS)

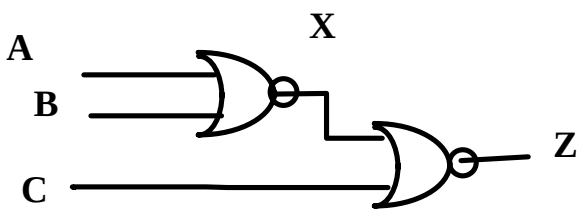
Critical races o dinamic hazards

Hasta el momento hemos ignorado que las puertas lógicas tienen retardos.

La propagación de una señal de un bloque lógico al siguiente produce transiciones erróneas.

Un Glitching es cuando un nodo tiene varias transiciones en un ciclo antes de alcanzar el valor estable.

Ejemplo;



A, B y C llegan simultáneamente.

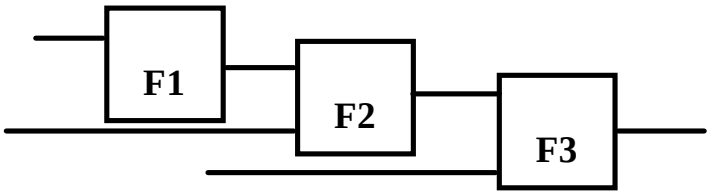
Puesto que existe en retardo de propagación en la primera NOR, la segunda evalúa inicialmente a un valor falso con el valor primitivo de x.

Esta primera evaluación consume potencia sin ejecutar ninguna función. Supongamos que pasamos 101 a 000 para las entradas A,B y C.

La salida Z debería permanecer a 0, pero sufre un hazard que realiza un consumo de potencia.

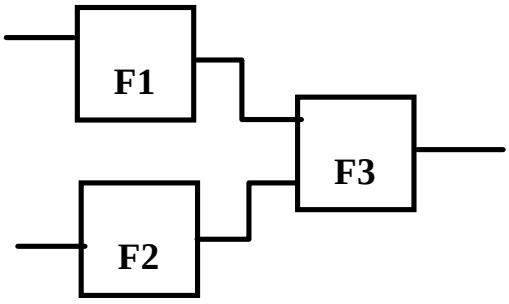
La aparición del Glitching en circuitos se debe a las diferencias de longitud de los caminos en la red. Si todas las señales cambian simultáneamente no se produce Glitching.

Ejemplo típico de red propensa a glitching



Todos los operadores F tienen la misma unidad de retardo y suponemos que todas las señales externas llegan simultáneamente.

Para evitar el problema seria suficiente igualar los tiempos de llegada de todas las señales de entrada a cada puerta.



Esto no siempre se puede hacer, pero casi siempre es suficiente que las longitudes sean aproximadamente iguales para que desaparezca el Glitching.

Conviene saber que aunque los Glitchings suelan ser parciales, contribuyen de manera significativa a la perdida de potencia.

Las cadenas largas de puertas suelen aparecer en estructuras lógicas importantes como sumadores y multiplicadores. El 25% de la potencia consumida en estos circuitos puede deberse a hazards.

Los glitchings no son problema en la lógica dinámica dado que en la fase de evaluación no existe camino a Vdd, no existiendo consumo.

9.9.4 CORRIENTES DE CORTOCIRCUITO

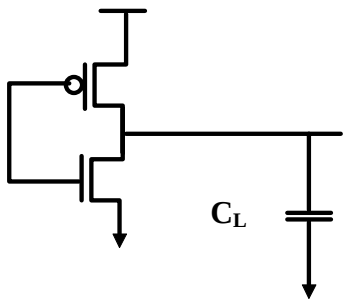
La potencia total disipada es función del tiempo que los transistores conducen y de los modos de operación de los transistores.

El calculo que vimos en su día dio como resultado;

$$P_{SC} = \frac{\beta}{12} (Vdd-2Vt)^3 \frac{T_R}{T_P}$$

Este calculo se hizo suponiendo que no existían capacidades parásitas.

Sea el inversor de la siguiente figura



Suponemos C_L cargada a 1.

Vamos a ver que ocurre en un inversor cuando la C_L es grande

Cuando Vin pasa de 0 a 1 la C_L tarda en descargarse luego en el instante inicial Vout permanece a 1; lego $V_{DS} \approx 0$ y I=0. Los dispositivos se cortan sin haber disipado corriente.

Caso en que la capacidad sea muy pequeña

cuando Vin pase rápidamente de 0 a 1 el Vout hace la transición de 0 a1 muy rápidamente; Isc = Mx

Conclusión:

Este consumo de potencia de cortocircuito se puede minimizar haciendo los tiempos de entrada mas pequeños que los de salida.

- **inconveniente**
  - ➢ no conviene que los tiempos de salida sean grandes
    - 4 porque relentiza el sistema.
    - 4 puede generar corrientes de cortocircuito en la siguiente fase

ν **Técnica de diseño**

- El consumo de potencia debido a cortos circuitos se minimiza igualando los tiempos de entrada y salida. No es la solución optima pero consigue limitar el consumo y no relentiza en exceso el funcionamiento
- Cuando Cl es muy pequeña predomina el consumo debido a la corriente de cortocircuito sobre el consumo debido a la carga de capacidades.

ν **Técnica de diseño**

- El impacto de consumo de cortocircuito se reduce al disminuir el Vdd debido a que las intensidades se hacen menores.
- En el caso extremo Vdd < Vtn + |Vtp| se elimina la disipación porque nunca conducen lo dos transistores simultáneamente.

- Aunque cumplir estas técnicas no es complicado, lo cierto es que se pasan por alto muy a menudo.

  Por ejemplo en el diseño de celdas estándar, para asegurar que las librerías cumplen las ligaduras de tiempo para un amplio rango de capacidades de carga todos los transistores de las celdas se sobredimensionan. Como resultado los tiempos de salida son menores que los tiempos de entrada. Esto produce un aumento del consumo de cortocircuito de un 50%.

9.10DISEÑO CMOS DE BAJA POTENCIA

Suponiendo controlado el consumo de potencia debido a:

- Corriente de cortocircuito.
- Glitching
- Las perdidas de corriente.

el principal consumo de potencia se debe a la carga y descarga de las capacidades

$$Pd= \frac{C_L \cdot Vdd^2}{Tp} = F_P \cdot C_L \cdot Vdd^2$$

Para un inversor

$$Pd= F \cdot P_{0\rightarrow 1} \cdot C_L \cdot Vdd^2$$

Para una puerta compleja

Suponiendo que el rendimiento del circuito deba permanecer cte. (frecuencia =cte) solo hay dos parámetros cuya modificación reduce el consumo:

Ceff

Vdd

Siendo Ceff el producto de la capacidad y de la actividad $P_{0\rightarrow 1}$.

9.10.1 MODIFICACION DEL POTENCIAL DE ALIMENTACION

- La forma más efectiva de reducir el consumo de potencia es reducir Vdd puesto que la disipación es proporcional a Vdd^2.
- La disminución del Vdd produce una mejora cuadrática del producto potencia-retardo, que es una de las medidas de calidad de un estilo lógico.
- Desafortunadamente el impacto sobre el retardo es muy importante.
    - Es especialmente malo cuando Vdd= Vtn + |Vtp| Los tiempos de subida y bajada son inversamente proporcionales a (Vdd-Vt) luego si la diferencia entre Vdd y Vt se hace pequeña los tiempos aumentan.
- Si se quiere mantener el rendimiento total del sistema se necesita algún tipo de compensación
- Una primera aproximación puede ser adoptar una tecnología con **potenciales umbrales bajos**. Esto significa que los dispositivos empiezan a conducir antes.
    - Desplaza la curva retardo la izquierda
    - **Problema**: Al disminuir Vt aumentan las intensidades **subumbral**, en su caso extremo si Vt es = 0 el dispositivo puede no cortarse nunca.
    - El efecto de las corrientes umbral es diferente según se trate de circuitos estáticos o dinámicos.
        - En los dispositivos dinámicos provoca el aumento de la frecuencia para evitar la perdida de información.
        - En los estáticos provoca perdidas de intensidad en las fases estáticas y márgenes de ruido reducidos.
    - Los potenciales umbral no deberían ser nunca inferiores a 0.5-0.6 V (aunque generalmente son superiores alcanzando el valor 1V).
    - Por otro lado el disminuir Vdd por debajo de 1.2-2V (2·Vt) es excesivamente costoso en términos de rendimiento.
    - En realidad estas conclusiones no son tan categóricas cuando se hace un estudio un poco más profundo.
        - En los dispositivos estáticos este consumo se ve compensado por la disminución de Vdd. Recordemos que el consumo estático viene dado por la expresión Ist·Vdd.

Veamos un ejemplo , para una tecnología de 0.25μ CMOS, las siguientes configuraciones de circuitos obtienen el mismo rendimiento;

Vdd=3V; Vt==0.7V

Vdd=0.45V; Vt=0.1V

La diferencia es que el consumo de potencia del segundo es 45 veces mas pequeño que el del primero.

- En el caso de los circuitos dinámicos El Vt permitido viene determinado por la frecuencia de operación mínima (para prevenir la perdida de información). En este caso el ahorro de potencia no es tan importante como en el caso estático y se limita a un factor de 8 cuando se mantiene el nivel de rendimiento (para un Vdd=1V; Vt=0.3V).

Corolario: Los circuitos estáticos son mas útiles para escalar el Vdd que los dinámicos.

- **SOLUCIONES ARQUITECTÓNICAS**
  - En la mayoría de los casos la modificación de los parámetros del proceso no está en las manos del diseñador.
  - La perdida de rendimiento debido a la disminución de Vdd debe tratarse por otros medios.
  - Métodos habituales en estos casos son las modificaciones de arquitectura. De todos es sabido que el rendimiento no esta ligado únicamente a la frecuencia del reloj.
  - Arquitecturas paralelas pueden provocar un incremento del área. Precio área se está reduciendo con el incremento de la tecnología.

9.10.2 REDUCCION DE LA CAPACIDAD EFECTIVA

Cuando el límite inferior del potencial V_{DD} es una ligadura externa que no se puede modificar o cuando la degradación del rendimiento debido a la disminución de V_{DD} es intolerable, la única opción es la disminución de Cef = $C_L \cdot P_{0\rightarrow 1}$. Esto se puede conseguir por dos medios:

- Reducir C_L
- Reducir $P_{0\rightarrow 1}$

ν **REDUCIR C_L**

- Es un objetivo global que ayuda también a mejorar el rendimiento del circuito.
- **Elección correcta de la lógica**
  - ∗ Muchos de los estilos lógicos producen una importante reducción de C_L, reduciendo el consumo
  - ∗ P.e.: el sumador Complementary Pass-Transistor Logic usa un 30% menos de potencia que el sumador CMOS estático.

- **Tamaño de los transistores**
  - ∗ La mayor parte de la capacidad se debe a las capacidades del transistor.
  - ∗ Se debe intentar que los transistores tengan tamaño mínimos siempre que sea posible o razonable. Esto afecta al rendimiento, mejorable mediante técnicas lógicas y de arquitectura.
  - ∗ El único caso en que los transistores deberían sobredimensionarse es cuando la capacidad de carga esta dominada por las capacidades extrínsecas al transistor
    - ➢ Fan-Out
    - ➢ Capacidades de hilos
  - ∗ Esto es lo contrario a lo que se hace al diseñar celdas estándar porque los transistores se hacen grandes para que se acomoden a un rango elevado de necesidades de carga y rendimiento.

∗ Vamos a analizar el caso de in inversor estático con una capacidad de carga CL compuesta de una capacidad interna y otra externa de tal manera que :

$C_L = C_{EX} + N \cdot C_{INT}$

Donde:

C_{INT} es la capacidad interna de un inversor de tamaño mínimo

N es indica el tamaño del inversor .Si N=1 los transistores son de tamaño mínimo.

Siendo $\alpha = C_{EXT}/C_{INT}$

Para $\alpha \leq 1$

Domina la capacidad intrínseca

cuando crece N (tamaño del transistor, crece el consumo).

Para $\alpha > 1$

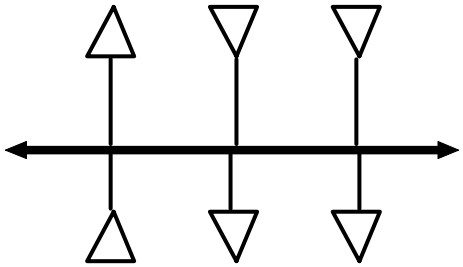
domina la capacidad extrínseca

para N grandes disminuye el consumo

- **Cambios de arquitectura**

∗ Otra aproximación para reducir la capacidad física es evitar en la medida de lo posible compartir recursos, como es el caso de una arquitectura común.

∗ Sea el siguiente BUS común.

