World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(REVIEW ARTICLE)

Check for updates

# FROM PMOS TO CFET: A COMPREHENSIVE SYSTEMATIC REVIEW AND COMPARATIVE ANALYSIS OF MOSFET-BASED DEVICE ARCHITECTURES FOR NEXT-GENERATION VLSI SYSTEMS

Pallavi Singh [1,*], Bhim Singh [2], Priyanka Gupta [3] and Ragini Rai [4]

[1] Department of Electronics Engineering, Kamla Nehru Institute of Technology, Sultanpur– 228118.
[2] Department of Electronics Engineering (VLSI), Maharana Pratap Institute of Technology-273015.
[3] Department of Electronics and communication Engineering, KIPM CET GIDA gorakhpur-273015.
[4] Department of Electronics and communication Engineering, Noida International University -203201.
* Corresponding Author

ORCID Details
Pallavi Singh: https://orcid.org/0009-0000-8984-6237
Ragini Rai: https://orcid.org/0009-0008-9731-1599

World Journal of Advanced Research and Reviews, 2026, 31(03), 756–766

Publication history: Received on 01 August 2026; revised on 07 September 2026; accepted on 09 September 2026

Article DOI: https://doi.org/10.30574/wjarr.2026.31.3.2326



## ABSTRACT

The relentless drive toward transistor miniaturization that has sustained Moore's Law for over five decades has forced the semiconductor industry through a succession of device-architecture transitions, each addressing the electrostatic and parasitic limitations of its predecessor. This paper presents a systematic review and comparative analysis of MOSFET-based device architectures spanning the foundational planar PMOS/NMOS transistor and complementary metal-oxide-semiconductor (CMOS) circuit, the FinFET/tri-gate transition commercialized in 2011, gate-all-around (GAA) nanowire and Nano sheet field-effect transistors now in high-volume manufacturing, the fork sheet architecture, and the emerging complementary FET (CFET). Technical literature and primary industry disclosures from IEEE Xplore, arXiv, and semiconductor foundries and research consortia (imec, Intel, TSMC, Samsung) covering 2011–2026 were reviewed and synthesized using a PRISMA-inspired methodology adapted for engineering and technology sources. Each architecture is evaluated against a consistent set of criteria: gate electrostatic control, short-channel-effect suppression, area and density scaling, power–performance trade-offs, and manufacturing readiness, and the findings are consolidated into a comparative summary table. The review finds that GAA nanosheet transistors have now entered high-volume production across three major foundries, while the CFET architecture — in which n-type and p-type devices are stacked vertically — is positioned as the leading contender to extend logic scaling into the sub-1 nm regime, contingent on resolving wafer-bonding, contact-yield, and thermal-crosstalk challenges. The paper concludes with an assessment of open research gaps and beyond-CFET directions, including two-dimensional channel materials and monolithic three-dimensional integration.

**Keywords:** MOSFET; PMOS; NMOS; CMOS; Finfet; Gate-All-Around; Nanosheet; Forksheet; Complementary FET (CFET); VLSI; Transistor Scaling; Semiconductor Device Architecture

## 1 INTRODUCTION

Since the invention of the point-contact transistor at Bell Laboratories in 1947 and the subsequent development of the metal-oxide-semiconductor field-effect transistor (MOSFET) in the early 1960s, the scaling of transistor dimensions has served as the principal engine of progress in very-large-scale integration (VLSI). Gordon Moore's 1965 observation that the number of components on an integrated circuit would double at a roughly constant rate became a self-fulfilling industrial roadmap that has guided decades of process-node transitions, from micrometer-scale devices to today's sub-2-nanometer logic [1].

The MOSFET exists in two complementary polarities: the p-channel device (PMOS), in which current is carried by holes, and the n-channel device (NMOS), in which current is carried by electrons. Early integrated circuits, including the first commercial microprocessors, were built using PMOS logic, since the PMOS fabrication process of the late 1960s was comparatively more tolerant of the mobile ionic contamination that limited the yield of contemporary NMOS lines. As process control matured through the 1970s, NMOS logic displaced PMOS because the higher mobility of electrons relative to holes in silicon delivers faster switching speeds for an equivalent device size. The eventual pairing of an NMOS pull-down network with a PMOS pull-up network in complementary MOS (CMOS) circuits — which draw negligible static current in either steady logic state — became the dominant logic family from the 1980s onward and remains the foundation of virtually every digital VLSI system produced today.

For roughly three decades, CMOS scaling followed the empirical rules articulated by Dennard et al., in which gate length, oxide thickness, and supply voltage were reduced together so that power density remained approximately constant from one generation to the next [2]. This “Dennard scaling” regime broke down during the mid-2000s as threshold and supply voltages approached practical floors, causing subthreshold leakage and static power dissipation to grow disproportionately with continued dimensional scaling. The resulting power-density ceiling motivated the industry's shift toward multi-core processor architectures and, at the device level, toward three-dimensional transistor structures capable of restoring the gate electrostatic control that planar geometries could no longer provide.

This paper traces that architectural evolution within a single, unified framework: from the planar PMOS/NMOS/CMOS transistor, through the FinFET (tri-gate) transition commercialized by Intel at the 22 nm node in 2011, to gate-all-around (GAA) nanowire and nanosheet transistors now in high-volume manufacturing at the 2 nm-class node, to the forksheet architecture proposed as an intermediate area-scaling booster, and finally to the complementary FET (CFET), in which n-type and p-type devices are monolithically or sequentially stacked in the vertical dimension. The objective is to provide VLSI researchers, device engineers, and graduate students with a single comparative reference that synthesizes the electrostatic operating principles, historical timeline, and current (2024–2026) manufacturing status of each architecture, and to identify the open research questions that remain as the industry approaches the CFET transition.

The remainder of the paper is organized as follows. Section 2 describes the review methodology. Section 3 revisits the fundamentals of PMOS, NMOS, and CMOS operation. Section 4 outlines the scaling theory and short-channel effects that motivate each architectural transition. Section 5 presents the evolutionary trajectory of device architectures from planar MOSFET through CFET. Section 6 discusses backside power delivery as a complementary design-technology co-optimization (DTCO) lever. Section 7 consolidates a comparative analysis across architectures. Section 8 surveys emerging beyond-CFET directions. Section 9 identifies open challenges and research gaps, Section 10 offers a synthesis discussion, and Section 11 concludes the paper.

