



# UNIVERSIDAD AUTONÓMA DEL ESTADO DE HIDALGO

---

---

## INSTITUTO DE CIENCIAS BÁSICAS E INGENIERÍA

*Área Académica de Computación y Electrónica Licenciatura en Ingeniería  
en Electrónica*

**Diseño de una ALU de 32 Bits Derivada del CI 74181 para  
Operaciones MAC en Inteligencia Artificial sobre Nodo NSFET de 3  
nm**

### **T E S I S**

***PARA OBTENER EL GRADO DE: INGENIERO EN ELECTRÓNICA***

**P R E S E N T A:**  
***ANTHONY ALEJANDRO DORANTES MEJIA***

***DIRECTOR DE TESIS:***  
***DR. JOSÉ LUIS GONZÁLEZ VIDAL***

***CODIRECTOR DE TESIS:***  
***DR. HEBERTO GÓMEZ POZOS***

**PACHUCA DE SOTO, HGO.**

**JULIO 2025**



Universidad Autónoma del Estado de Hidalgo  
Instituto de Ciencias Básicas e Ingeniería  
*School of Engineering and Basic Sciences*

Mineral de la Reforma, Hgo., a 26 de noviembre de 2025

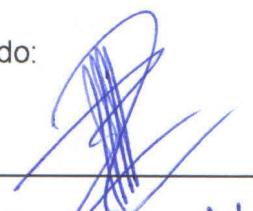
Número de control: ICBI-D/3004/2025  
Asunto: Autorización de impresión.

**MTRA. OJUKY DEL ROCÍO ISLAS MALDONADO**  
**DIRECTORA DE ADMINISTRACIÓN ESCOLAR DE LA UAEH**

Con Título Quinto, Capítulo II, Capítulo V, Artículo 51 Fracción IX del Estatuto General de nuestra Institución, por este medio, le comunico que el Jurado asignado al egresado de la Licenciatura en Ingeniería en Electrónica **Anthony Alejandro Dorantes Mejía**, quien presenta el trabajo de titulación “**Diseño de una ALU de 32 Bits Derivada del CI 74181 para Operaciones MAC en Inteligencia Artificial sobre Nodo NSFET de 3 nm**”, ha decidido, después de revisar fundamento en lo dispuesto en el Título Tercero, Capítulo I, Artículo 18 Fracción IV; dicho trabajo en la reunión de sinodales, **autorizar la impresión del mismo**, una vez realizadas las correcciones acordadas.

A continuación, firman de conformidad los integrantes del Jurado:

**Presidente:** Ing. Emmanuel Gutiérrez Rojas


**Secretario:** Dr. Heberto Gómez Pozos

**Vocal:** Dr. José Luis González Vidal

**Suplente:** Ing. Dafnis Efraín García Ángeles

Sin otro particular por el momento, reciba un cordial saludo.

Atentamente  
“Amor, Orden y Progreso”

Mtro. Gabriel Vergara Rodríguez  
Director del ICBI



GVR/YCC

Ciudad del Conocimiento, Carretera Pachuca-Tulancingo Km. 4.5 Colonia Carboneras, Mineral de la Reforma, Hidalgo, México. C.P. 42184  
Teléfono: 771 71 720 00 Ext. 40001  
direccion\_icbi@uaeh.edu.mx, vergarar@uaeh.edu.mx

“Amor, Orden y Progreso”



2025



uaeh.edu.mx

## **Agradecimientos**

*A mis padres, por brindarme la oportunidad de estudiar, por su apoyo incondicional y por motivarme a concluir esta etapa académica. Su esfuerzo y confianza han sido fundamentales para alcanzar esta meta.*

*Al Dr. José Luis González Vidal, asesor de tesis, por su orientación, tiempo y valioso acompañamiento durante el desarrollo de este trabajo. Su guía fue esencial para la correcta realización del proyecto.*

*Agradezco al Laboratorio de Microelectrónica del CEDAI-UAEH.*

*A Microwind y ni2designs, por proporcionar la licencia del software utilizado en este trabajo, lo cual hizo posible el desarrollo y validación de los resultados obtenidos.*

*A la Universidad Autónoma del Estado de Hidalgo (UAEH), por brindarme la formación académica y los recursos necesarios para consolidar mis conocimientos, así como por ser el espacio que me permitió crecer tanto profesional como personalmente.*

*A mis compañeros y amigos, en especial a Noé y Joana, por compartir esta etapa con entusiasmo, colaboración y amistad. El ambiente vivido durante la carrera hizo que el camino fuera más ameno y enriquecedor, dejando recuerdos valiosos que acompañarán siempre esta experiencia.*

*Y a una pequeña nube, por su ayuda y compañía discreta.*

*Finalmente, agradezco a todas las personas que, directa o indirectamente, hicieron posible la realización de este trabajo y la culminación de esta etapa académica.*

# Resumen

Este trabajo aborda el diseño, análisis y validación de una unidad aritmético-lógica (ALU) de 32 bits orientada a operaciones Multiply–Accumulate (MAC), desarrollada a nivel transistor bajo una arquitectura basada en tecnología NSFET en un nodo de 3 nm. La propuesta se construye a partir de bloques combinacionales y secuenciales cuidadosamente diseñados, con el propósito de ejecutar de manera eficiente aquellas operaciones que resultan esenciales en entornos de procesamiento neuronal, donde el desempeño, la precisión y el consumo energético son factores críticos.

Para sustentar el diseño, se llevó a cabo el modelado detallado de los dispositivos involucrados, así como el análisis de parámetros fundamentales como la corriente de drenador, el voltaje umbral y el comportamiento eléctrico intrínseco de los NSFET. Estos elementos permitieron describir con claridad la funcionalidad del bloque MAC, el cual integra multiplicadores, acumuladores y sumadores optimizados para maximizar el rendimiento y minimizar la latencia en aplicaciones de cálculo matricial.

Asimismo, se estudió el comportamiento estructural del sistema y la integración lógica necesaria para consolidar una unidad coherente con los requisitos de arquitecturas modernas de aceleración. Este enfoque permitió validar tanto la factibilidad del diseño como su alineación con los estándares actuales de cómputo.

En conjunto, el proyecto establece una base sólida para la construcción de módulos de aceleración más complejos orientados a redes neuronales y evidencia el potencial de los dispositivos NSFET para mejorar el desempeño y el control electrostático en tecnologías de nodos avanzados.

**Palabras clave:** Arquitecturas de cómputo de alto rendimiento, Circuito integrado, Control electrostático, Integración lógica, Multiplicadores digitales, Nanoelectrónica, NSFET, Operaciones MAC (Multiply–Accumulate), Procesamiento neuronal, Unidad aritmético-lógica (ALU).

# Abstract

This work addresses the design, analysis, and validation of a 32-bit arithmetic-logic unit (ALU) oriented toward Multiply-Accumulate (MAC) operations, developed at the transistor level under an architecture based on NSFET technology in a 3 nm node. The proposal is built from carefully designed combinational and sequential blocks, with the purpose of efficiently executing those operations that are essential in neural processing environments, where performance, accuracy, and energy consumption are critical factors. To support the design, detailed modeling of the devices involved was carried out, as well as analysis of fundamental parameters such as drain current, threshold voltage, and the intrinsic electrical behavior of NSFETs. These elements made it possible to clearly describe the functionality of the MAC block, which integrates multipliers, accumulators, and adders optimized to maximize performance and minimize latency in matrix computation applications. Likewise, the structural behavior of the system and the logical integration necessary to consolidate a unit consistent with the requirements of modern acceleration architectures were studied. This approach made it possible to validate both the feasibility of the design and its alignment with current computing standards. Overall, the project establishes a solid foundation for the construction of more complex acceleration modules geared toward neural networks and demonstrates the potential of NSFET devices to improve performance and electrostatic control in advanced node technologies.

**Keywords:** High-performance computing architectures, Integrated circuit, Electrostatic control, Logic integration, Digital multipliers, Nanoelectronics, NSFET, MAC (Multiply–Accumulate) operations, Neural processing, Arithmetic–logic unit (ALU)

# Índice general

## Contenido

<i>Resumen</i> .....	3
<i>Abstract</i> .....	4
<i>Índice general</i> .....	5
<i>Índice de figuras</i> .....	7
<i>Índice de tablas</i> .....	9
<b>1. CAPÍTULO 1 10</b>	
<b>MARCO TEÓRICO .....</b>	<b>10</b>
1.1 <i>Antecedentes</i> .....	11
1.1.1 <i>Invención del transistor bipolar</i> .....	12
1.1.2 <i>Desarrollo de los primeros MOSFET</i> .....	13
1.1.3 <i>Primeros circuitos integrados</i> .....	14
1.1.5 <i>Ley de Moore</i> .....	19
1.1.6 <i>Intel 4004, el primer microprocesador comercial</i> .....	20
1.1.7 <i>FinFET: la respuesta tecnológica al desafío de la miniaturización</i> .....	21
1.1.8 <i>NSFET: La actualidad (Por ahora)</i> .....	22
1.2 <i>Problemática</i> .....	23
1.3 <i>Fundamentos teóricos</i> .....	24
1.3.1 <i>Circuito Integrado (IC)</i> .....	24
1.3.1.1 <i>Tipos de Circuitos Integrados</i> .....	24
1.3.1.2 <i>Clasificación de los CI</i> .....	26
1.3.1.3 <i>Ventajas de los Circuitos Integrados</i> .....	27
1.3.1.4 <i>Desafíos de los Circuitos Integrados</i> .....	28
1.3.2 <i>Transistor de Efecto de Campo (FET)</i> .....	29
1.3.2.1 <i>Componentes principales de un transistor FET</i> .....	29
1.3.2.2 <i>Características de los Transistores FET</i> .....	31
1.3.3 <i>Tecnología MOSFET</i> .....	33
1.3.3.1 <i>Ventajas de los MOSFET</i> .....	35
1.3.3.2 <i>Desafíos en el diseño MOSFET</i> .....	35
1.3.4 <i>Tecnología FinFET</i> .....	37
1.3.4.1 <i>Ventajas de los FinFET</i> .....	38
1.3.4.2 <i>Desafíos en el diseño de FinFET</i> .....	38
1.3.5 <i>Tecnología NSFET</i> .....	39
1.3.4.1 <i>Ventajas de NSFET</i> .....	40
1.3.4.2 <i>Desafíos en el diseño de NSFET</i> .....	40
1.3.4.3 <i>Funcionamiento del Transistor NSFET</i> .....	41
1.3.5 <i>CI 74181 (4 bits): Unidad Aritmético Lógica (ALU)</i> .....	42
1.3.5.1 <i>Entradas y Salidas</i> .....	42
1.3.5.2 <i>Funcionamiento de la Unidad Aritmético lógica</i> .....	43
1.3.6 <i>Unidad Multiply-Accumulate (MAC)</i> .....	50
1.3.6.1 <i>Implementación de la unidad MAC</i> .....	50
1.3.6.1.1 <i>Acumulador</i> .....	51
1.3.6.1.2 <i>Multiplicadores</i> .....	51

1.3.6.1.3	<i>Sumadores</i> .....	52
1.3.6	<i>Unidades de Procesamiento Neuronal (NPU)</i> .....	53
1.3.6.2	<i>Arquitectura de la NPU</i> .....	54
1.3.6.2.1	<i>Arquitectura Von Neumann</i> .....	55
1.3.6.3	<i>Framework de la NPU</i> .....	57
1.4	<i>Objetivos</i> .....	60
2.	<i>CAPÍTULO 2</i> 61	
	<i>Metodología</i> .....	61
2.1	<i>Diagrama Esquemático</i> .....	62
2.2	<i>Cálculos</i> .....	67
2.3	<i>Parámetros de simulación</i> .....	79
3.	<i>CAPÍTULO 3</i> 90	
	<i>Layout y funcionamiento</i> .....	90
4.	<i>CAPÍTULO 4</i> 123	
	<i>Conclusiones y recomendaciones</i> .....	123
5.	<i>GLOSARIO</i> 127	
6.	<i>BIBLIOGRAFÍA</i> 131	
7.	<i>ANEXO 1 PRODUCTOS</i> 134	

# Índice de figuras

<i>Ilustración 1 Primer transistor de punto de contacto.</i> .....	13
<i>Ilustración 2 Apollo Guidance Computer (AGC).</i> .....	15
<i>Ilustración 3 CI SN510A.</i> .....	16
<i>Ilustración 4 CI SN514.</i> .....	16
<i>Ilustración 5 Microporcesador COSMAC 1802.</i> .....	18
<i>Ilustración 6 Ley de Moore 1970 – 2010 (Fuente: Wikipedia).</i> .....	19
<i>Ilustración 7 Layout de la MPU Intel 4004.</i> .....	20
<i>Ilustración 8 Evolución de la arquitectura de transistores desde Planar FET hasta GAAFET, mostrando la mejora en rendimiento y eficiencia energética (Fuente: Samsung).</i> .....	22
<i>Ilustración 9 Encapsulado de la ALU 74181.</i> .....	24
<i>Ilustración 10 Circuito integrado programable de señal mixta GreenPAK SLG46824.</i> .....	25
<i>Ilustración 11 Transistor MOSFET.</i> .....	31
<i>Ilustración 12 Estructura del MOSFET (Fuente: Microwind).</i> .....	33
<i>Ilustración 13 Curvas características del MOSFET.</i> .....	34
<i>Ilustración 14 Transistor MOSFET NMOS (Microwind).</i> .....	35
<i>Ilustración 15 Transistor FinFET NMOS de 4 Fins (Microwind).</i> .....	37
<i>Ilustración 16 Estructura del FinFET (Fuente: Microwind).</i> .....	37
<i>Ilustración 17 Estructura del NSFET (Fuente: Microwind).</i> .....	39
<i>Ilustración 18 Transistor NSFET NMOS 3 Nanosheets (Microwind).</i> .....	40
<i>Ilustración 19 Compuerta NOT NSFET Corte vertical (Microwind).</i> .....	41
<i>Ilustración 20 Diagrama esquemático de la ALU 74181 (4 BITS).</i> .....	47
<i>Ilustración 21 Chip 74181.</i> .....	48
<i>Ilustración 22 Diagrama a bloques de Unidad MAC.</i> .....	50
<i>Ilustración 23 Diagrama de Flujo de la Unidad MAC.</i> .....	52
<i>Ilustración 24 Arquitectura de la NPU de Von Neumann.</i> .....	56
<i>Ilustración 25 Marco de software típico de la NPU.</i> .....	57
<i>Ilustración 26 Esquemático de la ALU 74181 de 4bits (DSCH).</i> .....	62
<i>Ilustración 27 Esquemático de la ALU de 8 bits usando 2 bloques del 74181 (DSCH).</i> .....	63
<i>Ilustración 28 Esquemático de la ALU de 16 bits usando 2 bloques de 8 bits (DSCH).</i> .....	64
<i>Ilustración 29 Esquemático de la ALU de 32 bits usando 2 bloques de 16 bits (DSCH).</i> .....	65
<i>Ilustración 30 Esquemático teórico de la ALU de 64 bits usando dos bloques de 32 bits. (DSCH).</i> .....	66
<i>Ilustración 31 Simulación AND de 2 entradas del layout (0.65 V, 0.5 GHz), (Microwind).</i> .....	79
<i>Ilustración 32 Simulación AND de 3 entradas del layout (0.65V, 200Hz), (Microwind).</i> .....	80
<i>Ilustración 33 Simulación AND de 4 entradas del layout (0.65V, 100MHz), (Microwind).</i> .....	80
<i>Ilustración 34 Simulación AND de 5 entradas del layout (0.65V), (Microwind).</i> .....	81
<i>Ilustración 35 Simulación NAND de 2 entradas del layout (0.65V, 0.5 GHz), (Microwind).</i> .....	81
<i>Ilustración 36 Simulación NAND de 3 entradas del layout (0.65V, 200MHz), (Microwind).</i> .....	82
<i>Ilustración 37 Simulación NAND de 4 entradas del layout (0.65V, 100MHz), (Microwind).</i> .....	82
<i>Ilustración 38 Simulación NAND de 5 entradas del layout (0.65V), (Microwind).</i> .....	83
<i>Ilustración 39 Simulación OR de 2 entradas del layout (0.65V, 0.5 GHz), (Microwind).</i> .....	83
<i>Ilustración 40 Simulación NOR de 2 entradas del layout (0.65V, 0.5 GHz), (Microwind).</i> .....	84

<i>Ilustración 41 Simulación NOR de 3 entradas del layout (0.65V, 200 MHz), (Microwind).....</i>	84
<i>Ilustración 42 Simulación NOR de 4 entradas del layout (0.65V, 100 MHz), (Microwind).....</i>	85
<i>Ilustración 43 Simulación XOR de 2 entradas del layout (0.65V, 0.7 - 2 GHz), (Microwind).....</i>	85
<i>Ilustración 44 Simulación NOT del layout (0.65V, 1 GHz), (Microwind).....</i>	86
<i>Ilustración 45 Simulación ALU de 4 bits del layout (0.65V, 5.11 GHz), (Microwind).....</i>	87
<i>Ilustración 46 Simulación ALU 8 bits del layout (0.65V), (Microwind).....</i>	88
<i>Ilustración 47 Compuerta AND de 2 entradas NSFET 2D (3nm). .....</i>	91
<i>Ilustración 48 Modelo 3D (3-Nanosheet NSFET 3 nm).....</i>	92
<i>Ilustración 49 Compuerta AND de 3 entradas NSFET 2D (3nm) .....</i>	93
<i>Ilustración 50 Modelo 3D (3-Nanosheet NSFET 3 nm).....</i>	94
<i>Ilustración 51 Compuerta AND de 4 entradas NSFET 2D (3nm). .....</i>	95
<i>Ilustración 52 Modelo 3D (3-Nanosheet NSFET 3 nm).....</i>	96
<i>Ilustración 53 Compuerta AND de 5 entradas NSFET 2D (3nm). .....</i>	98
<i>Ilustración 54 Modelo 3D (3-Nanosheet NSFET 3 nm).....</i>	98
<i>Ilustración 55 Compuerta NAND de 2 entradas NSFET 2D (3nm).....</i>	99
<i>Ilustración 56 Modelo 3D (3-Nanosheet NSFET 3 nm).....</i>	100
<i>Ilustración 57 Compuerta NAND de 3 entradas NSFET 2D (3nm).....</i>	101
<i>Ilustración 58 Modelo 3D (3-Nanosheet NSFET 3 nm).....</i>	102
<i>Ilustración 59 Compuerta NAND de 4 entradas NSFET 2D (3nm).....</i>	103
<i>Ilustración 60 Modelo 3D (3-Nanosheet NSFET 3 nm).....</i>	104
<i>Ilustración 61 Compuerta NAND de 5 entradas NSFET 2D (3nm).....</i>	106
<i>Ilustración 62 Modelo 3D (3-Nanosheet NSFET 3 nm).....</i>	106
<i>Ilustración 63 Compuerta OR de 2 entradas NSFET 2D (3nm). .....</i>	107
<i>Ilustración 64 Modelo 3D (3-Nanosheet NSFET 3 nm).....</i>	107
<i>Ilustración 65 Compuerta NOR de 2 entradas NSFET 2D (3nm) .....</i>	108
<i>Ilustración 66 Modelo 3D (3-Nanosheet NSFET 3 nm).....</i>	108
<i>Ilustración 67 Compuerta NOR de 3 entradas NSFET 2D (3nm). .....</i>	109
<i>Ilustración 68 Modelo 3D (3-Nanosheet NSFET 3 nm).....</i>	109
<i>Ilustración 69 Compuerta NOR de 4 entradas NSFET 2D (3nm). .....</i>	110
<i>Ilustración 70 Modelo 3D (3-Nanosheet NSFET 3 nm).....</i>	111
<i>Ilustración 71 Compuerta XOR de 2 entradas NSFET 2D (3nm). .....</i>	112
<i>Ilustración 72 Modelo 3D (3-Nanosheet NSFET 3 nm).....</i>	112
<i>Ilustración 73 XOR Corte Transversal NSFET.....</i>	113
<i>Ilustración 74 Compuerta NOT NSFET 2D (3nm). .....</i>	114
<i>Ilustración 75 Modelo 3D (3-Nanosheet NSFET 3 nm).....</i>	114
<i>Ilustración 76 Layout de la ALU de 4 bits 74181 (NSFET 3 nm). ....</i>	115
<i>Ilustración 77 Modelo 3D parcial de la ALU 74181 de 4 bits (Microwind NSFET 3-Nanosheets 3 nm).....</i>	116
<i>Ilustración 78 Layout de la ALU de 8 bits usando dos bloques del CI 74181 (NSFET 3 nm). ...</i>	117
<i>Ilustración 79 Modelo 3D parcial de la ALU de 8 bits (Microwind NSFET 3-Nanosheets 3 nm). .....</i>	118
<i>Ilustración 80 Layout de la ALU de 16 bits usando 4 bloques del CI 74181 (NSFET 3 nm).....</i>	119
<i>Ilustración 81 Modelo 3D parcial de la ALU de 16 bits (Microwind NSFET 3-Nanosheets 3 nm). .....</i>	120
<i>Ilustración 82 Layout de la ALU de 32 bits usando 8 bloques del CI 74181 (NSFET 3 nm).....</i>	121
<i>Ilustración 83 Modelo 3D parcial de la ALU de 32 bits (Microwind NSFET 3-Nanosheets 3 nm). .....</i>	122

# Índice de tablas

<i>Tabla 1 Modos de selección de la ALU 74181.....</i>	44
<i>Tabla 2 Funciones de F de la ALU 74181. ....</i>	49
<i>Tabla 3 Correspondencia lógica de AND de 2 entradas. ....</i>	91
<i>Tabla 4 Correspondencia lógica de AND de 3 entradas. ....</i>	93
<i>Tabla 5 Correspondencia lógica de AND de 4 entradas. ....</i>	95
<i>Tabla 6 Correspondencia lógica de AND de 5 entradas. ....</i>	97
<i>Tabla 7 Correspondencia lógica de NAND de 2 entradas.....</i>	99
<i>Tabla 8 Correspondencia lógica de NAND de 3 entradas.....</i>	101
<i>Tabla 9 Correspondencia lógica de NAND de 4 entradas.....</i>	103
<i>Tabla 10 Correspondencia lógica de NAND de 5 entradas.....</i>	105
<i>Tabla 11 Correspondencia lógica de OR de 2 entradas. ....</i>	107
<i>Tabla 12 Correspondencia lógica de NOR de 2 entradas. ....</i>	108
<i>Tabla 13 Correspondencia lógica de NOR de 3 entradas. ....</i>	109
<i>Tabla 14 Correspondencia lógica de NOR de 4 entradas. ....</i>	110
<i>Tabla 15 Correspondencia lógica de XOR de 2 entradas. ....</i>	112
<i>Tabla 16 Correspondencia lógica de NOT. ....</i>	114

# **Capítulo 1**

## **MARCO TEÓRICO**

## 1.1 Antecedentes

El acelerado avance de la inteligencia artificial y el aprendizaje profundo ha impulsado el desarrollo de unidades de procesamiento especializadas que optimizan la ejecución de operaciones matemáticas fundamentales, como las operaciones de multiplicación y acumulación. Entre estos bloques, la Unidad Multiplicación-Acumulación (Multiply–Accumulate, MAC) es esencial, dado que constituye la base para el cálculo de productos punto y convoluciones en redes neuronales artificiales (Ranganath, Jay y Siva, 2016), (Chen, Krishna, Emer, & Sze, 2017; Yu et al., 2019).

Durante las últimas décadas, el diseño de unidades (MAC) para hardware de procesamiento digital, ha evolucionado desde implementaciones simples en procesadores generales hasta bloques dedicados en aceleradores de hardware, como las Unidades de Procesamiento Neuronal (NPU). Investigaciones previas han explorado diferentes arquitecturas para optimizar la eficiencia energética, el área y la velocidad de las unidades MAC, utilizando diversas tecnologías de fabricación (Teja, 2024), (Chen et al., 2017; Yu et al., 2019).

Paralelamente, el desarrollo tecnológico llevó inicialmente a la adopción de la tecnología FinFET en la fabricación de circuitos integrados a escala nanométrica. Los FinFET mejoraron el control electrostático del canal y redujeron fugas, traduciéndose en mayor eficiencia energética y desempeño (Hisamoto, 2000; Colinge, 2016; González et al., 2021; Narendra, Rajendran, & Bayoumi, 2008; Tschanz et al., 2012).

No obstante, a partir del nodo de 3 nm, la industria ha migrado hacia transistores compuerta rodeada (Gate-All-Around, GAA) basados en nano láminas (nanosheets), conocidos como NSFETs, que permiten un control aún más robusto del canal al envolver 360° la compuerta. Esta estructura tridimensional mejora la escalabilidad, reduce la variabilidad y permite ajustar el rendimiento y el consumo energético mediante la modulación del ancho de los transistores (Mukesh, 2023; Cho et al., 2020; Wu et al., 2019).

A la par, gran parte de estos avances se ha desarrollado utilizando herramientas y procesos de diseño propietarios y cerrados al público, lo que limita tanto el acceso como la reproducibilidad académica. En respuesta a ello, la comunidad de investigación ha promovido el desarrollo de flujos de diseño y tecnologías open-source, como el proceso SKY130 y herramientas como Magic, Xschem y OpenLane, que permiten diseñar, simular y validar circuitos integrados sin restricciones comerciales (Ziesler, 2020; SkyWater Technology Foundry, s.f.).

No obstante, en el presente trabajo se hace uso de herramientas con licencia académica como Microwind y Multisim, que ofrecen un entorno accesible y especializado para universidades y centros de investigación. En particular, Microwind resulta fundamental para la exploración de tecnologías avanzadas como CMOS, FinFET y NSFET, lo que abre la posibilidad de trabajar con nodos de última generación.

El diseño se centra específicamente en una ALU de 32 bits derivada del histórico CI 74181, con la novedad de implementarse sobre un nodo NSFET de 3 nm, lo cual representa un salto hacia la microelectrónica de escala nanométrica avanzada. La integración en una unidad MAC completa se proyecta como un trabajo a futuro, aprovechando esta base para su aplicación en aceleradores de inteligencia artificial.

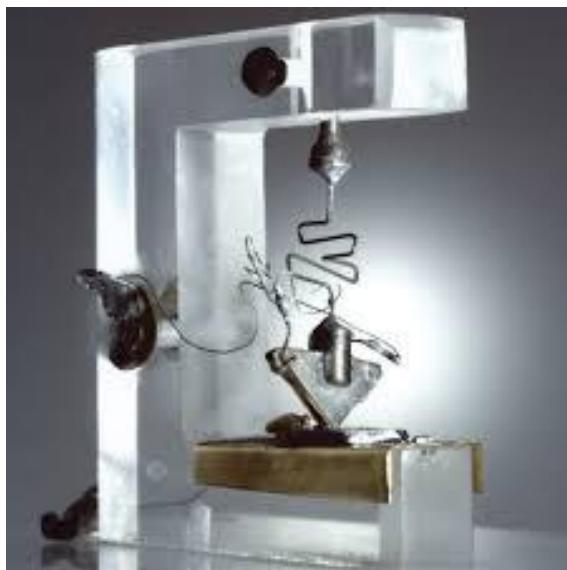
Sin embargo, para dimensionar la relevancia de esta propuesta en el contexto de la evolución tecnológica, es necesario remontarse a los orígenes de la microelectrónica. En particular, la invención del transistor bipolar en 1947 marcó el punto de partida de todo el desarrollo posterior en circuitos integrados, desde los primeros dispositivos discretos hasta las arquitecturas de nanómetros actuales.

### **1.1.1 Invención del transistor bipolar**

El transistor se logró demostrar con éxito el 23 de diciembre de 1947 en los Laboratorios Bell, ubicados en Murray Hill, Nueva Jersey. Estos laboratorios eran el centro de investigación de American Telephone and Telegraph (AT&T). Los tres inventores del transistor fueron William Shockley, John Bardeen y Walter Brattain, aunque cada uno tuvo un rol muy distinto en su desarrollo. Shockley llevaba más de diez años trabajando en la teoría del dispositivo; logró perfeccionarla, pero tras ocho años de intentos no consiguió construir un modelo funcional. Entonces, Bardeen y Brattain fueron llamados para encargarse de la parte de ingeniería y desarrollo, y lograron crear un transistor funcional en apenas dos años, lo que sorprendió a Shockley. Como supervisor, Shockley también compartió el reconocimiento. El modelo que Bardeen y Brattain desarrollaron se conoció como el transistor de “punto de contacto”. Más adelante, Shockley diseñó un transistor “bipolar”, más avanzado que el de punto de contacto, y lo reemplazó. Por esto, se puede decir que el transistor fue, en gran medida, fruto del trabajo de Shockley, aunque la ingeniería inicial fue obra de Bardeen y Brattain.

El transistor bipolar eliminó los delicados y problemáticos contactos puntuales, haciendo el dispositivo mucho más confiable. Más adelante, los transistores comenzaron a fabricarse con silicio, un material abundante y resistente a la corrosión gracias a una fina capa de dióxido de silicio que lo protege.

En 1954, Texas Instruments, en Dallas, Texas, inició la producción comercial de transistores de unión para radios portátiles. Poco después, la compañía japonesa Sony adquirió los derechos para producir transistores y rápidamente se convirtió en líder del mercado. Durante la década de 1960, Sony empezó a fabricar televisores utilizando transistores en lugar de tubos de vacío, lo que marcó el inicio del fin de la era de los tubos de vacío (Watkins, 2022).



*Ilustración 1 Primer transistor de punto de contacto.*

### 1.1.2 Desarrollo de los primeros MOSFET

En 1959, John Atalla y Dawon Kahng, de los Laboratorios Bell, crearon el primer transistor de efecto de campo (FET) con puerta aislada, superando un problema conocido como “estados superficiales”, que impedían que los campos eléctricos penetraran correctamente en el semiconductor (Hofstein & Heiman, 1963).

Al investigar capas de dióxido de silicio generadas por calentamiento, descubrieron que estos estados podían minimizarse significativamente en la interfaz entre el silicio y su óxido. Así nació una estructura en planar o de “sándwich” compuesta por metal (M - puerta), óxido (O - aislamiento) y silicio (S - semiconductor), conocida como MOSFET, o simplemente MOS (Hofstein & Heiman, 1963).

Como el dispositivo inicial era lento y no resolvía ninguna necesidad urgente en el sistema telefónico, su desarrollo no continuó inmediatamente. No obstante, en un memorando de 1961, Kahng resaltó su gran potencial, mencionando su “facilidad de fabricación y la

posibilidad de usarlo en circuitos integrados". Otros investigadores, en Fairchild y RCA, si vieron su potencial (Hofstein & Heiman, 1963), (Dawon, Electric Field Controlled Semiconductor Device, 1960).

En 1960, Karl Zaininger y Charles Meuller fabricaron un transistor MOS en RCA, mientras que CT Sah, de Fairchild, construyó un tetrodo controlado por MOS. Siguiendo este camino, en 1962 Fred Heiman y Steven Hofstein desarrollaron un dispositivo integrado experimental que contenía 16 transistores en RCA (Hofstein & Heiman, 1963).

La región conductora de un transistor MOS puede ser de tipo p (dispositivo de "canal p") o de tipo n (dispositivo de "canal n"). Los transistores de canal n son más rápidos que los de canal p, aunque su fabricación resulta más complicada (Dawon, Silicon-silicon dioxide field induced surface devices, 1961), (Sah, 1961).

Los primeros dispositivos MOS llegaron al mercado en 1964. General Microelectronics (GME 1004) y Fairchild (FI 100) lanzaron transistores de "canal p", pensados para aplicaciones lógicas y de conmutación, mientras que RCA presentó un transistor de "canal n" (3N98), diseñado para la amplificación de señales. Gracias a su menor tamaño y bajo consumo de energía en comparación con los transistores bipolares, hoy en día más del 99 % de los microchips utilizan transistores MOS (Dawon, Electric Field Controlled Semiconductor Device, 1960), (Dawon, Silicon-silicon dioxide field induced surface devices, 1961), (Sah, 1961), (Hofstein & Heiman, 1963).

### 1.1.3 Primeros circuitos integrados

Los primeros circuitos integrados eran bastante limitados: funcionaban con relativa lentitud, sustituían solo unos cuantos componentes y su costo era mucho más alto que el de los transistores discretos. Por ello, en sus inicios solo resultaban atractivos para aplicaciones aeroespaciales y militares, donde su bajo consumo de energía y su reducido tamaño justificaban estas desventajas (Bush, 1964), (Hall, 1965).

En 1961, Harvey Cragon, colega de Jack Kilby, construyó una "Computadora Electrónica Molecular" de demostración para la Fuerza Aérea de Estados Unidos. Con este proyecto mostró que 587 circuitos integrados de Texas Instruments podían reemplazar a 8500 transistores y otros componentes que cumplían la misma función en un diseño tradicional (Bush, 1964), (Hall, 1965).

A partir de 1961 comenzaron a diseñarse los dispositivos *Fairchild Micrologic* (un hito de 1960), que se usaron en computadoras como la *AC Spark Plug MAGIC* y la *Martin*

*MARTAC 420*. Sin embargo, el proyecto más relevante en esta etapa fue la *Computadora de Guía Apolo* (AGC) de la NASA. Diseñada por el MIT en 1962 y fabricada por Raytheon, cada unidad de la AGC integraba alrededor de 4000 circuitos “Tipo G” (puertas NOR de 3 entradas). Con un requerimiento de 200 000 circuitos en total, y un costo de entre 20 y 30 USD por cada uno, la AGC se convirtió en el mayor consumidor de circuitos integrados hasta 1965 (Bush, 1964), (Hall, 1965).

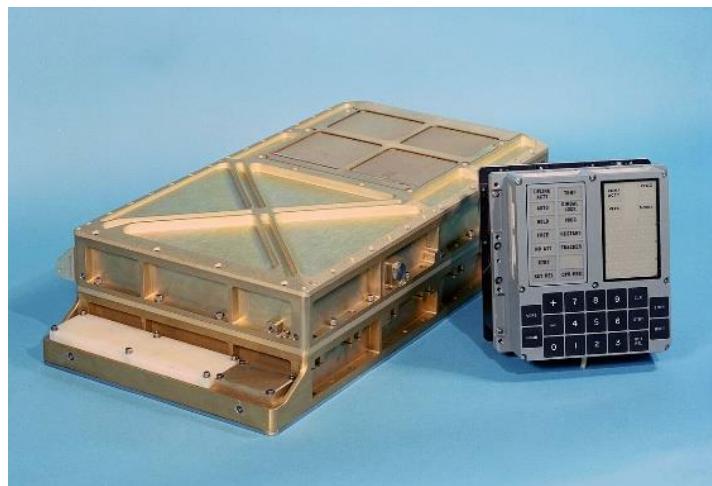


Ilustración 2 *Apollo Guidance Computer* (AGC).

El ingeniero Bob Cook diseñó la Serie 51 DCTL, la primera familia de circuitos integrados planares de Texas Instruments, con el objetivo de cumplir las especificaciones de bajo consumo exigidas por la Computadora de Aspecto Óptico de la Sonda de Monitoreo Interplanetario (*IMP*) de la NASA. Gracias a circuitos como los SN510A y SN514, utilizados como contadores binarios, biestables y circuitos inhibidores, el satélite *IMP* se convirtió en 1963 en el primero en llevar circuitos integrados al espacio (Bush, 1964), (Hall, 1965).



Ilustración 3 CI SN510A.

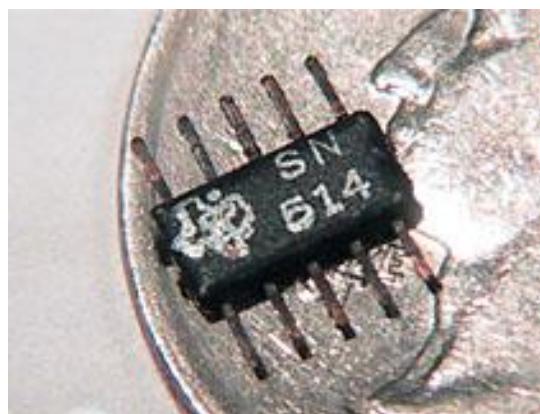


Ilustración 4 CI SN514.

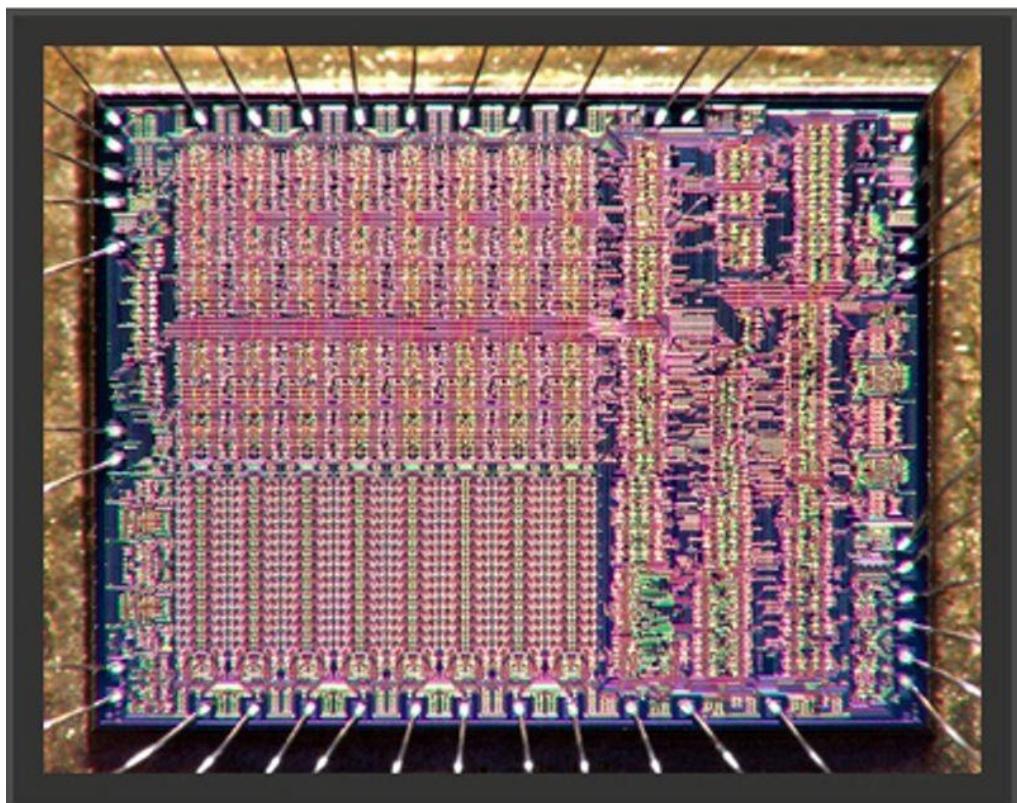
Un año antes, en 1962, *Texas Instruments* había conseguido un contrato con la *División Autonética de North American Aviation* para desarrollar 22 circuitos personalizados destinados al sistema de guiado del misil *Minuteman II*. Empresas como *Clevite* y *Westinghouse* también trabajaron en circuitos para este proyecto, que hacia 1965 superó al programa Apolo de la NASA como el mayor consumidor individual de circuitos integrados (*Texas Instruments*, 1961), (*Faulkner & Gurzi*, 1962), (*Martin-Marietta Company*, 1962) (*Bush*, 1964), (*Hall*, 1965).