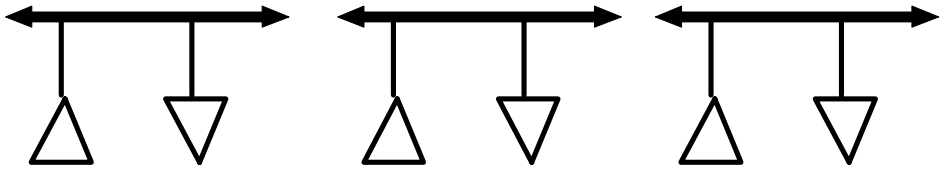


∗ Se conectan a un bus único varios drivers y varios receivers

∗ Esta configuración tiene una capacidad de BUS muy elevada y probablemente una longitud también grande

∗ Desde el punto de vista del consumo de potencia, es preferible la estructura de varios BUSES punto a punto, efecto negativo sobre el área del chip.

∗



9.11ELECCIÓN DEL ESTILO LÓGICO

- Una vez vistos los diferentes estilos de diseño que existen hay que decidir cual de ellos se utiliza.
- La elección de un estilo de diseño viene determinada por la ligadura principal del sistema, que puede ser una de las siguientes:
    - ∗ Facilidad de diseño
    - ∗ Robustez frente a los ruidos
    - ∗ Área
    - ∗ Velocidad
    - ∗ Consumo de potencia
- Es importante darse cuenta que ninguno de los estilos de diseño optimiza todas estas ligaduras simultáneamente.

ν **LA CMOS COMPLEMENTARIA**

- Es la mejor opción en la mayoría de los circuitos
- Ventaja
    - ∗ Robusta a los ruidos.
    - ∗ No disipa potencia estática
    - ∗ Esto tiene como consecuencia que el diseño está libre de problemas de comportamiento lógico, y además las hace fáciles de automatizar.
- Desventajas
    - ∗ Para puertas con un Fan-in elevado tiene importantes penalizaciones tanto en área como en rendimiento.

ν **LÓGICA PSEUDONMOS,**

- Ventaja
  - ∗ Que es simple y rápida
- Desventaja
  - ∗ Tiene reducidos márgenes de ruido
  - ∗ Consumo de potencia estático
    - ∗ Este se puede reducir mucho colocando un PMOS de carga muy pequeño y otro en paralelos que se active solo cuando sea necesario
  - La más útil para puertas NOR de Fan-in elevado.
  - Se suele utilizar mucho en ROMS ,PLAS y los circuitos de Carry-look ahead en los sumadores

ν **LÓGICA DE PUERTAS DE TRANSMISIÓN**

- Se pueden conseguir velocidades elevadas de funcionamiento si las estructuras se limitan a unas pocas puertas de transmisión serie
- Se suele utilizar para funciones booleanas complejas en los que se tienen que minimizar o el área o el consumo
- No existen muchas herramientas automáticas orientadas a esta lógica. La recomendación sería comparar la densidad, velocidad, potencia y facilidad de diseño con el circuitos cmos complementario equivalente para justificar su uso.
- En algunas ocasiones los fabricantes proporcionan tecnologías de bajos potenciales umbral para la utilización de este estilos de diseño

ν **LA LÓGICA DINÁMICA**

- Facilita la implementación de puertas pequeñas y rápidas
- Los problemas de distribución de carga hacen duro y difícil el diseño .
- Las corrientes de leackage fuerzan al refresco periódico de la información lo que limita el rendimiento de estas lógicas.
- **LÓGICA DOMINÓ**
  - ∗ Se debe usar para aplicaciones de bajo consumo o alta velocidad
  - ∗ Se debe tener cuidado con los efectos de la distribución de carga
  - ∗ Si no se desea la simulación exhaustiva de las puertas a nivel de circuito, una vez que se conocen las capacidades después de su extracción postlayaut mejor que no se utilice esta técnica
  - ∗ Se debe tener cuidado con el estudio de los tiempos de precarga, un mas estudio puede hacer que esta lógica pierda sus ventajas de rendimiento frente a la lógica

- Lo cierto es que la tendencia actual es al mayor uso de la lógica CMOS complementaria debido a la facilidad de implementación que proporcionan la herramientas automáticas
- A continuación aparece una tabla con las características de según los distintos estilos para una puerta NAND de 4 entradas

Estilo	**Proporcional**	**W estática**	**Nº de transistores**	**Area (μ2)**	**Retardo(ns)**
Complementario	no	no	8	533	0.61
pseudo-NMOS	si	si	5	288	1.49
CPL	no	no	14	800	0.75
Dinámico NP	no	no	6	212	0.37

10 DISEÑO SECUENCIAL

Los circuitos combinacionales tiene la característica que las salidas dependen única y exclusivamente de las entradas. Los circuitos secuenciales se caracterizan en cambio porque las salidas dependen tanto de las entradas como de la historia del sistema, por lo tanto los sistemas secuenciales deben guardar la historia del sistema por lo tanto necesitan elementos de memoria

Los elementos de memoria se pueden descomponer en estáticos cuando se utiliza una realimentación positiva para almacenar el dato, y dinámicos cuando se utilizan las capacidades parásitas para almacenar la información

por otro lado los sistemas pueden ser síncrono - un reloj coordina todo el sistema - o asíncronos.

10.1TIPOS DE SISTEMAS SECUENCIALES

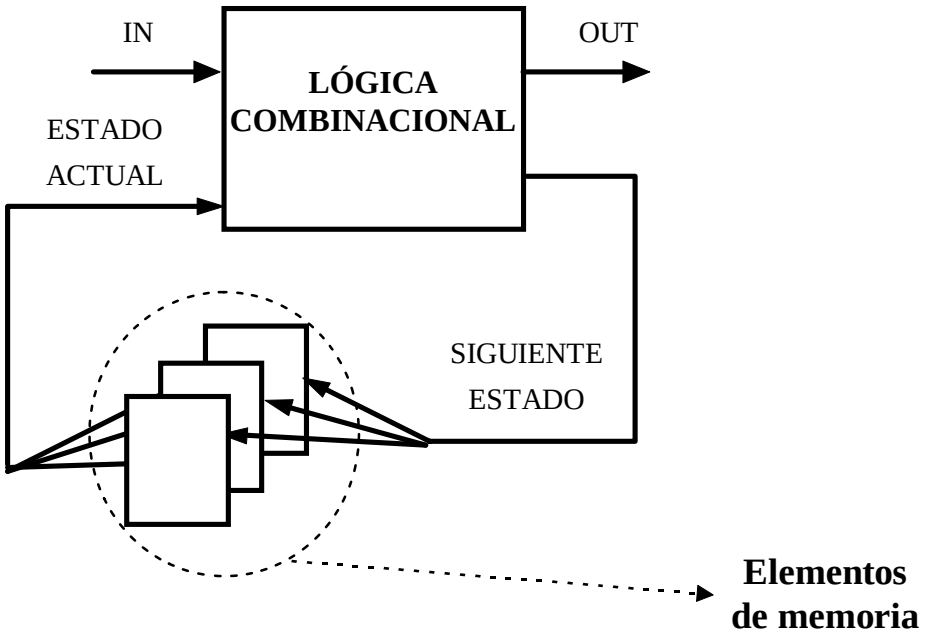
Existen dos formas típicas de usar los almacenamientos

La máquinas de estados finitos

Los Pipe-line

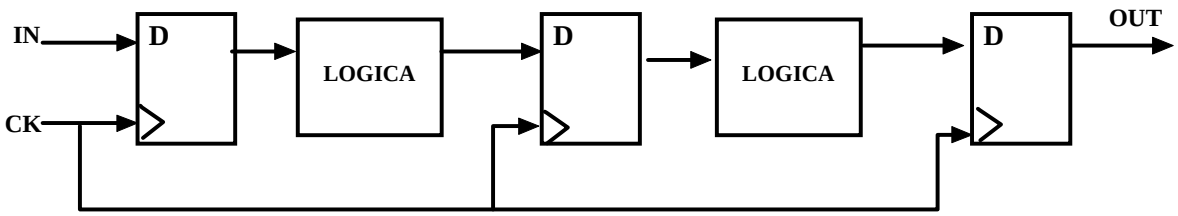
ν **La Maquina de estados finitos (FSM)**

- Conjunto de entradas que alimentan un bloque combinacional que da como resultado una salida y el siguiente estado.
- Se caracteriza por la realimentación.
- Cuando la señal de reloj pasa de 0 a 1 , el siguiente estado pasa as er el actual cargándose en los elementos de memoria. A continuación a través de la lógica combinacional se genera el siguiente estado
- Cuando la señal out es estable, se puede generar la nueva carga de CK.
- El mínimo tiempo en que las salidas y el siguiente estado están estables determina la máxima frecuencia de trabajo.
- Se suele utilizar para unidades de control.



ν **Pipe-Line**

- El procesamiento de un dato se puede dividir en etapas para acelerar el procesamiento.
- Utiliza elementos de memoria para capturar la salida de proceso al final de cada circuito de CK. Tantos ciclos de CK como etapas
- Superposición de procesamiento
- No necesita que acabe el proceso de un dato para comenzar con el siguiente.
- En este caso no hay realimentación.
- Se suele utilizar en Data-Path.



ν LAS ESTRATEGIAS DE SINCRONIZACIÓN

- Están definidas por los conjuntos de CK que se usan para almacenar y acceder a sus contenidos.
- La selección de una determinada estrategia de sincronización influye en:
  - Numero de transistores por unidad de memoria
  - Número de señales de CK.
- Estas decisiones influyen en:
  - Tamaño C.I.
  - Consumo de potencia.

10.2 TIEMPOS RELEVANTES EN LA CARGA DE UN DISPOSITIVO

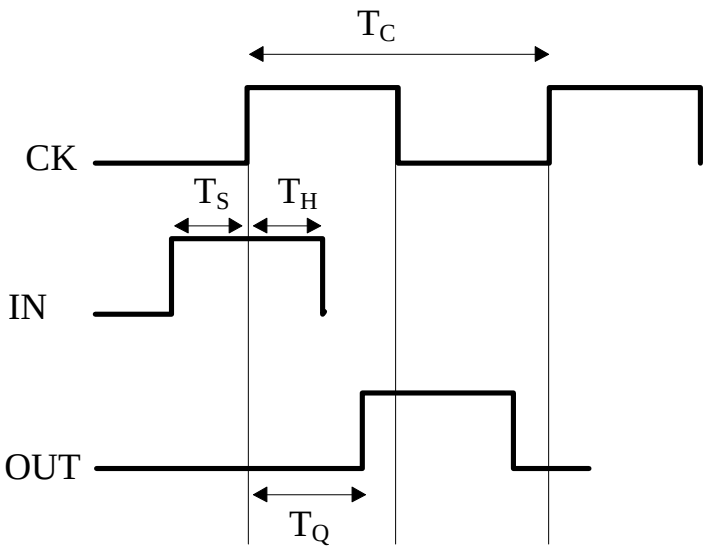
Tiempo de Set-up (T_S) . el tiempo que el dato debe estar estable a la entrada de la puerta de memoria antes que llegue la CK. (Ts).

Tiempo de Hold (T_H) Tiempo que debe estar estable el dato después de que llegue la señal de CK.

Tiempo de clock-to-q (T_Q) Tiempo desde que llega le señal de CK hasta que la salida Q esta estable

Tiempo de ciclo (T_C). Tiempo entre dos subidas de CK consecutivas

Las celdas de memoria se pueden clasificar en caches y registros según sea la señal que produce la carga.



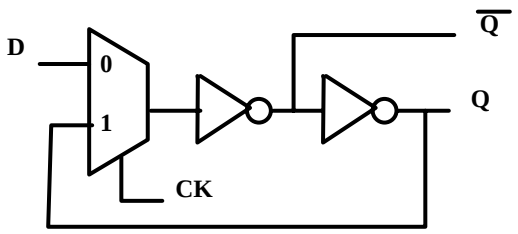
10.3ELEMENTOS DE MEMORIA

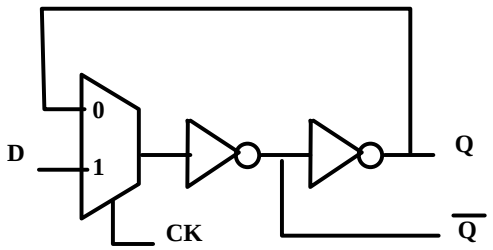
- Son de dos tipos latch y registros
- **Latch**

  Carga por nivel.

  Mientras el CK=1 la entrada se comunica a la salida (latch transparente).

  El dato debe estar estable durante todo el CK=1.





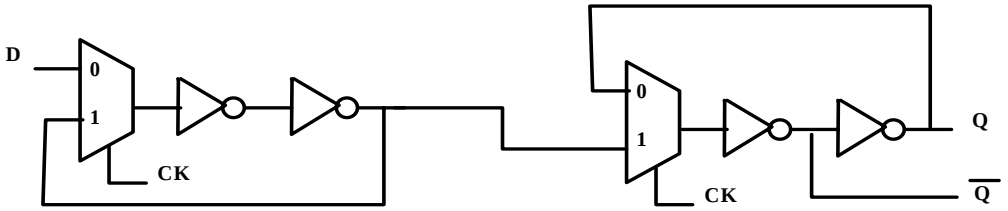
- **Registro**

  Disparo por flanco.