## 2 REVIEW METHODOLOGY

This review adopts a structured methodology inspired by the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework, adapted for an engineering and technology literature base that spans peer-reviewed journals, conference proceedings, and primary technical disclosures from semiconductor foundries and research consortia. A conventional clinical-style PRISMA flow is not directly applicable to a fast-moving, largely proprietary technology domain; accordingly, the adaptation described below prioritizes source triangulation and recency over meta-analytic pooling of quantitative outcomes.

### 2.1 Sources and Search Strategy

Sources consulted include IEEE Xplore (IEEE Transactions on Electron Devices, the International Electron Devices Meeting (IEDM), and the IEEE/JSAP Symposium on VLSI Technology and Circuits), the arXiv preprint repository, and technical publications and press disclosures issued by imec, Intel Foundry, TSMC, and Samsung Semiconductor, covering January 2011 (the year of the first commercial FinFET) through mid-2026. Search terms combined “PMOS,” “NMOS,” “CMOS scaling,” “FinFET,” “tri-gate,” “gate-all-around,” “nanosheet transistor,” “nanowire FET,” “forksheet,”

"complementary FET," "CFET," "monolithic 3D integration," and "backside power delivery," restricted to device-architecture and VLSI-technology contexts.

### 2.2 Inclusion and Exclusion Criteria

- Included: primary research or technical-disclosure sources describing a specific transistor architecture's structure, fabrication flow, or measured/claimed electrical performance.
- Included: sources reporting node-level density, power, or performance figures relative to a named predecessor node.
- Included: English-language sources published or last updated between 2011 and 2026.
- Excluded: sources limited to circuit- or system-level applications with no device-architecture content.
- Excluded: marketing material lacking any technical specification, and duplicate reporting of a single disclosure across multiple outlets — in such cases, the earliest or most authoritative (foundry or research-institute) source was retained.

### 2.3 Synthesis Approach

Because the underlying technology is largely proprietary and rapidly evolving, this review favors triangulation across foundry technical disclosures (Intel, TSMC, Samsung), independent research-institute publications (imec), and peer-reviewed IEEE literature, rather than reliance on any single vendor's marketing claims. Reported percentage improvements in density, performance-per-watt, and power are presented as originally disclosed by the source organization, relative to the baseline node that source specifies, and should be read as vendor-reported figures rather than independently verified third-party benchmarks.

## 3 FUNDAMENTALS OF MOSFET-BASED SWITCHING

### 3.1 The MOS Structure and Threshold Behaviour

A MOSFET consists of a gate electrode separated from a semiconductor channel by a thin insulating dielectric, with source and drain terminals at either end of the channel. Applying a gate-to-source voltage beyond a device-specific threshold voltage inverts the semiconductor surface beneath the gate, forming a conductive channel between source and drain. Below threshold, the device is nominally off, though a small subthreshold leakage current still flows; above threshold, the device conducts in either a linear (resistive) or saturation regime depending on the applied drain voltage. Gate electrostatic control — the degree to which the gate, rather than the source and drain terminals, governs the channel potential — is the central design variable that every architecture discussed in this paper attempts to improve upon its predecessor.

### 3.2 PMOS and NMOS

PMOS transistors conduct via holes and turn on when the gate is held sufficiently negative relative to the source, while NMOS transistors conduct via electrons and turn on when the gate is held sufficiently positive relative to the source. Because electron mobility in silicon is roughly two to three times higher than hole mobility, an NMOS device delivers substantially more drive current than a PMOS device of identical channel dimensions; PMOS transistors are consequently sized wider in balanced logic gates to equalize the pull-up and pull-down drive strengths.

### 3.3 CMOS Logic

A CMOS inverter pairs a PMOS pull-up transistor with an NMOS pull-down transistor so that, for either logic input, exactly one of the two devices is off, blocking any direct conduction path between the supply rails. This property gives CMOS near-zero static power dissipation in the ideal case, with power consumption dominated instead by the dynamic energy expended each time the output switches. This favorable power characteristic, rather than any raw speed advantage, is the primary reason CMOS displaced NMOS-only and PMOS-only logic families as integration density — and the resulting static power burden of always-on pull-up or pull-down networks — grew through the 1980s.

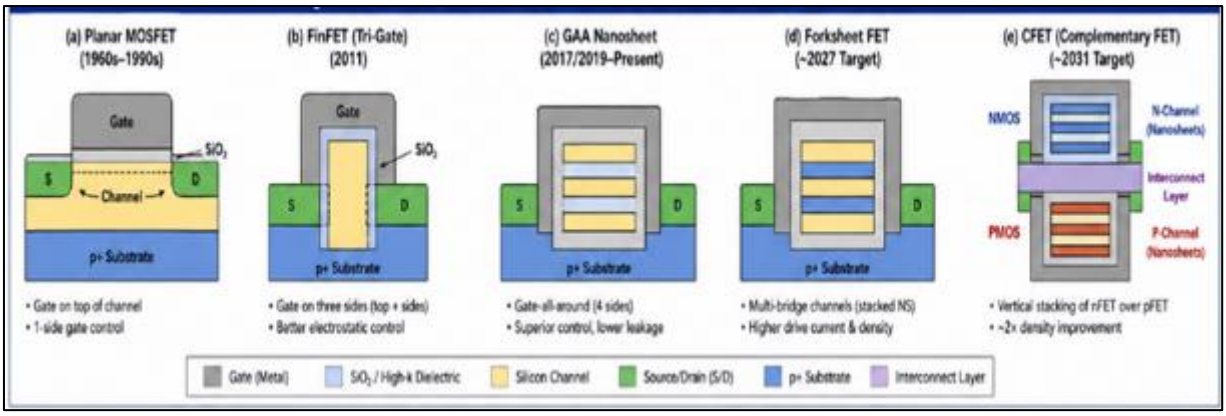



**Figure 1: Evolution of MOSFET device architectures from planar MOSFET to FinFET, GAA nanosheet, forksheet, and CFET, illustrating the progressive increase in gate-to-channel electrostatic control and the transition from lateral to vertically stacked complementary devices.**