## 1.1.4 Configuración complementaria MOS

En 1963, C. T. Sah y Frank Wanlass, del laboratorio de investigación y desarrollo de Fairchild, presentaron un circuito lógico que combinaba transistores MOS de “canal p” y de “canal n” en una configuración complementaria simétrica. Este diseño lograba un consumo de energía cercano a cero en modo de espera. Wanlass patentó esta idea, que más tarde se conocería como tecnología CMOS (Chih Tang, 1988).

Posteriormente, los Laboratorios de Investigación de RCA y su planta en Somerville fueron pioneros en la producción de circuitos integrados basados en CMOS, comercializados bajo el nombre COS/MOS, inicialmente para aplicaciones aeroespaciales y, más tarde, en el ámbito comercial. En 1965, Gerald Herzog lideró el desarrollo de circuitos lógicos y de memoria CMOS para un proyecto de la Fuerza Aérea (Chih Tang, 1988).

Tres años después, RCA presentó una memoria estática de 288 bits y lanzó la reconocida familia de dispositivos lógicos CD4000. En 1975, la compañía introdujo el microprocesador COSMAC 1802, fabricado con un proceso CMOS de compuerta de silicio y geometría cerrada que reducía las fugas eléctricas. Este procesador sería el precursor de millones de controladores utilizados en automóviles de Chrysler (Chih Tang, 1988).



*Ilustración 5 Microprocesador COSMAC I802.*

Las primeras aplicaciones de gran volumen de la tecnología CMOS aparecieron en productos electrónicos portátiles alimentados por batería, como relojes digitales y equipos de medición, donde el bajo consumo era más importante que la velocidad. Hacia 1978, gracias a la reducción del tamaño de las estructuras y al uso de la compuerta de silicio, la tecnología CMOS alcanzó un rendimiento comparable al de los circuitos bipolares y MOS convencionales. A medida que la miniaturización permitió integrar cientos de miles de transistores en un solo chip, CMOS se consolidó como la mejor alternativa para controlar la densidad de potencia en los sistemas integrados (Chih Tang, 1988), (Knox Basset, 2002).

## 1.1.5 Ley de Moore

En 1965, Gordon Moore, entonces director de investigación y desarrollo en Fairchild Semiconductor, elaboró un análisis interno en el que representó la cantidad de componentes por circuito integrado de menor costo entre 1959 y 1964. Al proyectar la tendencia observada, predijo que para 1975 los circuitos integrarían aproximadamente 65,000 componentes por chip, lo que implicaba una duplicación anual en la densidad de integración (Moore, 1964).

Esta predicción se convirtió en una profecía autocumplida, consolidándose como uno de los principios de la industria de los semiconductores. A partir de entonces, los avances tecnológicos se orientaron a cumplir con lo que se conocería como la Ley de Moore, desafiando continuamente a los ingenieros a lograr innovaciones anuales que mantuvieran ese ritmo de crecimiento.

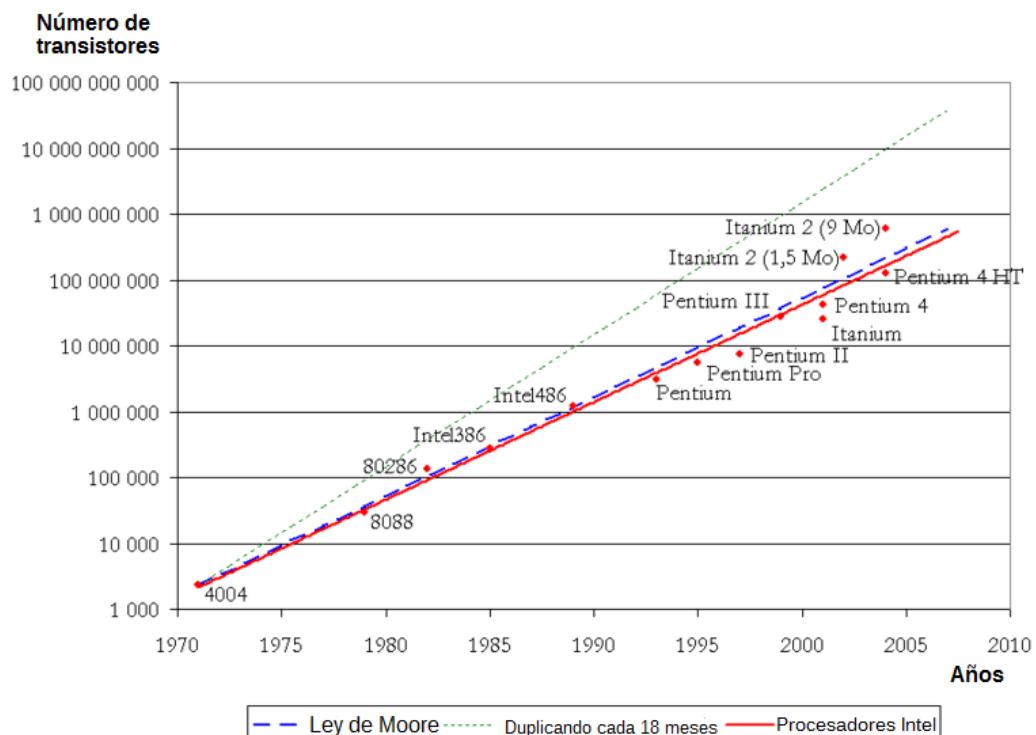


Ilustración 6 Ley de Moore 1970 – 2010 (Fuente: Wikipedia).

## 1.1.6 Intel 4004, el primer microprocesador comercial

En 1971, Ted Hoff y Stanley Mazor concibieron el primer procesador integrado de Intel, el 4004, un dispositivo de 4 bits desarrollado como parte del conjunto MCS-4 Micro Computer Chip Set. Ese mismo año, Federico Faggin, con la colaboración de Masatoshi Shima, aplicó su experiencia en la tecnología MOS de compuerta de silicio para integrar alrededor de 2,300 transistores en un encapsulado económico de 16 pines.

Faggin también supervisó el diseño del Intel 8008, un procesador de 8 bits desarrollado por Hal Feeney y presentado en 1972 (Augarten, 1983).

El 8008 había sido diseñado originalmente para la empresa CTC (más tarde Datapoint), y aunque Texas Instruments llegó a fabricar prototipos bajo la denominación TMX1795, estos nunca se comercializaron. Estos dispositivos, conocidos como microprocesadores o MPU (Microprocessor Unit), requerían circuitos de apoyo y periféricos adicionales para formar un sistema completo. En contraste, las soluciones de un solo chip diseñadas para tareas específicas dieron origen a los microcontroladores, marcando un hito tecnológico en 1974 (Augarten, 1983).

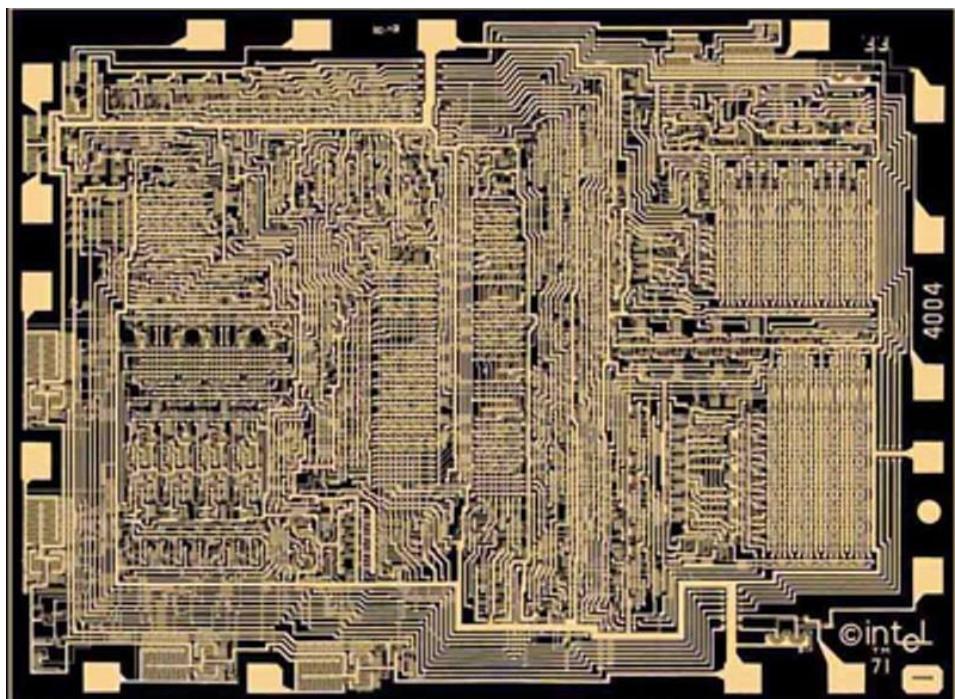


Ilustración 7 Layout de la MPU Intel 4004.

## **1.1.7 FinFET: la respuesta tecnológica al desafío de la miniaturización**

El origen de los transistores FinFET se remonta a 1987, cuando Hieda y sus colaboradores presentaron una arquitectura pionera que mostraba un mejor desempeño de conmutación y una menor sensibilidad al efecto de polarización, en comparación con los MOSFET convencionales. Posteriormente, en 1989, Hisamoto y su equipo introdujeron el primer transistor de doble compuerta, denominado DELTA (Fully Depleted Lean-channel Transistor). Este avance despertó gran interés debido a su reducción significativa del fenómeno Short Channel Effect (SCE), lo que representó una mejora notable frente a los transistores planos tradicionales (Karimi, Fardoost, & Javanmard, 2024).

Estos desarrollos sentaron las bases conceptuales para la tecnología FinFET, que puede considerarse como la evolución natural de los transistores multicompuerta hacia estructuras tridimensionales. A finales de la década de 1990, investigadores de la Universidad de California en Berkeley propusieron un diseño que reemplazaba la arquitectura planar del MOSFET por una configuración tridimensional con aletas (Fin). Este enfoque buscaba maximizar el control del canal, reducir las corrientes de fuga y superar las limitaciones que presentaban los transistores planos al reducir su tamaño (Karimi, Fardoost, & Javanmard, 2024).

La estructura FinFET utiliza una región de canal con forma de aleta, que permite un control electrostático más efectivo y una mayor eficiencia energética, al tiempo que mantiene un alto rendimiento incluso a escalas nanométricas.

Durante los primeros años de la década de 2000, empresas líderes como Intel, IBM y TSMC iniciaron programas intensivos de investigación y desarrollo para perfeccionar y comercializar la tecnología FinFET, consolidando así su lugar como la evolución avanzada de los transistores y marcando el inicio de una nueva era en el diseño de semiconductores. Sin embargo, transcurrió casi una década antes de que los chips basados en tecnología FinFET comenzaran a producirse en masa. Los primeros en hacerlo fueron los de Intel, en el año 2011. Los transistores FinFET comenzaron a incorporarse cuando la industria de los semiconductores avanzó hacia nodos tecnológicos con geometrías inferiores a los 25 nm (Perry, 2020).

## 1.1.8 NSFET: La actualidad (Por ahora)

Los NSFET (Nanosheet FET o Gate-All-Around FET) representan la evolución más reciente en la tecnología de transistores de efecto de campo, desarrollados como sucesores de los FinFET para mantener el ritmo de miniaturización por debajo de los 5 nm.

Su origen se remonta a investigaciones de la década de 2010, cuando los límites físicos del FinFET comenzaron a hacerse evidentes debido a fugas de corriente y pérdida de control del canal. En 2017, IBM, Samsung e Intel anunciaron de forma independiente sus prototipos de transistores tipo GAA (Gate-All-Around) basados en nanosheets apiladas horizontalmente. Este diseño permite envolver completamente el canal con la compuerta, mejorando el control electrostático, reduciendo la corriente de fuga y permitiendo ajustar dinámicamente el ancho del canal para optimizar rendimiento o consumo energético (IBM, 2021).

En 2022, Samsung Foundry se convirtió en el primer fabricante en producir chips comerciales con esta tecnología en su nodo 3 nm GAA, marcando el inicio de la era industrial de los NSFET. Hoy, tanto Intel (20A) como TSMC (2 nm) planean utilizar variaciones del mismo principio para sus futuras generaciones (Samsung, 2022), (Intel, 2021).

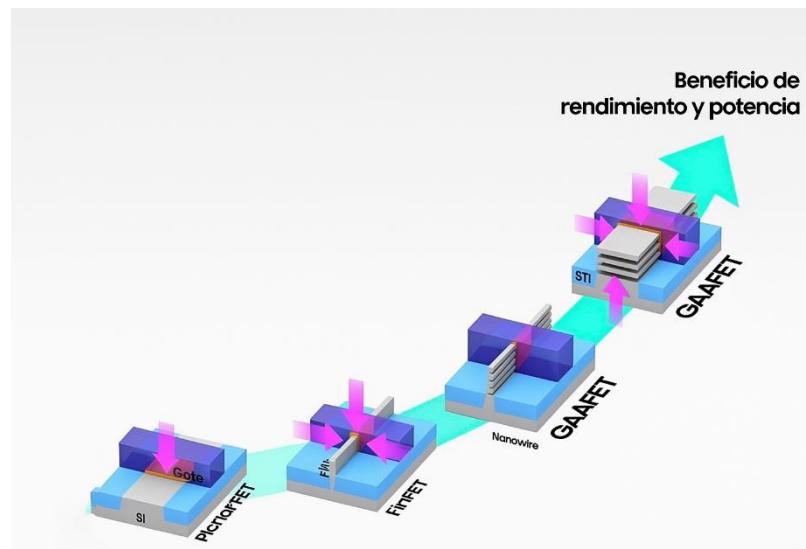


Ilustración 8 Evolución de la arquitectura de transistores desde Planar FET hasta GAAFET, mostrando la mejora en rendimiento y eficiencia energética (Fuente: Samsung).

## 1.2 Problemática

En los últimos años, el escalado de los tradicionales transistores de efecto de campo de metal-óxido-semiconductor (MOSFET) planos ha enfrentado serias limitaciones, como el efecto de canal corto, fugas de corriente y variabilidad en los procesos de fabricación. Además, la continua miniaturización y la demanda de dispositivos cada vez más compactos requieren mayores velocidades de conmutación y reducción en el consumo de potencia, lo que a su vez demanda una densidad de transistores por unidad de área cada vez mayor. Estas necesidades han impulsado el desarrollo de los FinFETS hasta ahora los NSFET (Nanowire o Nanosheet FETs), que ofrecen un control más preciso del canal, reducen las fugas de corriente y permiten seguir escalando hacia nodos de 3 nm y menores (Calderón, González y Raygoza, 2025).

En este contexto, el presente trabajo se centra en el diseño de una ALU de 32 bits derivada del CI 74181, implementada sobre tecnología NSFET de 3 nm. La elección de este nodo permite aprovechar las ventajas mencionadas: mayor densidad de transistores, menor consumo energético y mejor control eléctrico, lo que constituye una base sólida para futuros desarrollos, como la integración en unidades MAC completas orientadas a inteligencia artificial. De esta forma, el diseño no solo responde a las limitaciones de la miniaturización de MOSFETs tradicionales, sino que también contribuye al avance de arquitecturas digitales de alto rendimiento y eficiencia energética (Calderón, González y Raygoza, 2025).

# 1.3 Fundamentos teóricos

## 1.3.1 Circuito Integrado (IC)

Un circuito integrado (CI) es una estructura electrónica en la que múltiples componentes como transistores, diodos, resistencias y capacitores están fabricados sobre una única pieza de material semiconductor, generalmente silicio, mediante procesos de fotolitografía y depósito y grabado de los diferentes compuestos químicos. Estos componentes están interconectados eléctricamente para realizar una función electrónica específica. Combina múltiples componentes electrónicos, como transistores, resistencias y condensadores, en un único chip semiconductor. Sirve como bloque de construcción de los sistemas electrónicos modernos, proporcionando funcionalidad y capacidad de procesamiento en un formato compacto y eficiente (Lenovo, 2024).

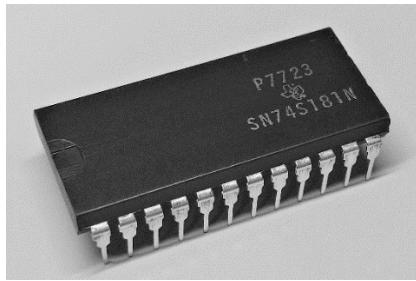


Ilustración 9 Encapsulado de la ALU 74181.

### 1.3.1.1 Tipos de Circuitos Integrados

Los circuitos integrados se pueden clasificar en diferentes tipos según sus funciones y diseños.

- **Circuito Integrados Digitales**

Los circuitos integrados digitales operan mediante señales binarias, las cuales presentan únicamente dos estados: alto (1 lógico) y bajo (0 lógico). Este tipo de dispositivos constituye el núcleo fundamental de los sistemas informáticos y de comunicación modernos. Son ampliamente utilizados en microprocesadores, microcontroladores, puertas lógicas y memorias digitales, como la RAM y la ROM (Zhang, 2025).

- ***Circuitos Integrados Analógicos***

Los circuitos integrados analógicos se utilizan para manejar señales continuas. La salida cambia con la variación de la señal de entrada. Entre sus aplicaciones más comunes se incluyen amplificadores de audio, amplificadores de RF, sensores de temperatura y gestión de energía.

En la electrónica de circuitos integrados analógicos, los componentes de circuitos integrados se utilizan para procesar diversas señales del mundo real, como sonido, luz y temperatura (Zhang, 2025).

- ***Circuitos integrados de señal mixta***

Los circuitos integrados de señal mixta combinan en un mismo chip componentes analógicos y digitales, permitiendo así el procesamiento simultáneo de ambos tipos de señales. Este tipo de integración resulta fundamental en sistemas donde las señales del entorno real deben ser convertidas, interpretadas o transmitidas de forma eficiente.

Entre sus aplicaciones más comunes se encuentran los convertidores analógico-digitales (ADC), los convertidores digital-analógicos (DAC), los generadores de reloj y las interfaces de sensores. Gracias a su versatilidad, estos microchips son ampliamente utilizados en dispositivos de audio, convertidores de datos y sistemas de comunicación inalámbrica, donde es necesario manejar de manera conjunta señales continuas y discretas para garantizar un funcionamiento preciso y estable (Zhang, 2025).

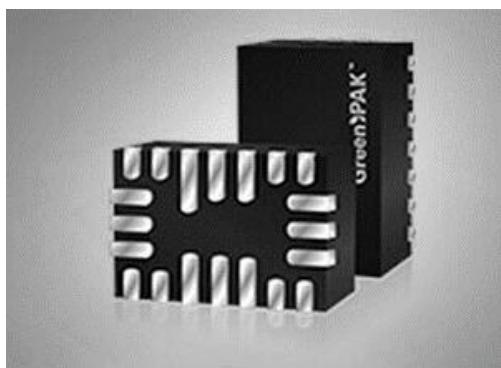


Ilustración 10 Circuito integrado programable de señal mixta GreenPAK SLG46824.

### 1.3.1.2 Clasificación de los CI

Los CIs pueden clasificarse según el número de componentes integrados:

- SSI (Small Scale Integration): contiene pocos componentes.
- MSI (Medium Scale Integration): contiene cientos de componentes.
- LSI (Large Scale Integration): contiene miles de componentes.
- VLSI (Very Large Scale Integration): contiene millones de componentes.

Los circuitos integrados (CI) se clasifican en tres categorías principales según las técnicas empleadas en su fabricación: circuitos integrados de película delgada y gruesa, circuitos integrados monolíticos y circuitos integrados híbridos o multichip (ElProCus, 2013).

- ***Circuitos integrados de película delgada y gruesa***

En este tipo de dispositivos se emplean principalmente componentes pasivos, como resistencias y condensadores, mientras que los componentes activos —como transistores y diodos— se conectan de forma independiente. Estos circuitos representan una combinación entre elementos integrados y discretos, diferenciándose principalmente por el método de deposición de la película utilizada.

Su fabricación consiste en la deposición de capas conductoras sobre una superficie de vidrio o sobre un sustrato cerámico. Al modificar el grosor de dichas capas, se obtiene una resistividad variable que permite formar distintos componentes pasivos. En algunos casos, se utiliza la técnica de serigrafía para modelar el patrón del circuito sobre el sustrato, razón por la cual también se les conoce como circuitos integrados de película delgada impresa (ElProCus, 2013).

- ***Circuitos integrados monolíticos***

Los circuitos integrados monolíticos permiten integrar en un único chip de silicio los componentes activos, pasivos y las interconexiones necesarias. El término “monolítico” proviene del griego mono (uno) y lithos (piedra), haciendo referencia a su construcción en un solo bloque.

Este tipo de circuitos es el más utilizado en la actualidad debido a su bajo costo, fiabilidad y facilidad de producción a gran escala. Son empleados en una amplia variedad de

aplicaciones, como reguladores de voltaje, amplificadores operacionales, sistemas computacionales y receptores de radio AM (ElProCus, 2013).

- ***Circuitos integrados híbridos o multichip***

Los circuitos integrados híbridos o multichip agrupan en un solo encapsulado varios chips interconectados. En ellos coexisten componentes activos —como transistores o diodos— junto con elementos pasivos encapsulados, tales como resistencias o condensadores. Las interconexiones se realizan mediante prototipos metalizados que permiten la integración de múltiples módulos funcionales en un solo dispositivo.

Estos circuitos se emplean con frecuencia en aplicaciones de alta potencia, especialmente en etapas de amplificación que van desde 5 W hasta 50 W. En comparación con los circuitos monolíticos, los híbridos ofrecen un rendimiento superior y una mayor capacidad de manejo de energía (ElProCus, 2013).

### **1.3.1.3 Ventajas de los Circuitos Integrados**

Los circuitos integrados ofrecen múltiples beneficios en comparación con los sistemas basados en componentes discretos:

- Bajo consumo de energía: Su pequeño tamaño y alta integración reducen las pérdidas eléctricas.
- Tamaño compacto: Permiten implementar circuitos complejos en un espacio reducido.
- Menor costo: La producción en masa y el bajo uso de materiales disminuyen los costos de fabricación.
- Menor peso: Su estructura miniaturizada reduce significativamente el peso del sistema.
- Alta velocidad de operación: Su bajo retardo y mínima capacitancia parasitaria permiten commutaciones más rápidas.
- Alta fiabilidad: La baja densidad de conexiones externas reduce las probabilidades de fallo (ProEx, 2023).

Los circuitos integrados permiten fabricar miles de componentes en un solo chip, optimizando tanto la velocidad de operación como el costo y el tamaño del sistema.

### **1.3.1.4 Desafíos de los Circuitos Integrados**

Pese a sus numerosas ventajas, los circuitos integrados también presentan ciertas limitaciones técnicas:

- La disipación de calor es reducida, lo que puede ocasionar sobrecalentamiento ante corrientes elevadas.
- No es posible integrar transformadores o inductores dentro del chip.
- Su capacidad de manejo de potencia es limitada, generalmente inferior a 10 W.
- No se pueden implementar fácilmente configuraciones PNP de alta calidad.
- Poseen un rango de voltaje restringido y una limitada inmunidad al ruido (ProEx, 2023).

A pesar de estos desafíos, los circuitos integrados continúan siendo la base tecnológica de la microelectrónica moderna, y su evolución ha permitido el desarrollo de sistemas cada vez más potentes, compactos y eficientes.

## 1.3.2 Transistor de Efecto de Campo (FET)

Un transistor es un componente electrónico fundamental dentro de la electrónica moderna, cuya función principal consiste en amplificar o controlar el paso de señales eléctricas. Este dispositivo puede actuar como un interruptor o como un amplificador, regulando el flujo de corriente entre sus terminales en respuesta a los voltajes aplicados (Lenovo, 2023).

Gracias a su versatilidad y eficiencia, el transistor se ha convertido en la base de prácticamente todos los circuitos electrónicos actuales, desde los más simples hasta los sistemas integrados más complejos (Lenovo, 2023).

El término “efecto de campo” describe el principio fundamental de funcionamiento de estos dispositivos. En un transistor de efecto de campo (FET), un campo eléctrico controla el flujo de corriente a través de un canal semiconductor. El proceso se inicia cuando se aplica un voltaje al terminal de puerta, generando un campo eléctrico que actúa sobre una capa aislante y forma una región de agotamiento dentro del canal. Esta región modifica la cantidad de portadores de carga libres disponibles, ajustando así la conductividad del canal (Johnson, 2025).

Los FET destacan por su alta eficiencia, ya que permiten controlar la corriente de manera precisa sin requerir una corriente de entrada significativa.

### 1.3.2.1 Componentes principales de un transistor FET

El transistor de efecto de campo (FET) se compone de tres terminales principales: la fuente (Source), el drenador (Drain) y la compuerta (Gate). Estos elementos interactúan para controlar el flujo de corriente a través del canal mediante la aplicación de un campo eléctrico. En ciertos diseños, se incorpora un cuarto terminal denominado sustrato o bulk, el cual corresponde al material semiconductor base sobre el que se forma el canal de conducción. Este terminal adicional permite un control más preciso de las características eléctricas del dispositivo y del comportamiento del canal (Riordan, 2025).

- **Fuente (Source):**

Es el terminal por el cual ingresan los portadores de carga al canal del transistor. En un FET de “tipo n”, los portadores son electrones; mientras que en uno de “tipo p”, son huecos. La fuente actúa como el punto de referencia para el flujo de corriente dentro del dispositivo (Riordan, 2025).

- **Drenador (Drain):**

Es el terminal por donde salen los portadores de carga del canal hacia el circuito externo. La corriente que fluye entre la fuente y el drenador es la variable principal controlada por la tensión aplicada en la compuerta (Riordan, 2025).

- **Compuerta (Gate):**

Es el terminal de control del dispositivo. Se encuentra separado del canal por una fina capa de material aislante (normalmente óxido de silicio en el MOSFET), lo que permite que la puerta controle el flujo de corriente sin contacto directo. La tensión aplicada entre la puerta y la fuente genera un campo eléctrico que modifica la conductividad del canal, permitiendo o evitando el paso de corriente entre fuente y drenador (Riordan, 2025).

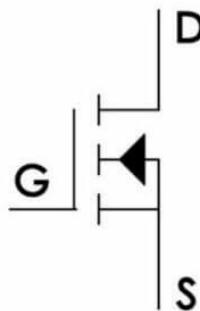
- **Canal de conducción (Channel):**

Es la región del semiconductor situada entre la fuente y el drenador, donde fluye la corriente controlada por la puerta. El canal puede ser de “tipo n” o “tipo p”, dependiendo del tipo de portadores mayoritarios. La formación o modulación del canal se produce al aplicar una tensión la cual altera la concentración de portadores en la superficie del semiconductor (Riordan, 2025).

- **Sustrato o cuerpo (Substrate / bulk):**

Es el material semiconductor base sobre el cual se fabrican las regiones del canal, fuente y drenador. En algunos diseños, el sustrato se conecta a la fuente para mantener un potencial común y evitar variaciones no deseadas del canal. Sin embargo, en estructuras más avanzadas (como FinFET o NSFET), el cuerpo puede estar completamente aislado o estructurado tridimensionalmente (Riordan, 2025).

MOSFET Canal N



MOSFET Canal P

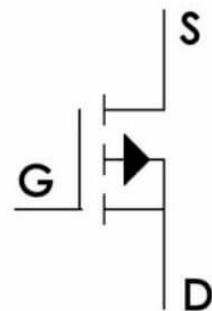


Ilustración 11 Transistor MOSFET.

### 1.3.2.2 Características de los Transistores FET

Estas características determinan su eficiencia en diversas aplicaciones y les otorgan ventajas significativas frente a otros dispositivos semiconductores.

- ***Impedancia de entrada alta***

Los FET presentan una impedancia de entrada que puede alcanzar valores del orden de miles de megaohmios a bajas frecuencias. Esta propiedad se debe a la unión de puerta polarizada en inversa, que actúa como un circuito abierto. Gracias a ello, los FET pueden muestrear señales sin interferir con la fuente de señal. En la práctica, es común observar impedancias de entrada de cientos o incluso miles de megaohmios, lo que los hace ideales para circuitos donde se requiere una mínima interacción con la fuente (Johnson, 2025).

- ***Operación controlada por voltaje***

A diferencia de los transistores bipolares de unión (BJT), los FET son dispositivos controlados por voltaje. El voltaje aplicado al terminal de la compuerta regula la corriente que circula entre la fuente y el drenador. La compuerta solo requiere una corriente mínima para cargar su capacitancia, generando un campo eléctrico que modifica la conductividad del canal. De esta manera, el FET puede modular el flujo de corriente con un consumo energético muy bajo (Johnson, 2025).

- ***Bajo consumo de energía***

El reducido consumo de energía de los FET los hace especialmente adecuados para dispositivos alimentados por batería. Una vez que la compuerta se carga o descarga, no requiere corriente adicional, a diferencia de los BJT, que necesitan una corriente de base constante. Esta eficiencia energética permite diseñar circuitos más compactos y con menor generación de calor, siendo los FET elementos clave en los sistemas de gestión de energía de dispositivos electrónicos portátiles (Johnson, 2025).

- ***Conducción unipolar (solo portadores mayoritarios)***

Los FET son dispositivos unipolares, lo que significa que utilizan únicamente un tipo de portador de carga: electrones en los canales tipo n o huecos en los canales tipo p. Esta característica fundamental simplifica su funcionamiento y su control eléctrico, favoreciendo una operación eficiente y estable en una amplia variedad de aplicaciones (Johnson, 2025).

- ***Respuesta en frecuencia y velocidad de conmutación***

Los FET destacan por su elevada velocidad de conmutación, lo que los hace ideales para circuitos digitales y aplicaciones de alta frecuencia. A diferencia de los BJT, no presentan problemas de almacenamiento de carga, lo que permite un funcionamiento más rápido y preciso. Sin embargo, a medida que aumenta la frecuencia, los componentes pasivos deben ser más pequeños, lo que puede incrementar las pérdidas por conmutación. Por ello, los diseñadores deben equilibrar cuidadosamente la eficiencia, el tamaño del circuito y la velocidad de operación (Johnson, 2025).

### 1.3.3 Tecnología MOSFET

La estructura metal-óxido-semiconductor (MOS) está formada por tres capas principales y dos terminales. Su estructura planar permite una integración densa en circuitos integrados, lo que ha sido fundamental para la evolución de la microelectrónica. En la base se encuentra un sustrato de silicio (Si), dopado con tipo impurezas tipo p o n. Encima de este sustrato se crece una capa de óxido de silicio ( $\text{SiO}_2$ ), la cual actúa como aislante gracias a sus propiedades dieléctricas y, entre cada material depositado se depositará otra capa de  $\text{SiO}_2$ , después se implantan las regiones tipo n o p para la compuerta y drenador, luego se deposita el polisilicio que conformará la compuerta; finalmente, sobre el óxido se deposita un metal, usualmente es aluminio (Al), que permite la conducción de corriente entre todos los transistores del circuito. En la parte inferior del dispositivo se agrega un contacto óhmico, que asegura una conexión eléctrica eficiente con el sustrato (Sedra & Smith, 2004).

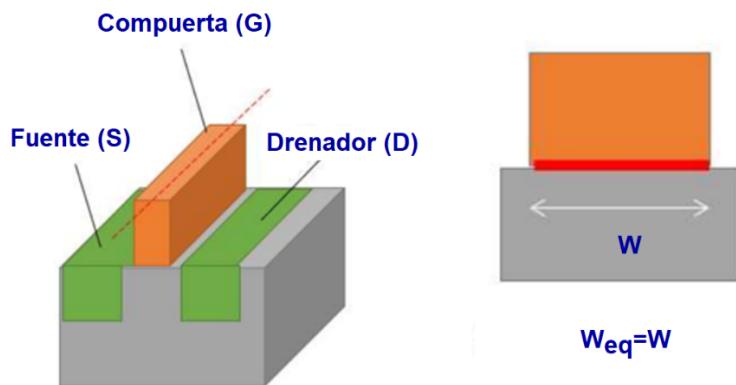


Ilustración 12 Estructura del MOSFET (Fuente: Microwind).

La estructura MOS actúa como un condensador de placas paralelas en el que la compuerta (G) y el sustrato o bulk (B) son las placas y el óxido el aislante. Este permite controlar la densidad de portadores en un semiconductor mediante la aplicación de un voltaje entre la compuerta y el sustrato. Cuando la tensión entre la compuerta y el sustrato  $V_{GB} = 0$ , no existe acumulación de cargas y los portadores en el semiconductor permanecen distribuidos de manera aleatoria, en un estado de equilibrio (Sedra & Smith, 2004).

Al incrementar la tensión positiva  $V_{GB} > 0$ , se genera un campo eléctrico que actúa sobre los portadores del semiconductor tipo p. Esto provoca la formación de una región de vaciamiento, caracterizada por una disminución de portadores libres, similar a lo que ocurre en una unión PN polarizada negativamente. A medida que el voltaje aumenta, la concentración de iones negativos en esta zona también crece (Sedra & Smith, 2004).

Cuando la tensión aplicada supera el voltaje de umbral, los iones presentes ya no son suficientes para compensar el campo eléctrico. En consecuencia, los electrones libres comienzan a acumularse cerca de la terminal positiva, dando lugar a un estado de *fuerte inversión*. Este proceso implica un cambio de polaridad del sustrato bajo la compuerta, formando un canal de electrones libres que permite el flujo de corriente eléctrica, mientras que los huecos permanecen en la región p+ cercana al fondo del sustrato.

Es importante destacar que la corriente en la compuerta  $i_G$  es prácticamente nula, ya que en continua la estructura se comporta como un condensador. Esto significa que la compuerta presenta una impedancia muy alta y no permite el paso de corriente continua. Por lo tanto, la estructura MOS se convierte en un dispositivo ideal para modular la conductividad del canal mediante la tensión aplicada, lo que constituye la base de operación de transistores y circuitos integrados (Sedra & Smith, 2004).

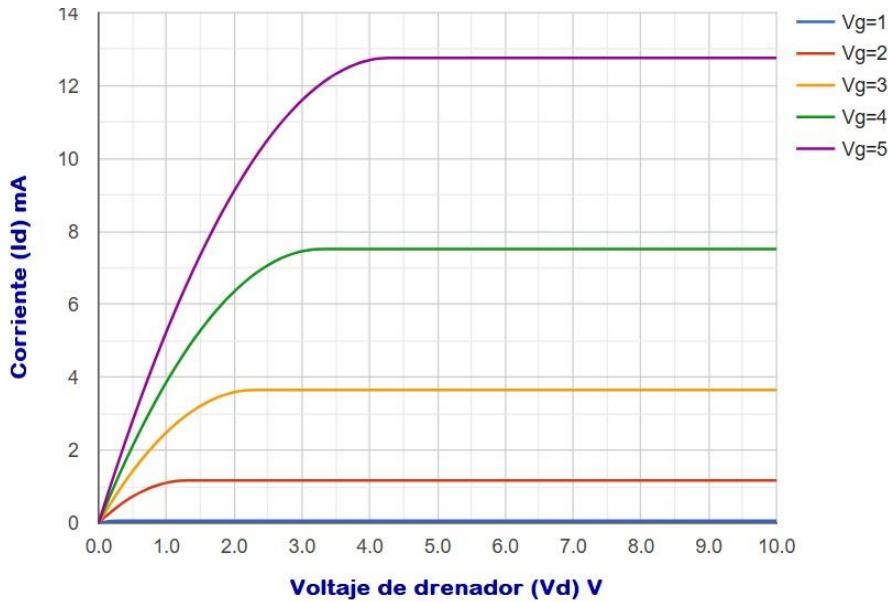


Ilustración 13 Curvas características del MOSFET.

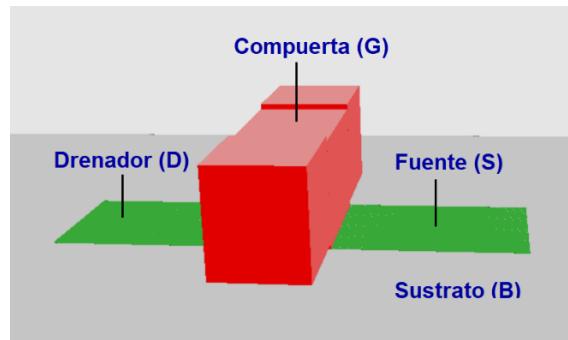


Ilustración 14 Transistor MOSFET NMOS (Microwind).