  El dato se carga justo en el instante CK=1.

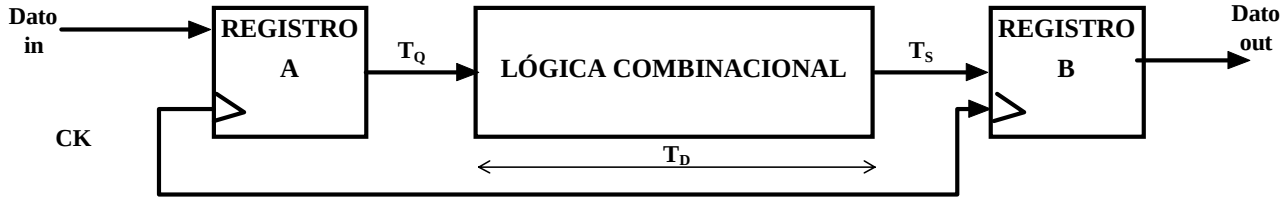
  Combinación de dos Latches (master-slave).

  El mas habitual es el D



10.4PIPELINE CON REGISTROS Y CON LATCHES

10.4.1CON REGISTROS:



$T_C = T_Q+T_D+T_S$

Donde:

TC es el tiempo de ciclo

TQ es el tiempo que clock to q

TD retardo a través de la lógica combinacional

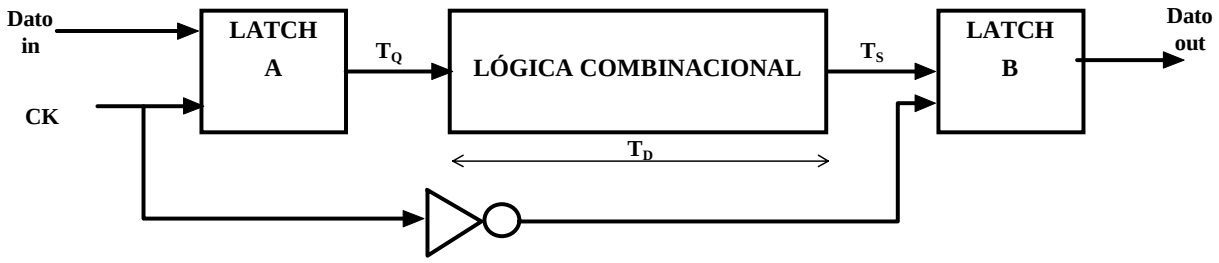
TS tiempo de set _up del registro B

La señal de reloj es la misma para los dos registros y esto facilita el diseño

Se colocan registros en serie con el flujo lógico para reducir el tiempo de ciclo Tc cuando Td es demasiado grande.

En ningún instante está comunicada directamente la entrada de datos con la salida de datos

10.4.2 CON LATCHES



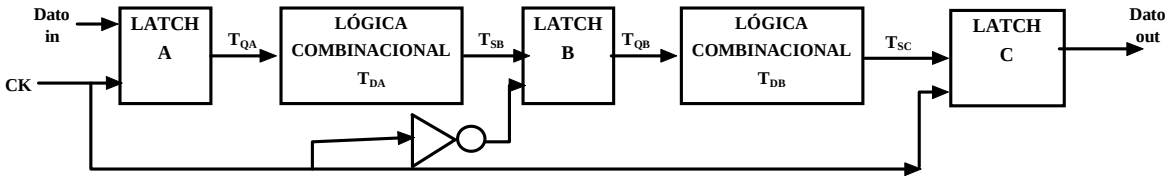
Utiliza aproximadamente la mitad del área

LA y LB están controladas por señales de CK complementarias

como una de las características del latch es la trasparencia cuando CK=1, esta inversión en la CK evita que se trasmitan a OUT las variaciones de IN.

Aísla la entrada de salida.

Vamos a suponer Tc<<Td entonces convendría fraccionar la lógica combinacional mediante un nuevo Latch resultando lo siguiente:



Notar que con menor numero de Latches que con registros se consigue minimizar Tc. Sin embargo el estudio de la sincronización se hace mucho mas complejo.

El primer bloque de lógica combinacional. recibe el dato T_{QA} más tarde la señal de reloj y lo debe entregar al Latch B T_{SB} antes de que baje la señal de reloj CK luego se debe cumplir

$$T_{QA} + T_{SB} + T_{DA} < T_{C1}$$

Siendo TC1 el tiempo de ciclo parcial para el primer bloque de lógica combinacional

Haciendo un estudio similar para el segundo bloque de lógica combinacional

$$T_{QB} + T_{SA} + T_{DB} < T_{C0}$$

como $T_{C1} = T_C/2$ Y $T_{C0} = T_C/2$ y justo en el limite para que se cumplan ambas condiciones:

$$T_{C1} = T_{DA} + T_{QA} + T_{SB}$$

$$T_{C0} = T_{DB} + T_{QB} + T_{SA}$$

Suponiendo los latch idénticos

$$T_{SB} = T_{SA}$$

$$T_{QB} = T_{GA}$$

$$T_C = T_{C1} + T_{C0}\ ;\ \mathbf{T_C = T_{DA} + T_{DB} + 2[T_Q + T_S]}$$

La estrategia con registros ocupa mas área, pero es mas sencilla de diseñar porque todos los registros se cargan con el mismo flanco.

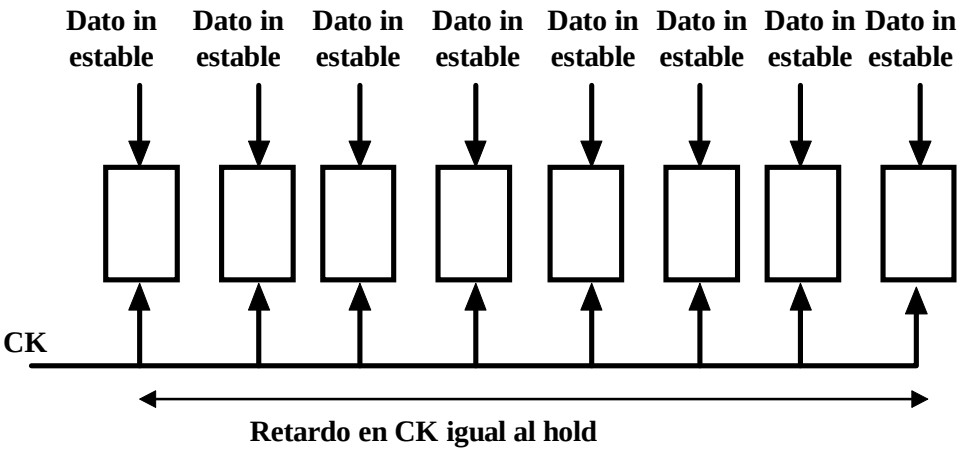
10.5TIEMPOS DE SET-UP Y DE HOLD

Tiempos de Hold

Tiempo que el dato de in debe estar estable después de haber subido la señal de reloj

Sirve para evitar que se cargue un dato erróneo.

Se debe al Retardo de tiempo que tiene la señal de CK en llegar al último elemento de memoria de un registro elemento de memoria.



Tiempo de Set-Up

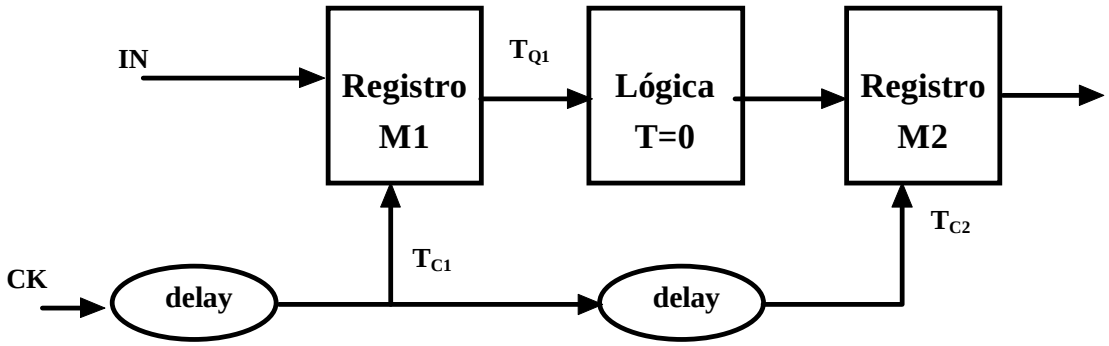
El tiempo que debe estar estable la señal de entrada antes de que llegue la señal de reloj

Se debe al tiempo que tarda el dato en llegar al último de elemento de memoria.

Si en un sistema síncrono no se cumplen los tiempos de Set-Up y de Hold se produce un mal funcionamiento del circuito.

Vamos a estudiar el problemas con un poco mas de detenimiento.

Sea la siguiente figura:



Suponemos que el bloque lógico tiene retardo = 0. Un dato no aparece a la entrada de M2 hasta que ha pasado un tiempo:

$$T_{C1} + T_{Q1}$$

desde que CK sube

Donde

T_{C1} es lo que tarda la señal de CK en transmitirse hasta M1

T_{Q1} es lo que tarda la IN en llegar a Q desde que llega T_{C1}.

Siendo T_{C2} el tiempo que tarda la señal CK en transmitirse hasta M2 ($T_{C2} > T_{C1}$).

ν **$T_{C2} > T_{C1} + T_{Q1}$**

el modulo M2 almacenaría un dato erróneo.

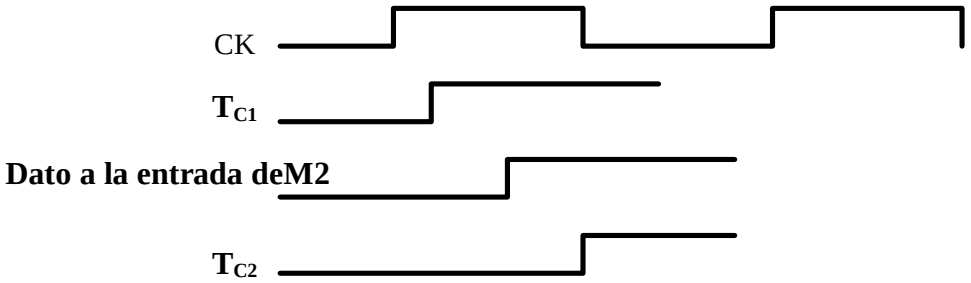
Violación del tiempo de Hold.

El dato no permanece el tiempo suficiente a la puerta M2 después de CK sube para que se cargue el dato correcto en M2.

La señal de reloj CK debería activar la carga de los registros M1 y m2 simultáneamente. Ahora bien, debido a los retardos esas cargas no se producen simultáneamente.

Como se carga primero M1 y el retardo de la señal de reloj es grande con respecto al de la transmisión de datos da tiempo a que el dato a la entrada de M2 se vea modificado; Carga un dato erróneo.

Las formas de onda serían la siguiente:



ν **$(T_{C1} + T_{Q1}) > T_C + T_{C2}$**

El registro M2 se carga antes de tener un dato correcto en la entrada, debido a que el retardo $T_{C1} + T_{Q1}$ es demasiado grande.

El tiempo que tarda en llegar el dato a la entrada de M2 es mucho mayor que el tiempo que tarda en llegar la señal de carga del siguiente dato;

Violacion del Set-Up.

Circuitos demasiado lentos para el ciclo de CK usado.

El reloj T_{C2} cambia mas rápido que el dato.

ν **SOLUCIONES**

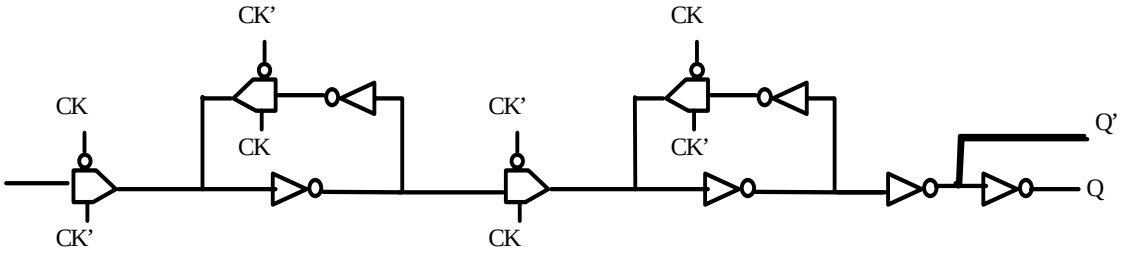
Estos dos problemas pueden evitarse poniendo retardos en las líneas adecuadas;

En la línea combinacional.