## 4 SCALING THEORY AND THE ONSET OF SHORT-CHANNEL EFFECTS

As gate length is reduced across successive technology generations, the gate progressively loses exclusive electrostatic control over the channel, and the electric fields associated with the source and drain terminals increasingly penetrate the channel region. This produces a family of short-channel effects: threshold-voltage roll-off, in which the threshold voltage decreases as gate length shrinks; drain-induced barrier lowering (DIBL), in which the source-to-channel potential barrier is reduced by the drain voltage, increasing off-state leakage; subthreshold-slope degradation away from the ideal room-temperature value of approximately 60 mV per decade of current, which raises the leakage current needed to sustain a given on-state current; and gate-induced drain leakage at high drain fields. Collectively, these effects raise static power dissipation and reduce the on/off current ratio available to a logic gate.

The breakdown of Dennard scaling during the 90–65 nm generations in the early-to-mid 2000s was driven jointly by gate-oxide tunneling leakage — partially addressed by the introduction of high-k gate dielectrics and metal gates at the 45 nm node — and by the practical floor on further threshold- and supply-voltage reduction imposed by subthreshold leakage. These constraints set the stage for the transition from single-gate planar transistors toward multi-gate architectures capable of wrapping the gate around more than one face of the channel, restoring electrostatic control without a matching reduction in supply voltage. The remainder of this review traces that transition through each successive device architecture.

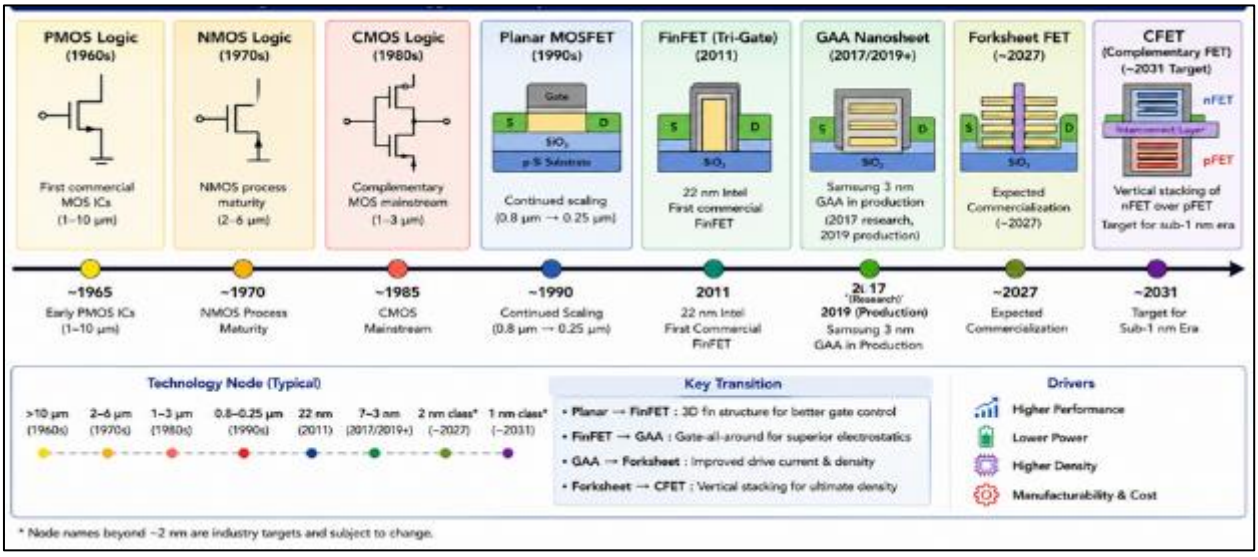



**Figure 2: Technology roadmap illustrating the evolutionary progression of MOSFET-based device architectures from early PMOS/NMOS/CMOS logic through planar MOSFET, FinFET, gate-all-around (GAA) nanosheet, forksheet, and complementary FET (CFET), with representative technology-node milestones and commercialization/research timelines.**