### 1.3.3.1 Ventajas de los MOSFET

- Alta eficiencia y rapidez: Los MOSFET ofrecen velocidades de conmutación elevadas y requieren una potencia de control mínima, lo que los hace ideales para aplicaciones de alta velocidad como fuentes de alimentación conmutadas y controladores de motores.
- Baja resistencia de encendido: Esto se traduce en menores pérdidas de energía durante la operación.
- Facilidad de fabricación: Su estructura planar simplifica los procesos de fabricación y reduce los costos.
- Escalabilidad hasta cierto límite: Funcionan bien en escalas mayores a 45–32 nm, permitiendo integración densa sin complicaciones de litografía extrema (Mendiola Oria, 2018), (Del Alamo, y otros, 2016).

### 1.3.3.2 Desafíos en el diseño MOSFET

- Efecto de canal corto: A medida que los transistores se miniaturizan, el control del canal se ve comprometido, aumentando las fugas de corriente.
- Variabilidad en procesos de fabricación: Las pequeñas variaciones en el proceso de fabricación pueden afectar el rendimiento del transistor.
- Limitaciones en escalabilidad: A medida que los nodos de proceso se reducen, los MOSFETs enfrentan desafíos significativos en términos de rendimiento y eficiencia.
- Efecto de canal corto en nodos nanométricos: Por debajo de 32–22 nm, el control del canal se debilita, aumentando fugas de corriente.

- Limitaciones de escalado: Los MOSFET planos pierden eficiencia en nodos avanzados (<14 nm), (Mendiola Oria, 2018), (Del Alamo, y otros, 2016).

### 1.3.4 Tecnología FinFET

A diferencia de los MOSFET planos, los FinFET presentan una estructura tridimensional en forma de aleta, donde la puerta rodea el canal por tres lados. Esta configuración permite un control mucho más preciso del flujo de corriente, minimizando las fugas y mejorando el comportamiento eléctrico incluso a escalas nanométricas. Gracias a ello, la evolución hacia FinFET marcó un hito en la microelectrónica, ya que posibilitó la continuación de la miniaturización de los dispositivos sin sacrificar rendimiento, velocidad ni eficiencia energética (González Vidal, 2021).

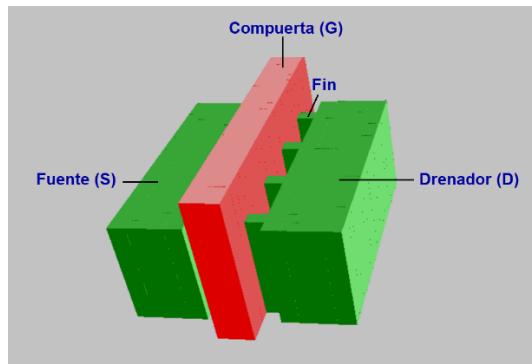


Ilustración 15 Transistor FinFET NMOS de 4 Fins (Microwind).

La necesidad de estructuras como los FinFET surge porque la reducción del tamaño de los transistores enfrenta un límite impuesto por la corriente de fuga en estado de corte, que incrementa el consumo de energía y limita el desempeño de las memorias. Esta corriente depende exponencialmente de la tensión de puerta, de modo que disminuir el margen entre los estados de apagado y encendido reduce significativamente la eficiencia del dispositivo (García Moreno, 2010).

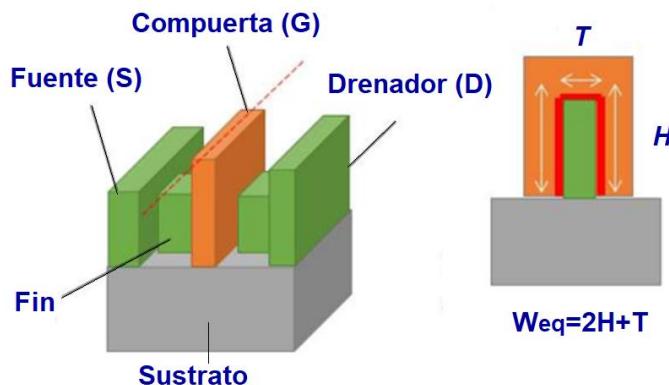


Ilustración 16 Estructura del FinFET (Fuente: Microwind).

Para enfrentar este problema, se desarrollaron estructuras avanzadas como los TFTs y los MOSFET de doble puerta (DGFET), que logran un mayor control del canal mediante capas de óxido adicionales. Dentro de los DGFET, los FinFET se destacan por su canal en forma de “fin” rodeado por la puerta en varios lados, lo que reduce las fugas y mejora el desempeño, consolidándolos como los candidatos más prometedores para la próxima generación de transistores (García Moreno, 2010).

#### **1.3.4.1 Ventajas de los FinFET**

- Mejor control del canal: La estructura en forma de aleta permite que la puerta controle el canal desde tres lados, reduciendo los efectos de canal corto.
- Reducción de fugas de corriente: El diseño tridimensional minimiza las fugas de corriente en estado de apagado.
- Alta densidad de corriente: Esto se traduce en una mayor velocidad de conmutación y rendimiento.
- Escalabilidad efectiva: Adecuados para nodos avanzados, hasta 7 nm o 5 nm, manteniendo eficiencia y estabilidad.  
Mejor consistencia y fiabilidad: La tridimensionalidad reduce la variabilidad en nanoescala (Mari, 2020).

#### **1.3.4.2 Desafíos en el diseño de FinFET**

- Mayor complejidad en el layout y el enrutamiento.
- Sensibilidad a variaciones de proceso y geometría.
- Necesidad de herramientas de diseño específicas para su correcta implementación.
- Complejidad de fabricación en nanometría: Requiere litografía avanzada y control preciso de la geometría de las aletas.
- Variabilidad de aletas a escala nanométrica: Pequeñas diferencias afectan significativamente el rendimiento.
- Consumo de área físico: La tridimensionalidad puede limitar la densidad total de transistores en nodos extremadamente pequeños (Mari, 2020).

### 1.3.5 Tecnología NSFET

El transistor de efecto de campo de nano placas o láminas (Nanosheet FET, NSFET) es esencialmente una evolución avanzada del FinFET y se basa en la arquitectura de compuerta rodeada (Gate-All-Around, GAA). Aunque la tecnología 3D del FinFET ha sido fundamental para la miniaturización de los transistores, presenta varios desafíos de diseño y fabricación, especialmente relacionados con la degradación de las características de canal corto. Algunos elementos, como los espaciadores de alta permitividad, pueden mejorar el desempeño, pero el FinFET enfrenta limitaciones: a medida que las tecnologías avanzan a nodos más pequeños, la altura de la aleta del FinFET aumenta mientras que su ancho se reduce, haciendo cada vez más difícil fabricar dispositivos pequeños con alta eficiencia (Kumar, Kumar Pal, & Yadav, 2019).

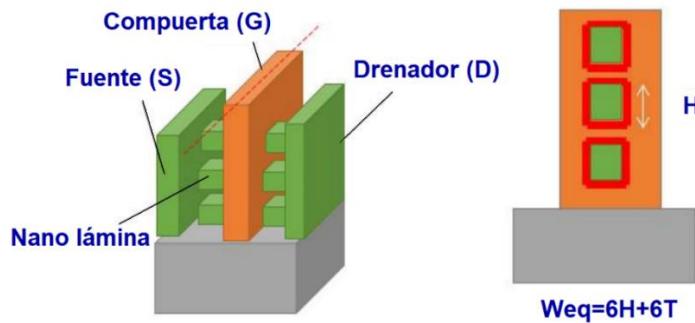


Ilustración 17 Estructura del NSFET (Fuente: Microwind).

En este contexto, los NSFET GAA surgen como una alternativa prometedora. A diferencia del FinFET, donde el canal es vertical y parcialmente controlado por la compuerta, en los NSFET el canal es horizontal y está completamente envuelto por la compuerta, de ahí el término Gate-All-Around o GAA. Esto permite un mejor control electrostático y un incremento de la corriente de conducción, reduciendo significativamente fugas y variaciones de umbral. Además, al emplear varias láminas de semiconductor como compuertas en paralelo, se forma un Multi-fin Nanosheet, que aumenta la densidad de corriente sin ampliar el área del dispositivo. Samsung ha desarrollado su versión de esta tecnología, conocida como Multi Bridge Channel (MBC) FET, implementando precisamente la arquitectura GAA para mejorar el rendimiento en nodos de 3 nm (Kumar, Kumar Pal, & Yadav, 2019).

Gracias a estas características, los NSFET GAA ofrecen un desempeño superior al FinFET, combinando un mejor control de la puerta con una mayor capacidad de conducción de corriente, manteniendo el tamaño compacto de los transistores modernos.

Su canal más ancho y la posibilidad de ajustar el número y ancho de las láminas hacen que los NSFET sean cada vez más relevantes en la industria de semiconductores de próxima generación (Kumar, Kumar Pal, & Yadav, 2019).

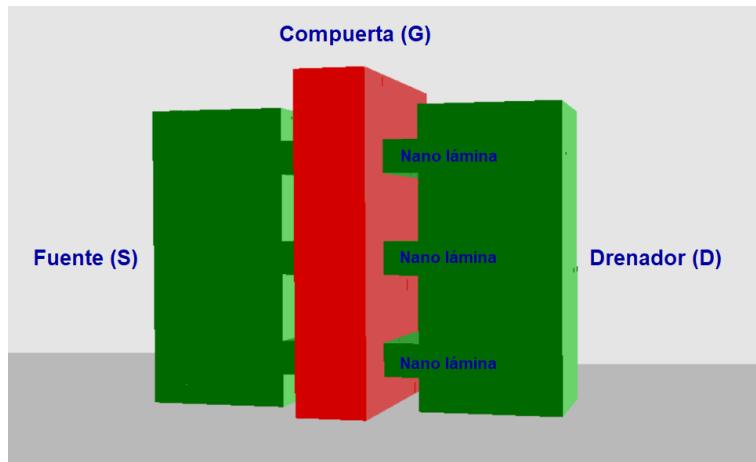


Ilustración 18 Transistor NSFET NMOS 3 Nanosheets (Microwind).

#### 1.3.4.1 Ventajas de NSFET

- Traduce en una mayor eficiencia energética.
- Escalabilidad: Los NSFETs son adecuados para nodos de proceso avanzados, permitiendo la fabricación de dispositivos más pequeños y potentes.
- Mayor densidad de transistores: La estructura de hojas facilita más transistores por área de chip a nanoescala.
- Control superior del canal a nanoescala extremo: Adecuado para nodos de 5 nm a 3 nm y más allá, reduciendo fugas al mínimo (Mukesh & Zhang, 2022).

#### 1.3.4.2 Desafíos en el diseño de NSFET

- Complejidad en la fabricación: La fabricación de NSFETs es más compleja que la de FinFETs, requiriendo técnicas avanzadas de litografía.
- Variabilidad en el proceso de fabricación: Las pequeñas variaciones en el proceso de fabricación pueden afectar el rendimiento del transistor.
- Desafíos en la integración: La integración de NSFETs en circuitos existentes puede presentar desafíos técnicos.
- Variabilidad en nodos sub-5 nm: Pequeñas desviaciones afectan rendimiento y consistencia.
- Integración compleja con tecnologías existentes: Requiere adaptación de procesos y diseños para ser eficiente (Mukesh & Zhang, 2022).

### 1.3.4.3 Funcionamiento del Transistor NSFET

El NSFET está diseñado para trabajar a escalas extremadamente pequeñas, donde mantener el control del canal se vuelve un desafío. Su funcionamiento se basa en un principio sencillo: usar nano láminas (nanosheets) muy delgadas de material semiconductor y rodearlas completamente con la compuerta (GAA) para controlar con precisión el paso de corriente. En operación, el transistor cuenta con tres terminales principales: fuente, drenador y compuerta. La corriente fluye entre la fuente y el drenador, pero solo cuando la compuerta aplica un voltaje suficiente para “activar” los nanosheets.

Cuando se aplica un voltaje positivo en la compuerta (en un dispositivo “tipo n”), el campo eléctrico penetra alrededor de cada nanosheet por todos sus lados. Este campo induce una capa de inversión en cada lámina, permitiendo que los electrones se muevan a lo largo del canal. Dado que los nanosheets están apilados verticalmente, cada uno contribuye de manera simultánea al flujo total de corriente. Esto permite aumentar la conducción sin que el dispositivo crezca en tamaño horizontal. En ausencia de voltaje en la compuerta, el sistema actúa de forma inversa. La envoltura completa de la compuerta bloquea la formación de la capa de inversión y evita el movimiento de portadores, manteniendo el transistor apagado.

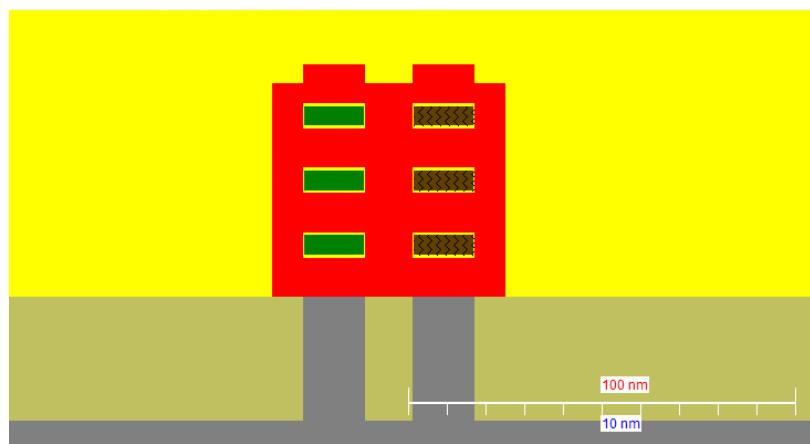


Ilustración 19 Compuerta NOT NSFET Corte vertical (Microwind).

## 1.3.5 CI 74181 (4 bits): Unidad Aritmético Lógica (ALU)

La Unidad Aritmeticológica (ALU) es el núcleo esencial de cualquier procesador digital, encargada de ejecutar operaciones aritméticas —como suma, resta, multiplicación y división y las funciones lógicas, tales como AND, OR, XOR y NOT, sobre los datos binarios procesados por la Unidad Central de Procesamiento (CPU). Su diseño determina en gran medida la eficiencia, velocidad y capacidad de cálculo del sistema (Texas Instruments, 1972).

El circuito integrado 74181 es una ALU de 4 bits que implementa 16 funciones lógicas posibles con dos variables. Además, sus funciones aritméticas incluyen suma y resta, tanto con acarreo como sin él. Este CI puede trabajar con datos activos-altos, donde un nivel lógico alto representa 1, o con datos activos-bajos, donde un nivel lógico bajo representa 0 (Texas Instruments, 1972).

### 1.3.5.1 Entradas y Salidas

La ALU 74181 cuenta con varias señales de control y datos:

- $S_0$  a  $S_3$ : cuatro entradas de selección que determinan la función a ejecutar, ya sea lógica o aritmética.
- $M$ : selecciona entre operaciones lógicas o aritméticas.
- $C$ : acarreo de entrada, utilizado en operaciones aritméticas para sumar o restar con el valor anterior.
- $A$  y  $B$ : entradas de datos de 4 bits que representan los operandos.
- $F$ : salida numérica de la operación seleccionada.
- $P$  y  $G$ : señales de propagación y generación de acarreo, utilizadas para implementar un sumador con acarreo anticipado (carry look-ahead adder), que puede construirse mediante uno o varios chips 74182, optimizando la velocidad de cálculo de sumas complejas.

Esta configuración permite que el 74181 ejecute operaciones aritméticas y lógicas de manera rápida y eficiente, siendo ampliamente utilizado en minicomputadoras y sistemas de procesamiento digital de la época.

### 1.3.5.2 Funcionamiento de la Unidad Aritmético lógica

La función principal de la unidad aritmético lógica (ALU) 74181 consiste en realizar cálculos aritméticos y operaciones lógicas entre dos operandos de 4 bits, denominados  $A$  y  $B$ , generando como salida un resultado también de 4 bits ( $F$ ). Además, la ALU produce diversas señales de estado como *Carry*, *Overflow*, *Zero* y *Negative* que permiten al sistema interpretar los resultados o encadenar varios módulos en operaciones de mayor complejidad.

#### Bloque Fundamental de 4 bits

##### *Entradas principales*

- $A_0 - A_3$ : Operando A (4 bits).
- $B_0 - B_3$ : Operando B (4 bits).
- $S_0 - S_3$ : Líneas de selección de operación (4 bits).
- **M**: Modo de operación
  - $M = 0 \rightarrow$  Operaciones **aritméticas** (suma, resta, etc.)
  - $M = 1 \rightarrow$  Operaciones **lógicas** (*AND*, *OR*, *XOR*, etc.)
- $C_n$ : Bit de acarreo de entrada (*Carry in*).

##### *Salidas principales*

- $F_0 - F_3$ : Resultado de 4 bits.
- $C_n + 4$ : Acarreo de salida.
- **P (Propagate)** y **G (Generate)**: Señales que indican la propagación o generación de acarreo, utilizadas para la conexión en cascada con otras ALUs.

## Funcionamiento interno

La ALU 74181 integra dos bloques funcionales principales: una red lógica y una red aritmética, que operan de forma conjunta dependiendo del modo de operación seleccionado.

### Bloque lógico:

Cuando  $M = 1$ , las salidas se generan aplicando funciones booleanas entre  $A$  y  $B$  según las líneas  $S_0 - S_3$ . Por ejemplo:

Código ( $S_3 - S_0$ )	Operación lógica
0	$\text{NOT } A$
1001	$A \text{ XOR } B$
1010	$A \text{ OR } B$
1111	$A \text{ AND } B$

Tabla 1 Modos de selección de la ALU 74181.

### Bloque aritmético:

Cuando  $M = 0$ , el circuito interno implementa una **suma condicional** mediante compuertas XOR y una red de acarreo. Se pueden realizar operaciones como:

- Suma:  $A + B$
- Resta:  $A - B$  (*usando el complemento de B*)
- Incremento / Decremento:  $A \pm 1$

Las señales **P** y **G** controlan la propagación del acarreo, permitiendo conectar varias ALUs 74181 de manera eficiente sin pérdida funcional.

## Expansión del diseño a 32 bits.

El 74181 procesa únicamente 4 bits por módulo. Para construir una ALU de mayor capacidad (8, 16 o 32 bits), se emplea la configuración en cascada tipo *ripple-carry*, en la cual varios circuitos 74181 se conectan secuencialmente.

En este esquema, el acarreo de salida de un bloque  $C_n + 4$  se enlaza directamente al acarreo de entrada del siguiente  $C_n$ . De esta manera, el flujo de acarreos se propaga desde el bit menos significativo hacia el más significativo.

$$C_0 \rightarrow ALU_0 \rightarrow C_4 \rightarrow ALU_1 \rightarrow C_8 \rightarrow ALU_2 \rightarrow \dots \rightarrow C_{28} \rightarrow ALU_7 \rightarrow C_{32}$$

### Asignación de segmentos de bits

Cada módulo 74181 procesa un grupo de 4 bits:

- $ALU_0: A_0 - A_3, B_0 - B_3 \rightarrow F_0 - F_3$
- $ALU_1: A_4 - A_7, B_4 - B_7 \rightarrow F_4 - F_7$
- $ALU_2: A_8 - A_{11}, B_8 - B_{11} \rightarrow F_8 - F_{11}$
- ...
- $ALU_7: A_{28} - A_{31}, B_{28} - B_{31} \rightarrow F_{28} - F_{31}$

### Líneas de control y datos

- Las líneas de control  $S_3 - S_0$  y el modo  $M$  se aplican en paralelo a todos los módulos.
- Las entradas  $A$  y  $B$  se distribuyen por bloques de 4 bits.
- Las salidas  $F$  de cada módulo se concatenan para formar el resultado total  $F_{31} \dots F_0$ .

## Señales de estado global

Una vez interconectadas las ALUs, se obtienen las siguientes señales globales:

- **$C_{32}$** : Acarreo final de la cadena.
- ***Zero***: Detectado mediante una compuerta NOR sobre todas las salidas  $F_0 - F_{31}$ .
- ***Signo (Negative)***: Corresponde al bit más significativo del resultado  $F_{31}$ .
- ***Overflow (V)***: Calculado mediante la relación:

$$V = (A_{31} \wedge B_{31} \wedge \neg F_{31}) \vee (\neg A_{31} \wedge \neg B_{31} \wedge F_{31})$$

## Consideraciones de diseño

El retardo total del sistema crece de forma lineal con el número de módulos interconectados:

$$t_{total} \approx \left(\frac{N}{4}\right) \cdot t_{propagación}$$

donde  $t_{propagación}$  representa el tiempo de propagación típico de un módulo individual. Aunque la latencia se incrementa al aumentar el número de bits, esta arquitectura ofrece gran simplicidad, modularidad y facilidad de comprobación experimental.

El diseño escalable de la ALU mediante la conexión en cascada de módulos 74181, uniendo las señales de acarreo  $C_n$  y  $C_{n+4}$ , permite extender su capacidad a 8, 16 o 32 bits de manera ordenada y modular. Esta configuración, basada en propagación secuencial de acarreos, resulta adecuada para aplicaciones didácticas y de demostración, donde la claridad del funcionamiento y la facilidad de montaje son

prioritarias. Posteriormente, este mismo diseño puede optimizarse mediante técnicas de anticipación de acarreo (Carry Lookahead) sin alterar la disposición estructural básica.

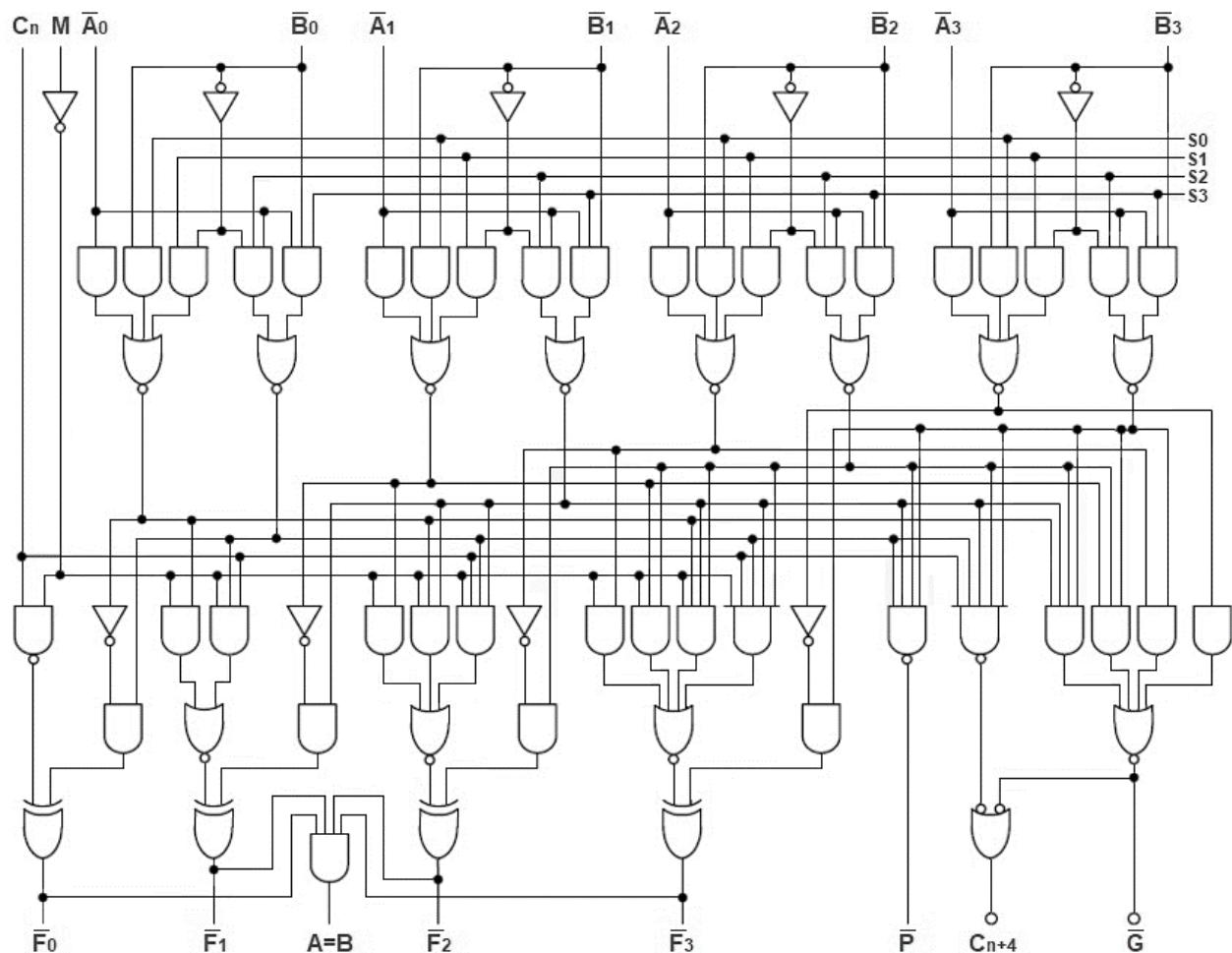


Ilustración 20 Diagrama esquemático de la ALU 74181 (4 BITS)

La ilustración 20 muestra la fotografía del circuito integrado de la ALU 74181 de 4 bits.

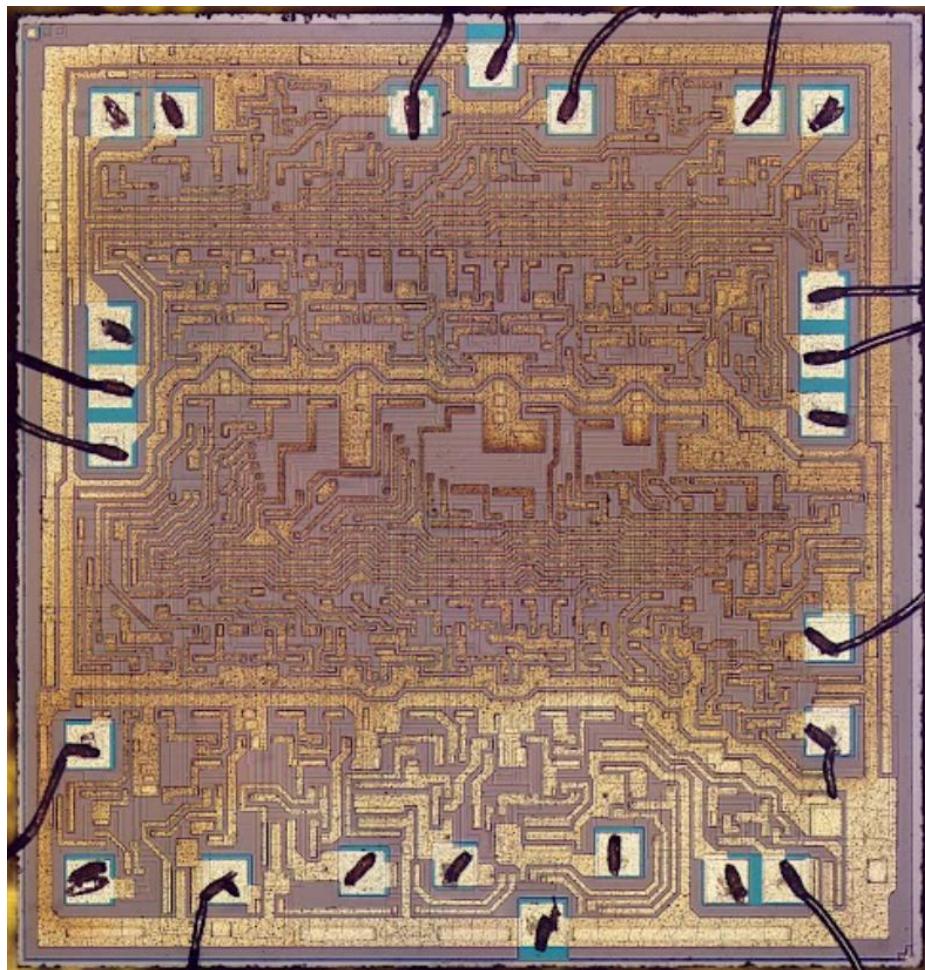


Ilustración 21 Chip 74181

### 1.3.5.3 Tabla de Funciones de la Salida $F$

La salida  $F$  depende de la combinación de las entradas de selección ( $S_0$ – $S_3$ ) y de la señal  $M$ , y puede representarse mediante operaciones lógicas y aritméticas:

- $AND$ : denotada como un producto entre bits.
- $OR$ : denotada con el signo  $+$ .
- $XOR$ : denotada con  $\oplus$ .
- $NOT$ : denotada con barra superior.
- *Suma y resta aritmética*: representadas mediante los símbolos  $+$  y  $-$ , incluyendo o no el acarreo de entrada según corresponda.

El circuito integrado 74181 permite seleccionar 16 funciones lógicas y varias funciones aritméticas de 4 bits según las entradas de selección  $S_0$ – $S_3$  y la señal  $M$  (operación lógica o aritmética), (Shirriff's, 2017).

La salida  $F$  se determina como:

$S_3$	$S_2$	$S_1$	$S_0$	$M$	<b>Función de <math>F</math></b>
0	0	0	0	0	$A + B$ (suma aritmética)
0	0	0	1	0	$A - B$ (resta aritmética)
0	0	1	0	0	$A + B + C$ (suma con acarreo)
0	0	1	1	0	$A - B - C$ (resta con acarreo)
0	1	0	0	1	$A \text{ AND } B$
0	1	0	1	1	$A \text{ OR } B$
0	1	1	0	1	$A \text{ XOR } B$
0	1	1	1	1	$\text{NOT } A$
1	0	0	0	1	$A \text{ AND NOT } B$
1	0	0	1	1	$A \text{ OR NOT } B$
1	0	1	0	1	$A \text{ XOR NOT } B$
1	0	1	1	1	$\text{NOT } B$
1	1	0	0	1	$A \text{ NAND } B$
1	1	0	1	1	$A \text{ NOR } B$
1	1	1	0	1	$A \text{ XNOR } B$
1	1	1	1	1	0 (cero lógico)

Tabla 2 Funciones de  $F$  de la ALU 74181.

- $M = 0$  indica operaciones aritméticas, donde el acarreo de entrada  $C$  puede afectar la operación.
- $M = 1$  indica operaciones lógicas, sin influencia del acarreo.
- Las operaciones aritméticas utilizan las señales  $P$  y  $G$  para implementar sumadores con acarreo anticipado mediante chips 74182, aumentando la velocidad de cálculo.
- Las operaciones lógicas incluyen  $\text{AND}$ ,  $\text{OR}$ ,  $\text{XOR}$ ,  $\text{NOT}$ ,  $\text{NAND}$ ,  $\text{NOR}$  y combinaciones de negación de entradas.

Con esta tabla, el 74181 puede funcionar como un bloque modular en la ALU, capaz de realizar operaciones aritméticas y lógicas de manera eficiente, y se puede integrar en unidades MAC o coprocesadores vectoriales para aplicaciones de procesamiento digital y redes neuronales (Texas Instruments, 1972).

## 1.3.6 Unidad Multiply-Accumulate (MAC)

Las aplicaciones modernas han generado el incremento de las exigencias computacionales en el procesamiento digital de señales (DSP), la visión por computadora y la inteligencia artificial (IA) impulsó el desarrollo de unidades más especializadas capaces de realizar operaciones matemáticas repetitivas de manera más eficiente.

De esta necesidad surgió la Unidad de Multiplicación y Acumulación (MAC, por sus siglas en inglés), considerada una extensión funcional de la ALU. Mientras que la ALU ejecuta operaciones aritméticas básicas, la MAC está optimizada para realizar de manera continua la operación:

$$Y = (A \times B) + C$$

donde  $A$  y  $B$  son los operandos de multiplicación y  $C$  es el valor acumulado previamente. Este proceso se repite de forma iterativa, siendo fundamental en aplicaciones como los filtros digitales, las transformadas rápidas de Fourier (FFT), las redes neuronales y las operaciones de convolución, pilares del procesamiento de señales y la inteligencia artificial moderna (IRJET, 2022).

### 1.3.6.1 Implementación de la unidad MAC

Desde una perspectiva de hardware, la unidad MAC se compone de tres bloques funcionales principales (ilustración 22):

- Multiplicador, que calcula el producto de los operandos.
- Sumador, que añade el resultado del producto al valor almacenado previamente.
- Acumulador, que conserva los resultados parciales y los retroalimenta para el siguiente ciclo.

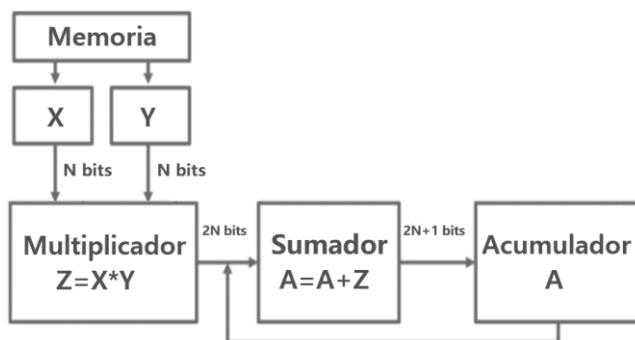


Ilustración 22 Diagrama a bloques de Unidad MAC.

El acumulador, el multiplicador y el sumador, cada uno de los cuales desempeña un papel crítico en el procesamiento digital de señales y en operaciones de inteligencia artificial (Frank & Ramesh, 2022).

### **1.3.6.1.1 Acumulador**

El acumulador es un registro especial que permite almacenar temporalmente los resultados parciales de las operaciones. Está compuesto por un registro de entrada, un registro de desplazamiento (shift register) y una unidad aritmética. Los registros de desplazamiento suelen implementarse con flip-flops tipo D configurados como entrada paralela y salida paralela (Parallel-In Parallel-Out, PIPO), lo que permite que los datos se transfieran simultáneamente en un solo ciclo de reloj, reduciendo el tiempo de retardo.

La Unidad Aritmética (ALU) realiza las operaciones sobre los valores almacenados en los registros de entrada, y el resultado se transfiere al acumulador a través del registro de desplazamiento. La salida del sumador se genera en paralelo, permitiendo que tanto las entradas como las salidas del acumulador se procesen de forma simultánea, lo que aumenta la eficiencia y la velocidad del circuito (Frank & Ramesh, 2022).

### **1.3.6.1.2 Multiplicadores**

Los multiplicadores son fundamentales en sistemas modernos de procesamiento digital de señales y en aplicaciones de inteligencia artificial. Su rendimiento impacta directamente en la eficiencia global del sistema. Para ser efectivos, los multiplicadores deben cumplir ciertas características:

- Alta precisión en los cálculos.
- Capacidad de operar a alta velocidad, minimizando el tiempo de retardo.
- Área de circuito optimizada, para reducir el espacio físico ocupado.
- Bajo consumo de energía.

Estas propiedades permiten que las unidades MAC procesen grandes volúmenes de datos de manera rápida y eficiente, algo crítico en operaciones como filtros digitales, FFT y convoluciones en redes neuronales. (Frank & Ramesh, 2022).

### 1.3.6.1.3 Sumadores

El sumador es un componente esencial para la ejecución de operaciones aritméticas, ya que su arquitectura determina en gran medida la velocidad y eficiencia del sistema. Entre los tipos más utilizados se encuentran: ripple carry adder (RCA), carry look-ahead adder (CLA), carry save adder y carry select adder (CSA), (Frank & Ramesh, 2022).

El carry select adder (CSA) se destaca por su rapidez, ya que utiliza señales de propagación de acarreo ( $P_i$ ) generadas mediante puertas XOR para reducir el tiempo de propagación del acarreo. Estas señales se procesan a través de puertas AND y se utilizan como bits de selección en un multiplexor, generando la salida de acarreo ( $Cout$ ) de manera eficiente. Este enfoque disminuye la ruta crítica del sumador y mejora la velocidad en comparación con los sumadores en cascada tradicionales, aunque su consumo de energía y retraso combinado pueden ser mayores que otros tipos de sumadores (Frank & Ramesh, 2022).

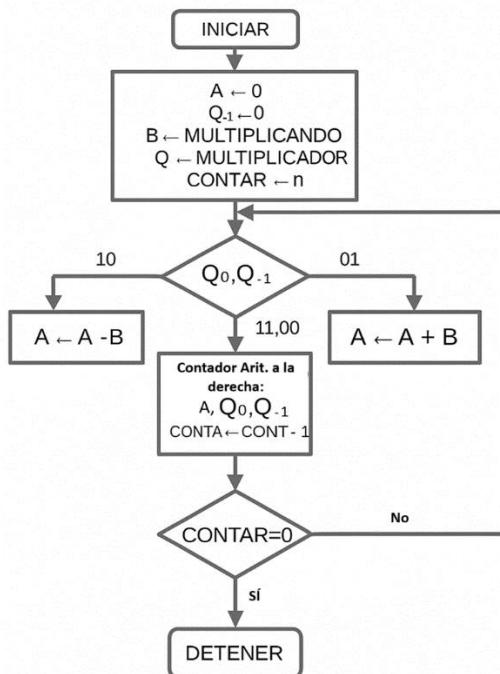


Ilustración 23 Diagrama de Flujo de la Unidad MAC.

Esta arquitectura permite ejecutar sumas sucesivas de productos sin reiniciar el cálculo, optimizando el rendimiento y reduciendo los tiempos de procesamiento en tareas intensivas. En muchas implementaciones, el resultado del acumulador se retroalimenta al sumador, posibilitando una secuencia continua de operaciones con mínima latencia (Frank & Ramesh, 2022).

En los procesadores modernos, las unidades MAC suelen integrarse directamente dentro de la ALU o formar parte de unidades aritméticas extendidas y coprocesadores vectoriales, como ocurre en los núcleos múltiples datos instrucciones sencillas (Single Instruction Multiple Data, SIMD) y en las unidades de procesamiento neuronal (Neural Processing Units, NPU). Estas arquitecturas están especialmente diseñadas para operaciones masivas de multiplicación y acumulación paralela, fundamentales en el entrenamiento y la inferencia de redes neuronales profundas (Dorantes, 2025).