En T_{C2}

10.6 ESTRUCTURAS DE MEMORIA DE UNA FASE DE RELOJ

10.6.1 REGISTRO ESTÁTICO DE UNA FASE DE RELOJ



Básicamente es el mismo que hemos estudiado hasta el momento.

Se suelen añadir buffers en la entrada y en la salida.

Las salidas Q y Q se toman del mismo punto de salida (antes se tomaba la salida complementada entre los dos latches).

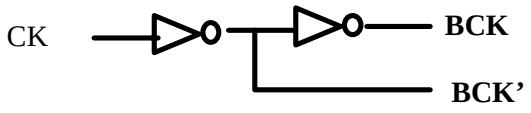
Es típico de Gate Arrays y celdas estándar.

En diseños custom es deseable reducir el numero de transistores, lo que se puede hacer según el tratamiento que se de al reloj.

ESTUDIO DE CK

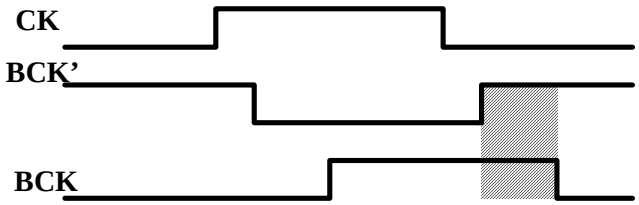
Para les celdas estándar

se utiliza un buffer de reloj para cada celda de memoria, como se puede ver en la figura;



Otra opción puede ser utilizar un único Buffer para todo el registro en lugar de utilizar uno para cada celda.

Este buffer de la señal de reloj puede dar problemas debido a los retardos que aparecen puede provocar una importante desviación entre BCK y BCK'.

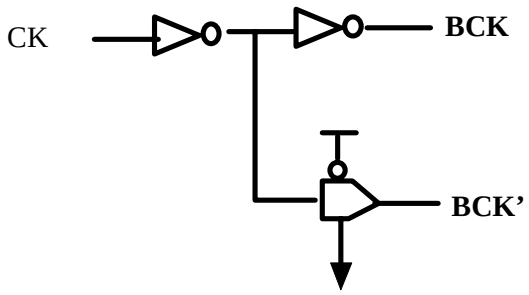


Como hay un periodo de tiempo en que ambas señales están a 1, todas las puertas de paso dejan pasar el dato y aparece una realimentacion, lo que produce la carga de un mal dato, no esta aislada la carga de la salida.

Solución:

Añadir una puerta de transmisión cuyos transistores tengan el mismo tamaño que los de los inversores.

Las puertas de paso deben estar siempre conduciendo.



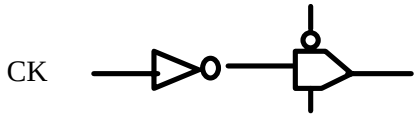
10.6.2 OTROS REGISTROS ESTÁTICOS

Guardan el estado cuando CK esta parado y se mantiene la alimentación, gracias a la realimentación.

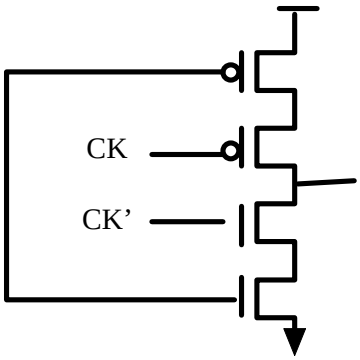
Como hemos visto los registros se componen de dos Latches, luego estudiando diferentes tipos de latches se ven diferentes tipos de registros.

10.6.2.1 Latch estático c^2MOS

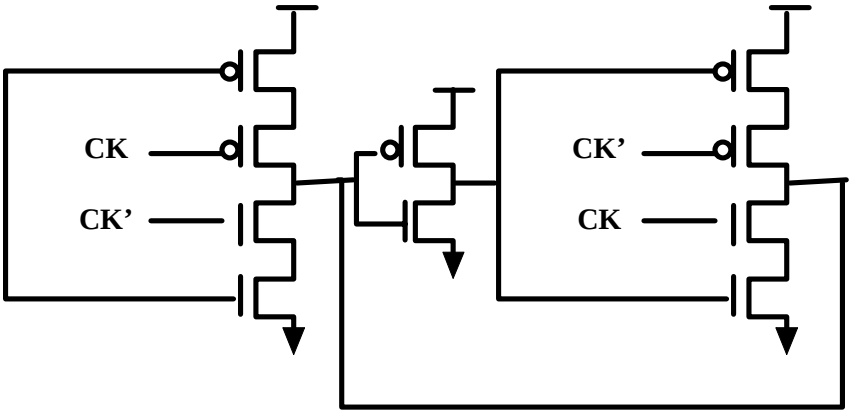
Partiendo del Latch estático cMOS de entrada bufferada Un inversor seguido de una puerta CMOS de paso



es equivalente a :



Sustituyéndolo en el esquema del latch se obtiene:



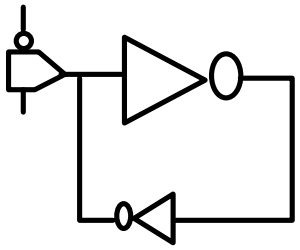
- La funcionalidad es la misma pero tienen muchas ventajas a la hora de diseñar el layout

- Elimina conexiones de metal.

- Menor área

- Diseño menos complejo.

10.6.2.2 Latch con solo una puerta de transmisión.

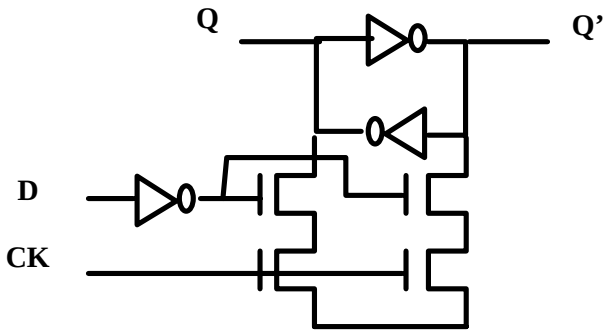


Las ganancias del inversor de realimentación deben ser mucho menores que las del otro inversor,

esto se consigue haciendo la longitud del canal mayor que la longitud mínima del proceso.

La puerta de paso debe tener una ganancia suficientemente grande como para vencer el valor que el inversor de realimentación pone a su salida.

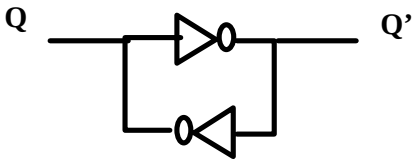
10.6.2.3 Celda RAM estática



CK =0,

caminos a tierra cortados.

El circuito equivalente conserva el dato.

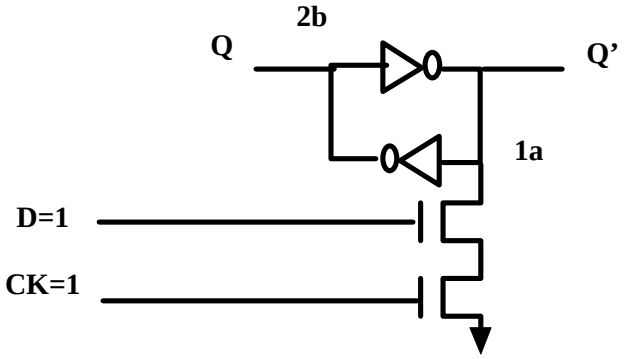


CK =1,

un camino a tierra abierto y otro cerrado.

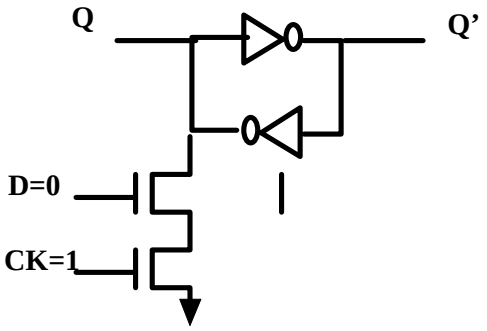
Dato = D = 1.

Un camino a tierra con la suficiente ganancia para poner un 0 en 1(a) y un 1 en 2(b).



Dato = D = 0.

Camino a tierra con la suficiente ganancia para poner un 0 en 2(b) y un 1 en 1(a).

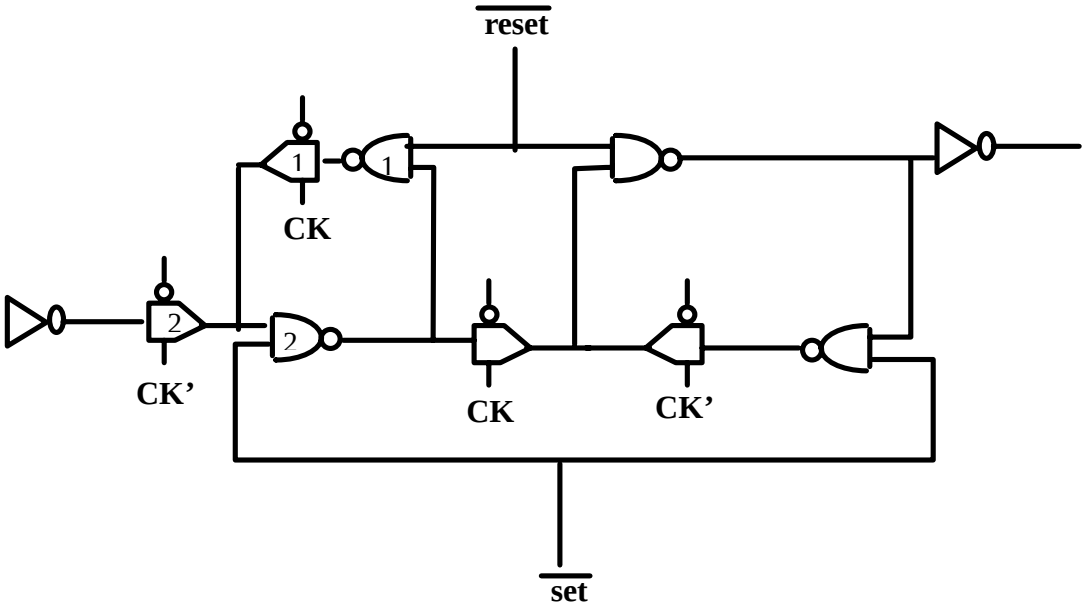


Características:

- Márgenes de ruido reducidos.
- Requiere de un cuidadoso diseño.
- Pequeños y rápidos.

10.6.2.4 Registros con Set y Reset asíncronos

Son registros en los que se pueden cargar 0's o 1's de manera sincrona o asincrona. En lugar de inversores utilizamos puertas NAND.



Funcionamiento.

Supongamos CK = 1.

Puerta 1 conduce.

Puerta 2 cerrada.

Si reset = 1 y set = 0

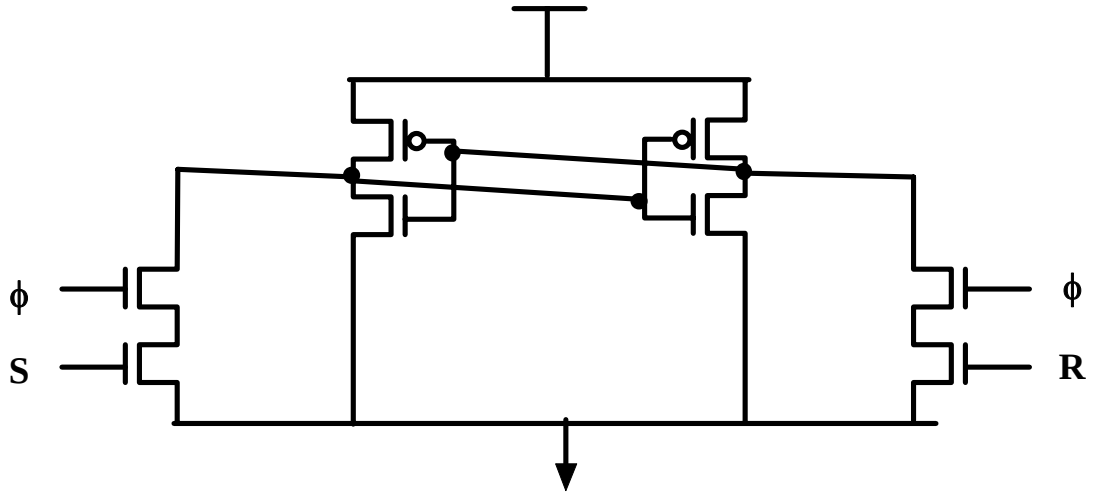
NAND1 (1'·X)'=1

NAND2 (1·0')'=Q=0

Si reset = 0 y set=1

NAND2 (1'·X)'=1=Q

10.6.2.5 registro con set y reset síncrono:



Que diferencia existe con el anterior?