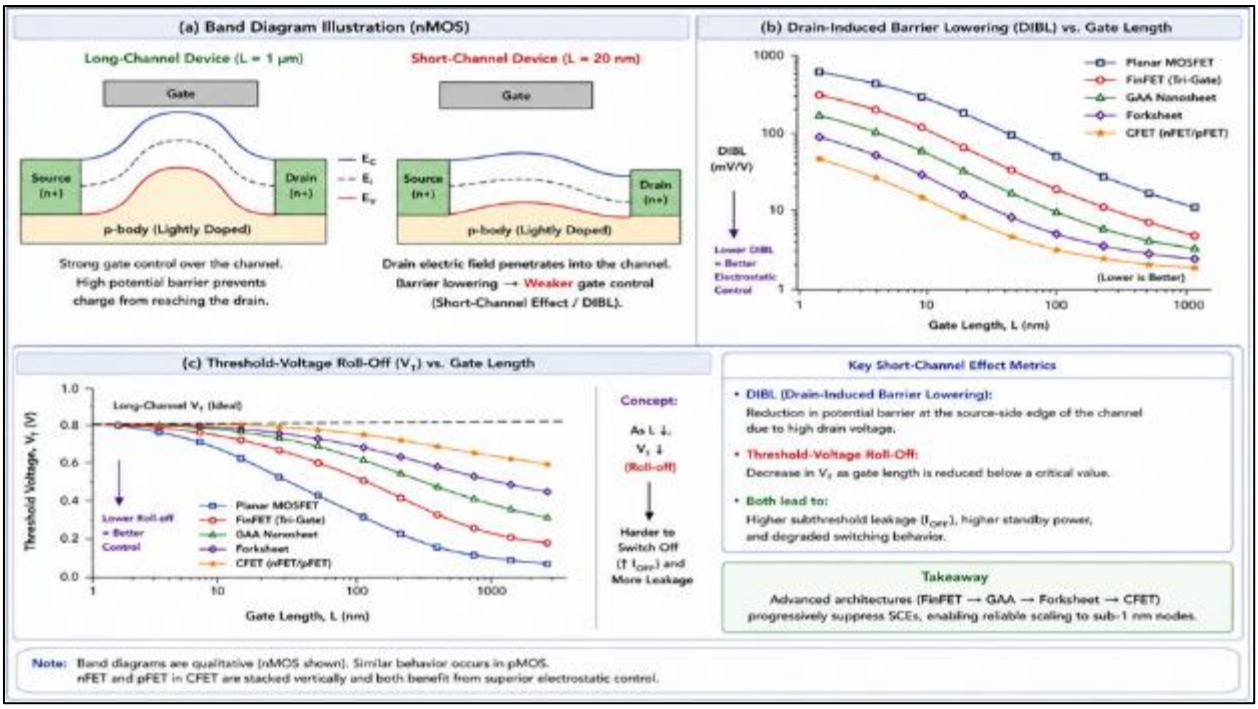



**Figure 3: Short-Channel Effects — DIBL and Threshold-Voltage Roll-Off**

## 5 EVOLUTIONARY TRAJECTORY OF DEVICE ARCHITECTURES

### 5.1 Planar Bulk and Fully-Depleted SOI MOSFET

The planar bulk MOSFET, with a single gate electrode atop a horizontal channel, was the workhorse architecture through the 32/28 nm technology generation. Its principal limitation is that the gate controls the channel from only one side (or, in fully-depleted silicon-on-insulator variants, effectively two sides), leaving the architecture increasingly susceptible to the short-channel effects described in Section 4 as gate length is reduced below roughly 28–30 nm. This deteriorating electrostatic margin, rather than any single manufacturing bottleneck, is what motivated the industry-wide transition to three-dimensional multi-gate transistors.

### 5.2 FinFET / Tri-Gate Architecture

The FinFET concept, in which the channel is raised into a thin vertical silicon “fin” so that the gate can wrap around three of its four sides, was explored in academic research from the late 1990s, with the term “FinFET” coined by researchers at the University of California, Berkeley, in the early 2000s [3]. Intel was the first company to commercialize the architecture, introducing what it branded “Tri-Gate” transistors at its 22 nm node; volume production began in late 2011, and the first commercial tri-gate processors — the Ivy Bridge family — reached the market in April 2012 [4],[5]. Competing foundries, including TSMC, Samsung, and GlobalFoundries, did not introduce their own FinFET processes until their 14 nm and 16 nm nodes in 2014–2015, a gap of more than two years behind Intel's initial commercialization [6]. At the 22 nm node, Intel's process achieved a transistor density of approximately 16.5 million transistors per square millimeter [5].

By wrapping the gate around three sides of the fin rather than one, tri-gate/FinFET architectures substantially improved subthreshold slope and DIBL relative to planar devices at equivalent gate lengths, allowing supply voltage to continue scaling down. FinFET subsequently became the industry workhorse across roughly a decade and a half of technology generations — 22 nm, 14 nm, 10 nm, 7 nm, 5 nm, and 3 nm — extended through fin-pitch scaling, multi-fin devices, channel strain engineering, and contact/via scaling, even as its fundamental three-sided gate geometry remained unchanged.

### 5.3 Gate-All-Around (GAA) Nanowire and Nanosheet FETs

As standard-cell heights and fin pitches continued shrinking toward the 3 nm and 2 nm-class nodes, the FinFET's fixed fin height began to cap further channel-width and drive-current scaling, and electrostatic control again became marginal at very short gate lengths. This motivated a transition to gate-all-around (GAA) architectures, in which the gate fully encircles the channel on all four sides. Two GAA channel geometries have been pursued industrially: vertically stacked horizontal nanosheets, which offer wider, higher-current channels per stack, and cylindrical nanowires, which are

narrower and require more vertical stacks to deliver equivalent current; the nanosheet geometry has become the industry's preferred implementation [11].

Samsung was the first foundry to commercialize a GAA architecture, branding its implementation "Multi-Bridge-Channel FET" (MBCFET) and beginning initial 3 nm production in June 2022 [7]. Samsung's own disclosures for the first-generation 3 nm process claimed up to a 35% reduction in area, a 30% improvement in performance, or a 50% reduction in power consumption relative to its 5 nm node [8]. Subsequent SF3E and SF2 process generations reportedly added further iso-power performance and power-efficiency gains together with variable nanosheet channel widths that can be tuned within a single standard cell for additional design flexibility [9],[10].

TSMC took a more conservative path, extending FinFET through three successive generations of its 3 nm family (N3, N3E, and N3P) before introducing its own GAA nanosheet technology at the N2 node, which entered volume production in the fourth quarter of 2025 [13],[14]. TSMC discloses that N2 delivers a 10–15% performance improvement at equal power, or a 25–30% power reduction at equal performance, together with roughly a 15% transistor-density increase for mixed logic/SRAM/analog designs (rising to approximately 20% for logic-only blocks), relative to N3E [13],[14].

Intel branded its own GAA implementation "RibbonFET" and paired it, for the first time in the industry at high volume, with a backside power-delivery network branded "PowerVia" on the Intel 18A node, which reached high-volume production in 2025 [15],[16]. Intel discloses up to a 15% improvement in performance-per-watt and roughly a 30% increase in transistor density relative to its prior Intel 3 node, attributing the gains to full-perimeter gate control from RibbonFET together with reduced voltage-droop and routing congestion enabled by PowerVia's nano-through-silicon-via backside power network [17],[18].

By 2026, GAA logic is, for the first time, shipping in volume from three foundries simultaneously — TSMC N2 nanosheets, Intel 18A RibbonFET with PowerVia, and Samsung SF2 MBCFET — although the process maturity, defect density, and external customer adoption reported for each differs substantially, and Samsung's first-generation 3 nm GAA process in particular faced widely reported yield challenges during its early ramp [10],[34].