La evolución de la microelectrónica hacia la nanoelectrónica ha permitido que tanto las ALU como las MAC se beneficien de tecnologías de transistor cada vez más avanzadas (Dorantes, 2025).

### **1.3.6 Unidades de Procesamiento Neuronal (NPU)**

Las Unidades de Procesamiento Neuronal (Neural Processing Units, NPU) son arquitecturas especializadas diseñadas para acelerar los cálculos asociados con redes neuronales y otros algoritmos de inteligencia artificial. Su funcionamiento se basa en la integración de un gran número de bloques MAC (Multiply–Accumulate) que operan en paralelo, optimizados para realizar operaciones matriciales y vectoriales con alta eficiencia (Dorantes, 2025).

Gracias a esta estructura paralelizada, las NPUs ofrecen una gran capacidad de cómputo y un consumo energético reducido, lo que las convierte en un componente esencial para el procesamiento inteligente en dispositivos móviles, servidores de alto rendimiento y sistemas de edge computing. Estas unidades permiten ejecutar modelos complejos de aprendizaje profundo directamente en el hardware, sin depender completamente de la nube, reduciendo la latencia y mejorando la privacidad de los datos (Qing, Decheng, & Dongxin, 2024).

En la actualidad, la investigación en torno a las NPUs se centra principalmente en dos objetivos: mejorar la precisión de inferencia y optimizar la eficiencia operativa. En el primer caso, se busca alcanzar un equilibrio entre el tiempo de procesamiento y la exactitud de los resultados, considerando las limitaciones de rendimiento y consumo energético del sistema. En cuanto a la eficiencia, se exploran técnicas de distribución dinámica de carga entre la CPU, GPU y NPU, con el propósito de maximizar el aprovechamiento del hardware y reducir los tiempos de respuesta en tareas de inteligencia artificial (Qing, Decheng, & Dongxin, 2024).

En conjunto, estas características posicionan a las NPUs como uno de los avances más relevantes en la evolución del cómputo moderno, permitiendo que los sistemas digitales ejecuten tareas cognitivas de forma más rápida, precisa y sostenible (Qing, Decheng, & Dongxin, 2024).

### **1.3.6.2 Arquitectura de la NPU**

Actualmente, existen diversos tipos de arquitecturas para las Unidades de Procesamiento Neuronal, sin embargo, la más común sigue siendo la arquitectura tipo Von Neumann. Cabe destacar que ciertos aceleradores, como la Unidad de Procesamiento Tensorial (TPU) de Google, difieren de este esquema al emplear una arquitectura de tipo matriz sistólica, orientada al procesamiento masivo y paralelo de datos (Russel, 2025).

De forma paralela, se están desarrollando nuevas arquitecturas que implementan matrices sistólicas avanzadas, procesamiento en memoria (in-memory computing) e incluso computación analógica basada en neuronas físicas. Estas aproximaciones emergentes prometen una reducción significativa del consumo energético y una mejora sustancial en el rendimiento, abarcando desde sistemas de inferencia compactos hasta redes neuronales de gran escala (Russel, 2025).

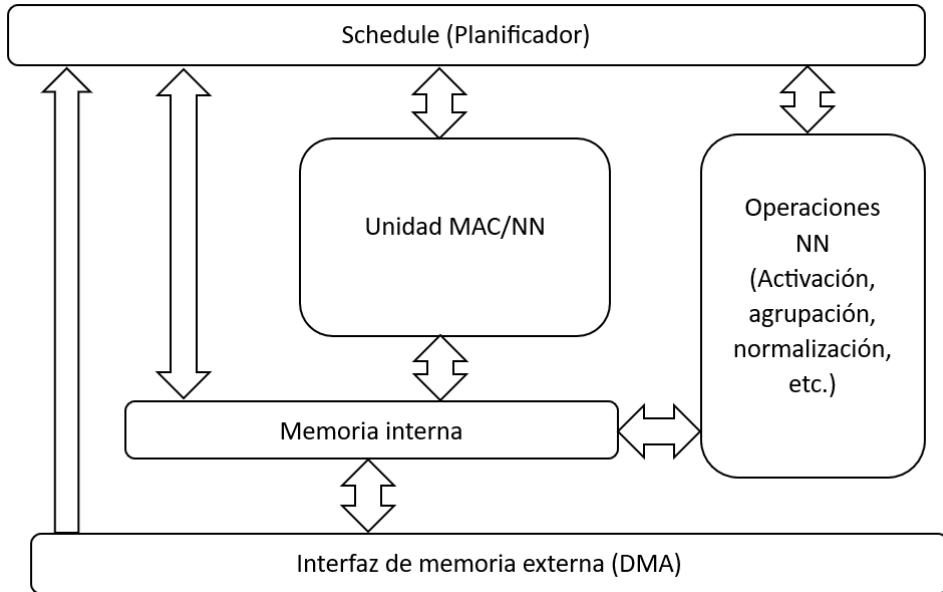
### **1.3.6.2.1 Arquitectura Von Neumann**

Aunque el modelo Von Neumann fue originalmente concebido para las Unidades Centrales de Procesamiento (CPU), sus principios fundamentales se aplican también en el diseño de las NPUs. Estas arquitecturas se componen principalmente de un núcleo de cómputo, un bus de memoria que interconecta memorias internas y externas (como SRAM y DDR), y una unidad de control o planificador de tareas (scheduler) que coordina las operaciones (Russel, 2025).

El núcleo de cómputo constituye el corazón de la NPU y está compuesto por una matriz de multiplicación-acumulación (MAC) o una matriz multiplicadora (MM). Este bloque es altamente eficiente en la ejecución de operaciones de multiplicación y acumulación a gran escala, esenciales para el cálculo de convoluciones y multiplicaciones matriciales, que son la base del funcionamiento de la mayoría de las redes neuronales modernas. Además, las NPUs integradas suelen incorporar soporte para operaciones comunes en redes neuronales, como activaciones, normalización y pooling, las cuales se ejecutan dentro del mismo circuito especializado para maximizar la eficiencia (Russel, 2025).

No obstante, la limitada capacidad de almacenamiento local restringe el tamaño de las redes neuronales que pueden ser aceleradas directamente dentro de la NPU, haciendo necesario recurrir al uso de memorias externas (Russel, 2025).

En el caso de redes neuronales de gran escala, como los modelos de lenguaje extensos (LLMs), se requiere un volumen masivo de almacenamiento tanto para el tamaño del modelo como para los datos intermedios. Debido a ello, las NPUs integradas en dispositivos móviles o embebidos no cuentan con la capacidad suficiente para manejar este nivel de complejidad, limitando su uso a redes de menor tamaño o a operaciones de inferencia específicas (Russel, 2025).



*Ilustración 24 Arquitectura de la NPU de Von Neumann.*

A un nivel general, esta arquitectura sigue un conjunto de instrucciones determinadas en tiempo de compilación, es decir, un programa. Dicho programa contiene comandos fundamentales, tales como: obtener datos desde una ubicación A, ejecutar un conjunto de operaciones algorítmicas sobre esos datos y, finalmente, escribir los resultados en una ubicación B.

El NPU incorpora soporte para una amplia variedad de operaciones propias de redes neuronales, incluyendo diversas variantes de convoluciones, multiplicaciones matriciales, normalizaciones, operaciones de agrupamiento (pooling), operaciones de memoria y funciones de activación, entre otras.

Los NPUs de este tipo suelen trabajar principalmente con operaciones en punto fijo, en lugar de punto flotante, aunque algunas implementaciones también ofrecen soporte para FP16. El uso de aritmética en punto fijo permite reducir tanto el área de silicio necesaria como el consumo de energía, en comparación con la aritmética en punto flotante. Para dimensionar esta diferencia, el área de silicio requerida por una unidad de multiplicación en punto flotante puede ser entre dos y seis veces mayor que la de una unidad en punto fijo, dependiendo del tamaño y la arquitectura del multiplicador (Russel, 2025).

### 1.3.6.3 Framework de la NPU

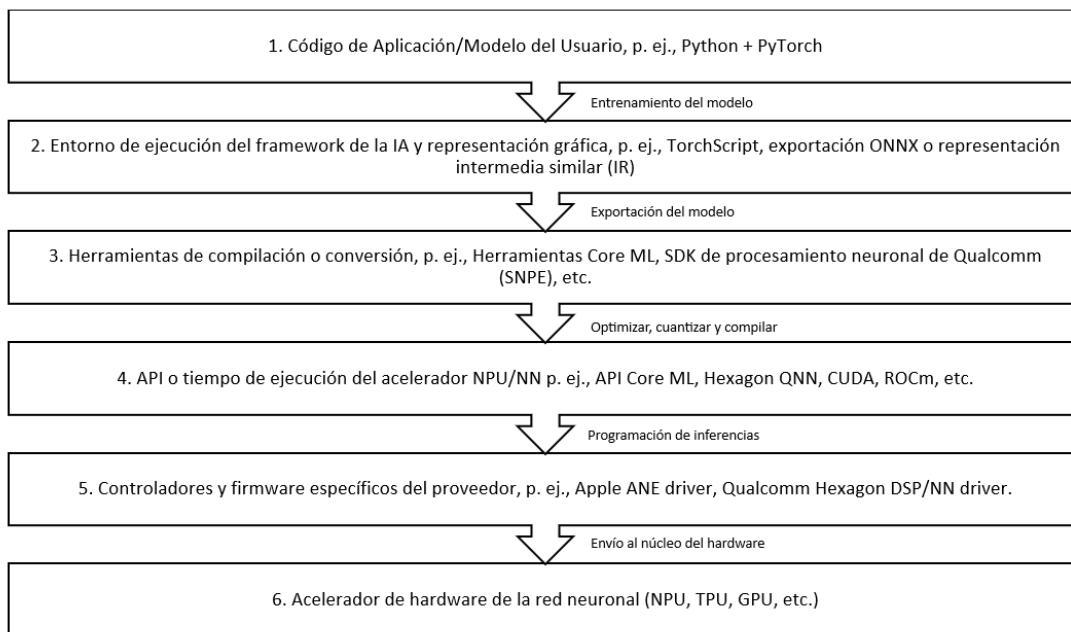


Ilustración 25 Marco de software típico de la NPU.

Es importante destacar que los núcleos NPU integrados en sistemas embebidos suelen operar con tipos de datos y profundidades de bits diferentes (por ejemplo, punto fijo de 8 bits) respecto a las redes neuronales entrenadas en GPU en centros de datos, las cuales utilizan normalmente aritmética en punto flotante de 32 bits (Russel, 2025).

Tal como se observa en el diagrama anterior, esta discrepancia provoca que el comportamiento durante la inferencia de la red (etapa 6) difiera del de la red original entrenada (etapa 1). No obstante, estas diferencias pueden ser mitigadas o incluso eliminadas mediante el uso de técnicas modernas de cuantización y reentrenamiento de redes neuronales, las cuales permiten adaptar los modelos a las limitaciones de precisión y formato de los NPUs embebidos sin una pérdida significativa en el rendimiento o la exactitud del modelo (Russel, 2025).

#### **1.3.6.4 Aplicaciones de la NPU**

Actualmente, sus aplicaciones más comunes se encuentran en los siguientes campos:

- Imágenes y video: clasificación, segmentación, mejora fotográfica y análisis de contenido visual.
- Voz: asistentes inteligentes, traducción de lenguaje, detección de palabras clave y supresión de ruido.
- Biometría: reconocimiento facial y dactilar, así como detección de gestos.
- Realidad aumentada y virtual (AR/VR): mapeo tridimensional, comprensión de escenas, detección de objetos, seguimiento de manos y de la mirada.
- Robótica: localización y mapeo simultáneo (SLAM), detección de defectos en manufactura y análisis de datos de sensores (Russel, 2025).

Entre las aplicaciones emergentes, se destacan la compresión de video asistida por inteligencia artificial y la detección de malware o amenazas directamente en el dispositivo, mejorando la seguridad sin requerir conexión constante a la nube (Russel, 2025).

En la actualidad, las NPUs embebidas no se utilizan ampliamente para la inferencia de modelos de lenguaje de gran escala (LLM), como ChatGPT. Este tipo de aplicaciones se ejecutan casi exclusivamente en centros de datos, debido a sus altas demandas de memoria y procesamiento. Incluso en sistemas como Apple Intelligence, las consultas simples se procesan localmente en el dispositivo, mientras que las más complejas se derivan hacia la nube para su ejecución (Russel, 2025).

### **1.3.6.5 Capacidades de la NPU**

- Aceleración masiva
- Alta eficiencia energética
- Limitación por operaciones de hardware
- Baja flexibilidad computacional
- Dependencia de CPU y GPU
- Coste en transferencia de datos
- Optimización para inferencia
- Sensibilidad a nuevas arquitecturas de red
- Consumo energético reducido (Russel, 2025).

Independientemente de la arquitectura empleada en una Unidad de Procesamiento Neuronal (NPU), su rendimiento final está determinado por las operaciones de hardware implementadas en su núcleo. Las NPUs están diseñadas para acelerar de manera masiva las operaciones de redes neuronales, pero esta optimización implica una reducción en la flexibilidad computacional. A medida que las redes neuronales evolucionan e incorporan nuevas capas y tipos de operaciones, la capacidad efectiva de la NPU tiende a disminuir, ya que su arquitectura no siempre contempla dichas innovaciones (Russel, 2025).

Para mantener la compatibilidad y asegurar que las redes neuronales más recientes puedan ejecutarse en dispositivos embebidos, las NPUs suelen trabajar en conjunto con núcleos CPU y GPU, los cuales asumen aquellas operaciones que la NPU no puede procesar directamente. Sin embargo, este intercambio de datos entre los distintos núcleos de cómputo conlleva un costo en tiempo de inferencia y en consumo energético, afectando la eficiencia global del sistema (Russel, 2025).

## **1.4 Objetivos**

### **Objetivo general**

Diseñar una Unidad Lógica Aritmética (ALU) de 32 bits para aplicaciones Multiply-Accumulate (MAC) de procesamiento neuronal, a nivel transistor, utilizando tecnología NSFET en nodo de 3nm, con el fin de validar su funcionamiento y explorar los beneficios de esta arquitectura en términos de eficiencia y escalabilidad.

### **Objetivos específicos**

- Analizar los fundamentos teóricos de la tecnología NSFET y su aplicación en el diseño de circuitos integrados orientados a procesamiento neuronal.
- Diseñar circuitos esquemáticos de compuertas lógicas a nivel transistor, como bloques básicos para la construcción de la ALU.
- Implementar el layout de la ALU con herramientas de diseño comerciales, siguiendo las buenas prácticas de diseño en tecnología NSFET.
- Simular el funcionamiento eléctrico de las compuertas como de la ALU utilizando herramientas como Microwind y DSCH, validando su correcto desempeño en operaciones de multiplicación y acumulación.
- Evaluar los resultados obtenidos en cuanto a área, consumo de potencia, energético y comportamiento lógico, identificando las ventajas o limitaciones del diseño propuesto.

# **Capítulo 2**

## **Metodología**

## 2.1 Diagrama Esquemático

### a. ALU base de 4 bits (74181)

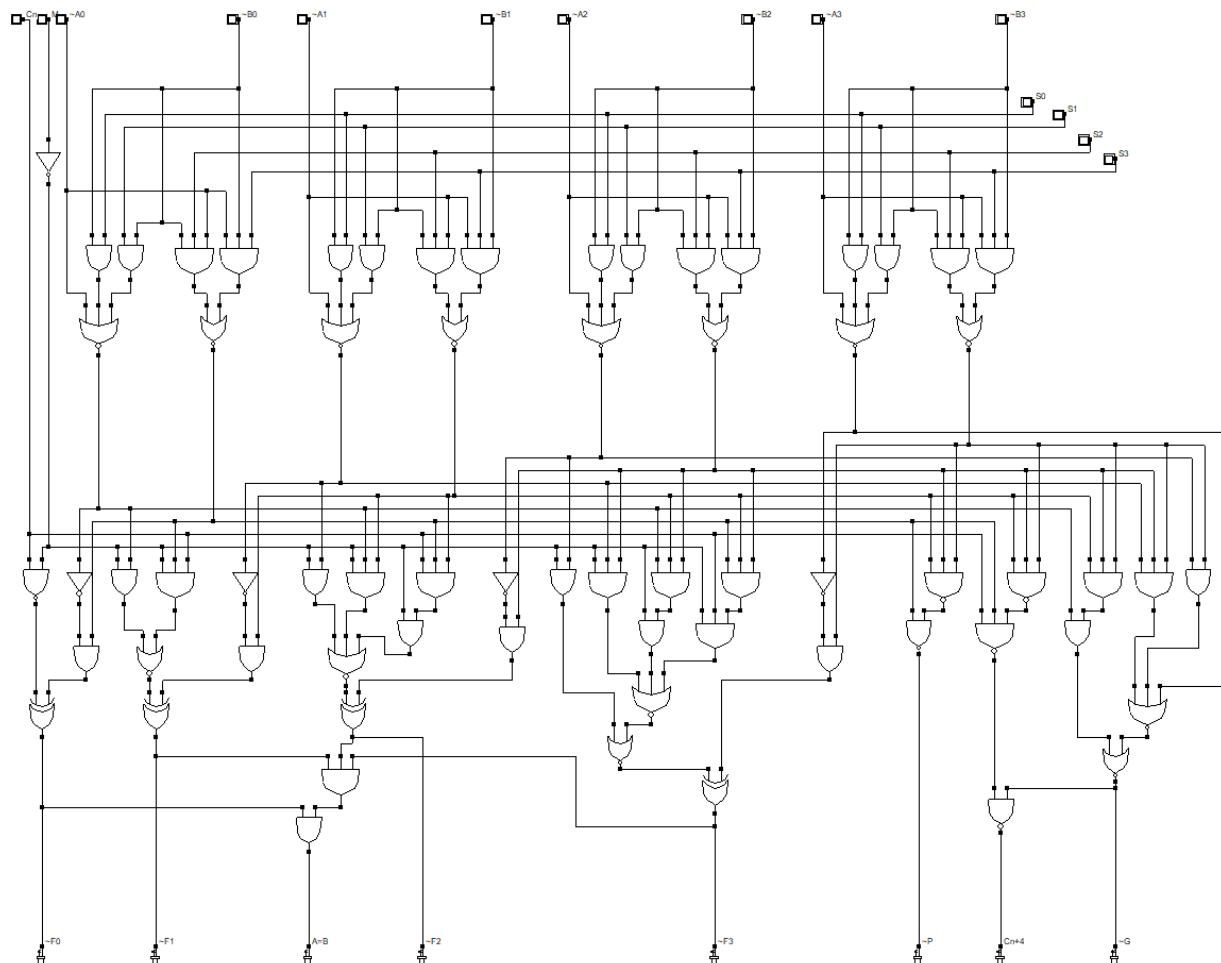


Ilustración 26 Esquemático de la ALU 74181 de 4bits (DSCH).

## b. ALU de 8 bits (2 bloques de 74181)

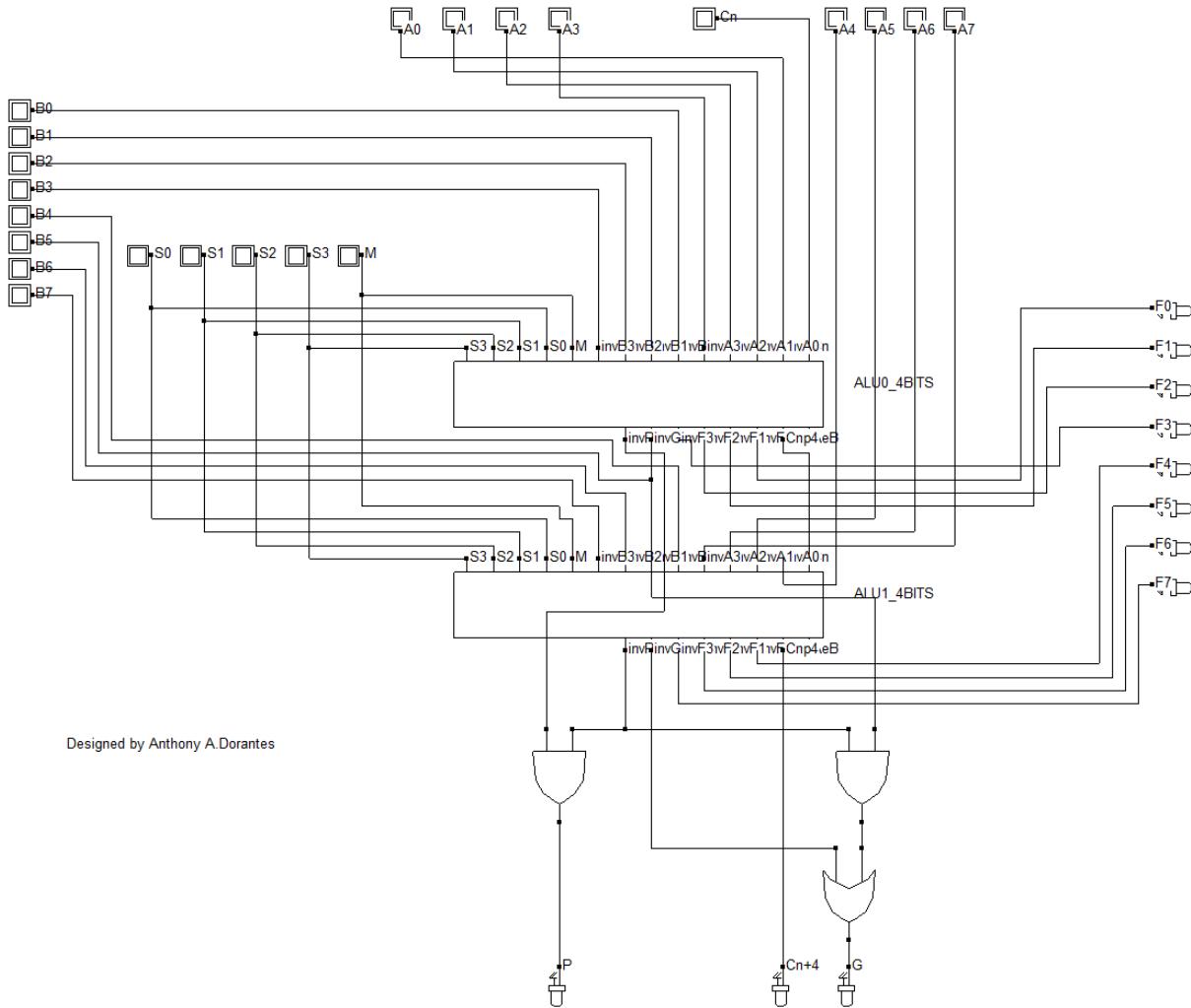


Ilustración 27 Esquemático de la ALU de 8 bits usando 2 bloques del 74181 (DSCH).

### c. ALU de 16 bits (2 bloques de 8 bits)

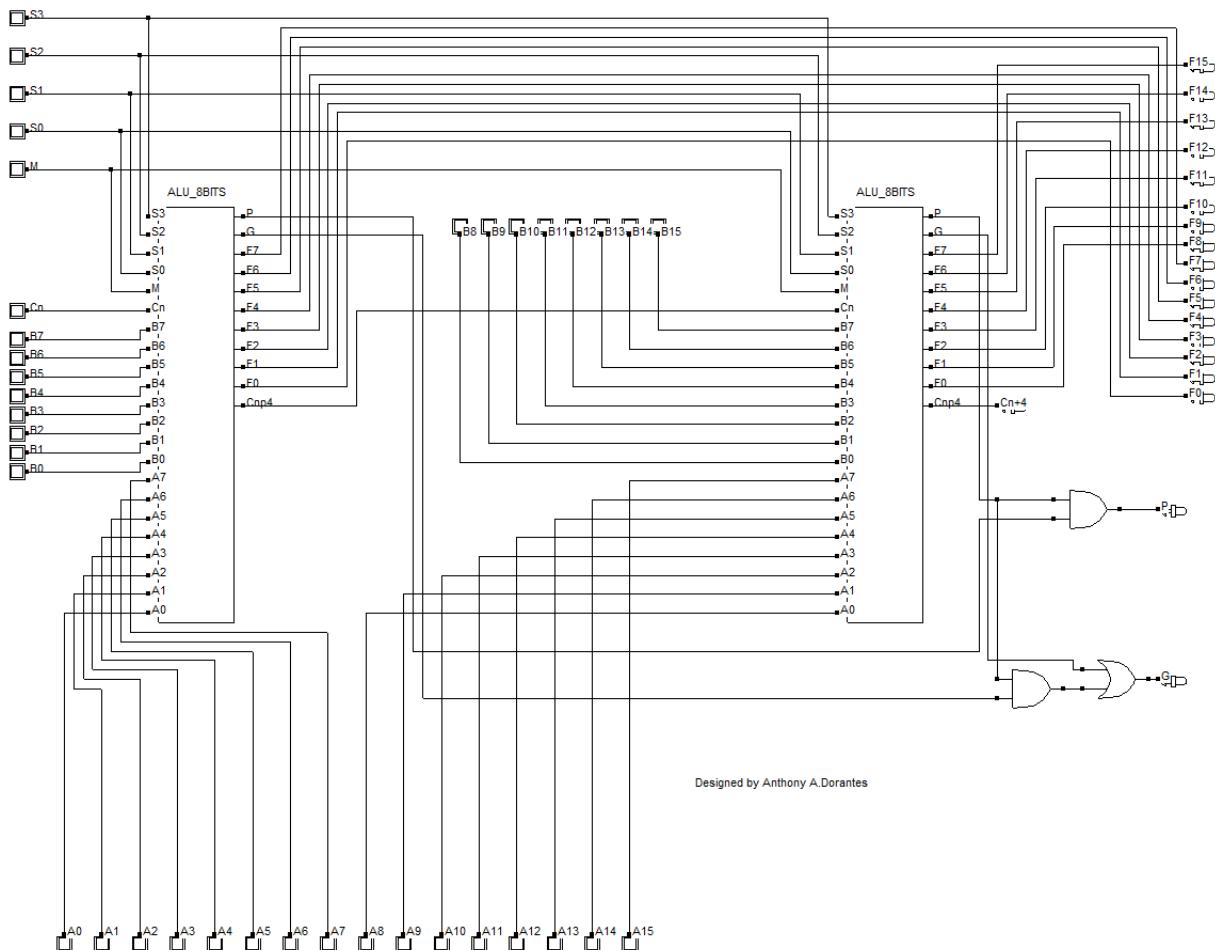
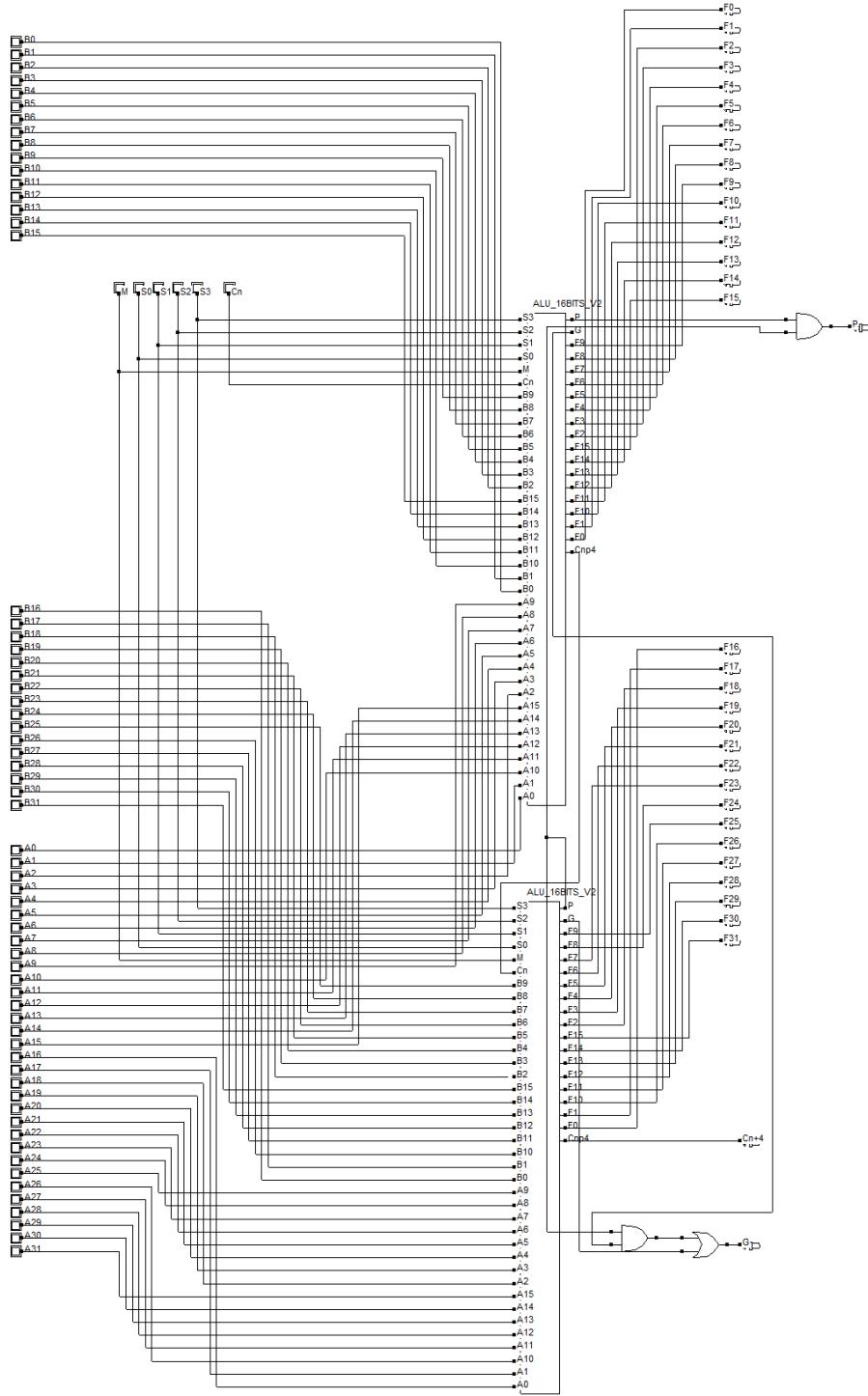


Ilustración 28 Esquemático de la ALU de 16 bits usando 2 bloques de 8 bits (DSCH).

## d. ALU de 32 bits (usando 2 bloques de 16 bits)



Designed by Anthony A. Dorantes

Ilustración 29 Esquemático de la ALU de 32 bits usando 2 bloques de 16 bits (DSCH)

e. ALU de 64 bits (usando dos bloques de 32 bits).

Nota: Debido a las limitaciones del software, fue posible construir un esquemático gráfico de la ALU de 64 bits (tomando como base dieciséis bloques 74181 o dos bloques de 32 bits); sin embargo, este no pudo ser simulado ni compilado, por lo que se concluye como un diseño teórico.

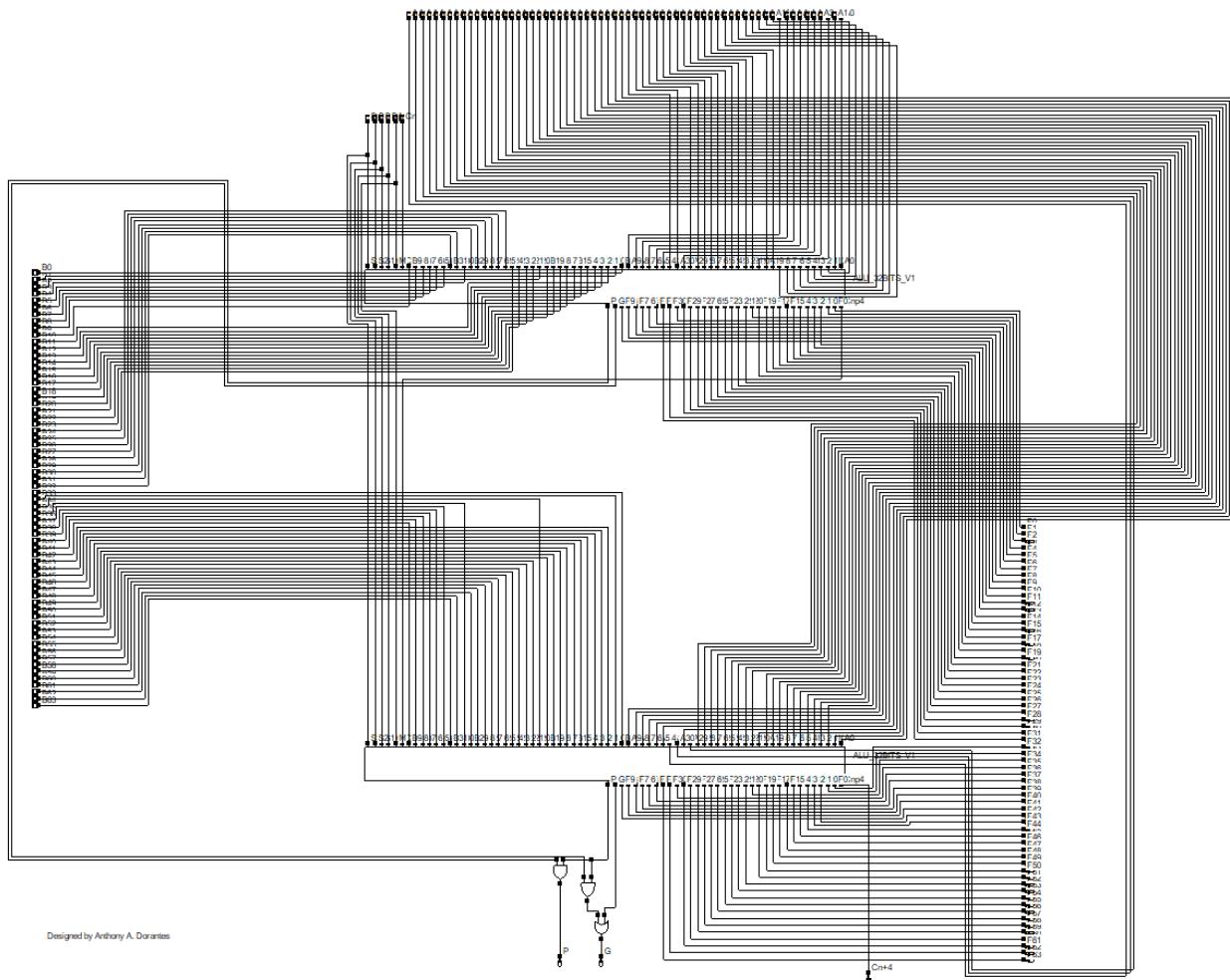


Ilustración 30 Esquemático teórico de la ALU de 64 bits usando dos bloques de 32 bits. (DSCH).

## 2.2 Cálculos

Cálculos comparativos entre NMOS y PMOS para MOSFET (20 nm), y NSFET (3 nm).

- **NSFET (3 nm, 3 nanosheets)**: VDD = 0.65 V,  $V_{GS} = 0.65 V$ ,  $V_{DS} = 0.65 V$
- **MOSFET (20 nm High-K / Metal-Gate)**: VDD = 0.90 V,  $V_{GS} = 0.90 V$ ,  $V_{DS} = 0.90 V$

**Importante:** Todos los números que siguen son *ejemplos representativos* para cálculos comparativos.

**Movilidad  $\mu$**  (electrones): valores de ejemplo extraídos de rangos mostrados en textos y artículos de modelos MOSFET/FinFET/ GAA. En tecnologías avanzadas la movilidad efectiva aumenta y/o varía por estrés y geometría (Sze & Lee, 2012), (Tsividis & McAndrew, 2011), (Dasgupta & Hu, 2020), por eso usamos:

- NSFET:  $\mu_n = 250 \text{ cm}^2/\text{V} \cdot \text{s}$  (*electrones*).  $\mu_p = 100 \text{ cm}^2/\text{V} \cdot \text{s}$ . (*huecos*)
- MOSFET:  $\mu_n = 550 \text{ cm}^2/\text{V} \cdot \text{s}$  (*electrones*).  $\mu_p = 250 \text{ cm}^2/\text{V} \cdot \text{s}$ . (*huecos*)

**Capacitancia del dieléctrico de compuerta por área  $C_{ox}(F/m^2)$ :** Calculable como  $\varepsilon_{ox}/t_{ox}$  valores de ejemplo para nodos avanzados provistos por literatura de procesos (BSIM/Foundries) se usa:

- NSFET (3 nm típico) =  $5 \times 10^{-2} F/m^2$
- MOSFET (20 nm típico) =  $5 \times 10^{-2} F/m^2$

(Sze & Lee, 2012), (Tsividis & McAndrew, 2011), (Dasgupta & Hu, 2020).

**Doping  $N_A$  (densidad de átomos dopantes tipo p),  $N_D$  (densidad de átomos dopantes tipo n) y  $n_i$  (concentración intrínseca de portadores en silicio puro):**

- NSFET:  $N_A = 1 \times 10^{21} m^{-3}$ ,  $N_D = 2 \times 10^{21} m^{-3}$
- MOSFET:  $N_A = 1 \times 10^{23} m^{-3}$ ,  $N_D = 1.5 \times 10^{23} m^{-3}$
- $n_i = 1.5 \times 10^{16} m^{-3}$

Valores típicos de dopado y la densidad intrínseca del silicio a 300 K.

(Sze & Lee, 2012), (Tsividis & McAndrew, 2011), (Dasgupta & Hu, 2020).

Usados en literatura para ilustrar umbrales de diseño (Dasgupta & Hu, 2020).