1. En este el Set-Reset es síncrono.
2. En el otro había tres entradas (dato, set, reset) en este solo hay dos.

10.6.3 REGISTROS DINÁMICOS

Las celdas de memoria son muy robustas y poco sensibles a las distorsiones en las líneas de señal.

Su principal desventaja es su complejidad, sobre todo en estructuras computacionales tales como caminos de datos segmentados. El gran tamaño de los registros es un factor dominante

Cuando la frecuencia de trabajo de una memoria o registro va a ser elevada, para reducir el tamaño y la complejidad se pueden reducir los transistores de realimentación y las puertas de transmisión. El dato se guarda en la capacidad de salida.

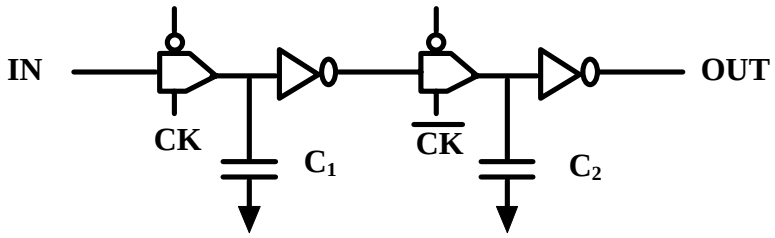
Como los capacitores no son ideales tienen perdidas.

- El dato solo se guarda una cantidad de tiempo (milisegundos).
- Para conservar la señal correcta es necesario un refresco.

Para que la lectura no destruya el dato es necesario que los dispositivos tengan altas impedancias de entrada.

- Esta condición es típica de los CMOS
- Difícil en los bipolares (solo estáticos).

10.6.3.1 Registro de desplazamiento dinámico



Cuando CK = 1:

La primera puerta conduce la segunda cortada.

La C_1 se carga con el valor de entrada.

La C_2 esta aislada con el ultimo valor.

Cuando CK = 0;

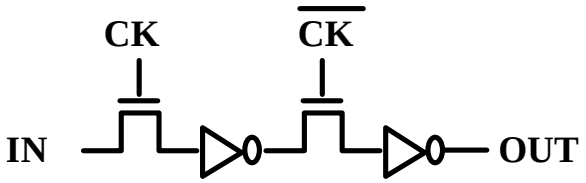
La primera puerta se corta

la segunda conduce el dato cargado en C1, se traslada a C1.

¿Porque pintamos las capacidades a la entrada de los inversores y no a la salida?

Porque son capacidades que quedan aisladas de la fuente.

Este registro se puede implementar con transistores de paso en vez de utilizar puertas de paso.



ventaja

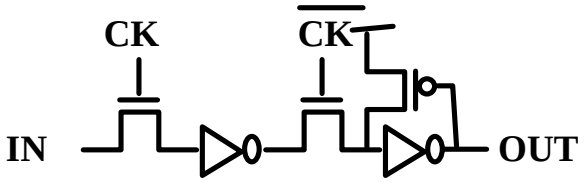
Ahorra área.

Problema,

mala transmisión del 1,

corriente en fase estática que produce la descarga de las capacidades.

Solución:



En general, el principal inconveniente de los registros de desplazamiento de una sola fase de reloj es la posibilidad de que los relojes se solapen;

CK = CK’ = 1. Queda transparente el registro.

Conviene saber que es difícil impedir que dos señales de reloj se solapen cuando tienen mucho camino que recorrer.

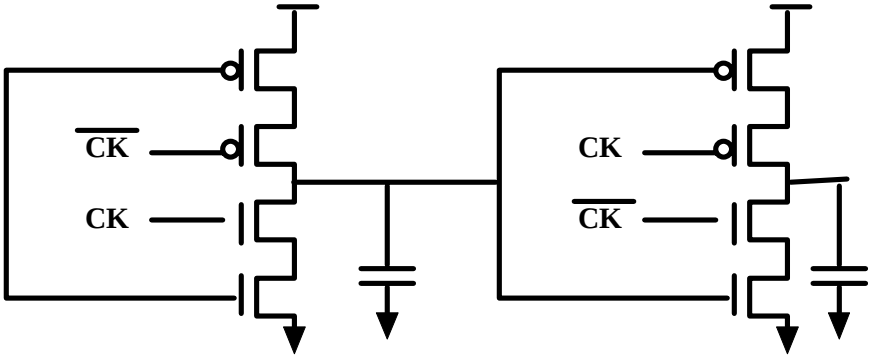
Solución,

generación interna al registro.

Conviene darse cuenta de que el numero de transistores es inferior al de un registro estático.

Con los registros dinámicos C2MOS que se explican a continuación se consigue mejorara notablemente el diseño del layout.

10.6.3.2 Registro C^2MOS



Soluciona el problema del solapamiento y el del Layout.

ν **Modo de operación.**

- CK = 1.

  El primer C^2MOS actúa como un inversor → Fase evaluación.

  El segundo se encuentra en alta impedancia→ fase Hold

- CK = 0.

  Se invierten los papeles

  El valor almacenado en G se propaga a la salida.

En definitiva el circuito actúa como un biestable de master-slave de carga por flanco

ν **Ventaja:**

No se ve afectado su comportamiento por el solapamiento de las señales de reloj (siempre que los tiempos de subida y de bajada sean lo suficientemente rápidos).

Demostración:

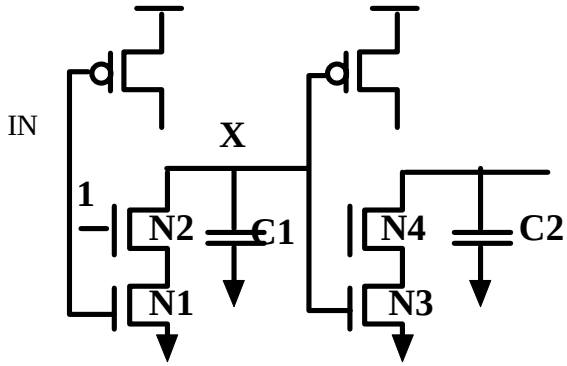
El solapamiento es imposible,

si $\phi = \phi' = 1$ solo Pull-Down conduce.

Si $\phi = \phi' = 0$ solo conduce Pull-Up.

Para que exista propagación a la salida es necesario que existan Pull-Up seguidos de Pull-Down o viceversa.

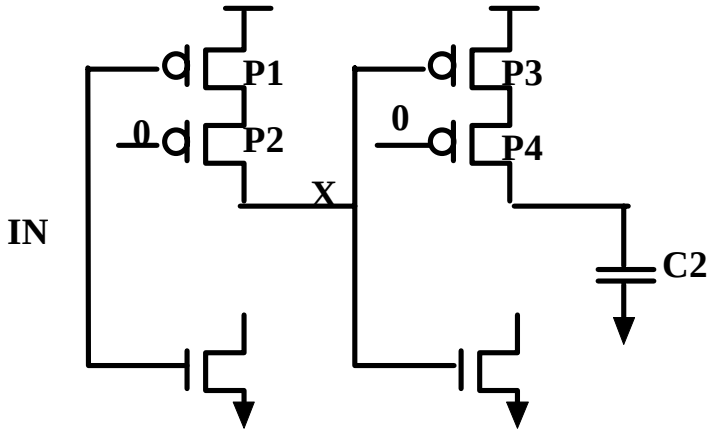
La capacidad intermedia esta aislada del de entrada, no existe transparencia.



CK=CK'=1

Si IN=1→X=0→ N3 cortado→ C2 aislado

sSi IN==→N1 cortado



CK=CK'=0

si in=0→X=1→P3→ C2 aislado

si in=1 P1 cortado

Solo cuando los tiempos de subida y de bajada son demasiado lentos se producen problemas al estar conduciendo todos los transistores por lo que puede aparecer un camino entre la entrada y la salida produciendo errores.

La conducción para que el funcionamiento sea correcto, los tiempos de subida y de bajada sean menores que el retardo multi0plicado por 5.

El Latch c2MOS es util para operaciones de alta velocidad.

En estos casos es casi imposible evitar que las señales se solapen.

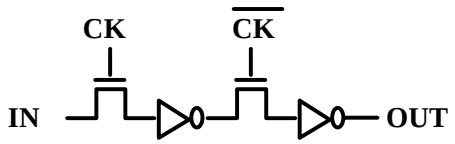
Al necesitar menos contactos el Lay-Out es mas compacto.

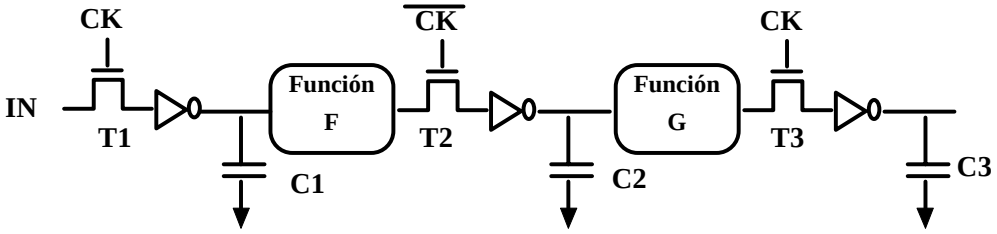
También es valido para las dos fases de reloj.

10.7 ESTRUCTURAS LÓGICAS DE UNA FASE DE RELOJ

Sirven para implementar pipesImplementar pipelines

Partiendo de un registro de desplazamiento dinámico se puede realizar la siguiente implementación de un Pipe Line:





Los módulos F y G se implementan en lógica combinacional estática.

En el instante *T* CK = 1.

T1 conduce→ C1 se carga con el valor de la entrada

T2 cortado→ C2 almacena el valor cargado en el instante *T*-1.

Las etapas F y G están aisladas.

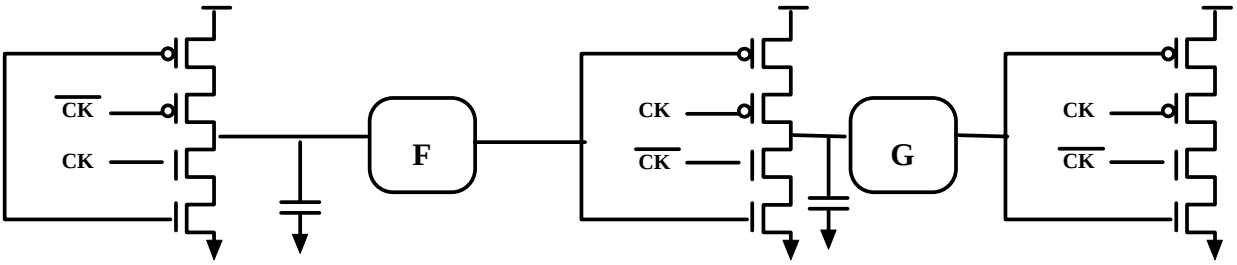
Problema:

¿Que ocurre cuando CK = CK’ = 1?.

Las etapas dejan de estar aisladas y se produce mal funcionamiento.

Solución

Parte del problema se puede solucionar utilizando lógica C^2MOS para implementar el Pipe Line.



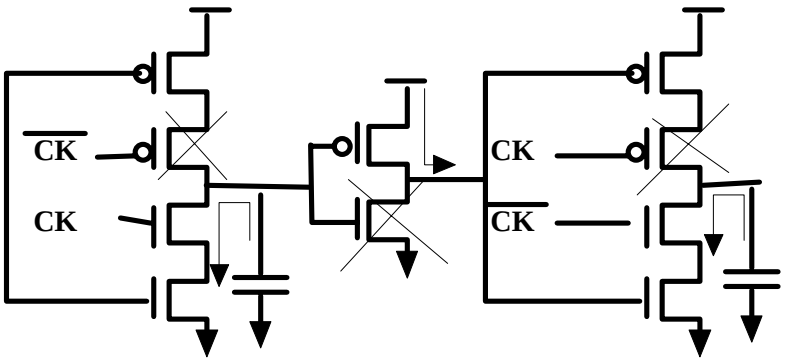
Evita el problema de los solapamientos siempre que la función F (implementada con lógica estática) no este invertida.

Esta condición no aparecía cuando hablábamos de los registros dinámicos C^2MOS.

Cuando CK = CK' = 1 los C^2MOS se reducen a mallas de Pull-Down.

El único caso en que la señal se puede transmitir es cuando la lógica combinacional es invertida (alteran Pull-Up Pull Down).

Ejemplo:



Suponemos IN = 1

F = inversor

C_1 se descarga;→ 0 a la puerta inversor, →1 a la puerta del segundo C^2MOS. → 0 a la salida del segundo C^2MOS. No esta aislada la entrada de la salida: **error**.

10.8LÓGICA NORA

En lugar de utilizar lógica combinacional para implementar el Pipe line utilizamos lógica dinámica.

NORA = NO RACE LOGIC y esta es precisamente su principal característica, que asegura el aislamiento de los nodos en las fases de alta impedancia.

Se utiliza para implementar PipeLines rápidos.