### 5.4 Forksheet FET

The forksheet architecture, proposed by imec first as an SRAM-scaling booster in 2017 and later as a logic standard-cell scaling enabler in 2019, is a variant of the nanosheet transistor that inserts a dielectric wall between the n-type and p-type devices before gate patterning [19]. Because this wall serves as an etch-stop layer during the work-function-metal patterning step, it permits a substantially tighter n-to-p spacing than a conventional GAA nanosheet device can achieve, which in turn allows a wider effective channel width — and thus higher drive current — within the same standard-cell height, or a smaller cell height at equivalent drive current [20].

imec's first ("inner-wall") forksheet demonstration at the 2021 Symposium on VLSI Technology and Circuits achieved dual work-function metal gates at a 17 nm n-to-p spacing — about 35% of the spacing then achievable in state-of-the-art FinFET technology — at a 22 nm gate length, with a subthreshold swing of 66–68 mV/decade comparable to co-integrated GAA nanosheet devices fabricated on the same wafer [21]. Because the inner dielectric wall must be extremely thin (8–10 nm) to reach aggressive standard-cell heights near 90 nm, imec subsequently proposed an "outer-wall" forksheet variant that relocates the dielectric wall to the boundary between adjacent standard cells — separating same-polarity devices in neighboring cells rather than opposite-polarity devices within a single cell — which eases the manufacturability constraint at some cost in area-scaling potential [22],[23]. The forksheet is positioned as a non-disruptive bridge technology that reuses most of the nanosheet process flow, intended to extend nanosheet-class scaling toward imec's designated "A10" logic node before a transition to CFET becomes necessary [20],[21].

### 5.5 Complementary FET (CFET)

In a CFET, the n-type and p-type transistors of a standard CMOS cell are stacked vertically on top of one another rather than placed side by side, removing the n-to-p spacing entirely from standard-cell-height considerations. The area recovered by eliminating this spacing can be reinvested in wider, higher-current channels, or used to push standard-cell track heights down to four tracks or fewer — a scaling lever unavailable to any lateral (side-by-side) architecture, including the nanosheet and forksheet [24].

Two integration schemes are being pursued for CFET. In monolithic CFET (mCFET), the bottom and top device tiers are built sequentially within a single continuous process flow on one wafer, achieving a tighter vertical n-to-p separation and lower parasitic capacitance; in sequential CFET (sCFET), a separately processed bottom-tier wafer is completed first and a top-tier layer is subsequently bonded onto it via a low-temperature layer-transfer process. imec's design-technology co-optimization (DTCO) benchmarking indicates that monolithic CFET occupies less area and outperforms its sequential counterpart, which suffers a rise in effective parasitic resistance and capacitance; sequential CFET, in

exchange, decouples the thermal budgets of the two device tiers, which eases fabrication of a high-quality top-tier channel [24],[25].

The first monolithic integration of a three-dimensional CFET on a 300 mm production wafer was reported in 2020, pairing a PMOS FinFET bottom device with an NMOS nanosheet top device [28]. imec subsequently demonstrated, for the first time, electrically functional CMOS CFET devices with both bottom and top source/drain contacts, integrated at an 18 nm gate length, 60 nm gate pitch, and 50 nm vertical n-to-p separation, initially with both contacts patterned from the wafer's front side [26]. A key follow-on advance relocated bottom-contact formation to the wafer's back side, which raised the survival rate of the fragile top-tier device from roughly 11% to 79% in imec's reported process — addressing one of the architecture's principal yield-limiting fabrication steps [27].

imec's technology roadmap targets CFET introduction around the "A7" (0.7 nm-class) logic node, succeeding GAA nanosheet and forksheet devices, and imec, Intel, TSMC, and the Samsung-IBM collaboration all reported CFET progress at the 2024 IEEE International Electron Devices Meeting [27],[30]. Industry commentary suggests CFET-based commercial products are unlikely to reach volume production before approximately 2031, given the magnitude of the remaining integration, thermal, and cost challenges, and notes that the industry retains fallback scaling options should CFET ultimately fail to reach production [30].

A DTCO study presented by imec at the 2025 IEDM examined the "performance boosters" needed to hold per-transistor performance constant while area-scaling monolithic CFET across the A7, A5, and A3 nodes. The study concludes that at the A7 node, ring-oscillator performance comparable to a nanosheet-class node can be achieved chiefly by minimizing gate parasitic capacitance — for example, by reducing the gate area and the conductor area facing the gate — while at the more aggressive A3 node, mixing different channel crystal orientations for the n-type and p-type tiers becomes necessary to sustain performance. This heterogeneous-orientation integration is enabled by a newly demonstrated "embedded middle dielectric isolation" process module, and channel-orientation optimization alone was reported to raise drive current by up to 20% [2],[29].

## 6 BACKSIDE POWER DELIVERY AND DESIGN-TECHNOLOGY CO-OPTIMIZATION

Independent of the transistor architecture itself, the routing of power and signal wiring has become an increasingly significant constraint on further density and performance scaling. Traditional front-side power delivery shares the same metal layers used for signal routing, producing wiring congestion that limits both density and IR-drop performance as pitches shrink. Backside power delivery addresses this by relocating coarse-pitch power rails and bump connections to the wafer's back side, connected to each standard cell through nano-scale through-silicon vias, freeing the front side entirely for signal routing.

Intel's implementation, PowerVia, was introduced together with RibbonFET on the Intel 18A node and is reported to reduce worst-case dynamic voltage droop by as much as 10× while enabling up to an 11% block-level area compaction in routed designs [15]. TSMC's analogous "Super Power Rail" backside power scheme is planned for introduction with its A16 node, integrated with nanosheet transistors to improve logic density, routing efficiency, and IR-drop performance while preserving gate density and layout footprint relative to front-side delivery [12]. Because backside power delivery addresses the interconnect and power-network axis of scaling rather than the transistor's own electrostatic geometry, it functions as an orthogonal DTCO "booster" that can, in principle, be combined with any of the transistor architectures discussed in Section 5 — including future CFET generations, for which imec's DTCO studies already model booster techniques such as channel-orientation mixing alongside the gate-parasitic-minimization approach used with FinFET, nanosheet, and forksheet devices [2],[29].