**$W/L$  (efectivo):** se obtiene como suma del ancho efectivo de cada elemento (fin o nanosheet) dividido entre la longitud de canal. Para que el cálculo sea reproducible se ha fijado un  $W/L$  por estructura de ejemplo y luego multiplicado por el número de estructuras (3 nanosheets), (Dasgupta & Hu, 2020).

$$W = \text{Ancho efectivo}$$
$$L = \text{Longitud de Compuerta}$$

- **NSFET**

$$W_{sheet} = 2(t_{sheet} + h_{sheet})$$

$$W_{eff} = N \cdot W_{sheet}$$

$$\frac{W}{L} = \frac{W_{eff}}{L}$$

Donde:

$t_{sheet}$  = espesor vertical de cada nanosheet (nm).

$h_{sheet}$  = dimensión lateral del sheet (nm)

$N$  = Número de nanosheets.

Spacers = Distancia entre cada nanosheet (típico 1 nm). Esto se le resta a  $t_{sheet}$

Valores típicos de layout.

- $t_{sheet} = 16 \text{ nm}$  (sin spacers).
- $h_{sheet} = 48 \text{ nm}$ .
- $L = 3 \text{ nm}$ .
- $N = 3 \text{ nanosheets}$ .
- Spacers entre cada nanosheets:  $2 \cdot 1 = 2 \text{ nm}$ .

Calculando:

$$t_{sheet} = \frac{16 \text{ nm} - 2 \text{ nm}}{3} = 4.66 \text{ nm.}$$

$$W_{sheet} = 2(4.66 \text{ nm} + 48 \text{ nm}) = 105.3 \text{ nm}$$

$$W_{eff} = 3 \cdot 105.3 \text{ nm} = 316 \text{ nm}$$

$$\frac{W}{L} = \frac{316 \text{ nm}}{3 \text{ nm}} = 105.3$$

- **MOSFET**

Donde:

- $L = 20 \text{ nm.}$
- $W = 100 \text{ nm}$

$$\frac{W}{L} = \frac{100 \text{ nm}}{20 \text{ nm}} = 5$$

**NOTA:** Estas cifras son consistentes con ejemplos de diseño en la literatura técnica (Sze & Lee, 2012), (Tsividis & McAndrew, 2011), (Dasgupta & Hu, 2020).

## Constantes y valores usados (para reproducibilidad)

- $T = 300K \rightarrow kT/q = 0.02585 V$
- $q = 1.602 \times 10^{-19} C$
- $\epsilon_0 = 8.854 \times 10^{-12} F/m$
- $\epsilon_{Si} = 11.7\epsilon_0 = 1.036 \times 10^{-10} F/m$
- $\epsilon_{ox} = 3.9\epsilon_0$

Valores de ejemplo:

$$t_{ox} = 1 \times 10^{-9} m, \quad r_{ch} = 1.5 \times 10^{-9} m$$

(Sze & Lee, 2012), (Tsividis & McAndrew, 2011), (Dasgupta & Hu, 2020).

## Fórmulas Usadas

### Potencial de Sustrato (Sustrato/Bulk) (para sustrato tipo p)

$$\phi_{bn,NMOS} = \frac{kT}{q} \ln \left( \frac{N_A}{n_i} \right)$$

$$\phi_{bn,PMOS} = -\frac{kT}{q} \ln \left( \frac{N_D}{n_i} \right)$$

- **NSFET:**

$$\phi_{bn,NMOS} = 0.02585 \ln\left(\frac{1 \times 10^{21} m^{-3}}{1.5 \times 10^{16} m^{-3}}\right)$$

$$\phi_{bn,NMOS} = 0.287 V$$

$$\phi_{bn,PMOS} = -0.02585 \ln\left(\frac{2 \times 10^{21} m^{-3}}{1.5 \times 10^{16} m^{-3}}\right)$$

$$\phi_{bn,PMOS} = -0.305 V$$

- **MOSFET:**

$$\phi_{bn,NMOS} = 0.02585 \ln\left(\frac{1 \times 10^{23} m^{-3}}{1.5 \times 10^{16} m^{-3}}\right)$$

$$\phi_{bn,NMOS} = 0.424 V$$

$$\phi_{bn,PMOS} = -0.02585 \ln\left(\frac{1.5 \times 10^{23} m^{-3}}{1.5 \times 10^{16} m^{-3}}\right)$$

$$\phi_{bn,PMOS} = -0.416 V$$

$$\phi_{bn} = \phi_F$$

## Voltaje de Umbral

- **NSFET:**

$$V_{Th,NMOS} = V_{FB} + 2\phi_F + \frac{\sqrt{4\epsilon_{Si}qN_A(2\phi_F)}}{C_{ox}}$$

$$V_{Th,PMOS} = V_{FB} - |2\phi_F| - \frac{\sqrt{4\epsilon_{Si}qN_D(2\phi_F)}}{C_{ox}}$$

Se considera  $\phi_F = \phi_{bn}$  y  $V_{FB} = 0$  como valor de ejemplo neutro.

Sustitución:

$$V_{Th,NMOS} = 0 + 2(0.287 \text{ V}) + \frac{\sqrt{4(11.7)(8.85 \times 10^{-12} \text{ F/m})(1.6 \times 10^{-19} \text{ C})(1 \times 10^{21} \text{ m}^{-3})(2)(0.287 \text{ V})}}{5 \times 10^{-2} \text{ F/m}^2}$$

$$V_{Th,NMOS} \approx 0.578 \text{ V}$$

$$V_{Th,PMOS} = 0 - 2(|-0.305 \text{ V}|) - \frac{\sqrt{4(11.7)(8.85 \times 10^{-12} \text{ F/m})(1.6 \times 10^{-19} \text{ C})(2 \times 10^{21} \text{ m}^{-3})(2)(0.287 \text{ V})}}{5 \times 10^{-2} \text{ F/m}^2}$$

$$V_{Th,PMOS} \approx -0.615 \text{ V}$$

- **MOSFET:**

$$V_{Th,NMOS} = V_{FB} + 2\phi_{bn} + \frac{\sqrt{2\epsilon_{Si}qN_A(2\phi_{bn} + V_{SB})}}{C_{ox}}$$

$$V_{Th,PMOS} = V_{FB} - |2\phi_{bn}| - \frac{\sqrt{2\epsilon_{Si}qN_D(|2\phi_{bn}| - V_{SB})}}{C_{ox}}$$

Se considera  $\phi_F = \phi_{bn}$  y  $V_{FB} = 0$  como valor de ejemplo neutro.

Sustitución:

$$V_{Th,NMOS} = 0 + 2(0.424 \text{ V}) + \frac{\sqrt{2(11.7)(8.85 \times 10^{-12} \text{ F/m})(1.602 \times 10^{-19} \text{ C})(1 \times 10^{23} \text{ m}^{-3})(2 \cdot 0.424 \text{ V} + 0)}}{5 \times 10^{-2} \text{ F/m}^2}$$

$$V_{Th,NMOS} \approx 0.882 \text{ V}$$

$$V_{Th,PMOS} = 0 - |2(-0.416 \text{ V})| - \frac{\sqrt{2(11.7)(8.85 \times 10^{-12} \text{ F/m})(1.602 \times 10^{-19} \text{ C})(1.5 \times 10^{23} \text{ m}^{-3})(|2 \cdot -0.416 \text{ V}| - 0)}}{5 \times 10^{-2} \text{ F/m}^2}$$

$$V_{Th,PMOS} \approx -0.872 \text{ V}$$

## Corriente $I_{DS}$

- **NSFET:**

$$I_{DS,NMOS} = \frac{\mu_{eff} C_{ox}}{2} \left( \frac{W}{L} \right) (V_{GS} - V_{Th,NMOS})^2$$

$$I_{DS,PMOS} = \frac{\mu_{eff} C_{ox}}{2} \left( \frac{W}{L} \right) (V_{GS} - |V_{Th,PMOS}|)^2$$

Usando:

- $\mu_{eff,n} = 250 \text{ cm}^2/\text{V} \cdot \text{s} = 0.0250 \text{ m}^2/\text{V} \cdot \text{s}$
- $\mu_{eff,p} = 100 \text{ cm}^2/\text{V} \cdot \text{s} = 0.0100 \text{ m}^2/\text{V} \cdot \text{s}$

(Sze & Lee, 2012), (Tsividis & McAndrew, 2011), (Dasgupta & Hu, 2020).

$$I_{DS,NMOS} = \frac{0.0250 \text{ m}^2/\text{V} \cdot \text{s} \times (5 \times 10^{-2} \text{ F/m})}{2} \cdot 105.3 \cdot (0.65 \text{ V} - 0.578 \text{ V})^2$$

$$I_{DS,NMOS} \approx 341.17 \mu\text{A}$$

$$I_{DS,PMOS} = \frac{0.0100 \text{ m}^2/\text{V} \cdot \text{s} \times (5 \times 10^{-2} \text{ F/m})}{2} \cdot 105.3 \cdot (0.65 \text{ V} - |-0.615|)^2$$

$$I_{DS,PMOS} \approx -32.24 \mu\text{A}$$

Con dimensiones iguales ( $W/L$ ) del NMOS, el PMOS da una corriente de saturación aproximadamente 89.7% menor (Considerando el valor absoluto).

Para poder igual la corriente del NMOS en el PMOS, debemos modificar  $W/L$  del mismo, esto para poder hacer, por ejemplo, un inversor simétrico:

Así que:

$$I_{DS,NMOS} = 313.33 \mu A$$

$$I_{DS,PMOS} = -32.24 \mu A$$

$$\frac{W}{L_{PMOS}} = 105.3 \cdot \frac{313.33 \mu A}{|-32.24 \mu A|} \approx 105.3 \cdot 9.71 \approx 1023.37$$

Sustituyendo con la nueva  $W/L$ :

$$I_{DS,PMOS} = \frac{0.0100 \cdot (5 \times 10^{-2})}{2} \cdot 1023.37 \cdot (0.65 - |-0.615|)^2$$

$$I_{DS,PMOS} \approx -313.40 \mu A$$

Con estos ajustes igualamos  $I_{DS}$  tanto del PMOS como del NMOS, por lo que ahora podemos construir un inversor simétrico.

- **MOSFET:**

$$I_{DS,NMOS} = \mu C_{ox} \frac{W}{L} \frac{(V_{GS} - V_{Th})^2}{2} (1 + \lambda V_{DS})$$

$$I_{DS,PMOS} = -\mu C_{ox} \frac{W}{L} \frac{(V_{GS} - |V_{Th}|)^2}{2} (1 + \lambda V_{DS})$$

Usando condiciones óptimas:

- $\mu_{eff,n} = 550 \text{ cm}^2/\text{V} \cdot \text{s} = 0.0550 \text{ m}^2/\text{V} \cdot \text{s}$
- $\mu_{eff,p} = 250 \text{ cm}^2/\text{V} \cdot \text{s} = 0.0250 \text{ m}^2/\text{V} \cdot \text{s}$
- $\lambda \approx 0.10$  Típico en nodo CMOS 20 nm.

(Sze & Lee, 2012), (Tsividis & McAndrew, 2011), (Dasgupta & Hu, 2020).

$$I_{DS,nmos} = 0.0550 \cdot (5 \times 10^{-2}) \cdot 5 \cdot \frac{(0.90 V - 0.882 V)^2}{2} \cdot (1 + 0.10 \cdot 0.90 V)$$

$$I_{DS,nmos} \approx 2.427 \mu A$$

Considerando la misma  $W/L$  del NMOS

$$I_{DS,pmos} = -0.0250 \cdot (5 \times 10^{-2}) \cdot 5 \cdot \frac{(0.90 - |-0.872|)^2}{2} \cdot (1 + 0.10 \cdot 0.90)$$

$$I_{DS,pmos} = -2.67 \mu A$$

Los resultados muestran que el NSFET de 3 nm trabaja con mucha más eficiencia que el MOSFET de 20 nm. Sus umbrales más bajos y su estructura Gate-All-Around le dan un control del canal mucho más firme, así que necesita menos voltaje para activarse y puede mover mucha más corriente. Esto se refleja directamente en las cifras: el NMOS del NSFET supera los 300  $\mu$ A, mientras que el MOSFET planar apenas llega a unos cuantos microamperes. Y aunque el PMOS del NSFET arranca siendo más débil, basta ajustar W para que alcance prácticamente la misma fuerza que el NMOS, como se hace normalmente en diseños CMOS modernos.

En cambio, el MOSFET tradicional tiene un canal más dopado y umbrales más altos, lo que naturalmente frena la corriente y limita su rendimiento. Además, como aquí usamos un modelo clásico y no uno cuántico (que es el adecuado para tamaños tan pequeños), las corrientes del NSFET deben verse como una aproximación para entender la tendencia general. Aun así, la conclusión es clara: la geometría GAA ofrece un control del canal mucho mejor y permite operar a voltajes más bajos sin perder potencia, mientras que el MOSFET planar se queda atrás por las limitaciones propias de su estructura.

## 2.3 Parámetros de simulación

Para las siguientes simulaciones se realizó sobre el mismo layout ya que Microwind permite simular el propio diseño, en todas las compuertas se considera VDD=0.65V sobre NSFET de 3nm.

### Compuerta AND de 2 entradas

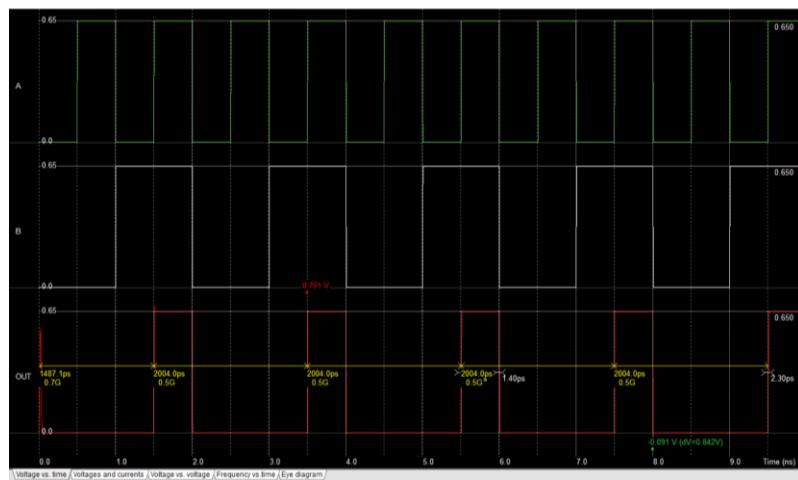


Ilustración 31 Simulación AND de 2 entradas del layout (0.65 V, 0.5 GHz), (Microwind).

## Compuerta AND de 3 entradas

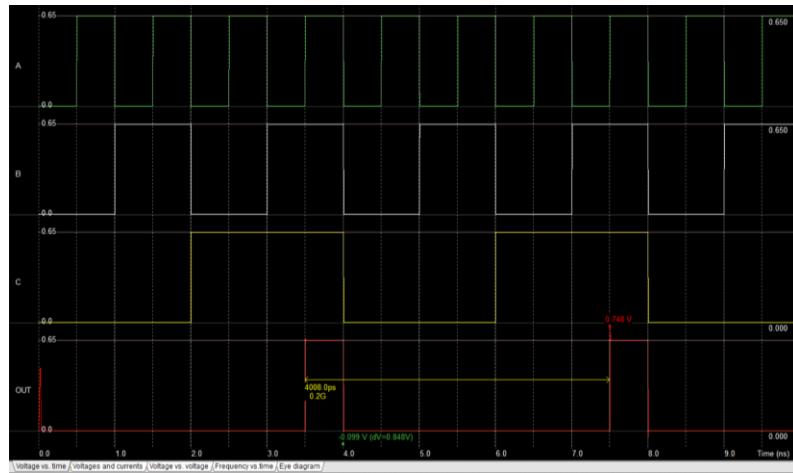


Ilustración 32 Simulación AND de 3 entradas del layout (0.65V, 200Hz), (Microwind).

## Compuerta AND de 4 entradas

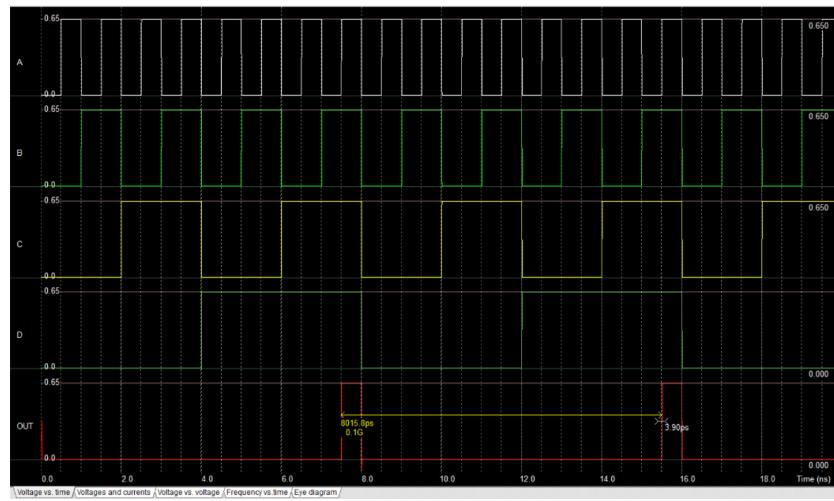


Ilustración 33 Simulación AND de 4 entradas del layout (0.65V, 100MHz), (Microwind).

## Compuerta AND de 5 entradas

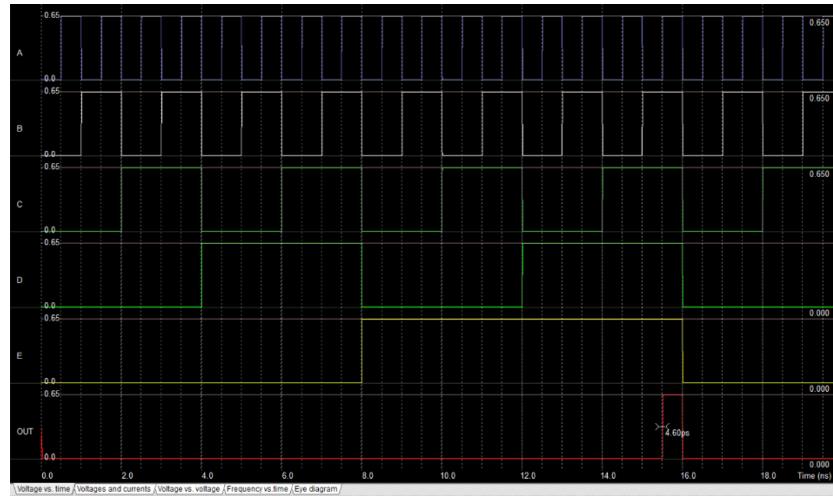


Ilustración 34 Simulación AND de 5 entradas del layout (0.65V), (Microwind).

## Compuerta NAND de 2 entradas

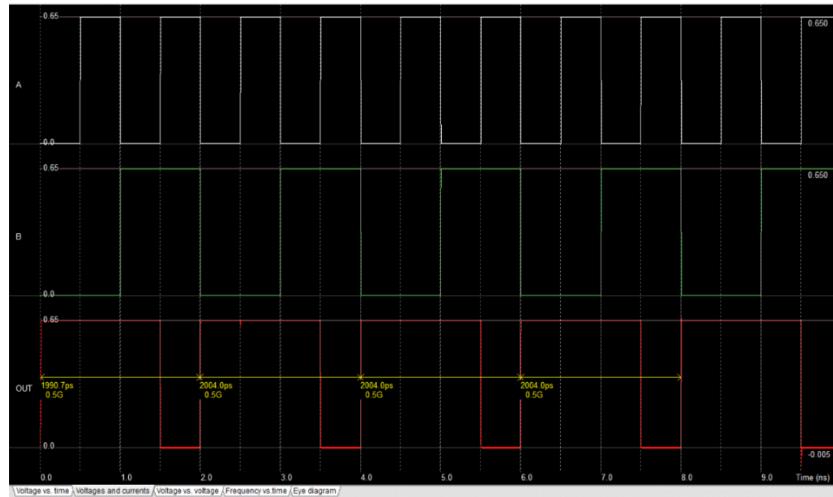


Ilustración 35 Simulación NAND de 2 entradas del layout (0.65V, 0.5 GHz), (Microwind).

## Compuerta NAND de 3 entradas

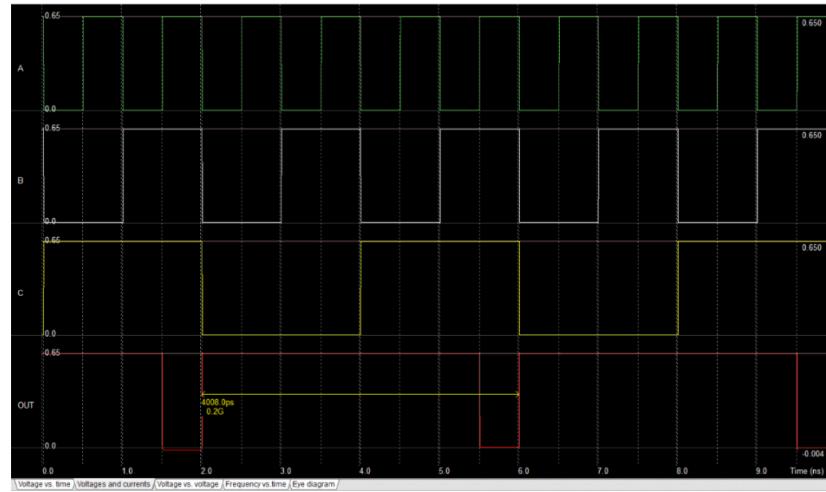


Ilustración 36 Simulación NAND de 3 entradas del layout (0.65V, 200MHz), (Microwind).

## Compuerta NAND de 4 entradas

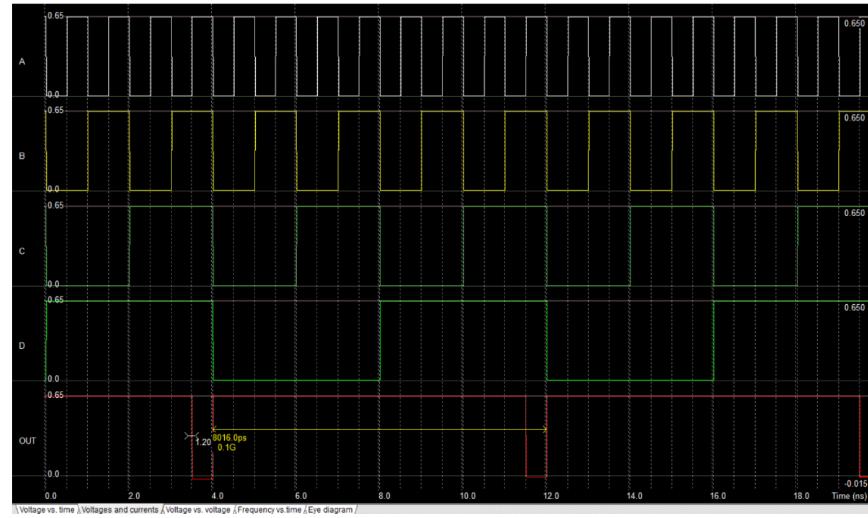


Ilustración 37 Simulación NAND de 4 entradas del layout (0.65V, 100MHz). (Microwind).

## Compuerta NAND de 5 entradas

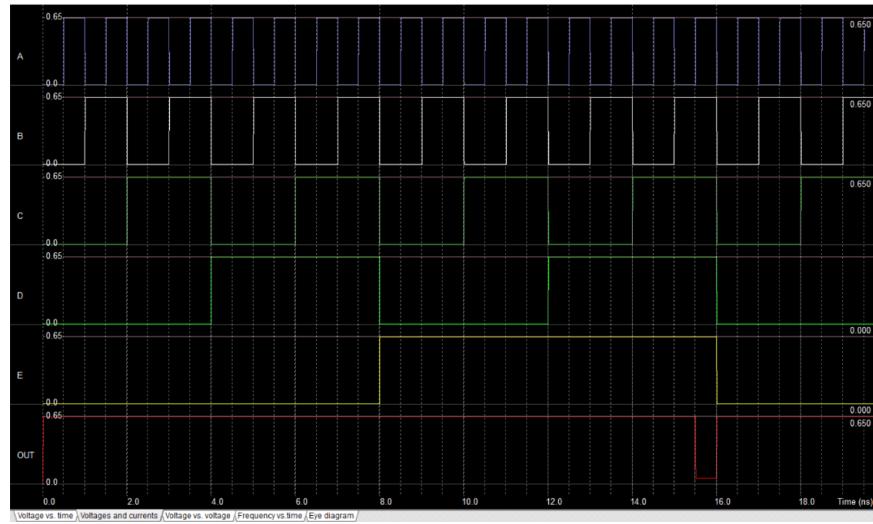


Ilustración 38 Simulación NAND de 5 entradas del layout (0.65V), (Microwind).

## Compuerta OR de 2 entradas

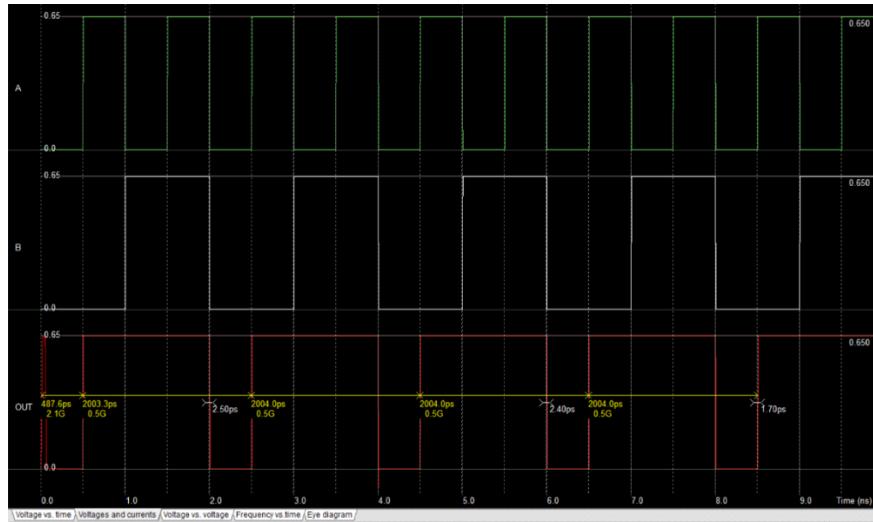


Ilustración 39 Simulación OR de 2 entradas del layout (0.65V, 0.5 GHz), (Microwind).

## Compuerta NOR de 2 entradas

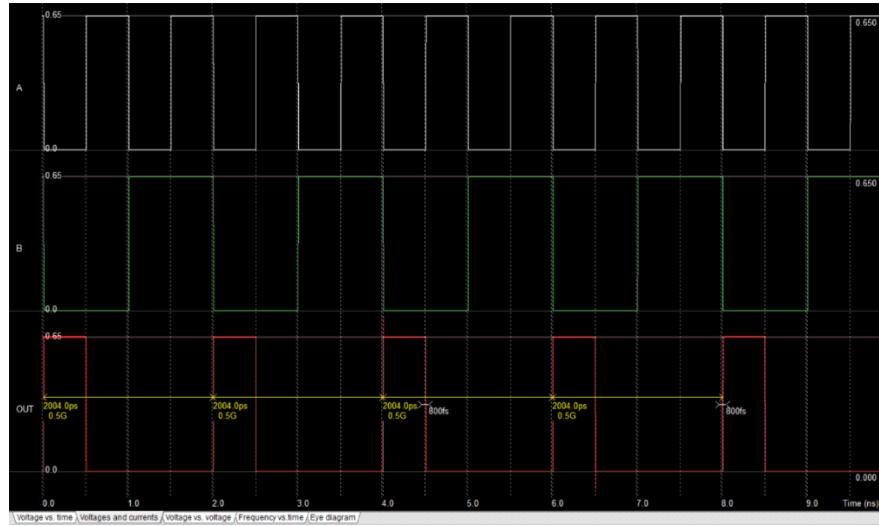


Ilustración 40 Simulación NOR de 2 entradas del layout (0.65V, 0.5 GHz), (Microwind).

## Compuerta NOR de 3 entradas

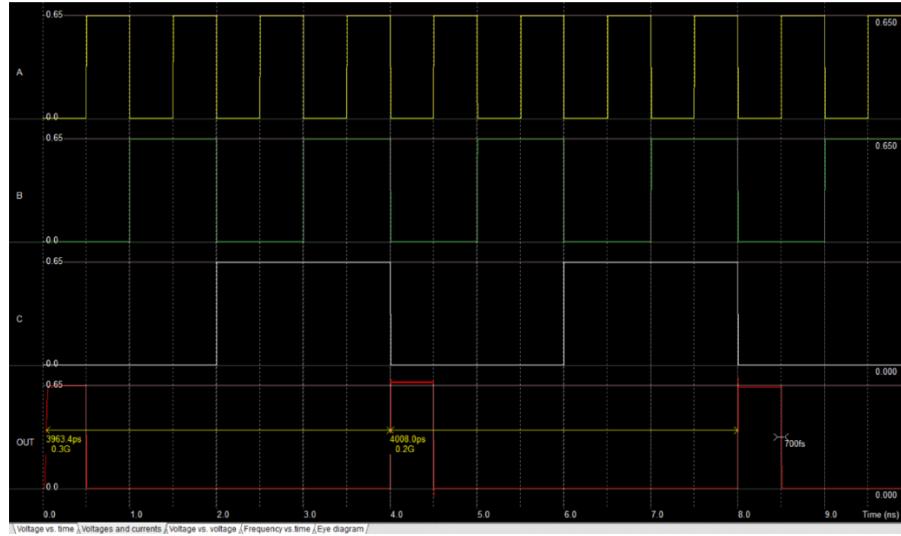


Ilustración 41 Simulación NOR de 3 entradas del layout (0.65V, 200 MHz), (Microwind).

## Compuerta NOR de 4 entradas



Ilustración 42 Simulación NOR de 4 entradas del layout (0.65V, 100 MHz), (Microwind).

## Compuerta XOR de 2 entradas

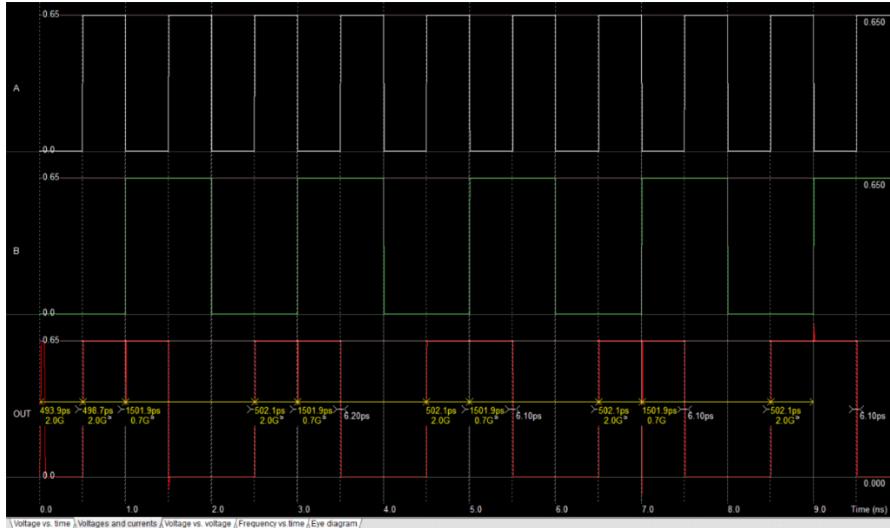


Ilustración 43 Simulación XOR de 2 entradas del layout (0.65V, 0.7 - 2 GHz), (Microwind).

## Compuerta NOT

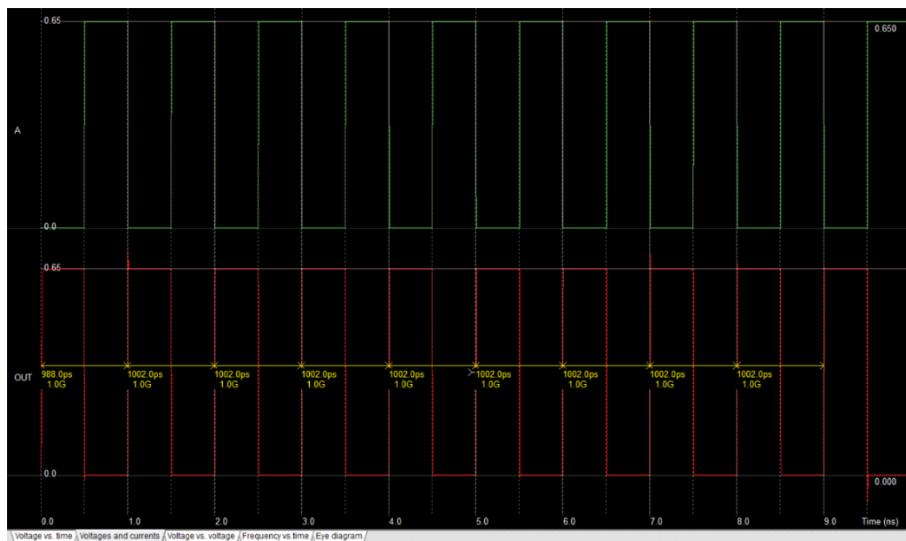


Ilustración 44 Simulación NOT del layout (0.65V, 1 GHz), (Microwind).

## ALU 74181 (4 bits)

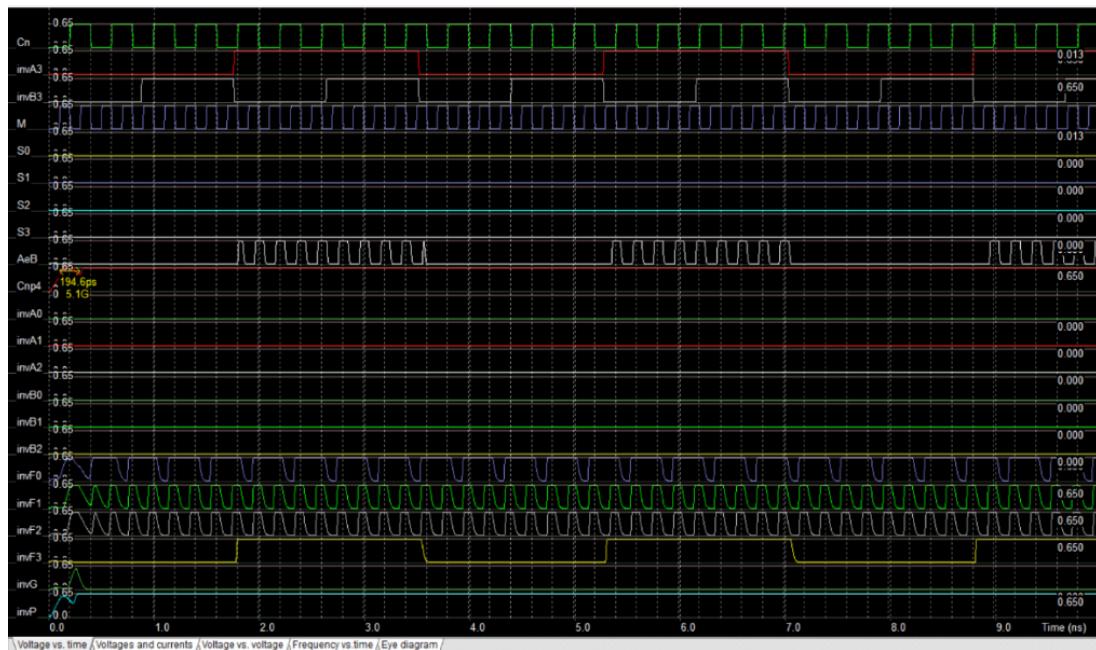


Ilustración 45 Simulación ALU de 4 bits del layout (0.65V, 5.11 GHz), (Microwind).

## ALU 74181 (8 bits)



Ilustración 46 Simulación ALU 8 bits del layout (0.65V), (Microwind).

Las simulaciones realizadas para cada una de las compuertas mostraron un comportamiento consistente: en ningún caso se observó caída de voltaje en los niveles lógicos de salida. Este resultado es coherente con las características del dispositivo NSFET, cuyo control electrostático mejorado y su arquitectura tipo nanosheet permiten mantener un canal estable y reducir las pérdidas internas. Gracias a ello, las compuertas conservan niveles lógicos bien definidos incluso en condiciones de conmutación continua, lo que refuerza la solidez del diseño implementado.

Por otra parte, se realizaron simulaciones funcionales de la ALU basada en la arquitectura del CI 74181 utilizando Microwind. En este entorno fue posible validar de manera completa la versión de 8 bits; sin embargo, al intentar escalar el diseño a 32 bits, la complejidad del circuito y el número de interconexiones superaron la capacidad de procesamiento del simulador. Esta limitación responde únicamente al tamaño del diseño y no afecta la fidelidad física de las simulaciones individuales ni la validez de los resultados obtenidos a nivel de compuertas.

Aun con estas restricciones, el layout de la ALU de 32 bits pudo realizarse, integrando manualmente cada bloque y asegurando la continuidad de las interconexiones, la coherencia y la correcta correspondencia con la arquitectura funcional prevista.

Este paso confirmó que el diseño es escalable y que su estructura se mantiene ordenada y consistente aun en una implementación de mayor magnitud. En conjunto, los resultados permiten concluir que el uso de tecnología NSFET garantiza niveles lógicos estables sin caída de voltaje, que la versión de 8 bits funciona correctamente en simulación completa, y que el diseño de 32 bits cuenta con una base estructural confiable para su validación futura en plataformas más robustas o dentro de flujos post-layout especializados.

# **Capítulo 3**

## **Layout y funcionamiento**

## 3.1 Layout

En esta sección se muestran los Layouts de las compuertas lógicas a base de NSFETS con tecnología (o nodo) de 3 nm utilizadas para el diseño de ALU de 32 bits en vista superior y 3D, además de su tabla de verdad.

### Compuerta AND de 2 entradas.

A	B	OUT
0	0	0
0	1	0
1	0	0
1	1	1

Tabla 3 Correspondencia lógica de AND de 2 entradas.