Existen dos posibilidades de implementación diferente:

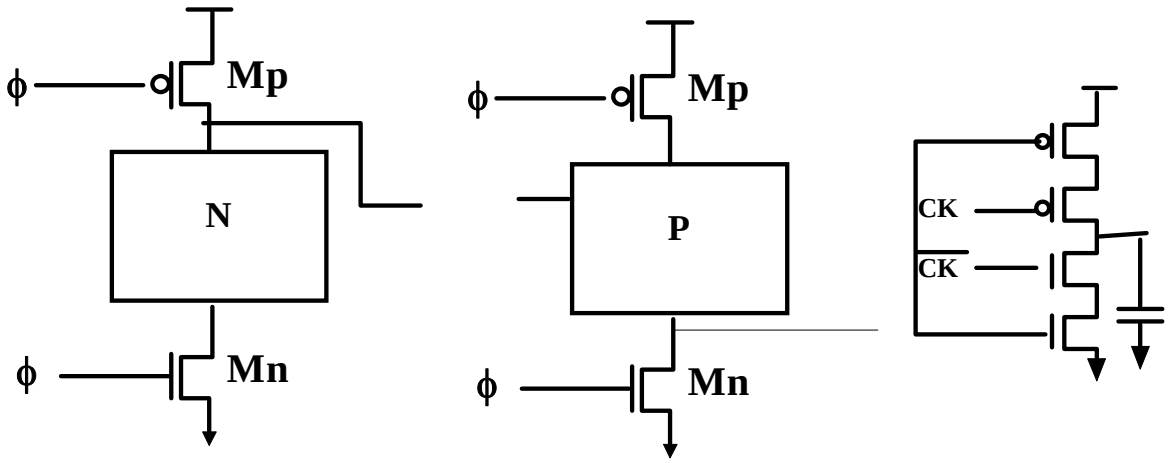
registros C^2MOS y lógica dinámica NP CMOS.

Registros C^2MOS y una mezcla de lógica estática y dinámica.

10.8.1ESTRUCTURA NORA N-P-C2MOS

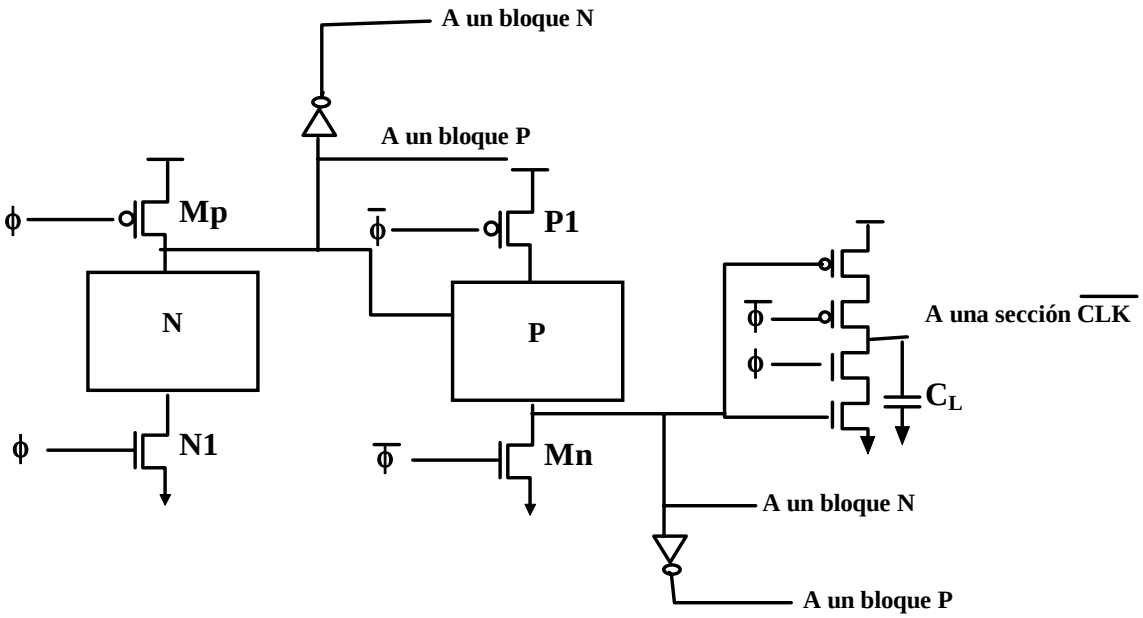
Optimiza la velocidad y la densidad a costa de un diseño mas complejo.

Combina secciones de lógica domino N-P con latches C^2MOS como etapas de salida.



Existen dos tipos e bloques diferentes, los CK y los notCK

ν BLOQUES CK



Las entradas tienen que venir de un bloque notCLK.

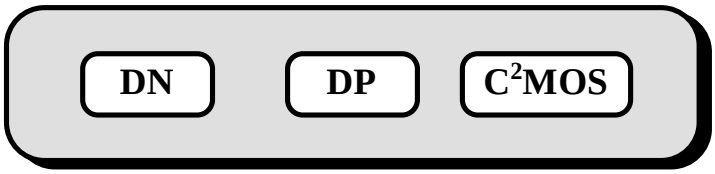
Las salidas van a un bloque notCLK.

Se evalúa durante CK=1.

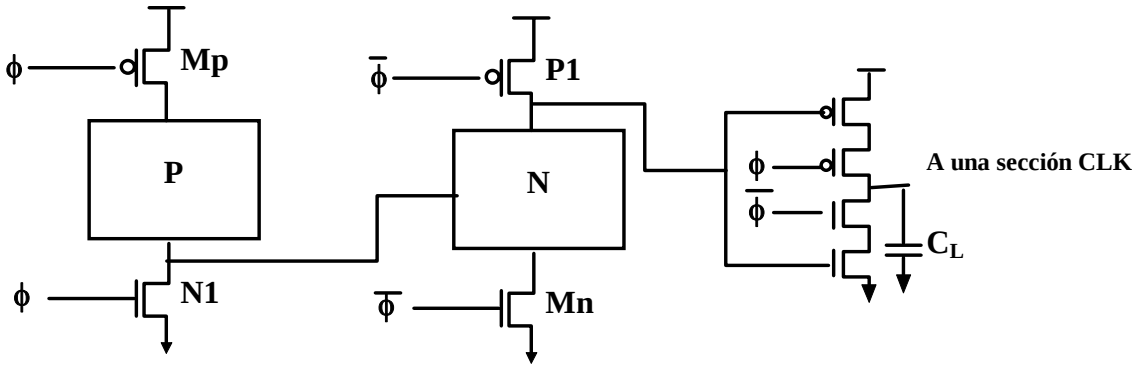
CK=0 N1 cortado→ P1 cortado C^2MOS cortado→CL aislado.

CK=1 se evalúa N se evalúa P se carga C_L

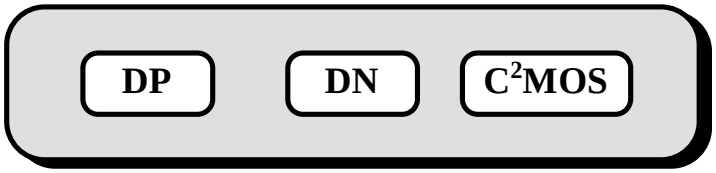
El esquema general del, bloque CLK es:



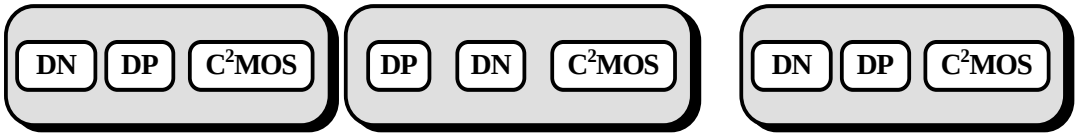
ν **BLOQUE NOTCLK**



Se evalúa durante CK = 0.



Para conseguir bloques en cascada se hace de la siguiente manera:



CLK	notCLK			
0	1	Evaluación	Precarga	Evaluación
1	0	Precarga	Evaluación	Precarg

Recordar. Gracias al Latch C^2MOS en las fases de precarga la salida de los bloques permanece constante y almacenada en las C_L debido a que no existe camino a tierra.

10.8.2 LÓGICA NOR MIXTA

El estilo de diseño ofrece un ancho rango de diseño:

Se pueden mezclar lógica dinámica con lógica estática.

Se pueden usar bloques dinámicos P o N tanto en cascada como en Pipe Line.

Esta libertad de diseño tiene como consecuencia que etapas extras de inversores puedan desaparecer.

Ejemplo de etapas extras: diseño PP domino y el NN domino.

REGLA DE DISEÑO

- **Reglas para la lógica dinámica.**

A los bloques N solo se les permite una transición 0 a 1 durante la fase de evaluación.

A los P de1 a 0 (notar que esto se extrae de la lógica domino).

- **Reglas C^2MOS.**

suponiendo lógica combinacional estática

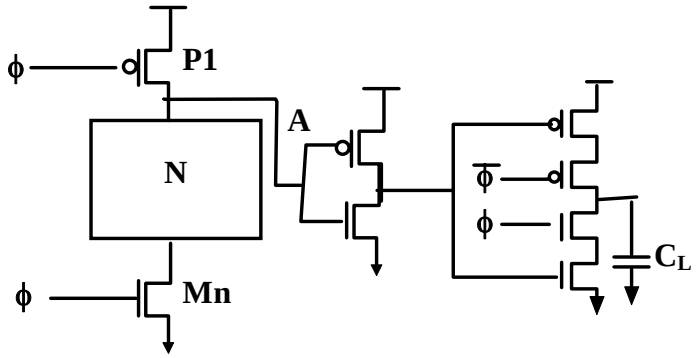
para evitar "races". el número de inversores entre latches C^2MOS debe ser par. Esto es equivalente a evitar las funciones invertidas entre Latches C^2MOS cuya problemática ya estudiaremos.

-**Regla C^2MOS**.

suponiendo Lógica combinacional mixta o dinámica.

La utilización de lógica dinámica tiene como consecuencia la ampliación de la primera regla C^2MOS.

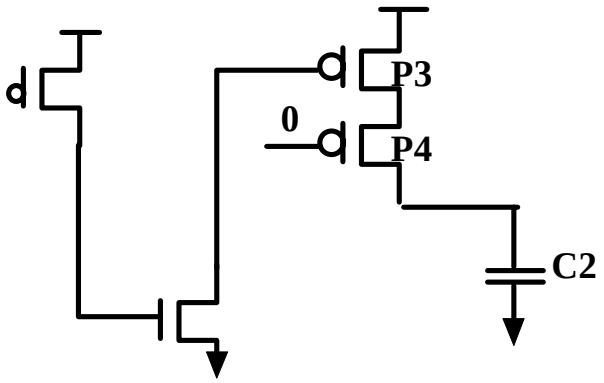
Sea el siguiente circuito:



Durante la precarga CK = 0 la salida OUT del C^2MOS esta aislada.

Vamos a suponer que CK = CK' = 0.

El nodo A se precarga a Vdd y el circuito equivalente quedaría:



La out se carga a Vdd, el valor almacenado se borra.

Este problema aparece porque el número de inversores estáticos entre el ultimo nodo dinámico y el Latch es impar, aparece un camino activo entre el nodo precargado y la salida.

REGLA:

El numero de inversores estáticos entre el ultimo bloque dinámico y el Latch C^2MOS debe ser par.

ν **REGLA C^2MOS REFORMULADA:**

El numero de inversores estáticos entre C^2MOS debería ser par en ausencia de nodos dinámicos.

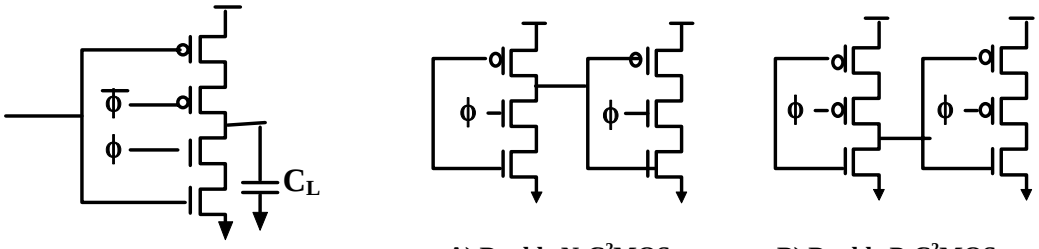
Si existen nodos dinámicos: El número de inversores estáticos entre el Latch C^2MOS y el nodo de la puerta dinámica debería ser par.

El numero de inversores estáticos entre la ultima puerta dinámica y el Latch también debería ser par.

10.8.3 TRUE SINGLE-PHASE CLOCKED LOGIC (TSPCL)

Es una Simplificación de la lógica NORA que se caracteriza por usar solo una fase de reloj, es decir se suprime el CK inverso.

Rediseño del Latch cMOS.