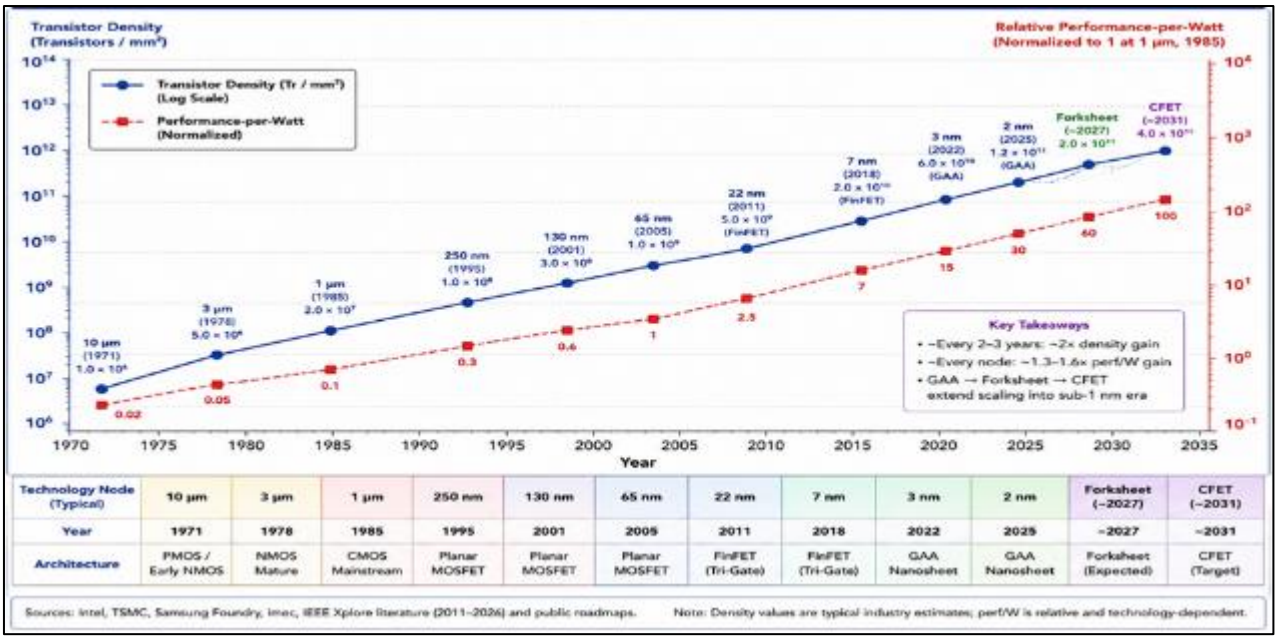
## 7 COMPARATIVE ANALYSIS

Table 1 consolidates the gate-control mechanism, representative introduction timeline, reported density or performance gains, and primary limitation of each architecture reviewed in Section 5, drawing on the foundry and research-institute disclosures cited throughout that section.

**Table 1: Comparative summary of MOSFET-based device architectures.**

| Architecture | Gate Electrostatic Control | Introduction / Representative Node | Reported Density or Performance Gain | Primary Limitation |
|---|---|---|---|---|
| Planar Bulk / FD-SOI MOSFET | Single- or double-sided gate | Baseline through ≥32/28 nm generation | N/A (reference baseline) | Severe short-channel effects below ~28 nm gate length |
| FinFET / Tri-Gate | Three-sided (wraps vertical fin) | 2011, Intel 22 nm (Ivy Bridge, 2012) [4],[5] | 16.5 million transistors/mm$^2$ at 22 nm [5]; sustained ~0.7× per-generation cadence through 3 nm | Fixed fin height caps further channel-width and drive-current scaling |
| GAA Nanosheet (MBCFET / RibbonFET / N2) | Four-sided, full gate wrap | 2022 Samsung 3 nm [7]; 2025 TSMC N2 and Intel 18A [13],[15] | TSMC N2 vs. N3E: +10–15% perf. @ iso-power, –25–30% power @ iso-perf., +15% density [13],[14]; Intel 18A vs. Intel 3: up to +15% perf./W, ~30% density [17] | Sheet width and drive current shrink as cell height scales further; early yield ramp challenges [10] |
| Forksheet FET | Forked tri-gate per polarity, dielectric-wall isolated | Proposed 2017/2019; demonstrated 2021, imec [19],[21] | 17 nm n–p spacing at 22 nm gate length, SS 66–68 mV/dec [21]; ~10% AC performance gain vs. nanosheet (simulation) [20] | Thin dielectric-wall manufacturability at aggressive cell heights [22],[23] |
| Complementary FET (CFET) | Four-sided per tier; tiers stacked vertically | First functional demo 2020–2024; roadmap target ~A7 node [27],[28],[30] | 18 nm gate length, 60 nm gate pitch, 50 nm n–p vertical separation; backside contact raised top-device survival from 11% to 79% [26],[27] | Top-tier thermal budget, bonding/contact yield, inter-tier thermal crosstalk [32],[33] |

Two patterns emerge from this comparison. First, each transition in Table 1 was triggered by the exhaustion of the previous architecture's electrostatic or geometric scaling margin: planar devices ran out of margin against short-channel effects around the 28 nm generation; FinFET's fixed fin height eventually capped further channel-width scaling; and nanosheet/forksheet devices, in turn, cannot remove the fundamental n-to-p lateral spacing that CFET eliminates by moving to a stacked topology. Second, historical device-level scaling followed a roughly consistent 0.7× dimensional-scaling cadence per generation, historically delivering on the order of a 15% performance improvement, a 50% area reduction, a 40% power reduction, and a 35% cost reduction at the device level [19]; the granular, incremental percentage gains now being reported for GAA and CFET generations (Sections 5.3 and 5.5) indicate that this cadence is increasingly sustained only through additional DTCO boosters — channel-orientation engineering, backside power delivery, and gate-parasitic minimization — layered on top of geometric scaling, rather than through geometric scaling alone.