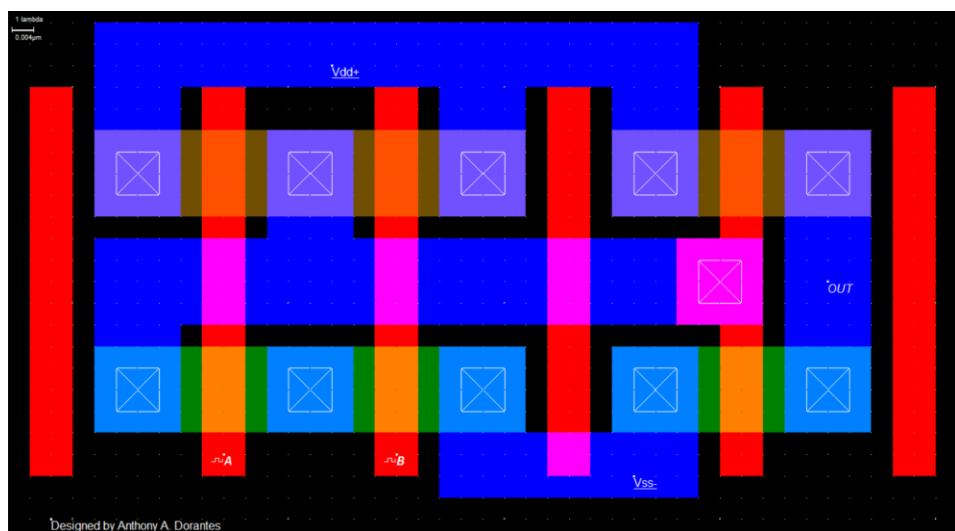


Ilustración 47 Compuerta AND de 2 entradas NSFET 2D (3nm).

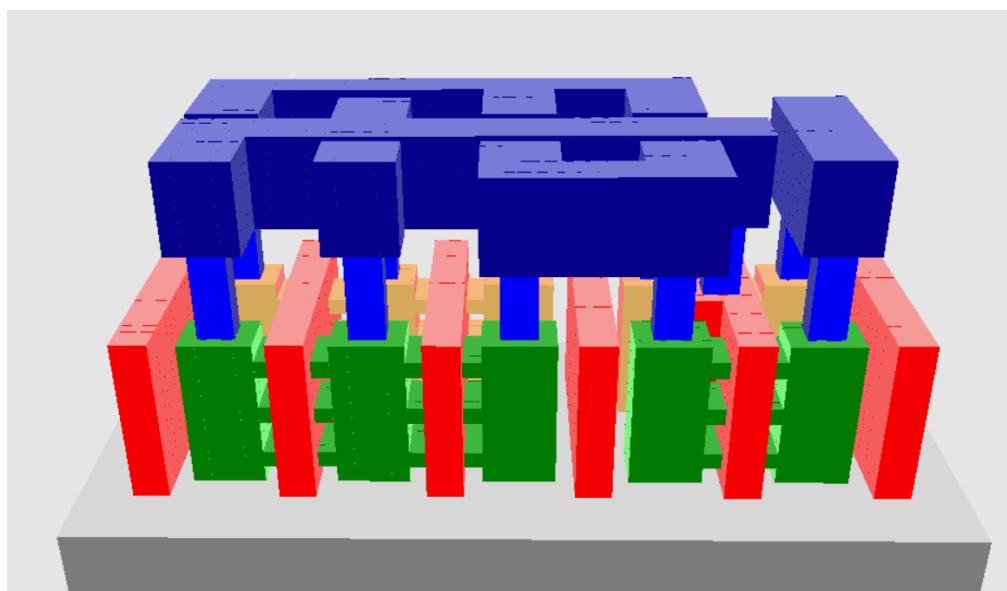


Ilustración 48 Modelo 3D (3-Nanosheet NSFET 3 nm).

## Compuerta AND de 3 entradas

A	B	C	OUT
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	0
1	1	0	0
1	1	1	1

Tabla 4 Correspondencia lógica de AND de 3 entradas.

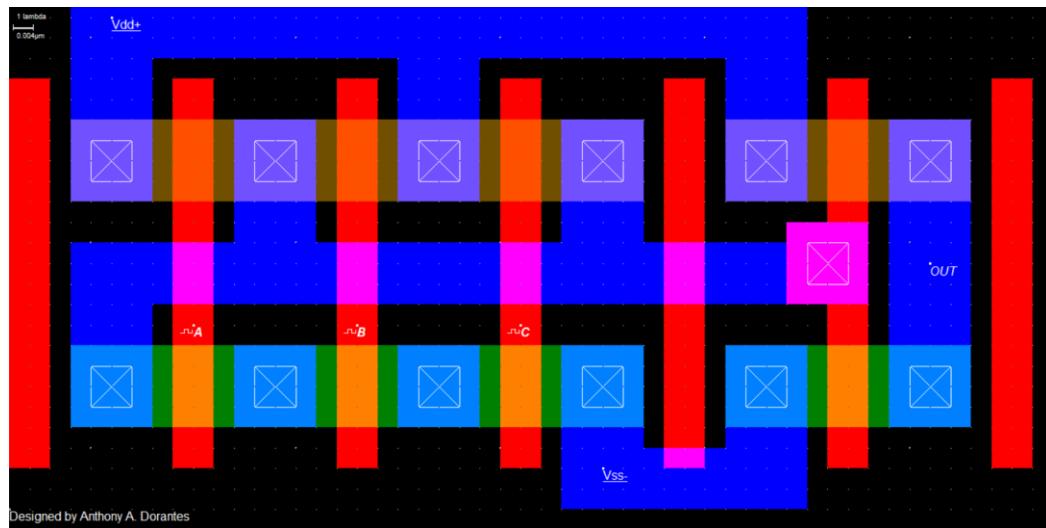
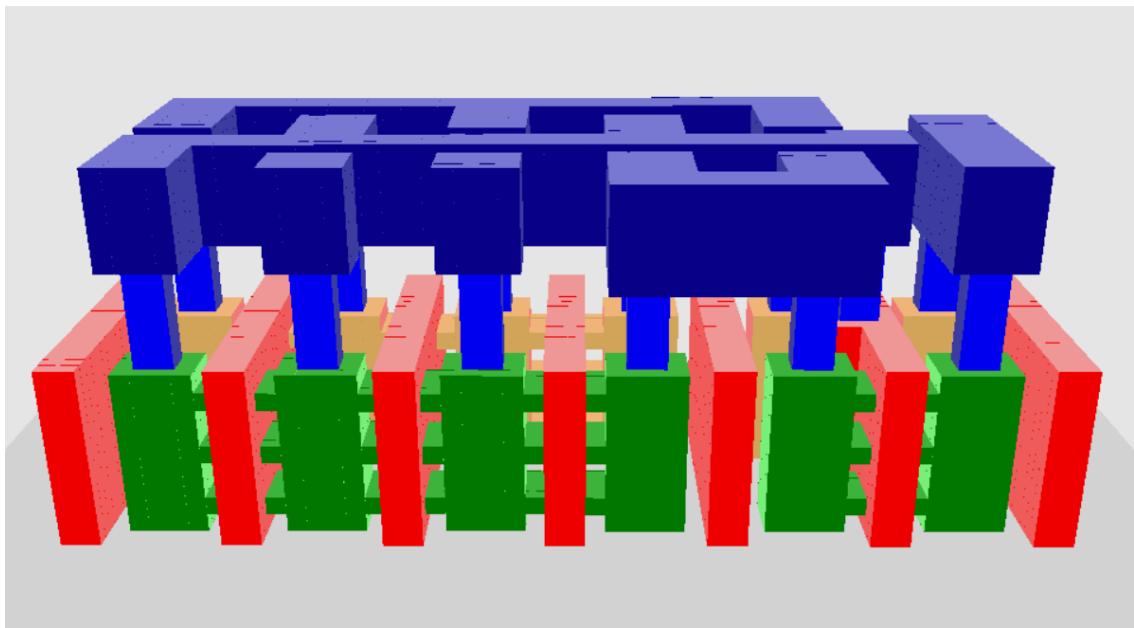


Ilustración 49 Compuerta AND de 3 entradas NSFET 2D (3nm)



*Ilustración 50 Modelo 3D (3-Nanosheet NSFET 3 nm)*

## Compuerta AND de 4 entradas

A B C D OUT

<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>
<b>0</b>	<b>0</b>	<b>0</b>	<b>1</b>	<b>0</b>
<b>0</b>	<b>0</b>	<b>1</b>	<b>0</b>	<b>0</b>
<b>0</b>	<b>0</b>	<b>1</b>	<b>1</b>	<b>0</b>
<b>0</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>0</b>
<b>0</b>	<b>1</b>	<b>0</b>	<b>1</b>	<b>0</b>
<b>0</b>	<b>1</b>	<b>1</b>	<b>0</b>	<b>0</b>
<b>0</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>0</b>
<b>1</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>
<b>1</b>	<b>0</b>	<b>0</b>	<b>1</b>	<b>0</b>
<b>1</b>	<b>0</b>	<b>1</b>	<b>0</b>	<b>0</b>
<b>1</b>	<b>0</b>	<b>1</b>	<b>1</b>	<b>0</b>
<b>1</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>0</b>
<b>1</b>	<b>1</b>	<b>0</b>	<b>1</b>	<b>0</b>
<b>1</b>	<b>1</b>	<b>1</b>	<b>0</b>	<b>0</b>
<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>

Tabla 5 Correspondencia lógica de AND de 4 entradas.

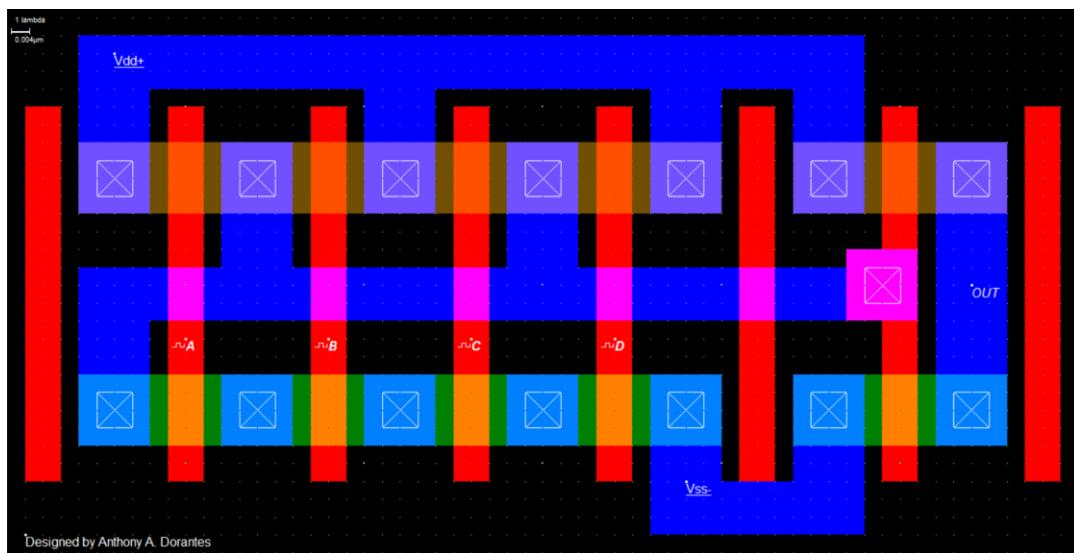
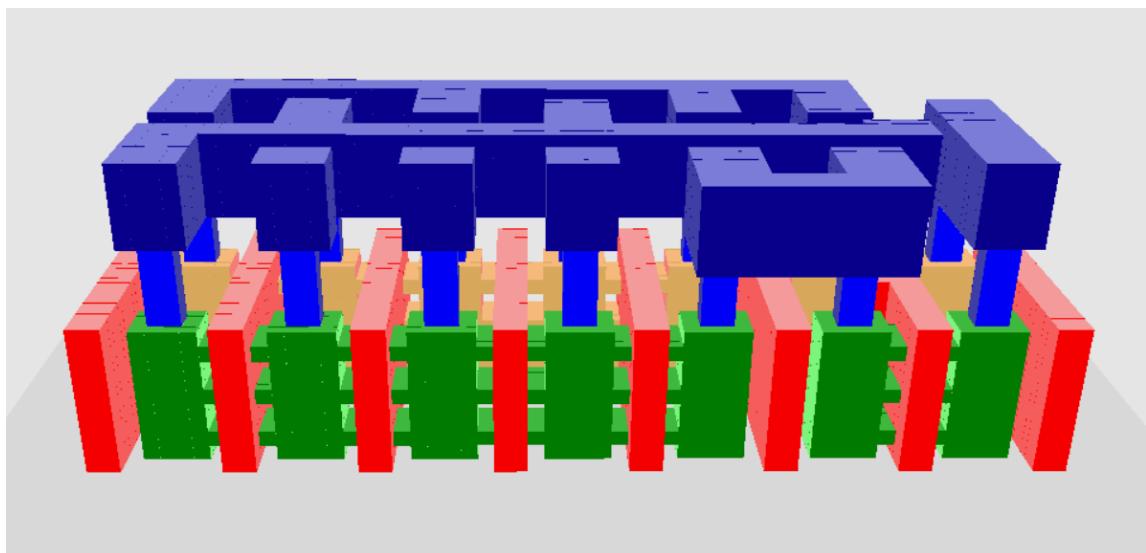


Ilustración 51 Compuerta AND de 4 entradas NSFET 2D (3nm).



*Ilustración 52 Modelo 3D (3-Nanosheet NSFET 3 nm).*

## Compuerta AND de 5 entradas

A	B	C	D	E	OUT
0	0	0	0	0	0
0	0	0	0	1	0
0	0	0	1	0	0
0	0	0	1	1	0
0	0	1	0	0	0
0	0	1	0	1	0
0	0	1	1	0	0
0	0	1	1	1	0
0	1	0	0	0	0
0	1	0	0	1	0
0	1	0	1	0	0
0	1	0	1	1	0
0	1	1	0	0	0
0	1	1	0	1	0
0	1	1	1	1	0
1	0	0	0	0	0
1	0	0	0	1	0
1	0	0	1	0	0
1	0	0	1	1	0
1	0	1	0	0	0
1	0	1	0	1	0
1	0	1	1	0	0
1	0	1	1	1	0
1	1	0	0	0	0
1	1	0	0	1	0
1	1	0	1	0	0
1	1	1	0	0	0
1	1	1	0	1	0
1	1	1	1	0	0
1	1	1	1	1	1

Tabla 6 Correspondencia lógica de AND de 5 entradas.

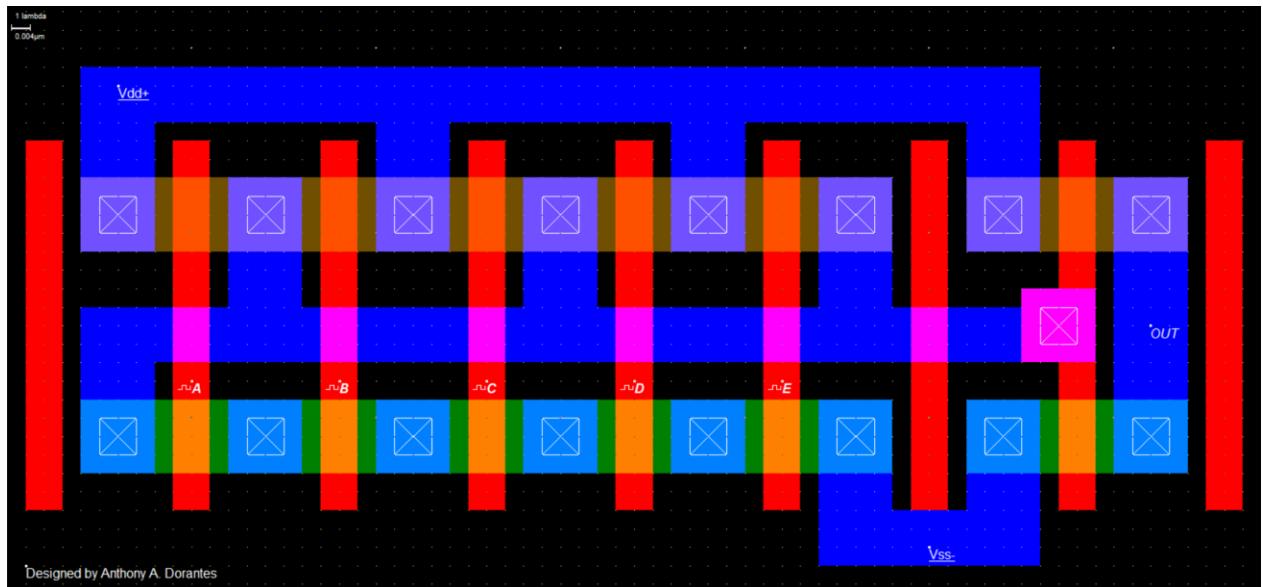


Ilustración 53 Compuerta AND de 5 entradas NSFET 2D (3nm).

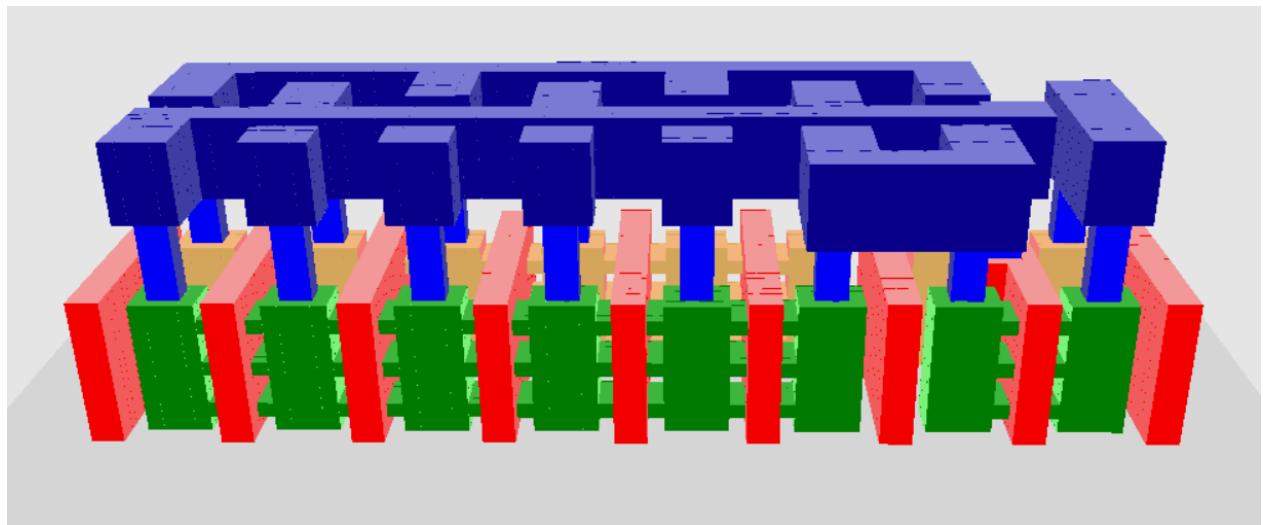


Ilustración 54 Modelo 3D (3-Nanosheet NSFET 3 nm).

## Compuerta NAND de 2 entradas.

A	B	OUT
0	0	1
0	1	1
1	0	1
1	1	0

Tabla 7 Correspondencia lógica de NAND de 2 entradas.

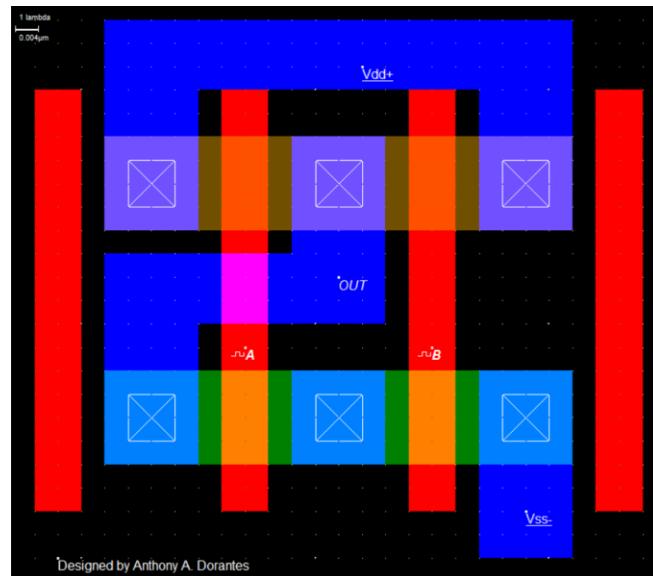
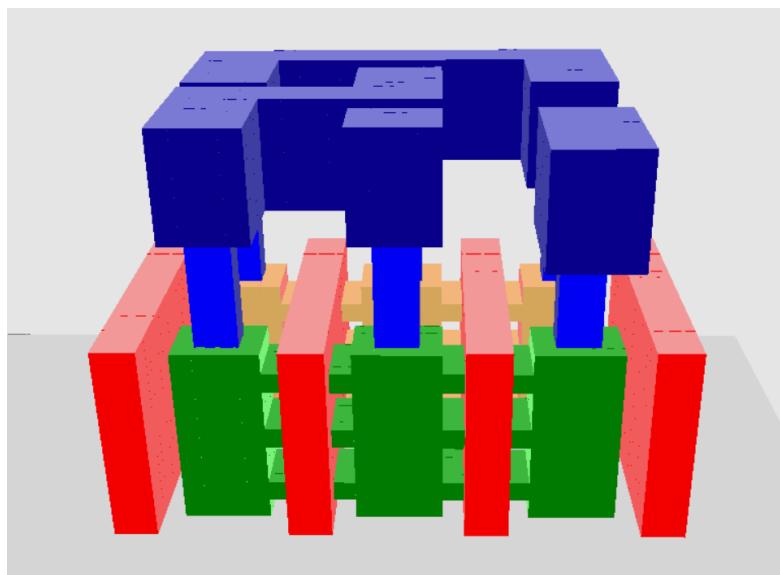


Ilustración 55 Compuerta NAND de 2 entradas NSFET 2D (3nm).



*Ilustración 56 Modelo 3D (3-Nanosheet NSFET 3 nm).*

## Compuerta NAND de 3 entradas.

A	B	C	OUT
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

Tabla 8 Correspondencia lógica de NAND de 3 entradas.

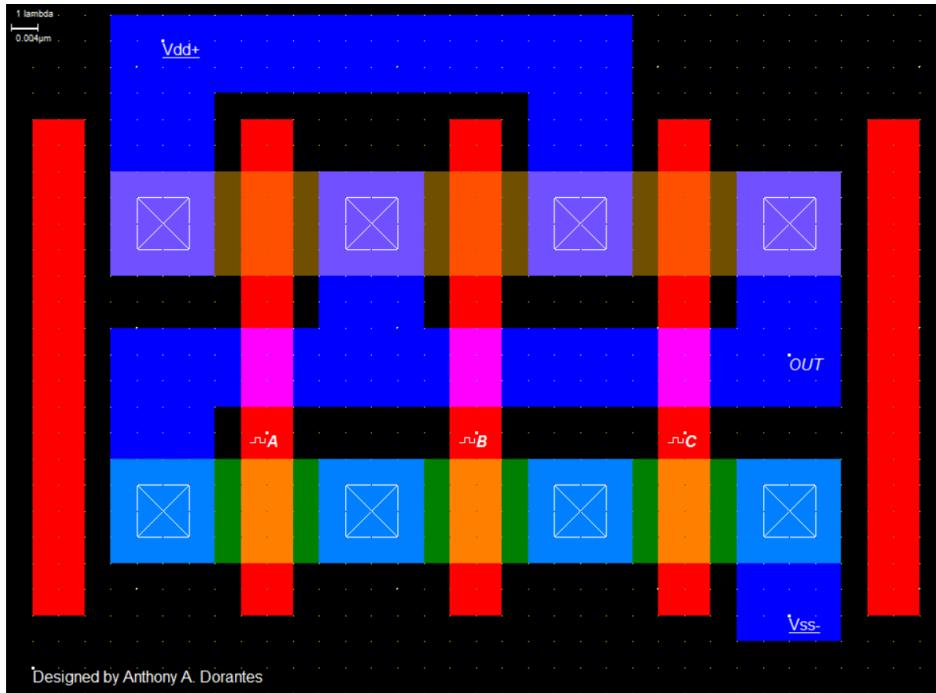
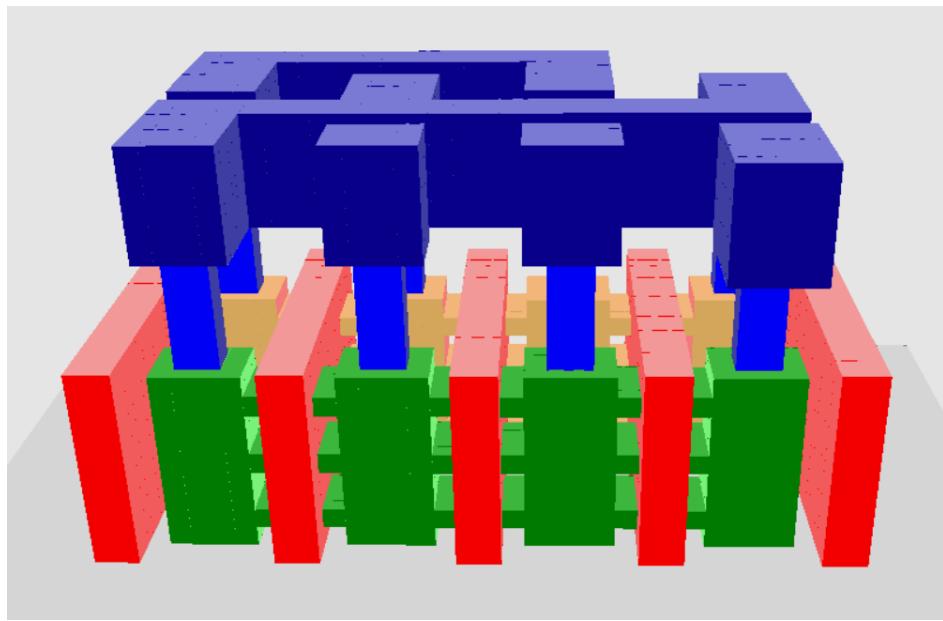


Ilustración 57 Compuerta NAND de 3 entradas NSFET 2D (3nm).



*Ilustración 58 Modelo 3D (3-Nanosheet NSFET 3 nm).*

## Compuerta NAND de 4 entradas.

A	B	C	D	OUT
0	0	0	0	1
0	0	0	1	1
0	0	1	0	1
0	0	1	1	1
0	1	0	0	1
0	1	0	1	1
0	1	1	0	1
0	1	1	1	1
1	0	0	0	1
1	0	0	1	1
1	0	1	0	1
1	0	1	1	1
1	1	0	0	1
1	1	0	1	1
1	1	1	0	1
1	1	1	1	0

Tabla 9 Correspondencia lógica de NAND de 4 entradas.

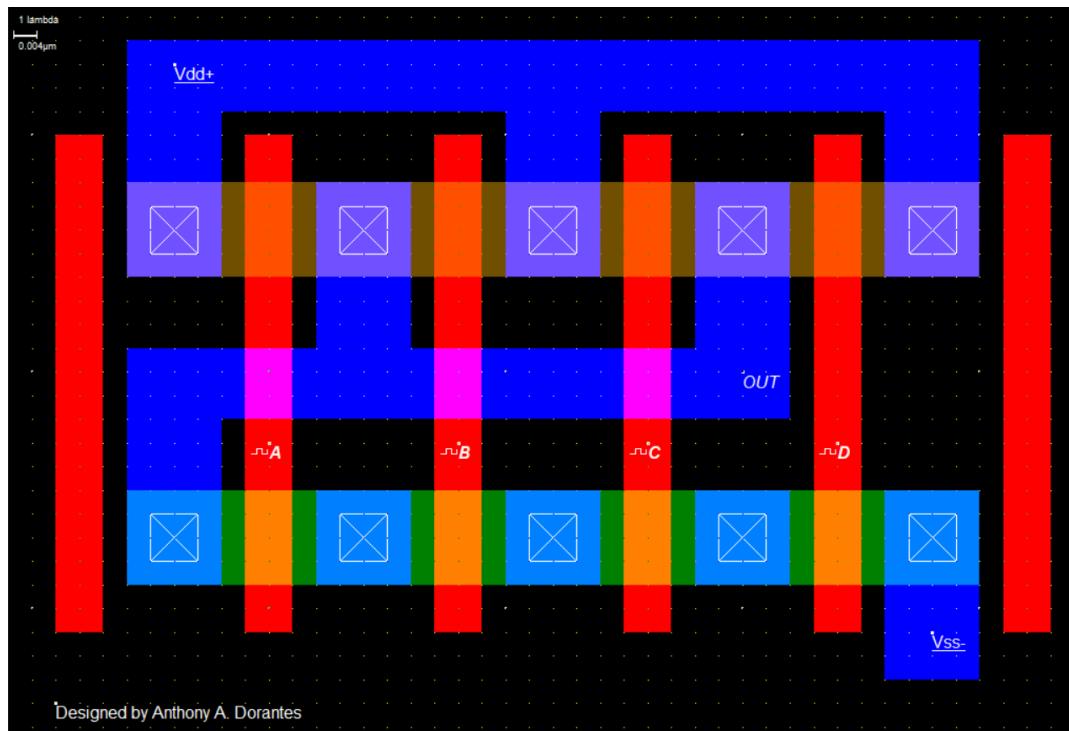


Ilustración 59 Compuerta NAND de 4 entradas NSFET 2D (3nm)

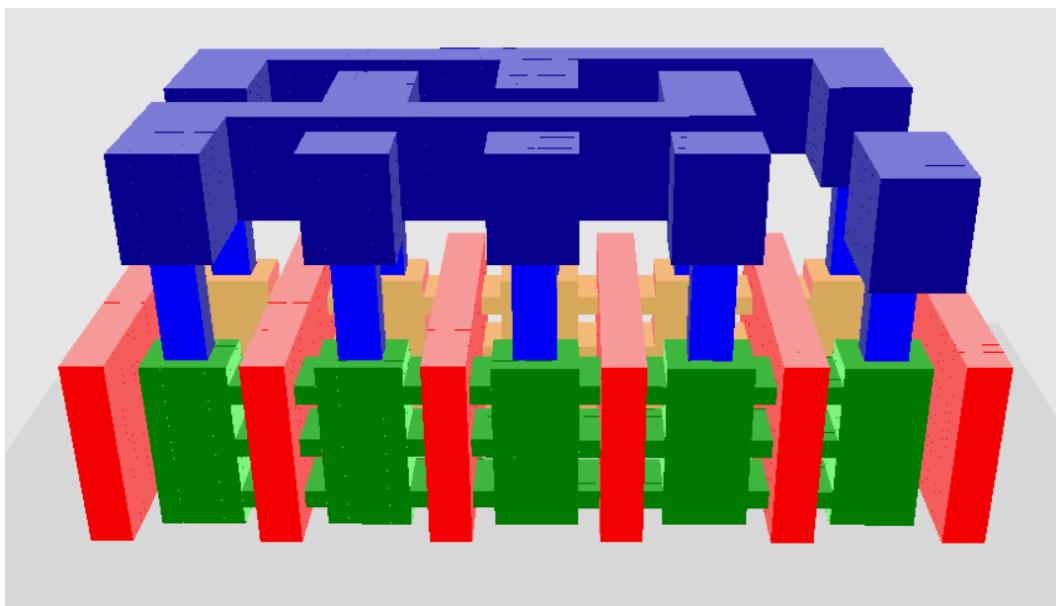


Ilustración 60 Modelo 3D (3-Nanosheet NSFET 3 nm)

## Compuerta NAND de 5 entradas.

A	B	C	D	E	OUT
0	0	0	0	0	1
0	0	0	0	1	1
0	0	0	1	0	1
0	0	0	1	1	1
0	0	1	0	0	1
0	0	1	0	1	1
0	0	1	1	0	1
0	0	1	1	1	1
0	1	0	0	0	1
0	1	0	0	1	1
0	1	0	1	0	1
0	1	0	1	1	1
0	1	1	0	0	1
0	1	1	0	1	1
0	1	1	1	0	1
0	1	1	1	1	1
1	0	0	0	0	1
1	0	0	0	1	1
1	0	0	1	0	1
1	0	0	1	1	1
1	0	1	0	0	1
1	0	1	1	1	1
1	1	0	0	0	1
1	1	0	0	1	1
1	1	0	1	0	1
1	1	0	1	1	1
1	1	1	0	0	1
1	1	1	0	1	1
1	1	1	1	0	1
1	1	1	1	1	0

Tabla 10 Correspondencia lógica de NAND de 5 entradas.

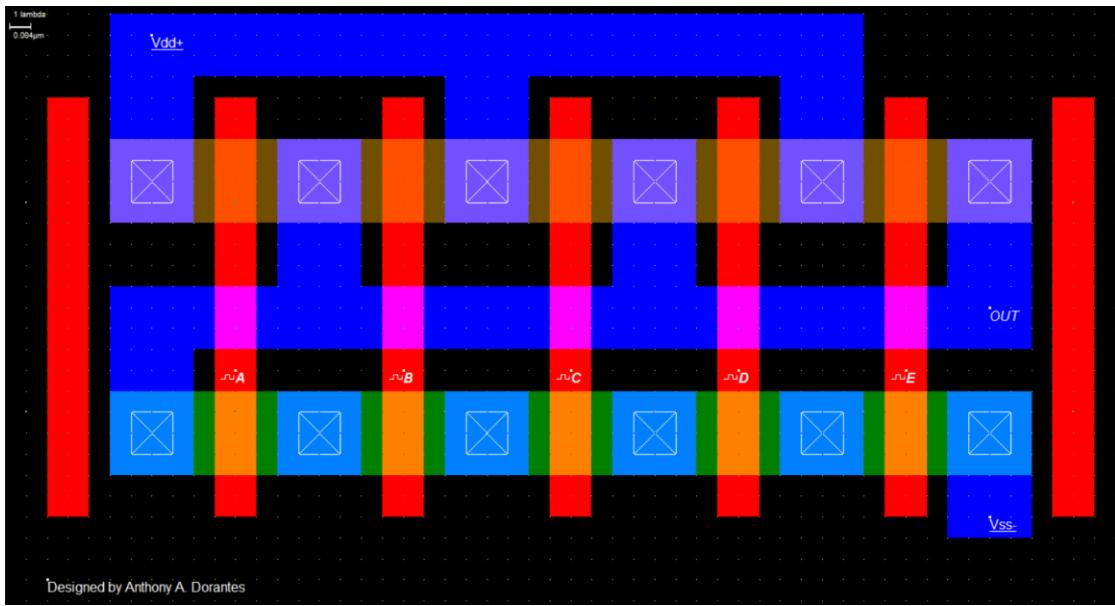


Ilustración 61 Compuerta NAND de 5 entradas NSFET 2D (3nm).

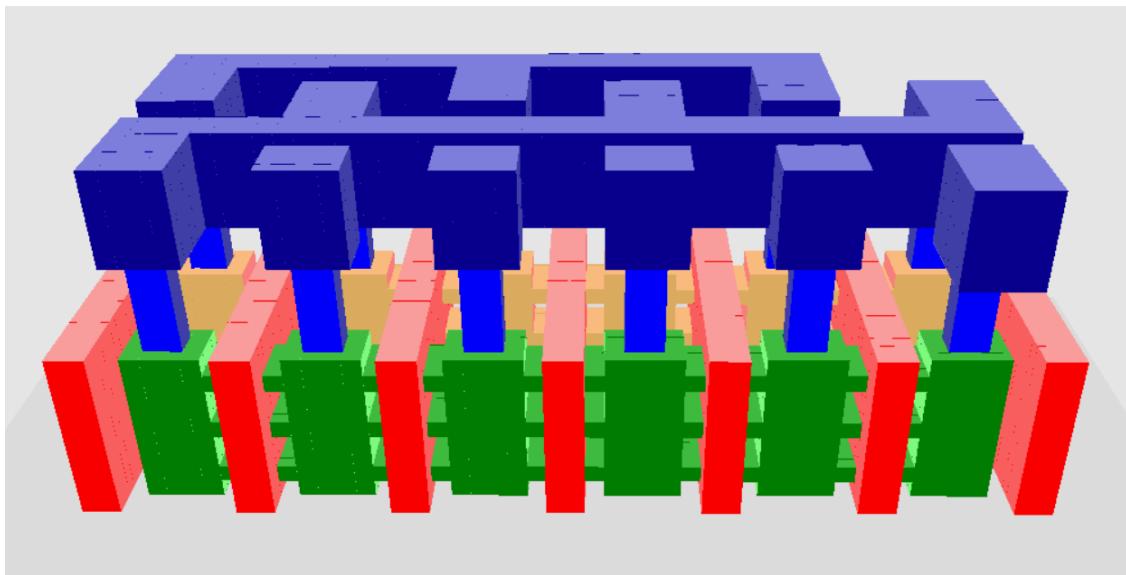


Ilustración 62 Modelo 3D (3-Nanosheet NSFET 3 nm).

## Compuerta OR de 2 entradas.

A	B	OUT
0	0	0
0	1	1
1	0	1
1	1	1

Tabla 11 Correspondencia lógica de OR de 2 entradas.