A) Double N-C²MOS **B) Double P-C²MOS**

a. Double n-c2MOS

b. Double P-c2MOS

Cuando $\phi = 1$.

Latch en el modo transparente de evaluación.

Dos inversores en cascada,→ no se produce inversión.

Cuando $\phi = 0$.

Ambos inversores están cortados,

no existe camino a tierra.

Comparado con el C^2MOS se eliminan las ligaduras del c2MOS.

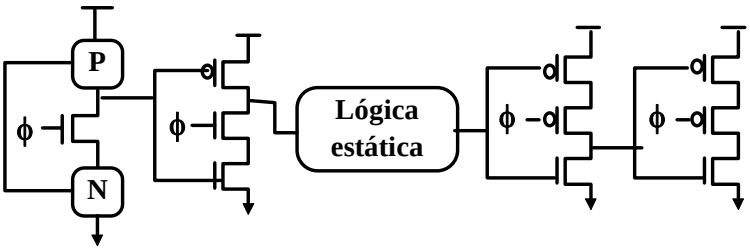
No hacen falta inversores extra entre dos Latches o entre al Latch y la lógica dinámica.

La lógica dinámica y estática se puede mezclar libremente.

Las funciones lógicas se pueden incluir dentro de los Latches:

N-C^2MOS

P-C^2MOS



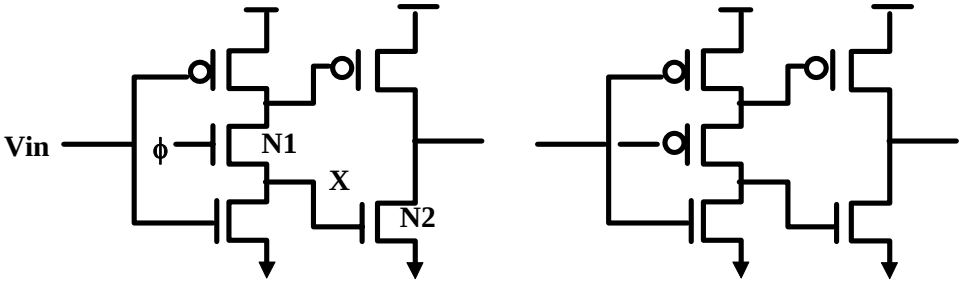
Para diseñar Pipe-Lines se debe alternar N-C^{2}MOS y P-C^{2}MOS.

Inconveniente:

Aumento del numero de transistores por Latch (6 en lugar de 4).

10.8.4 SPLIT- OUTPUT

Modificación de la lógica anterior



Solo el primer inversor esta controlado por el Ck,

ventajas:

Menor número de transistores.

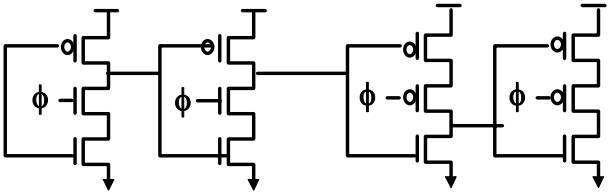
Disminuye la capacidad de carga del CK a la mitad.

desventaja:

No todos los nodos alcanzan valores de máxima dureza.

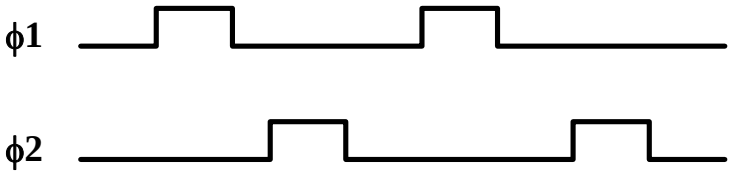
Por ejemplo. Dado que N1 conduce mal el 1 cuando Vin = 0, El nodo X toma un valor máximo Vdd-Vtn. El transistor N2 no se abre del todo lo que produce una degradación del rendimiento.

La implementaron de biestables de carga por flanco con esta lógica sería la siguiente


ϕ
ϕ
ϕ
ϕ

10.9DOS FASES DE CK

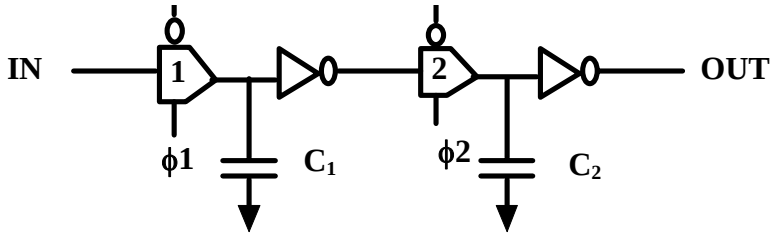
El problema mas importante que se plantea en al utilización de una fase de CK es la generación y distribución de señales en las que no se solapen CK y nCK.



$\phi1(t)\cdot\phi2(t)= 0; \forall\ t$

Es decir no se solapan las dos señales se reloj nunca.

Vamos a ver su funcionamiento para un registro de desplazamiento dinámico.



ϕ1= 1, ϕ2=0

La puerta de transmisión 1 conduce.

La puerta de transmisión 2 no conduce.

El dato de entrada se carga en la capacidad del primer Latch.

ϕ1=0, ϕ2=1

Se corta la puerta de transmisión 1.

Conduce la puerta de transmisión 2.

El dato cargado en la capacidad del Latch1 se transmite a Q.

ϕ1=1, ϕ2=1

No se da nunca. Si se da es por error.

ϕ1=0, ϕ2=0

Las capacidades de ambas Latchs permanecen con sus valores,

las puertas no conducen.

Las formas ϕ1 y ϕ2 dependen del circuito que se quiera sincronizar.

Si ϕ1 se usa como señal de precarga tiene que tener una duración que permita una precarga del nodo que sea el peor caso.

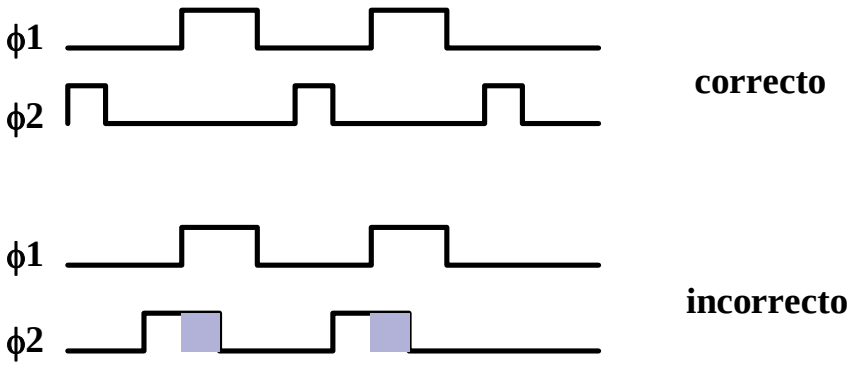
Caso típico una línea de bit de memoria RAM, por ser una línea muy larga.

El retardo entre los relojes se debe escoger para que no se solapen ni en el peor de los casos.

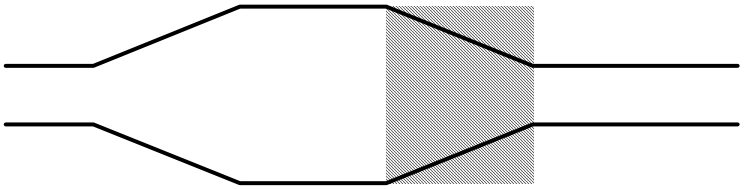
Los problemas de desviaciones en señales de CK se pueden deber a dos casos diferentes:

A. Las señales viajan a travas de diferentes retardos antes de llegar al Latch. El problema se produce cuando ambas señales están a 1, conducen las dos puertas de transmisión.

b. Los relojes no solapan pero sus tiempos de subida y de bajada son demasiado lentos.



La ϕ1 y la Fϕ tienen diferentes retardos y esto provoca solapamiento.



Al ser la subida y la bajada demasiado lentas existen zonas de las dos fases que solapan.

Los casos b y c provocan que se carguen en las capacidades del registro datos incorrectos.

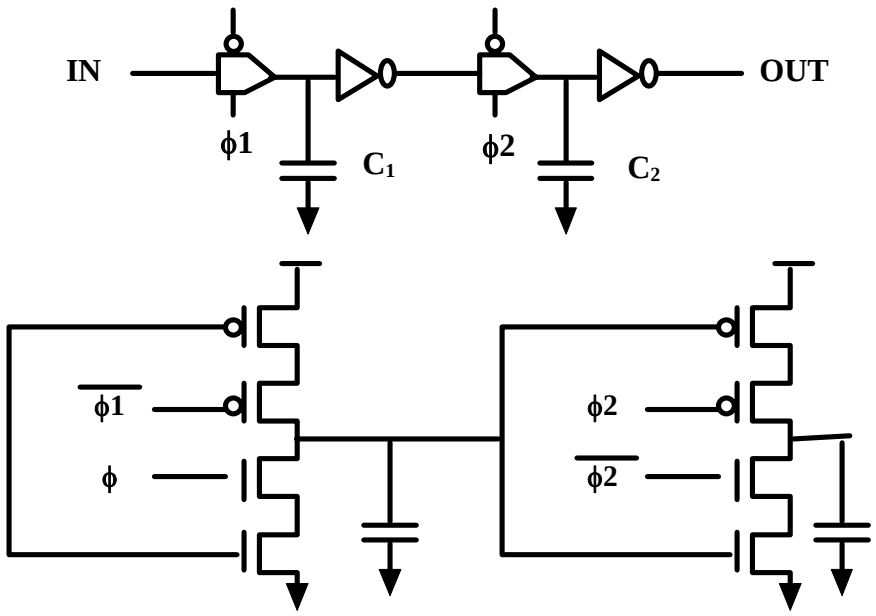
10.9.1 ESTRUCTURAS DE MEMORIA DE DOS FASES DE CK

una memoria de las fases es idéntica a una memoria de una fase:

$\phi 1$ ataca al maestro.

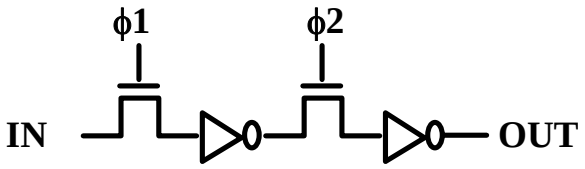
$\phi 2$ ataca al esclavo.

Ejemplos:



El principal inconveniente de estas celdas de memoria es el área que ocupan las cuatro líneas de reloj efectivas que se utilizan→ complejidad de rutado.

En algunos casos, como el ejemplo 1, esto se puede mejorar utilizando transistores de paso en lugar de puertas de paso, quedando el esquema de la siguiente manera;



Problemas:

El uno se degenera al pasar por cada transistor Vout = Vdd-Vtn.

Principal efecto relentizar las operaciones debido a que $V_G < V_{DD}$. Menores márgenes de ruido. Y aparece un consumo accidental de potencia estática debido a la conducción del transmisor P.

Si $|V_{TP}| < V_{TN}$

Condición de conducción de P:

$|V_{GS}| > |V_{TP}|$

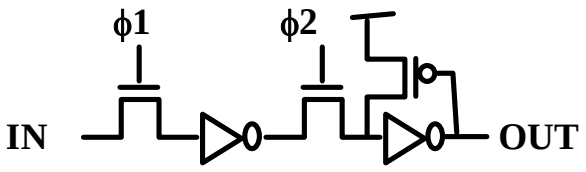
$V_{GS} = V_{DD} - V_{TN} - V_{DD}$

$|V_{TN}| > |V_{TP}|$

Cumple la condición de conducción.

Solución.

Añadir transistores realimentados para dotar al esquema de lógica totalmente restaurada



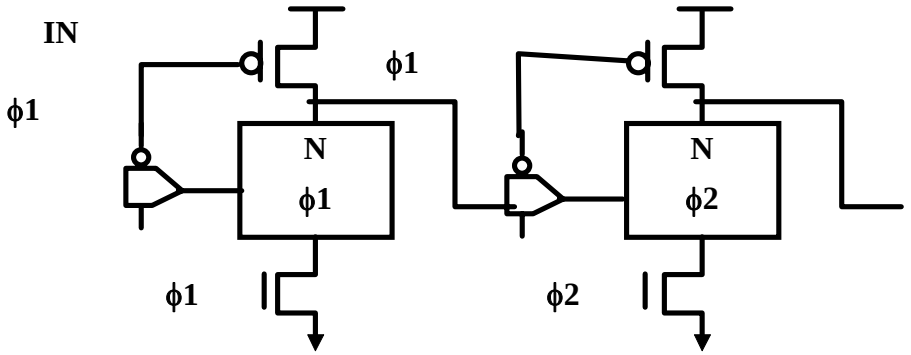
Ventaja.

Tiempos de subido mas rápidos debido a que tienen menor capacidad por la ausencia del transistor P en la puerta de paso.