| Technology Node (Typical) | 10 µm | 3 µm | 1 µm | 250 nm | 130 nm | 65 nm | 22 nm | 7 nm | 3 nm | 2 nm | Forksheet (~2027) | CFET (~2031) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Year | 1971 | 1978 | 1985 | 1995 | 2001 | 2005 | 2011 | 2018 | 2022 | 2025 | ~2027 | ~2031 |
| Architecture | PMOS / Early NMOS | NMOS Mature | CMOS Mainstream | Planar MOSFET | Planar MOSFET | Planar MOSFET | FinFET (Tri-Gate) | FinFET (Tri-Gate) | GAA Nanosheet | GAA Nanosheet | Forksheet (Expected) | CFET (Target) |

**Figure 4: Representative transistor-density and performance-per-watt scaling trajectory across successive MOSFET device architectures, illustrating the transition from planar MOSFETs to FinFET, GAA nanosheet, forksheet, and projected CFET architectures. The trend emphasizes the increasing contribution of device-architecture innovations and design-technology co-optimization (DTCO) to continued scaling.**

## 8 EMERGING BEYOND-CFET DIRECTIONS

### 8.1 Two-Dimensional Channel Materials

Two-dimensional (2D) semiconductors, particularly transition-metal dichalcogenides such as $MoS_2$, $WS_2$, and $WSe_2$, are being explored as channel materials whose atomic-scale thickness offers near-ideal electrostatic control and inherent immunity to the short-channel effects that constrain silicon channels as gate length shrinks [35]. The International Roadmap for Devices and Systems (IRDS) has listed 2D semiconductors as a leading candidate replacement channel material since 2018, with imec, TSMC, and Intel all pursuing dedicated 2D-semiconductor research programs [35]. The first transistor built from an exfoliated monolayer $MoS_2$ flake was demonstrated in 2011 [37], and by 2025 $MoS_2$-based CMOS circuits were reported to have entered preliminary pilot-scale fabrication, marking an early transition from laboratory research toward industrial evaluation [35]. Significant challenges remain before 2D-material transistors could reach production, including wafer-scale synthesis of large, defect-free films, high contact resistance at the metal–2D-semiconductor interface, and the difficulty of depositing high-quality gate dielectrics on a dangling-bond-free 2D surface [36],[37].

2D materials are also being investigated specifically as an enabler for future monolithic CFET and monolithic-3D generations, because their low-temperature (≤400 °C) processability is compatible with back-end-of-line integration directly onto a pre-fabricated CMOS wafer, in contrast to silicon channels, which generally require high-temperature growth steps that would damage underlying device tiers. This compatibility is motivating a shift in 2D-material integration research from front-end-of-line channel replacement toward back-end-of-line and monolithic-3D logic-on-logic stacking [38].

### 8.2 Monolithic Three-Dimensional Integration

CFET can be viewed as the first instance of a broader trend toward monolithic three-dimensional (M3D) integration, in which multiple active device tiers are stacked with fine-grained, high-density vertical interconnects rather than being placed side by side on a single plane. Extending the CFET's two-tier n-over-p stacking philosophy to additional logic or memory tiers, potentially using 2D-material channels for upper tiers to avoid thermally damaging lower silicon tiers, represents a longer-term trajectory beyond the single-pair CFET architectures discussed in Section 5.5 [38].

## 9 DISCUSSION

No single architecture reviewed here is unconditionally “best”; each represents the point at which the previous structure's geometric or electrostatic scaling margin was exhausted, and each remains in active industrial use

somewhere in the current technology-node portfolio. GAA nanosheet transistors are, as of 2024–2026, the production-proven mainstream architecture across the three leading logic foundries, while forksheet is positioned as a non-disruptive bridge that extends the nanosheet process flow without requiring the full complexity of vertical device stacking. CFET, by contrast, represents a topological rather than an incremental change: it substitutes vertical stacking for further lateral pitch-scaling, and its targeted adoption timeline — not before approximately 2031, on current industry commentary — suggests that nanosheet and forksheet variants, augmented by backside power delivery and other DTCO boosters, will remain the dominant production architectures through the remainder of this decade.

For VLSI system designers, the practical implication is that power-performance-area (PPA) gains are shifting from being dominated by raw device-geometry scaling toward being dominated by co-optimization across the device, interconnect, and power-delivery stack simultaneously — channel-orientation engineering, backside power rails, and gate-parasitic minimization are now reported as first-order contributors to node-over-node improvement alongside, rather than instead of, transistor architecture itself. This trend implies that close device–circuit–system co-design will matter at least as much as continued raw device scaling in sustaining the historical cost and performance trajectory associated with Moore's Law.

## 10 CONCLUSION

This review has traced the evolution of MOSFET-based device architectures from the foundational PMOS/NMOS/CMOS transistor through FinFET, gate-all-around nanosheet, forksheet, and complementary FET (CFET) structures, synthesizing technical literature and primary foundry disclosures spanning 2011 to 2026. Each transition has been driven by the need to restore gate electrostatic control lost to short-channel effects as the previous architecture's geometric scaling margin was exhausted. GAA nanosheet transistors, commercialized first by Samsung and now in high-volume production at TSMC and Intel as well, represent the current industrial state of the art, while CFET — which removes the n-to-p lateral spacing constraint entirely by stacking devices vertically — is the leading candidate to extend logic scaling into the sub-1 nm regime, contingent on resolving thermal-crosstalk, bonding-yield, and cost challenges that remain active research problems. This review is limited by its reliance on English-language, largely vendor- and research-institute-disclosed sources in a domain that evolves on a timescale of months; continued longitudinal review will be needed as CFET moves from research demonstration toward production over the remainder of this decade, and as two-dimensional channel materials and monolithic three-dimensional integration mature as potential successors beyond CFET.

## STATEMENTS ON COMPLIANCE WITH ETHICAL STANDARDS

### *Disclosure of conflict of interest*

The authors declare no conflicts of interest regarding this manuscript.

## REFERENCES

[1] G. E. Moore, “Cramming more components onto integrated circuits,” Electronics, vol. 38, no. 8, pp. 114–117, Apr. 1965.