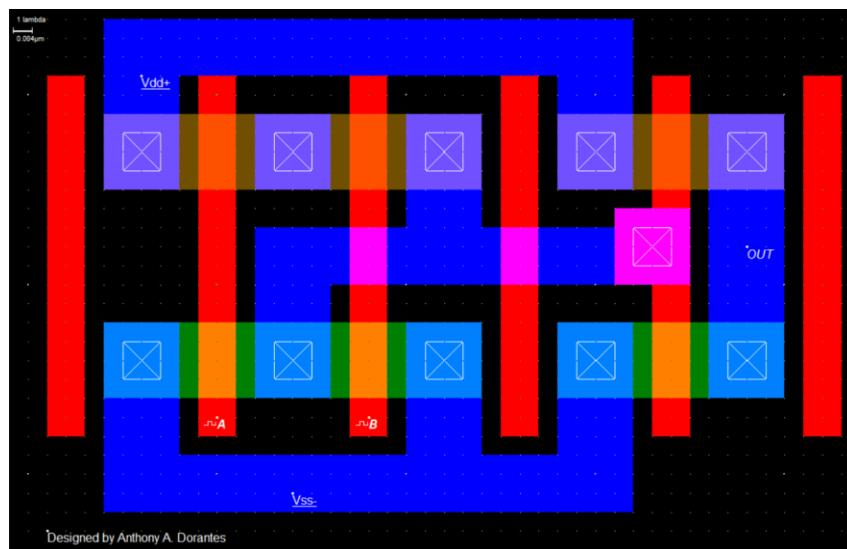


Ilustración 63 Compuerta OR de 2 entradas NSFET 2D (3nm).

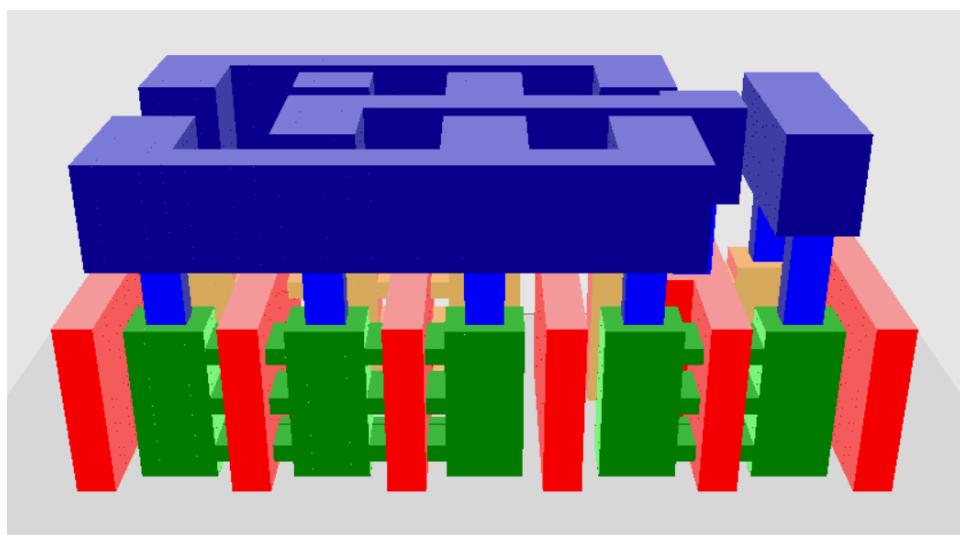


Ilustración 64 Modelo 3D (3-Nanosheet NSFET 3 nm).

## Compuerta NOR de 2 entradas.

A	B	OUT
0	0	1
0	1	0
1	0	0
1	1	0

Tabla 12 Correspondencia lógica de NOR de 2 entradas.

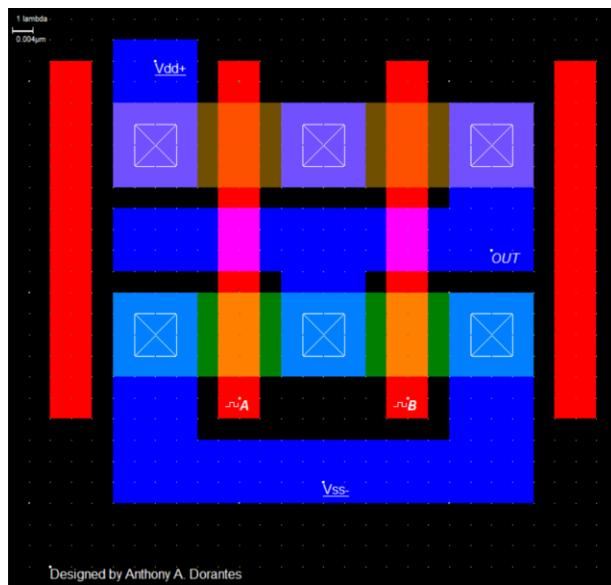


Ilustración 65 Compuerta NOR de 2 entradas NSFET 2D (3nm)

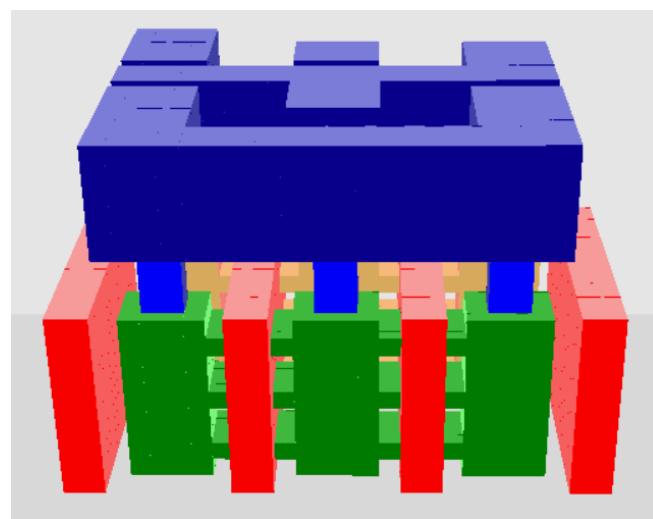


Ilustración 66 Modelo 3D (3-Nanosheet NSFET 3 nm)

## Compuerta NOR de 3 entradas.

A	B	C	OUT
0	0	0	1
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	0
1	1	0	0
1	1	1	0

Tabla 13 Correspondencia lógica de NOR de 3 entradas.

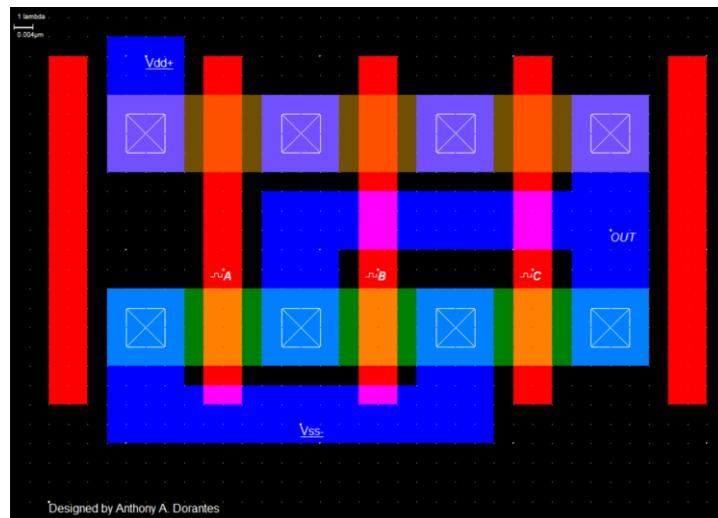


Ilustración 67 Compuerta NOR de 3 entradas NSFET 2D (3nm).

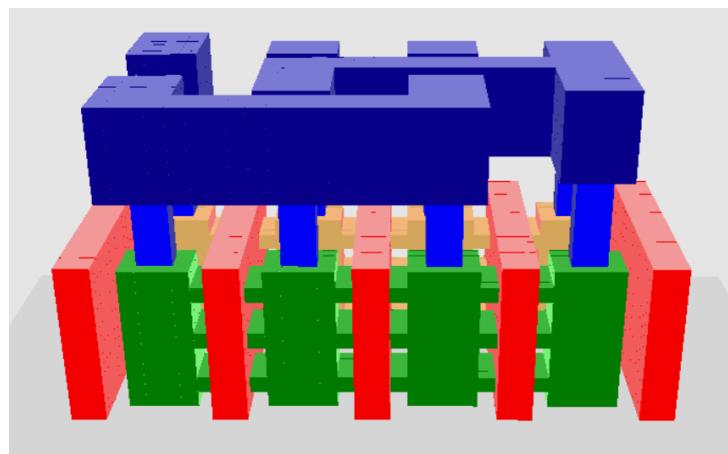


Ilustración 68 Modelo 3D (3-Nanosheet NSFET 3 nm).

### 3.1.1 Compuerta NOR de 4 entradas.

A	B	C	D	OUT
0	0	0	0	1
0	0	0	1	0
0	0	1	0	0
0	0	1	1	0
0	1	0	0	0
0	1	0	1	0
0	1	1	0	0
0	1	1	1	0
1	0	0	0	0
1	0	0	1	0
1	0	1	0	0
1	0	1	1	0
1	1	0	0	0
1	1	0	1	0
1	1	1	0	0
1	1	1	1	0

Tabla 14 Correspondencia lógica de NOR de 4 entradas.

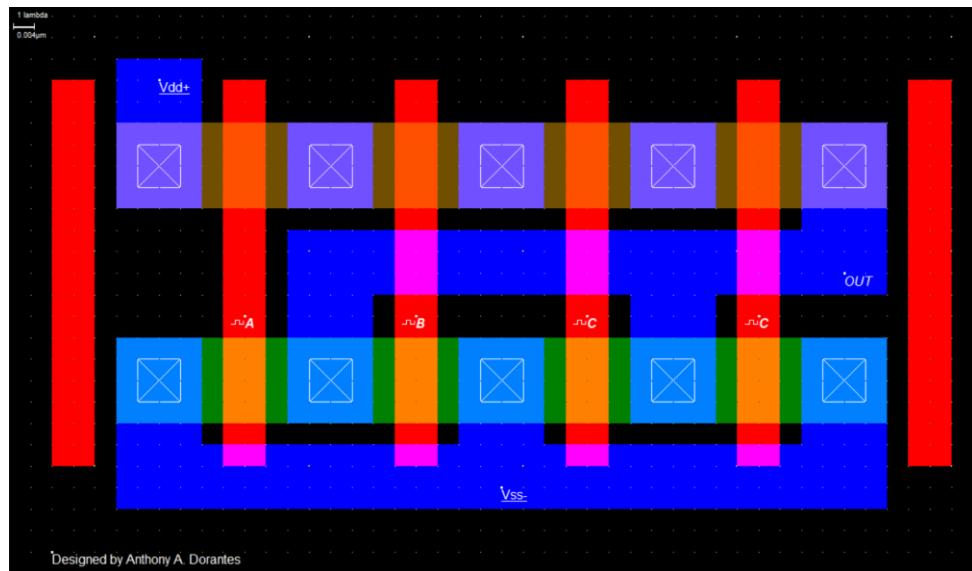


Ilustración 69 Compuerta NOR de 4 entradas NSFET 2D (3nm).

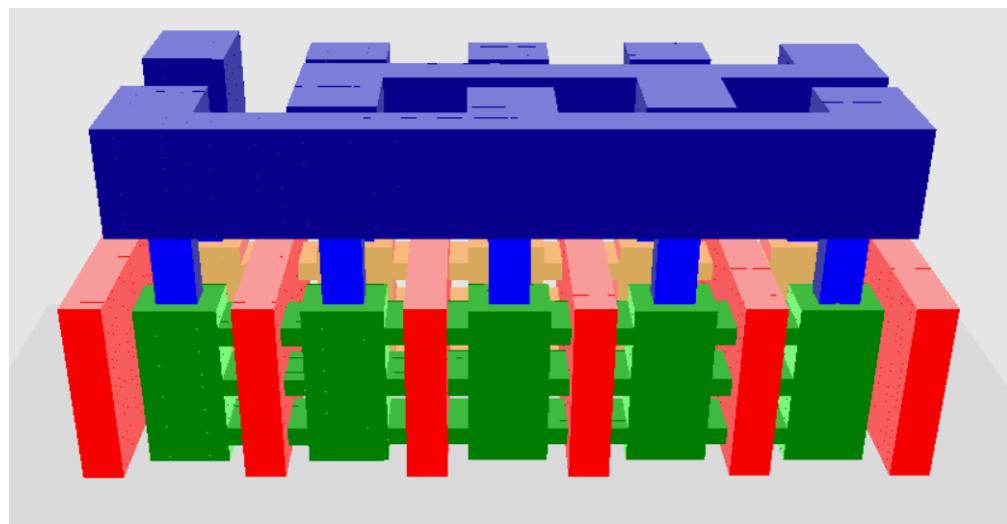


Ilustración 70 Modelo 3D (3-Nanosheet NSFET 3 nm).

## Compuerta XOR de 2 entradas.

A	B	OUT
0	0	0
0	1	1
1	0	1
1	1	0

Tabla 15 Correspondencia lógica de XOR de 2 entradas.

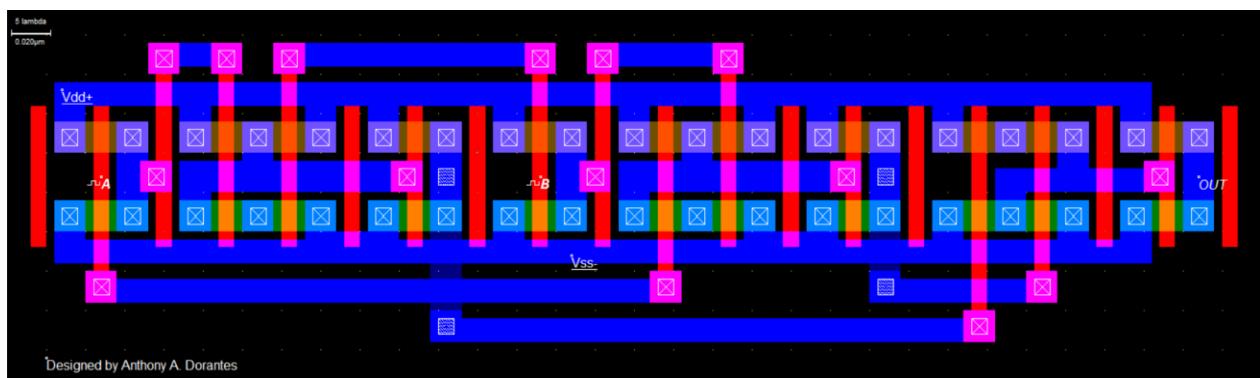


Ilustración 71 Compuerta XOR de 2 entradas NSFET 2D (3nm).

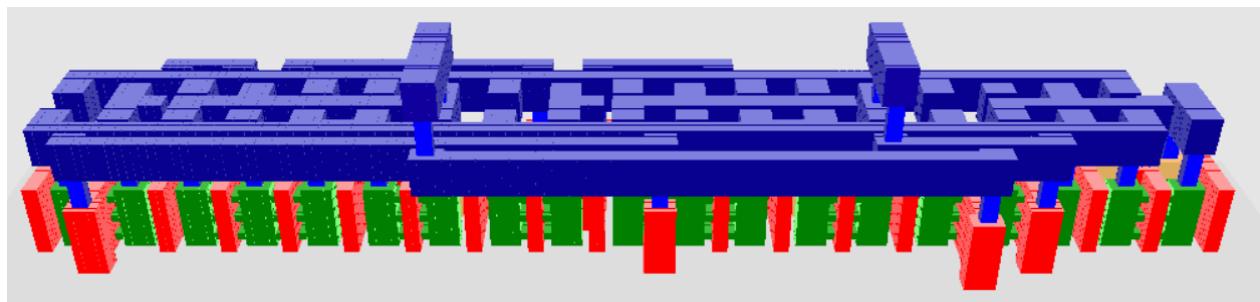


Ilustración 72 Modelo 3D (3-Nanosheet NSFET 3 nm).

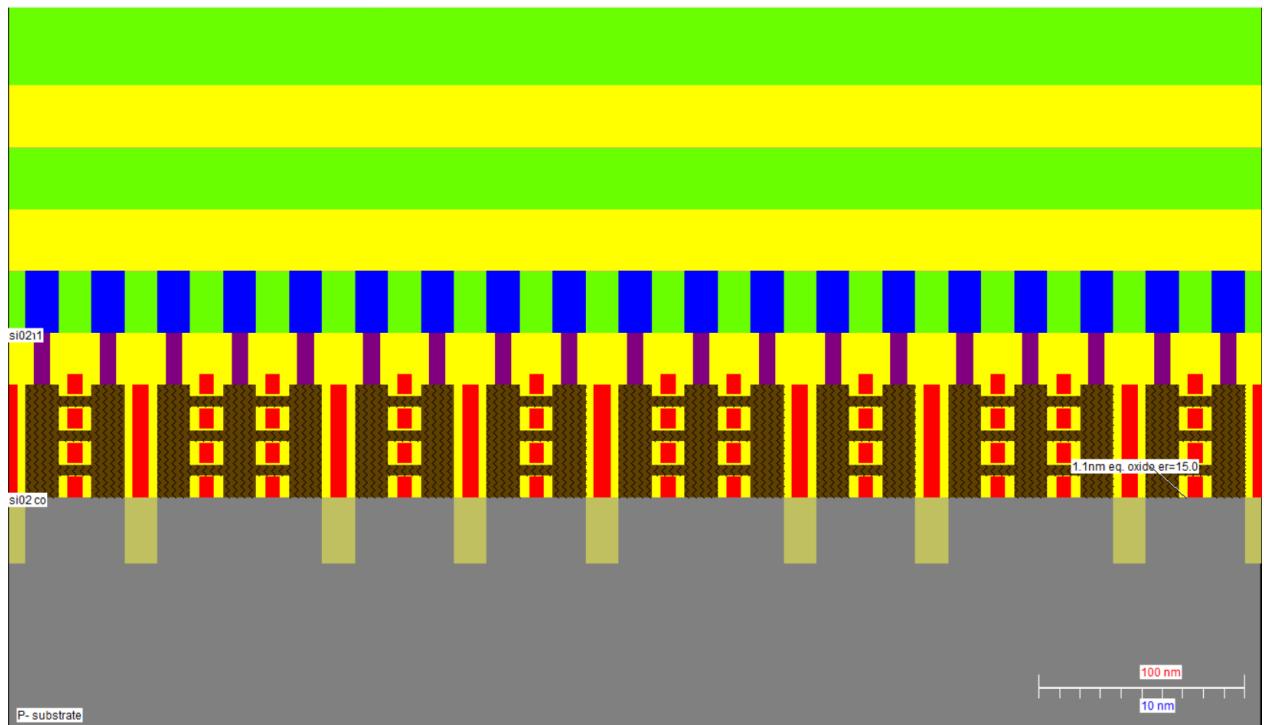


Ilustración 73 XOR Corte Transversal NSFET.

## Compuerta NOT.

A	OUT
0	1
1	0

Tabla 16 Correspondencia lógica de NOT.

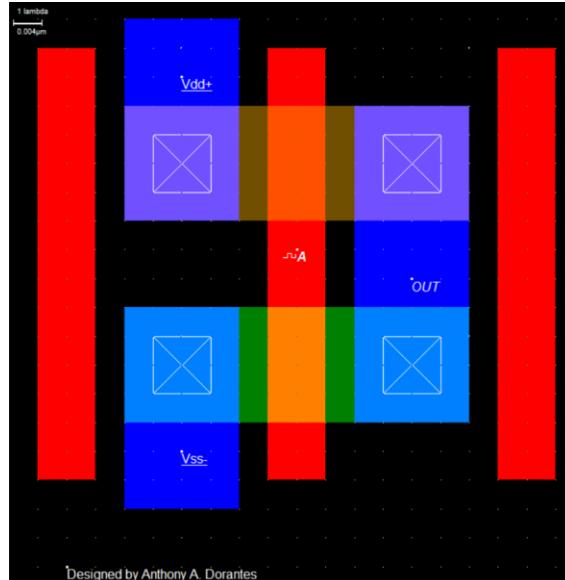


Ilustración 74 Compuerta NOT NSFET 2D (3nm).

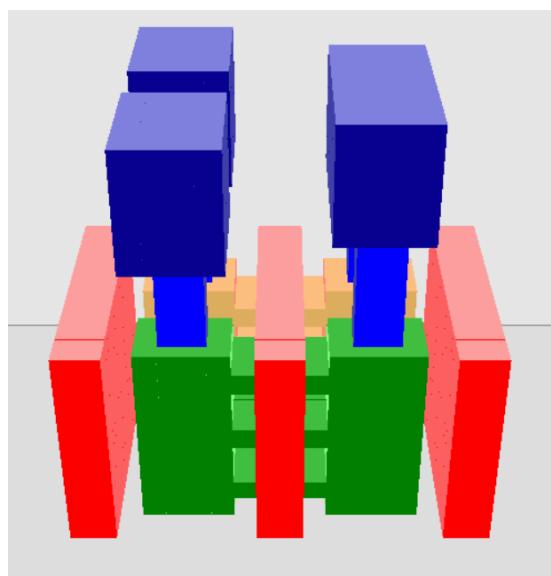


Ilustración 75 Modelo 3D (3-Nanosheet NSFET 3 nm).

## Layout ALU de 4 bits (CI 74181 bloque base)

La ilustración 76 muestra el layout de la ALU de 4 bits del 74181, y ocupa un área aproximada de  $2.2\mu\text{m} \times 14.9\mu\text{m}$ . Aproximadamente 1106 transistores.

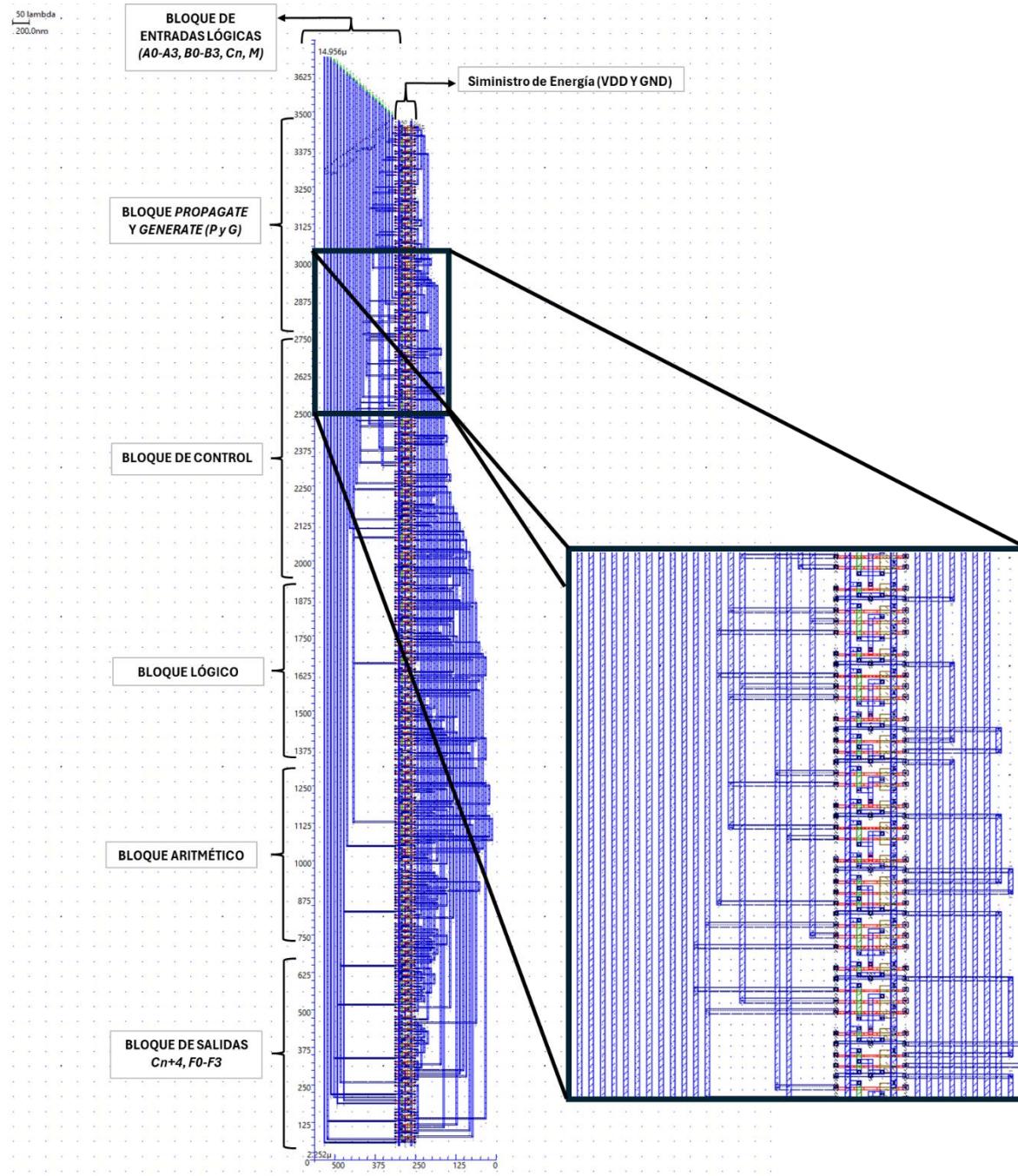


Ilustración 76 Layout de la ALU de 4 bits 74181 (NSFET 3 nm).

Nota: A partir de este punto, los modelos 3D presentados corresponden únicamente a una zona reducida del diseño, ya que, debido a las limitaciones del software, no fue posible realizar el renderizado tridimensional completo del circuito. Por ello, los layouts mostrados representan solo una parte parcial y demostrativa del diseño general.

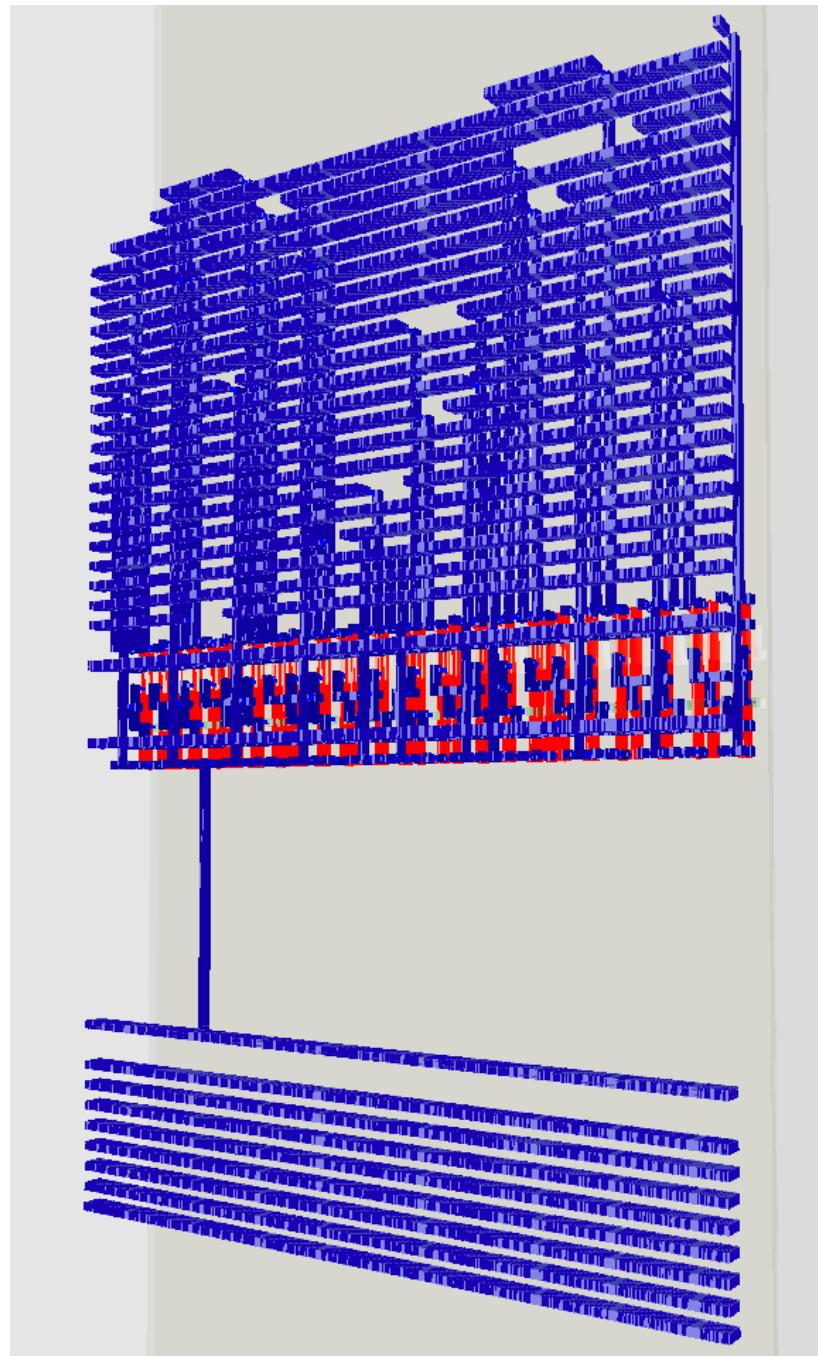


Ilustración 77 Modelo 3D parcial de la ALU 74181 de 4 bits (Microwind NSFET 3-Nanosheets 3 nm).

### 3.1.2 Layout ALU de 8 bits (2 bloques del CI 74181).

La ilustración 78 muestra el layout de la ALU de 8 bits del 74181, y ocupa un área aproximada de  $3.5\mu\text{m} \times 29.7\mu\text{m}$ . Aproximadamente 2290 transistores.

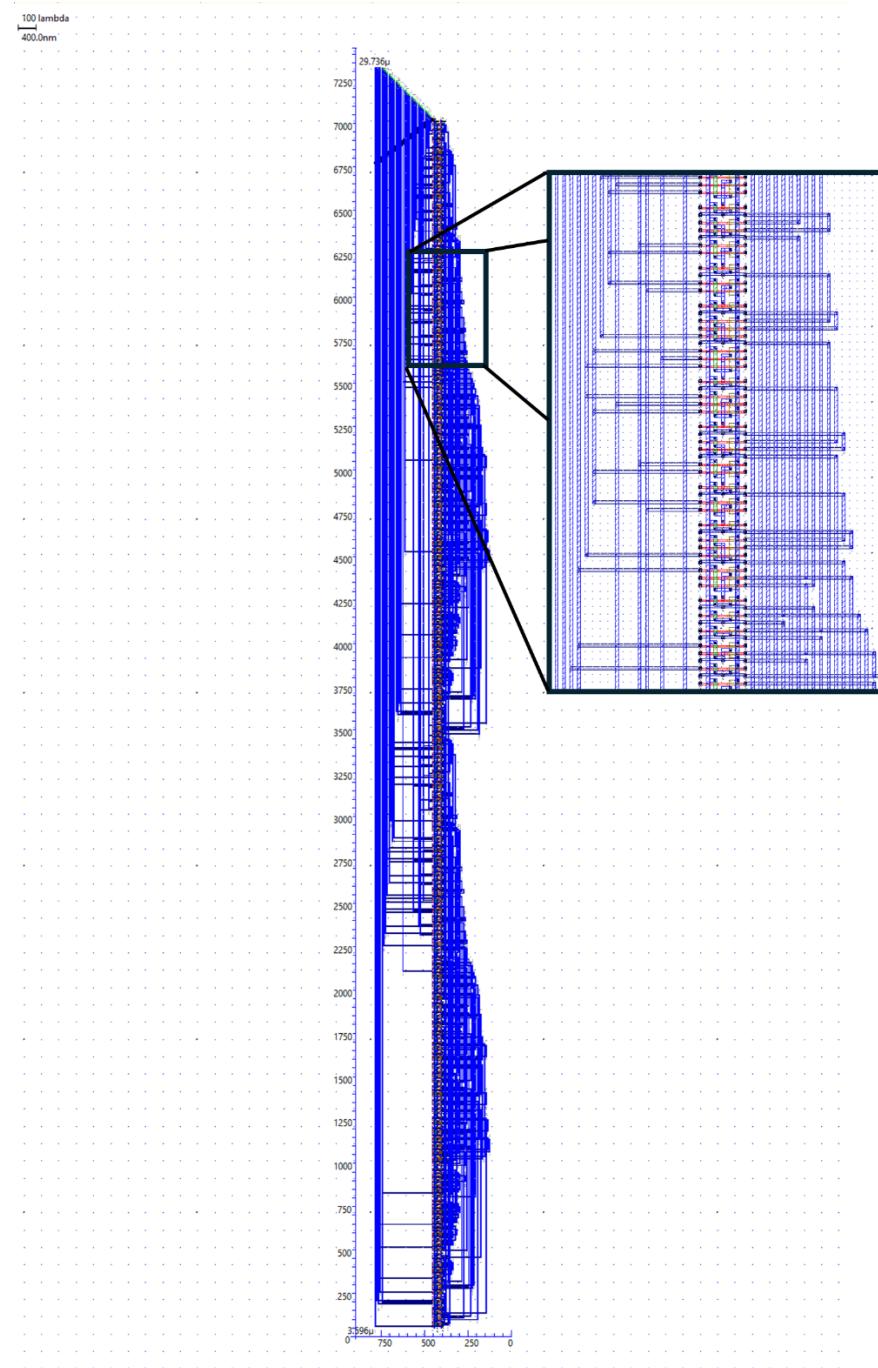


Ilustración 78 Layout de la ALU de 8 bits usando dos bloques del CI 74181 (NSFET 3 nm).

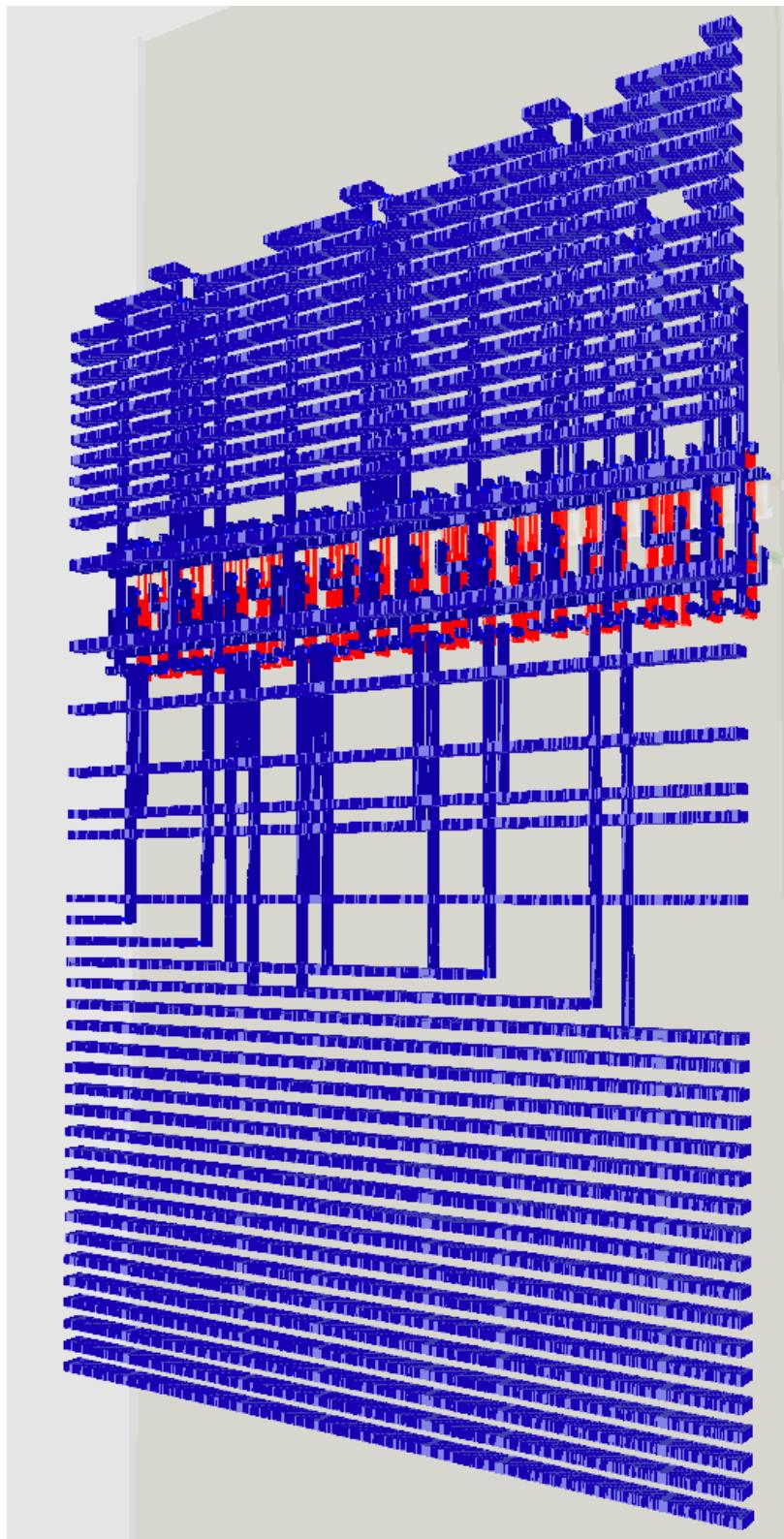


Ilustración 79 Modelo 3D parcial de la ALU de 8 bits (Microwind NSFET 3-Nanosheets 3 nm).

### 3.1.3 Layout de la ALU de 16 bits (4 bloques del CI 74181).

La ilustración 80 muestra el layout de la ALU de 16 bits del 74181, y ocupa un área aproximada de  $5.7\mu\text{m} \times 59.7\mu\text{m}$ . Aproximadamente 4,700 transistores.