10.10 ESTRUCTURAS LÓGICAS DE DOS FASES DE RELOJ

Igual que en el caso de una fase CK se pueden utilizar la lógica estática convencional junto con los elementos de memoria ya descritos.

En el caso de que se desee utilizar lógica dinámica se puede utilizar el siguiente esquema:



<u>1 Etapa:</u> Se precarga F1 = 1.

Se evalúa F2 = 1.

<u>2 Etapa:</u> Se precarga en F2

Se evalúa en F1.

En la fase de precarga la puerta de paso de entrada esta abierta. En la fase de evalúacion la puerta de paso esta cerrada, conserva en su salida el dato evalúado.

Porque no se coloca el inversor en lugar de la puerta?. Cuando se coloca el inversor no estamos sementando la lógica porque toda se evalúa al mismo tiempo, provoca problemas de falsas evaluaciones. Esto se ve mejor en el siguiente modulo.

10.11 ESTUDIO DEL RENDIMIENTO DE LOS CIRCUITOS SECUENCIALES. CLOCK SKEW (DESVIACIONES DE RELOJ)

En un circuito síncrono la señal de CK esta conectada a todos los registros, Flip-Flop y Latches, así como a los transistores de precarga y evaluación.

Este enorme fan-out actua como una carga capacitiva en la linea Deck. Este problema de fan-out se ve incrementado por la longitud del hilo que puede estar en todo el chip pudiendo tener una longitud de CM.

La carga total puede oscilar entre cientos de p.F.-nF.EJ: El micropocesador ALPHA 3.25nF. 40% de la carga capacitiva del chip.

Importante resistencia en serie aunque se rute en metal.

Como el retardo RC depende de la ????? señal de CK tienen las cargas en instantes diferentes, en función de su distancia a la fuente del ????.A esto se la llama “Clock Skew”, o desviacion del CK. Este problema puede afectar grandemente al rendimiento del sistema. Races:

trmin: minima propagacion del registro.

trmw: maxima propagacion del registro.

tlmin: minimo retardo en la lógica combinacional.

tlmx: maximo retardo en la lógica combinacional.

ti: retardo de interconexion.

t0: reloj local, tiene una desviacion ??? con el valor de referencia 0.

10.11.1 RETARDO MAXIMO DE PROPAGACION.

Tiempo que necesita la salida mas lenta en responder a los cambios en las entradas. Peor caso. Determina el periodo T.

RETARDO MINIMO DE PROPAGACION.

Tiempo que se tarda en la salida mas rapida en responder a los cambios de las entradas. Es el utilizado para estudiar las desviaciones de CK (Skew). Para todos los estudios vamos a suponer Setup = 0.

10.11.2 SINCRONIZACION DE UNA FASE CK.

Skew d=tF''-tF'.

Puede ser positiva o negativa dependiendo de la dirección de rutado y de la posición de la fuente de CK.

Suponiendo F' produce la caga en R1.

TIEMPO DE RETARDO MENO.

TF' + trmin +ti+ tlmin.

Si F'' retrasado respecto a F' puede ser que en el mismo ciclo se carguen en R2 datos correspondientes a otro ciclo, mas rápido el dato que la señal. Para que esto no ocurra se debe cumplir:

tF'' <= tF' + trmin = ti = tlmin

d <= trmin + ti + tlmin

Para fijar el periodo T se deben estudiar los mayores retardos:

tF' + trmx + ti+ tlmx

tiempo que tarda en estar la señal de datos conectada a la entrada de R2. el siguiente ciclo de reloj (la siguiente carga de F'' es de??? F''+T). Debe ser mayor que este valor para que los datos estén estables.

TF""+T >= tF' + trmx +ti+ Tlmx

$$T >= trmx + ti + tlmx - d$$

TF’’ + T instante en que llega el siguiente ciclo de reloj a R2, debe ser mayor que el tiempo mx que tarda la señal en estar estable.

d>0

El CK es rutado en la misma dirección que el flujo de datos a traves del pipe.

La desviacion debe ser controlada estrictamente y debe satisfacer la primera condición.

Si esta ligadura no se alcanza el circuito malfunciona independientemente del periodo de CK.

Una desviacion positiva aumenta el rendimiento del circuito (condición 2).

Cuidado grandes desviaciones provocan el incumplimiento de 1.

La señal de carga se desplaza en paralelo con el dato. No hace falta alargar el periodo hasta que el dato es estable, mientras el dato se estabiliza la señal de carga se va propagando.

d<0

El CK se ruta en la dirección opuesta del dato.

La condición 1 se cumple siempre puesto que los retrasos de propagacion son siempre positivos (trmin, ti, tl).

Desventaja:

Empeora el rendimiento al provocar el aumento de T(C2)

Los valores de las desviaciones en las señales de reloj estan determinadas por factores tecnologicos como los retardos de las interconexiones, registros, lógica.

Pero también influye en ellos la metodologia de diseño. En las celdas estandar se producen grandes desviaciones debido a la automatizacion del place&rout.

10.11.3 SINCRONIZACIÓN DE DOS FASES DE RELOJ.

Vamos a suponer F1 y F2 rutados en la misma dirección y tienen la misma desviacion relativa.

Suponer que se envía el dato y no se espera a que este estable para enviar la señal de reloj.

Mientras el dato esta estabilizándose la señal de reloj esta viajando.

Lo importante es que cuando llegue la señal de CK el dato ya este estable.

d<0

ocurre lo contrario.

La señal de CK tF'' es la primera en llegar por lo tanto es absolutamente obligatorio que la señal esta cuando llega tF''

Hay que retrasar la señal hasta que el dato este estable. EJ; Acequia con compuertas.

El señor encargado de abrir la siguiente compuerta no tiene que esperar a que haya llegado el agua, puede incluso acompañarla

Si la señal viaja en sentido contrario debe esperar hasta que llegue.

Sabiendo que:

TF1 y TF2 son los intervalos de tiempo que F1 y F2 permanecen en alta.

TF12 y TF21 son las separaciones entre F1 y F2.

TF12 y TF21 que se conocen como tiempos de no solapamiento deben ser positivos para que el circuito de la figura funcióne correctamente.

El nuevo dato se aplica en la lógica CL2 cuando sube F1' y F2.

El dato se carga en la capacidad de (0 en el modulo de memoria) de M2 cuando baja F1 y aísla la entrada de salida.

Debido a las desviaciones de F1' y F2 pueden solapar, transparencia, destrucción de los datos de M2 (races).

Para prevenir esto el solapamiento debe ser mas pequeño que el tiempo de propagación mas rápido:

d - TF12 <= trmin +ti +t??min

Comparando esta ecuación con 1 se puede ver que el tiempo de no solapamiento TF12 actua como un buffer que absorbe los efectos de la desviación

el problema degenra el rendimeinto (El aumento de TF12)

T+(d-TF12)>=trmx+ti+tlmin

Los efectos de la desviacion se pueden controlar aumentando el tiempo de no solapamiento.

SOLUCIÓN A LOS PROBLEMAS DE RELOJ

De lo anterior se pueden sacar las siguientes conclusiones.

Para evitar los problemas de desviacion:

1.Rutar el CK en el sentido opuesto de los datos.

2.Controlar los periodos de no solapamiento (solo valida en las segundas fases de CK.

Ambas soluciones empeoran los rendimientos de los circuitos.

además en muchos casos el flujo de datos no es unidirecciónal.

En general el reloj solo tiene desviaciones negativas en subcircuitos de tipo data-path.

A parte de las dos soluciones vistas se deben buscar algunas mas a nivel global del C.I.:

- Analisis y diseño de la red de distribucion de la señal de reloj.
- Notar que el valor absoluto de la desviacion entre el reloj local y la fuente global es indiferente.
- Lo que cuenta son las desviaciones entre las cargas de registros..

El diseñador puede controlar las desviaciones que aparecen en la red de reloj a traves de los siguientes parámetros:

- Materiela de interconexion.
- Forma de la red de distribucion.
- Drivers del CK y esquema de Buffers usado.
- Fan-Out.
- Tiempos de subida y de bajada del CK.

Una red de distribución que consigue minimizar la desviación es la red en “H”.

Es un modelo ideal y solo es útil para redes de Arrays-Regulares (todos los módulos idénticos). Todos los módulos a la misma distancia de la entrada del CK. Skew = 0.

Este modelo seria generalizable a un floorplan en el que se distribuye la señal de reloj a bloques funcionales de similar longitud.

Estos subloques a su vez deberían asegurar que la desviación en su interior es tolerable.

Cuando no son tan regulares se pueden utilizar otros recursos, que es lo que se explica a continuación.

En lo que a drivers y buffers de carga respecta existen dos aproximaciones:

A. Un buffer unico grande.

B. Un arbol de reloj.

A

Este método que se implementa con una cascada de inversores se usa para alimentar una línea de reloj global que alimenta todos los módulos.

Esta alternativa se utiliza en diseños con gran numero de modulos diversos y sin una estructura de rutado apreciable.

B. Arbol de relojes.

Los buffers intermedios aislan los relojes locales de los globales (reducir las cargas) y amplifica las señales de CK degeneradas en el camino. Produce mayor pendiente en los planos de CK. Debe tener dos características:

a. Misma profundidad a todas las ????.

b. Fan-Out equilibrado.

10.12 SINCRONIZACIÓN DE SISTEMA MEDIANTE PLL (PHASE LOCKED LOOP)

Generan relojes internos a partir de relojes externos.

Sea el sistema de la figura:

La tarjeta impresa contiene dos C.I. que se comunican intercambiando datos.

Un único reloj ataca a ambos C.I..

Estos circuitos en su interior generan las señales de CK necesarias.

En cada uno de los C.I. aparece una desviacion de las señales de C.K. respecto al generador basico, que no tienen que coincidir entre si.

Como ya se ha dicho el valor de la desviación varia de un chip a otro.

Es dificultoso intercambiar datos entre ambos C.I., puesto que las señales locales se activan en instantes de tiempo diferentes.

Solucion: PLL.

Cuando la señal de C.K. tiene que recorrer un camino largo desde que entra en el C.I. hasta qu llega a su destino, se producen importantes desviaciones en esta señal.

Estas desviaciones se producen, entre otras cosas, por los retardos RC.

Tenemos un C.I. con una señal CK externa que entra al C.I. a traves de un PAD Bufferado.

Se puede observar que realiza la sincronizacion de un registro a traves de un camino RC.

El retardo de salida del dato:

RETck = RETrc + RETbuffer + RETpad

con un dispositivo PLL queda:

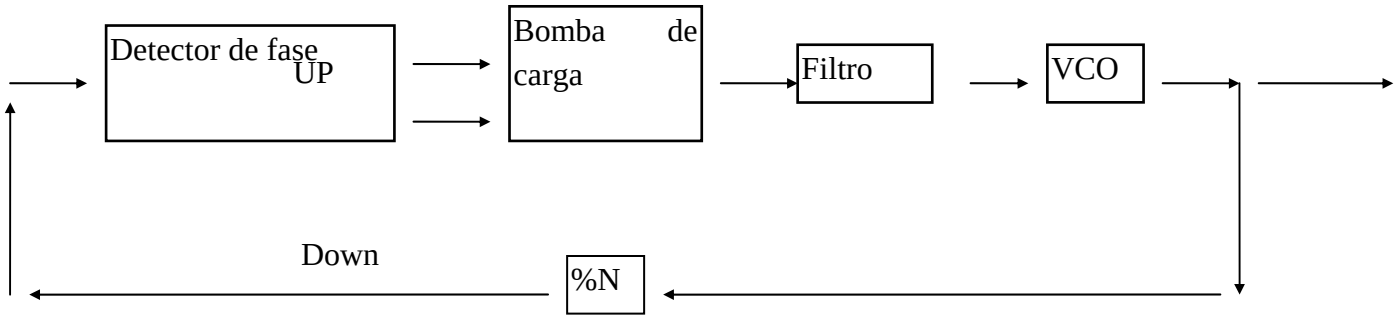
La misión del PLL es muestrear la señal CK interna y compararla con la externa generando una señal que este en fase con la externa.

De esta manera se elimina el retardo del Buffer de entrada y el de la linea RC.

El retardo del Buffer se elimina porque ya no se utiliza la señal externa directamente.

Vamos a ver como se reduce el retardo RC. Recordar que en este curso el problema es llenar la capacidad de la linea.

Esquema PLL



Detector de fase.

Detecta la diferencia entre la señal externa Ckex y la señal interna que sale de Vcu.

Bomba de carga.

El detector de fase la indica si va adelantado o retrasado respecto a la señal de referencia.

Si va retrasado suministra carga auxiliar

Si va adelantado elimina carga.

Filtro

Se utiliza para refinar los pulsos que se obtienen de la bomba.

VCO:

Voltage Controled Oscilator.

Apuntala la señal.

Genera tantas fases como sean necesarias.

Aumenta la frecuencia si es necesario.

%N.

Vuelve a dividir la frecuencia para compararla con el valor de referencia.