[2] R. H. Dennard, F. H. Gaensslen, H.-N. Yu, V. L. Rideout, E. Bassous, and A. R. LeBlanc, “Design of ion-implanted MOSFET's with very small physical dimensions,” IEEE J. Solid-State Circuits, vol. 9, no. 5, pp. 256–268, Oct. 1974.

[3] “Understanding FinFET Technology: A Comprehensive Guide,” Power Electronics News, Jun. 2025.

[4] D. James, “Intel Ivy Bridge unveiled — the first commercial tri-gate, high-k, metal-gate CPU,” in Proc. IEEE Custom Integrated Circuits Conf. (CICC), 2012, pp. 1–4.

[5] “22 nm process,” Wikipedia, The Free Encyclopedia, rev. Aug. 2025.

[6] “Intel 14A Defect Drop Rivals 22nm Era: Customers Now Asking for Capacity, Not Data,” Tech Times, Aug. 2026.

[7] Samsung Semiconductor, “Samsung Begins Chip Production Using 3nm Process Technology With GAA Architecture,” Samsung Semiconductor Newsroom, Jun. 2022.

[8] “From FinFET to GAA: Samsung's fab journey to 3 nm and 2 nm,” EDN, Oct. 2021.

[9] Samsung Semiconductor, “3nm GAA MBCFET™: A Cutting-Edge Advancement That Brings Unrivaled SRAM Design Flexibility,” Samsung Semiconductor Tech Blog, 2023.

[10] “Samsung's 3nm Process: MBCFET Gate-All-Around Transistors — Architecture, Variants, and Key Advantages Over FinFET,” Cyberraiden Technical Blog, Mar. 2026.

[11] “The GAA Transistor Revolution: The Ultimate Weapon in the 2nm Showdown Between Samsung, TSMC, and Intel,” Aminext, Sep. 2025.

[12] TSMC, “2nm Technology,” Taiwan Semiconductor Manufacturing Company Dedicated Foundry Technology Disclosure, 2026.

[13] “TSMC's 2nm N2 process node enters production this year, A16 and N2P arriving next year,” Tom's Hardware, Apr. 2025.

[14] “TSMC Quietly Begins Volume Production of 2nm-Class Chips,” SemiWiki, Dec. 2025.

[15] Intel Corp., “Intel 18A,” Intel Foundry Process Technology Disclosure, 2026.

[16] K. Fischer et al., “Intel 18A Platform Technology Featuring RibbonFET (GAA) and PowerVia for Advanced High-Performance Computing,” presented at the IEEE/JSAP Symp. VLSI Technol. Circuits, Jun. 2025.

[17] “Intel Reclaims the Silicon Throne: 18A Node Hits High-Volume Production, Ending a Five-Year Marathon,” FinancialContent, Jan. 2026.

[18] “Panther Lake And Clearwater Forest Showcase Intel 18A Process For Next-Gen Chips,” Forbes, Oct. 2025.

[19] J. Ryckaert, “Scaling CMOS beyond FinFETs: from nanosheets and forksheets to CFETs,” imec Magazine, Dec. 2019.

[20] imec, “Entering the nanosheet transistor era,” imec Technology Articles, 2021.

[21] imec, “First electrical demonstration of integrated forksheet devices to extend nanosheets beyond 2nm technology node,” imec Press Release, Jun. 15, 2021.

[22] imec, “Outer wall forksheet: bridging nanosheet and CFET,” presented at the IEEE Symp. VLSI Technol. Circuits, Jun. 2025.

[23] “Imec's next-gen high-speed chip transistor addresses manufacturing concerns,” Tom's Hardware, Jun. 2025.

[24] imec, “CFET (complementary FET),” imec Technology Articles, Jun. 2022.

[25] imec, “Towards a process flow for monolithic CFET transistor architectures,” imec Interview Article, Jun. 2023.

[26] “Imec shows progress in CFET development,” Bits&Chips, 2024.

[27] “imec shows working CFET devices for next-generation 0.7 nm chip designs,” EE News Europe, Jun. 2024.

[28] “First Monolithic Integration of 3D Complementary FET (CFET) on 300mm Wafers,” in Proc. IEEE Int. Electron Devices Meeting (IEDM), 2020.

[29] imec, “Scaling monolithic CFET across multiple logic technology nodes,” presented at the IEEE Int. Electron Devices Meeting (IEDM), 2025.

[30] M. Lapedus, “Imec, TSMC, Samsung-IBM Make Progress With CFETs,” Semiconductor Industry Newsletter, Jan. 2025.

[31] D. Yakimets, K. K. Bhuwalka, H. Wu, G. Rzepa, M. Karner, and C. Liu, “Inflection Points in GAA NS-FET to C-FET Scaling Considering Impact of DTCO Boosters,” IEEE Trans. Electron Devices, vol. 71, no. 4, pp. 2309–2314, Apr. 2024.

[32] S. Zhao, L. Cai, W. Chen, Y. He, and G. Du, “Self-Heating and Thermal Network Model for Complementary FET,” IEEE Trans. Electron Devices, vol. 69, no. 1, pp. 11–16, Jan. 2022.

[33] B. Vermeersch et al., “Self-Heating in iN8–iN2 CMOS Logic Cells: Thermal Impact of Architecture (FinFET, Nanosheet, Forksheet and CFET) and Scaling Boosters,” in Proc. IEEE Symp. VLSI Technol. Circuits, Honolulu, HI, USA, 2022, pp. 371–372.

[34] “GAAFET and Nanosheet Transistors: A Full Guide,” InsideDeepTech, 2026.

[35] “Progress of two-dimensional semiconductor transistors and integrated circuits,” Journal of the National University of Defense Technology, 2026.

[36] “2D Material Transistors (e.g., MoS2): Beyond Silicon?,” PatSnap Eureka, 2025.

[37] “Recent advances in CMOS-compatible synthesis and integration of 2D materials,” Nano Convergence, Springer Nature, Feb. 2025.

[38] “Challenges and prospects of 2D electronics for future monolithic complementary field-effect transistors,” PubMed Central, 2025.