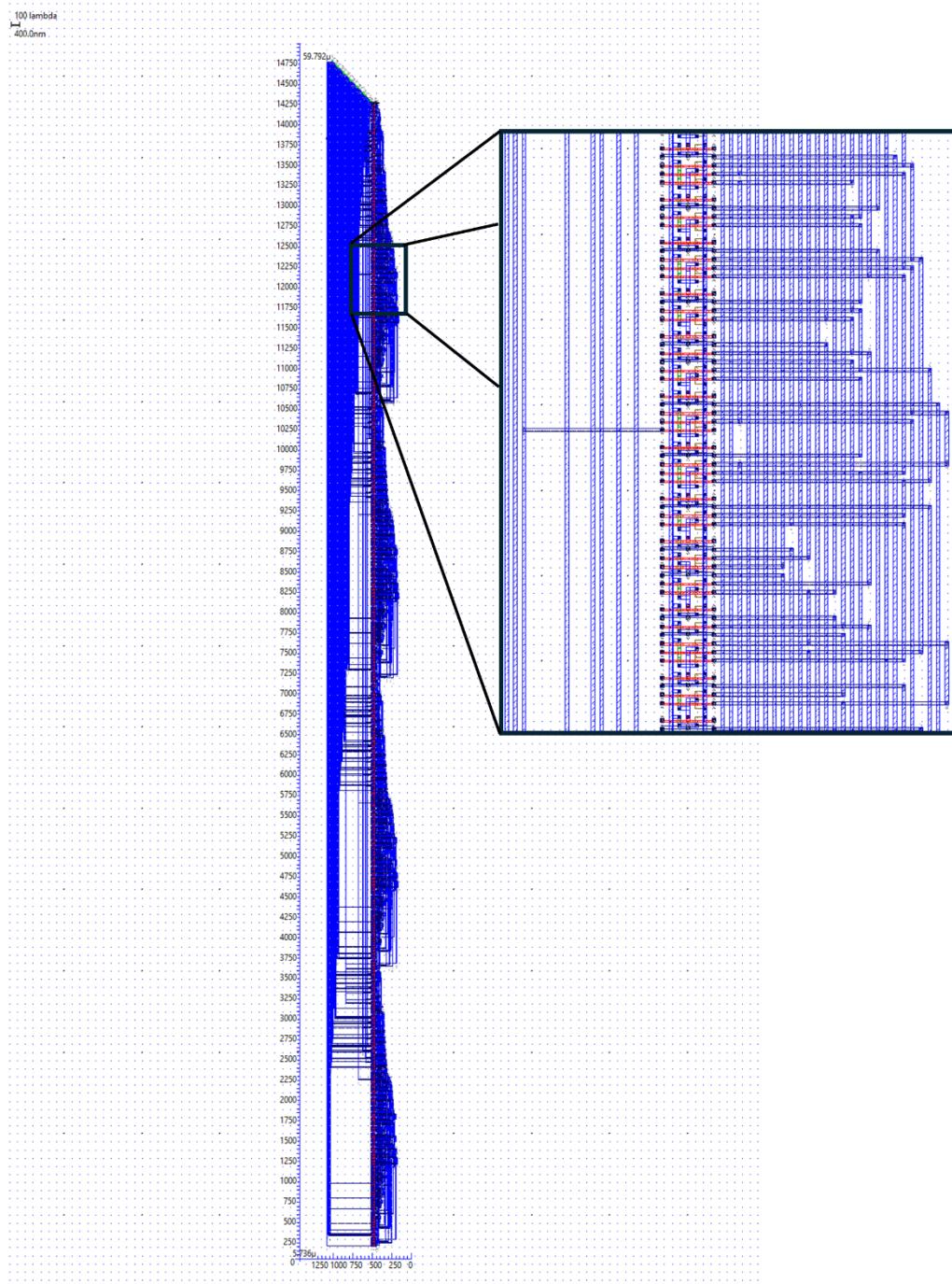


Ilustración 80 Layout de la ALU de 16 bits usando 4 bloques del CI 74181 (NSFET 3 nm).

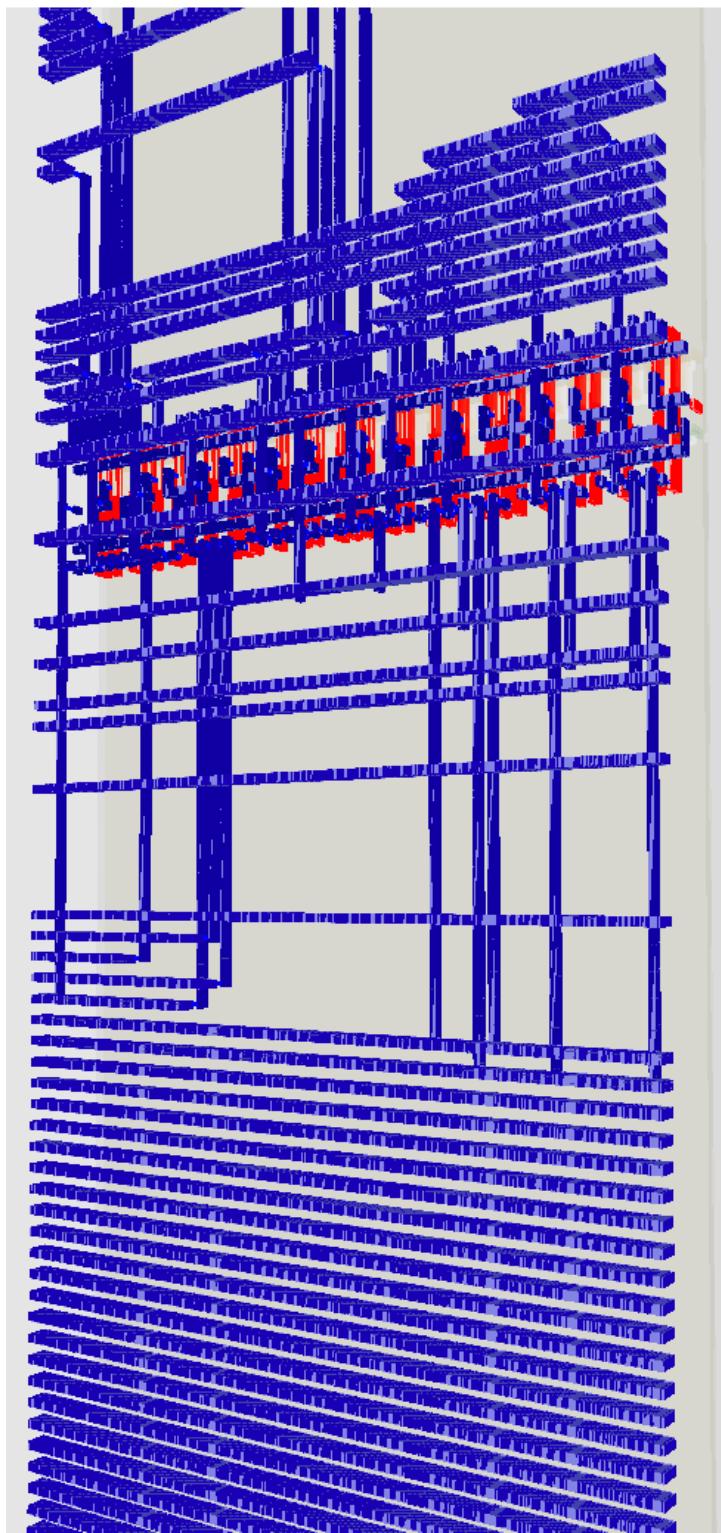


Ilustración 81 Modelo 3D parcial de la ALU de 16 bits (Microwind NSFET 3-Nanosheets 3 nm).

### 3.1.4 Layout de la ALU de 32 bits (8 bloques del CI 74181).

La ilustración 82 muestra el layout de la ALU de 32 bits del 74181, y ocupa un área aproximada de  $7.8\mu\text{m} \times 119.4\mu\text{m}$ . Aproximadamente 9,500 transistores.

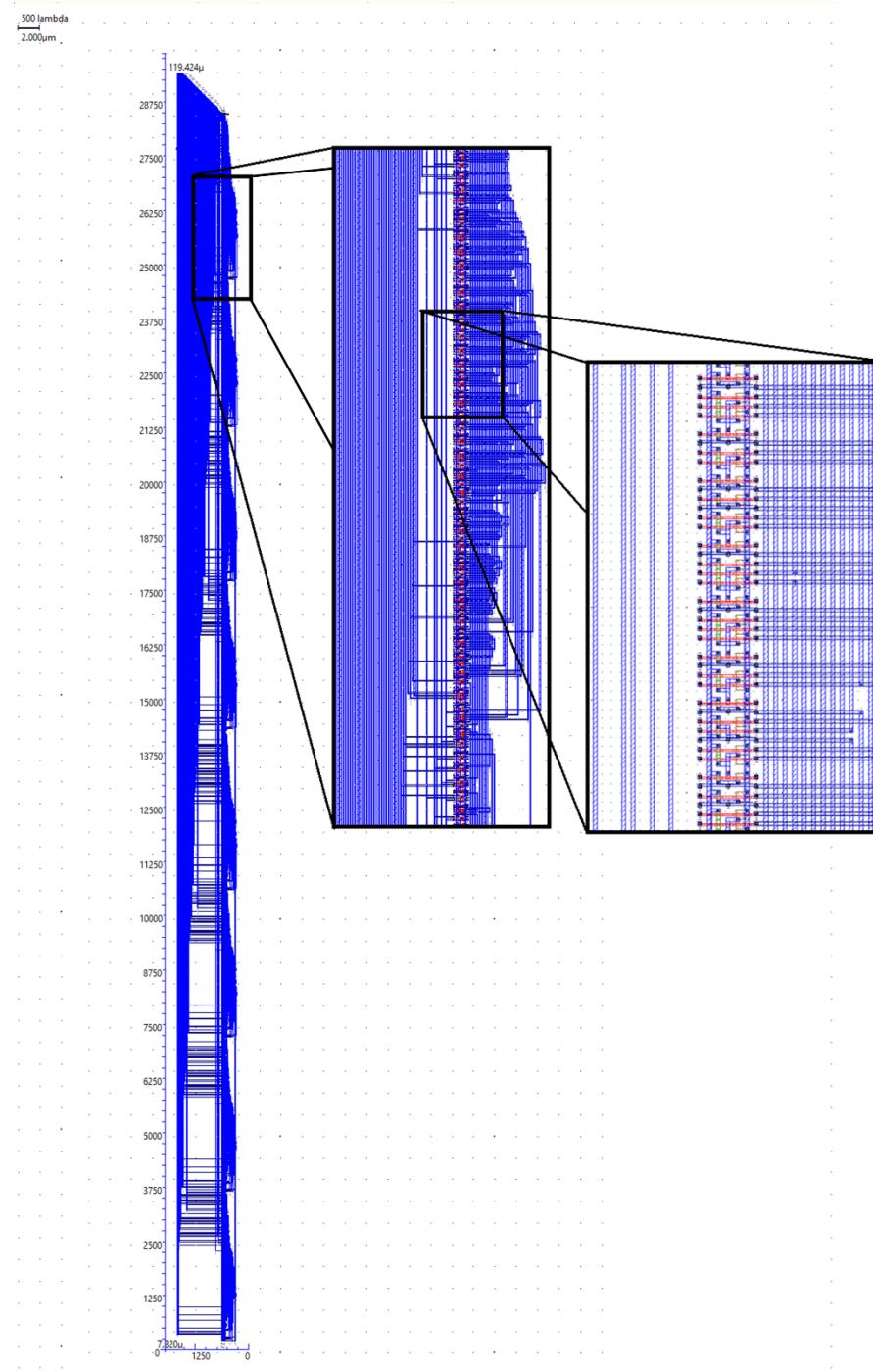


Ilustración 82 Layout de la ALU de 32 bits usando 8 bloques del CI 74181 (NSFET 3 nm).

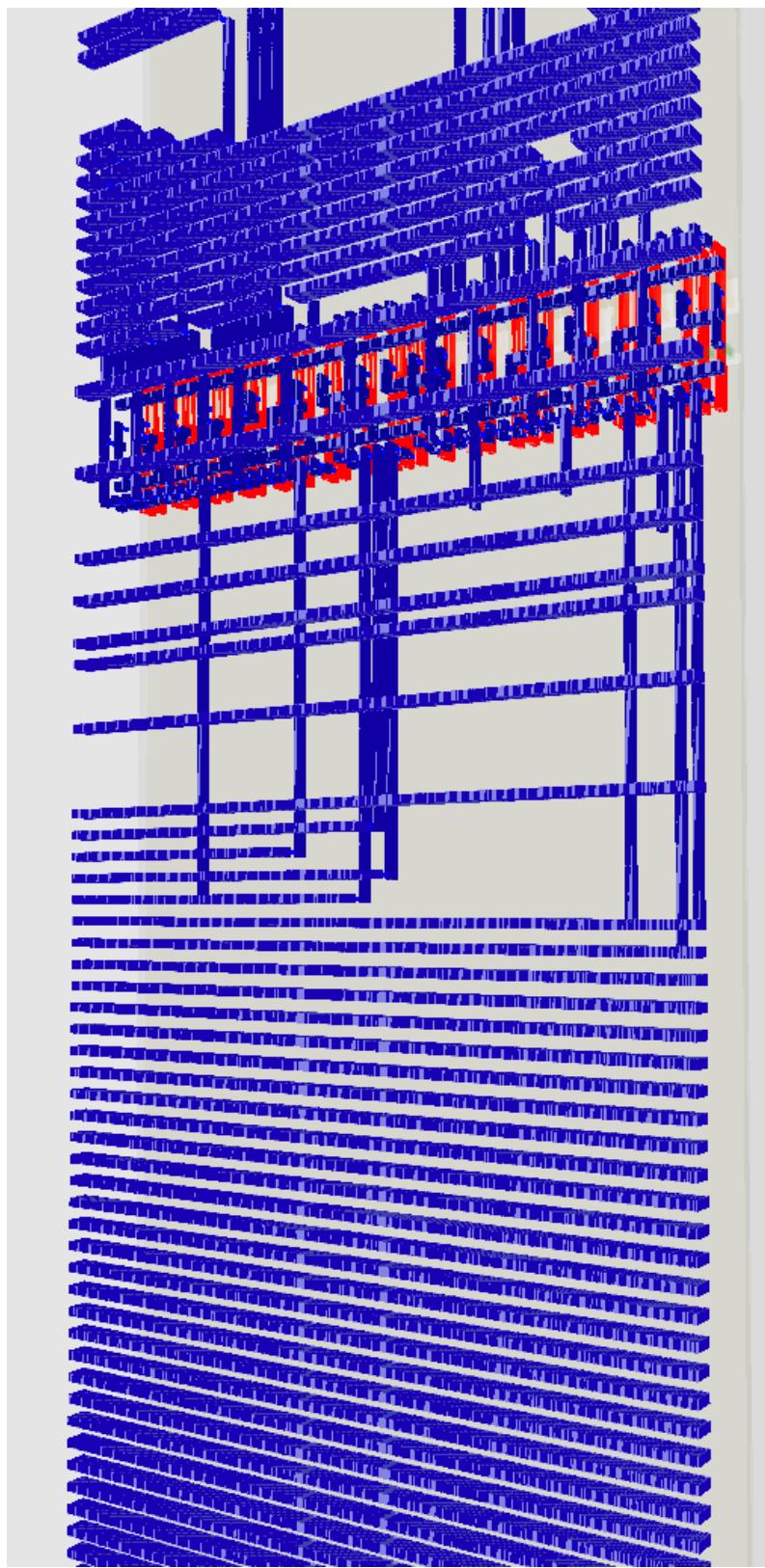


Ilustración 83 Modelo 3D parcial de la ALU de 32 bits (Microwind NSFET 3-Nanosheets 3 nm).

# Capítulo 4

## Conclusiones y recomendaciones.

El desarrollo de una ALU de 32 bits orientada a operaciones Multiply–Accumulate (MAC) utilizando tecnología NSFET de 3 nm permitió cumplir el objetivo general: demostrar que es posible trasladar una arquitectura clásica y modular hacia un entorno moderno, eficiente y adecuado para el procesamiento neuronal.

A lo largo del proyecto, el estudio de la tecnología NSFET resultó fundamental para comprender porque este tipo de transistores representa una evolución natural frente a arquitecturas tradicionales. Su estructura basada en nanosheets apiladas y su excelente control electrostático permiten operar con voltajes bajos y minimizar fugas de corriente, lo cual es esencial en aplicaciones donde el consumo energético y la estabilidad lógica son determinantes.

El diseño de compuertas a nivel transistor no solo funcionó como un requisito técnico, sino como un ejercicio que permitió comprender, desde su base más elemental, como se construyen y comportan los bloques que finalmente darían forma a la ALU completa. Este proceso consolidó la transición desde lo conceptual hacia lo funcional, manteniendo siempre una relación directa entre teoría e implementación.

La construcción del layout de la ALU de 32 bits confirmó que, incluso en nodos tan avanzados como los 3 nm, es posible preservar la filosofía modular del diseño basado en bloques de 4 bits sin perder integridad física ni violar las reglas del proceso. Esto demuestra que las arquitecturas clásicas pueden adaptarse a tecnologías actuales sin perder su esencia, siempre y cuando se respeten los principios de diseño físico a nivel nanométrico.

Las simulaciones realizadas en Microwind y DSCH permitieron verificar que, tanto las compuertas individuales como la ALU completa operarían correctamente en funciones de multiplicación y acumulación, lo cual valida la importancia del diseño como núcleo aritmético en aplicaciones de cómputo neuronal. Estas pruebas demostraron que el sistema es capaz de realizar operaciones que se utilizan ampliamente en inferencia y procesamiento matricial, reforzando su valor como posible componente dentro de arquitecturas MAC más complejas.

Finalmente, la evaluación del consumo, el área y el comportamiento lógico mostró con claridad las ventajas de la tecnología NSFET de 3 nm frente a aproximaciones tradicionales. La reducción de potencia, el aumento en densidad y la mejora en velocidad de conmutación se reflejan directamente en un desempeño más eficiente y estable, con menor latencia y mayor capacidad de integración. Estos resultados fortalecen la idea de que la migración hacia tecnologías como los nanosheets no solo es necesaria, sino estratégica para cumplir con las demandas actuales del cómputo de alto rendimiento y la inteligencia artificial.

En conjunto, el proyecto confirma que es posible llevar un diseño clásico como la ALU **74181 escalado a 32 bits y reorientado a operaciones MAC hacia la tecnología NSFET de 3 nm**, obteniendo beneficios tangibles en eficiencia, escalabilidad y desempeño. Este trabajo demuestra que la combinación entre arquitectura tradicional y tecnologías emergentes puede abrir un camino sólido hacia sistemas de procesamiento neuronal más compactos, rápidos y energéticamente eficientes.

El diseño desarrollado en este proyecto sienta una base robusta para la construcción de arquitecturas digitales más completas orientadas al procesamiento neuronal y a sistemas avanzados de integración. A partir de los resultados obtenidos, se identifican diversas líneas de trabajo futuro que permitirían ampliar las capacidades del diseño y explorar nuevas oportunidades tecnológicas en nodos de 3 nm basados en dispositivos NSFET.

## **Memorias especializadas**

- Diseñar celdas SRAM optimizadas para 3 nm.
- Incorporar memorias locales de baja latencia (scratchpads).
- Implementar memorias multi–puerto para accesos paralelos.

## **Codificadores y decodificadores digitales**

- Desarrollar codificadores y decodificadores en 3 nm basados en NSFET.
- Integrar decodificadores para control interno de memoria.
- Implementar codificadores de prioridad para sistemas de control.

## **Convertidores AD/DA (DAC)**

- Explorar convertidores AD de baja potencia (SAR o Flash).
- Diseñar DACs para sistemas mixtos.
- Evaluar linealidad y ruido en nodo de 3 nm.

## **Módulos digitales avanzados**

- Incorporar sumadores de alto rendimiento.
- Evaluar multiplicadores más complejos.
- Añadir divisores y módulos de raíz cuadrada.

## **Microarquitectura orientada a NPU**

- Diseñar una unidad de control para operaciones MAC.
- Implementar el procesamiento en línea (pipeline) interno para mejorar frecuencia.
- Integrar buses internos de alta velocidad.

## **Módulos para procesamiento neuronal**

- Desarrollar unidades de activación digital.
- Diseñar módulos acelerados para convolución y operaciones matriciales.
- Incluir soporte para cuantización (INT8, INT4, FP8).

## **Construcción de un SoC mínimo para IA**

- Integrar MAC, memoria, controladores y buses en un sistema funcional.
- Evaluar consumo energético y compararlo con tecnologías tradicionales.

# Glosario

ALU (Arithmetic Logic Unit)	<i>Unidad digital encargada de realizar operaciones aritméticas y lógicas dentro de un procesador o sistema digital.</i>
Arquitectura Flash (ADC)	<i>Convertidor A/D que utiliza comparadores en paralelo para obtener la conversión más rápida posible, a costa de mayor consumo y área.</i>
Arquitectura SAR	<i>Convertidor A/D que realiza una búsqueda binaria del valor digital a través de aproximaciones sucesivas.</i>
Buses de alta velocidad	<i>Interconexiones digitales capaces de transportar datos con grandes anchos de banda, destinadas a reducir cuellos de botella entre módulos.</i>
Capacitancia de compuerta (CGS)	<i>Capacitancia entre compuerta y surtidor de un transistor MOS o GAA que afecta la velocidad de conmutación.</i>
Capacitancia parásita	<i>Capacitancia no deseada presente entre conductores o terminales cercanos, generada por fenómenos geométricos o materiales.</i>
Celdas de memoria	<i>Estructuras digitales mínimas diseñadas para almacenar bits, tales como las celdas de SRAM o DRAM.</i>
CMOS	<i>Tecnología electrónica basada en el uso complementario de transistores NMOS y PMOS para implementar lógica de bajo consumo.</i>
Codificador binario	<i>Circuito digital que convierte una única entrada activa entre varias líneas en un código binario equivalente.</i>
Codificador de prioridad	<i>Codificador que selecciona la entrada activa de mayor prioridad y genera su representación binaria.</i>
Convertidor Analógico–Digital (AD)	<i>Circuito que transforma señales analógicas continuas en representaciones digitales discretas.</i>
Convertidor Digital–Analógico (DA)	<i>Circuito que convierte valores digitales en señales analógicas continuas.</i>
Convolución (operación)	<i>Operación fundamental del procesamiento neuronal donde un filtro (kernel) se desplaza sobre una entrada generando mapas de características.</i>

Corriente de drenador ( $I_{ds}$ )	<i>Corriente que fluye entre drenador y surtidor en un transistor bajo polarización.</i>
Corriente de fuga	<i>Corriente no deseada cuando el transistor está apagado; aumenta en nodos avanzados por efectos de escalamiento.</i>
CPU	<i>Unidad encargada de ejecutar instrucciones generales de propósito computacional.</i>
Cuantización	<i>Proceso de reducción de precisión numérica para acelerar cálculos en IA, generalmente mediante INT8, INT4 o FP8.</i>
DAC (Digital-to-Analog Converter)	<i>Convertidor Digital–Analógico (DA).</i>
Decodificador digital	<i>Circuito combinacional que convierte un código binario de entrada en la activación de una única salida entre varias líneas.</i>
Decodificador de memoria	<i>Unidad que traduce direcciones binarias en señales de selección para filas/columnas dentro de memorias SRAM o DRAM.</i>
Divisor Digital	<i>Bloque aritmético capaz de realizar operaciones de división en hardware mediante algoritmos como restaurativo, no restaurativo o SRT.</i>
Efectos de canal corto (SCE)	<i>Fenómenos presentes en transistores de longitud reducida, donde la compuerta pierde control electrostático sobre el canal, generando variaciones en <math>V_{th}</math>, DIBL y fugas.</i>
Epitaxia	<i>Crecimiento controlado de una capa cristalina sobre un sustrato semiconductor manteniendo alineación estructural.</i>
FinFET	<i>Transistor tridimensional en forma de aleta que mejora el control sobre el canal.</i>
Flip-Flop	<i>Elemento secuencial que almacena un bit y lo actualiza en el flanco del reloj. Fundamental para registros y pipelines.</i>
FP8	<i>Formato de punto flotante de baja precisión usado para acelerar inferencias en IA.</i>
Frecuencia de reloj	<i>Número de ciclos por segundo (Hz) a los que opera un circuito secuencial. Determina el rendimiento temporal del sistema.</i>
GAA (Gate-All-Around)	<i>Arquitectura de transistor donde la compuerta envuelve completamente al canal.</i>

INT4	<i>Formato de cuantización entero de 4 bits utilizado para aumentar velocidad y reducir consumo en IA.</i>
INT8	<i>Formato entero de 8 bits usado ampliamente en inferencia acelerada.</i>
Latch	<i>Elemento de almacenamiento sensible al nivel del reloj, mantiene un estado siempre que la señal de habilitación esté activa.</i>
Latencia	<i>Tiempo transcurrido entre el inicio de una operación y la obtención del resultado.</i>
Linealidad	<i>Medida que indica qué tan proporcional es la relación entre entrada y salida en un sistema, especialmente importante en ADC/DAC.</i>
Lógica combinacional	<i>Circuitos cuya salida depende únicamente de las entradas presentes.</i>
Lógica secuencial	<i>Circuitos cuya salida depende del estado previo almacenado, además de las entradas actuales.</i>
MOSFET (Metal–Oxide–Semiconductor FET)	<i>Transistor de efecto de campo ampliamente utilizado en circuitos digitales y analógicos.</i>
MAC (Multiply–Accumulate)	<i>Operación fundamental en redes neuronales que realiza simultáneamente una multiplicación y una acumulación</i>
Memoria multi–puerto	<i>Tipo de memoria que permite accesos simultáneos de lectura y/o escritura a través de múltiples puertos independientes.</i>
MPU	<i>Unidad de procesamiento que ejecuta instrucciones de manera secuencial y controla periféricos a nivel de software.</i>
Nanosheet	<i>Canal semiconductor de forma plana utilizado en la arquitectura Gate-All-Around (GAA), donde la compuerta lo rodea completamente.</i>
Nanoelectrónica	<i>Disciplina dedicada al diseño de dispositivos y sistemas electrónicos con escalas nanométricas.</i>
Nodo tecnológico	<i>Generación de fabricación de semiconductores que agrupa reglas de diseño, dimensiones mínimas y densidad (p. ej., 7 nm, 5 nm, 3 nm).</i>
NSFET (Nanosheet Field-Effect Transistor)	<i>Transistor basado en nanosheets apilados, con compuerta que rodea completamente cada hoja.</i>

NPU (Neural Processing Unit)	<i>Acelerador especializado para cargas neuronales que implementa operaciones MAC masivas, unidad de activación y estructuras paralelas optimizadas.</i>
Pipeline	<i>Técnica donde una operación se divide en etapas secuenciales, permitiendo ejecutar múltiples operaciones en paralelo mediante superposición temporal.</i>
Registro acumulador	<i>Registro utilizado para almacenar resultados parciales durante operaciones aritméticas o MAC.</i>
Ruido electrónico	<i>Variaciones aleatorias generadas por fenómenos térmicos o cuánticos que afectan señales analógicas y digitales.</i>
RTL	
Scratchpads	<i>Memoria interna de baja latencia, no jerárquica, usada por NPUs y GPUs para almacenar datos temporales de alto ancho de banda.</i>
SoC (System on Chip)	<i>Circuito integrado que contiene CPU, memoria, buses, periféricos y aceleradores dentro de un solo chip.</i>
SRAM	<i>Memoria volátil que almacena bits mediante biestables (latches) y ofrece tiempos de acceso muy rápidos.</i>
Sumador CLA (Carry Look-Ahead Adder)	<i>Sumador que acelera el cálculo del acarreo usando lógica anticipada basada en señales de generación y propagación.</i>
Sumador de alto rendimiento	<i>Aquellos sumadores diseñados para minimizar latencia y maximizar frecuencia mediante técnicas como CLA, CSKA, Ling o híbridos.</i>
Unidad de control	<i>Circuito que genera las señales de control necesarias para coordinar operaciones internas de un sistema digital.</i>
Unidad de raíz cuadrada	<i>Bloque aritmético que implementa algoritmos iterativos o aproximaciones para obtener raíces cuadradas en hardware.</i>
Voltaje umbral (V <sub>th</sub> )	<i>Voltaje mínimo necesario para crear un canal conductor en un transistor MOS o GAA.</i>

# Bibliografía

- Analog Devices. (2024). Educational glossary.  
<https://www.analog.com/en/education/glossary.html>
- Analog Devices. (2024). Flash ADC. <https://www.analog.com/en/education/glossary/flash-adc.html>
- Analog Devices. (2024). Linearity. <https://www.analog.com/en/education/glossary/linearity.html>
- Analog Devices. (2024). Noise. <https://www.analog.com/en/education/glossary/noise.html>
- Analog Devices. (2024). Analog-to-digital converter overview.  
<https://www.analog.com/en/education/glossary/analog-to-digital-converter.html>
- Analog Devices. (2024). Digital-to-analog converter overview.  
<https://www.analog.com/en/education/glossary/digital-to-analog-converter.html>
- Augarten, S. (1983). State of the art: A photographic history of the integrated circuit. Ticknor and Fields.
- Bush, E. (1964). The use of solid circuits in satellite instrumentation (NASA Technical Note D-1758).
- Cadence. (2024). Semiconductor glossary. <https://resources.cadence.com/blog/topic/glossary>
- Calderón Monroy, R. J., González-Vidal, J. L., & Raygoza-Panduro, J. J. (2025). Diseño del layout de una unidad aritmético lógica utilizando The Electric. Semestral Pädi, 12(24), 173–183.
- Cambridge Dictionary. (2024). Nanoelectronics.  
<https://dictionary.cambridge.org/dictionary/english/nanoelectronics>
- Chen, M.-L., Sun, X., Liu, H., Wang, H., Zhu, Q., S. W., ... Han, Z. (2020). A FinFET with one atomic layer channel. *Nature Communications*, 11, 1205. <https://doi.org/10.1038/s41467-020-15096-0>
- Chen, Y., Krishna, T., Emer, J. S., & Sze, V. (2017). Eyeriss: An energy-efficient reconfigurable accelerator for deep convolutional neural networks. *IEEE Journal of Solid-State Circuits*, 52(1), 127–138.
- Chih Tang, S. (1988). Evolution of the MOS transistor—From conception of VLSI. *Proceedings of the IEEE*, 76, 1295–1324.
- Colinge, J.-P. (2016). FinFETs and other multi-gate transistors (2nd ed.). Springer.
- Dasgupta, A., & Hu, C. (2020). BSIM-CMG compact model for IC CAD: From FinFET to GAA FET technology. *Journal of Microelectronics Manufacturing*.
- Del Alamo, J., Antoniadis, D., Lin, J., Lu, W., Alon, V., & Xin, Z. (2016). Nanometer-scale III-V MOSFETs. *Journal of the Electron Devices Society*, 1–10.
- ElProCus. (2013). Different types of integrated circuits. <https://www.elprocus.com/different-types-of-integrated-circuits/>
- Faulkner, A., & Gurzi, F. (1962). MAGIC – An advanced computer for spaceborne guidance systems. Spaceborne Computer Engineering Conference.
- Frank, A., & Ramesh, K. (2022). Design and analysis of multiplier accumulator circuit for DSP applications. RV College of Engineering.
- González-Vidal, J. L., et al. (2021). Nano transistores de efecto de campo tipo aleta para aplicaciones digitales. Pädi, 9, 99–105.

- HardZone. (2020). TSMC deja obsoleta a Intel y Samsung: sus 3 nm estarán listos en 2022.  
<https://hardzone.es/noticias/procesadores/tsmc-5-nm-3-nm-2021/>
- Hisamoto, D., Lee, W.-C., Kedzierski, J., Takeuchi, H., Asano, K., Kuo, C., ... Hu, C. (2000). FinFET—A self-aligned double-gate MOSFET. *IEEE Transactions on Electron Devices*, 47(12), 2320–2325. <https://doi.org/10.1109/TED.2000.887050>
- Hofstein, S., & Heiman, F. (1963). The silicon insulated gate field effect transistor. *Proceedings of the IEEE*, 51, 1190–1192.
- IBM. (2021). IBM unveils world's first 2 nm chip technology. <https://newsroom.ibm.com/2021-05-06-IBM-Unveils-Worlds-First-2-Nanometer-Chip-Technology,-Opening-a-New-Frontier-for-Semiconductors>
- imec. (2019). Nanosheet transistors. <https://www.imec-int.com/en/articles/nanosheet-transistors>
- Intel. (2021). Accelerating process innovation. <https://download.intel.com/newsroom/2021/client-computing/accelerating-process-innovation.pdf>
- IRJET. (2022). Efficient MAC architecture for high-speed DSP. *International Research Journal of Engineering and Technology*.
- Johnson, E. (2025). ¿Qué es un FET? PCBBasic.  
[https://www.pcbbasic.com/es/blog/what\\_is\\_a\\_fet.html](https://www.pcbbasic.com/es/blog/what_is_a_fet.html)
- Karimi, K., Fardoust, A., & Javanmard, M. (2024). Comprehensive review of FinFET technology. *Micromachines*.
- Knox Basset, R. (2002). To the digital age. Johns Hopkins University Press.
- Kumar, P., Kumar Pal, P., & Yadav, S. (2019). Analysis of nanosheet field effect transistor (NSFET) for device and circuit perspective. *WITCON ECE Conference Proceedings*, 2–3.
- Lenovo. (2023). ¿Qué es un transistor? <https://www.lenovo.com/mx/es/glosario/transistor/>
- Lenovo. (2024). Integrated circuit. <https://www.lenovo.com/us/en/glossary/integrated-circuit/>
- Mari, L. (2020). What is a FinFET? EEPower. <https://eepower.com/technical-articles/what-is-a-fin fet/>
- Martin-Marietta Company. (1962). MARTAC 420: A multipurpose digital control computer.
- Mendiola Oria, C. (2018). MOSFET: Qué es. Rincón de la Tecnología.  
<https://rincondelatecnologia.com/mosfet-que-es-ventajas-desventajas/>
- Micron. (2024). Memory glossary. <https://www.micron.com/support/glossary>
- Microchip. (2024). Technical glossary. <https://www.microchip.com/en-us/glossary>
- Moore, G. (1964). The future of integrated electronics. *Electronics*, 38(8).
- Mukesh, S., & Zhang, J. (2022). A review of the gate-all-around nanosheet FET process. *Electronics* (MDPI).
- NVIDIA. (2024). AI glossary. <https://www.nvidia.com/en-us/glossary/>
- NVIDIA Developer. (2024). Developer glossary. <https://developer.nvidia.com/glossary>
- ONNX. (2024). AI glossary. <https://onnx.ai/glossary/>
- Oxford University Press. (2024). Oxford dictionary of computing.  
<https://www.oxfordreference.com>
- Oxford University Press. (2024). Oxford dictionary of physics. <https://www.oxfordreference.com>
- Perry, T. (2020). How the father of FinFETs helped save Moore's law. *IEEE Spectrum*.  
<https://spectrum.ieee.org/how-the-father-of-finfets-helped-save-moores-law>
- ProEx. (2023). The advantages of integrated circuits.  
<https://www.proex1.com/blog/posts/2023/june/the-advantages-of-integrated-circuits-ics/>
- Qing, L., Decheng, Z., & Dongxin, W. (2024). High-performance Fourier transform algorithms based on the NPU. *Applied Sciences* (MDPI).

- Ranganath, L., Jay K. D., & P. Siva N. R. P. (2016). Design of MAC unit in artificial neural network architecture using Verilog HDL. SCOPES Conference.  
<https://doi.org/10.1109/SCOPES.2016.7955511>
- Riordan, M. (2025). Junction transistors. Encyclopædia Britannica.  
<https://www.britannica.com/technology/transistor/Junction-transistors>
- Russell, J. (2025). NPU landscape. Deep Render. <https://deeppreder.ai/blog/npu-landscape>
- Samsung. (2022). Samsung begins chip production using 3 nm process technology.  
<https://news.samsung.com/global/samsung-begins-chip-production-using-3nm-process-technology-with-gaa-architecture>
- Sedra, A., & Smith, K. C. (2004). Microelectronic circuits. Oxford University Press.
- Shirriff, K. (2017). Inside the vintage 74181 ALU: How it works.  
<https://www.righto.com/2017/03/inside-vintage-74181-alu-chip-how-it.html>
- SkyWater Technology. (2021). SkyWater 130 nm open source PDK.  
<https://github.com/google/skywater-pdk>
- Synopsys. (2024). Semiconductor glossary. <https://www.synopsys.com/glossary.html>
- Teja, L. D., Shoneeth, K., Siri, C. H., & Abhishek, T. (2024). Design and performance analysis of low-power multiply-accumulate (MAC) units. GCCIT Conference Proceedings.
- Texas Instruments. (1972). SN74181 datasheet. <https://doctor-pasquale.com/wp-content/uploads/2017/05/74181-ALU.pdf>
- Texas Instruments. (1961). A molecular electronic computer.
- Trombini, H., Marmitt, G. G., Alencar, I., Baptista, D. L., Reboh, S., Mazen, F., ... Grande, P. (2019). Unraveling structural and compositional information in 3D FinFET electronic devices. *Scientific Reports*, 9, 11629. <https://doi.org/10.1038/s41598-019-48117-0>
- Tsividis, Y., & McAndrew, C. (2011). Operation and modeling of the MOS transistor. Oxford University Press.
- Weste, N., & Harris, D. (2012). CMOS VLSI design. Pearson.
- Zhang, J. (2025). Circuitos integrados: estructura, tipos y aplicaciones. PCBASIC.  
[https://www.pcbasic.com/es/blog/integrated\\_circuit.html](https://www.pcbasic.com/es/blog/integrated_circuit.html)
- Ziesler, A., et al. (2020). Open-source tools for teaching and research in integrated circuit design. *Proceedings of the IEEE*, 108(4), 495–507.
- Yu, C., et al. (2019). A 10.5-TOPS/W deep neural network accelerator using low-voltage MAC array in 65 nm CMOS. *IEEE Journal of Solid-State Circuits*, 54(1), 217–228.

# Anexo 1 Productos

1.- Diseño de una Unidad Aritmético Lógica con NSFETs en nodo de 3 nm, (2025), Anthony Alejandro Dorantes Mejía y José Luis González Vidal , XXII Congreso Concurso de Diseño de Hardware— Software Inteligente y Microprocesadores. CC DIHSIM 2025B, Guadalajara, Jalisco, México, 27 y 28 de noviembre de 2025.