International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 10 Issue 4, Jul-Aug 2026 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

CMOS Technology Across Technology Nodes: A Comprehensive Review of Device Scaling, Architectural Evolution, and Future Directions

Tripti R Kulkarni, Vasudeva G

Dayananda Sagar Academy of Technology and Management, Bangalore, Karnataka, India

ABSTRACT

Complementary metal-oxide-semiconductor (CMOS) technology has served as the structural backbone of the semiconductor industry for more than five decades, evolving from micrometer-scale planar transistors to sub-2-nanometer three-dimensional architectures. This review synthesizes the trajectory of CMOS scaling across all major technology nodes, beginning with the empirical foundations laid by Moore's Law and Dennard's constant-field scaling theory, and progressing through the strain-engineered, high-k/metal-gate planar transistors of the 90-32 nm era, the tri-gate FinFET architectures that dominated the 22-7 nm nodes, and the gate-all-around (GAA) nanosheet and nanowire transistors now enabling the 5-2 nm regime. Particular attention is given to the enabling role of extreme ultraviolet (EUV) lithography, the persistent challenge of leakage power following the breakdown of Dennard scaling, and emerging post-CMOS strategies including complementary field-effect transistors (CFETs), monolithic 3D integration, and two-dimensional (2D) channel materials such as transition metal dichalcogenides. By critically examining device physics, process innovations, and the economic constraints that shape roadmap decisions, this review identifies electrostatic control, contact resistivity, and heterogeneous material integration as the dominant bottlenecks for future scaling and outlines the architectural pathways most likely to sustain transistor density growth beyond the 1 nm node.

KEYWORDS: *CMOS scaling; Moore's Law; Dennard scaling; FinFET; gate-all-around (GAA); nanosheet transistor; EUV lithography; high-k metal gate; complementary FET (CFET); 2D semiconductors.*

How to cite this paper: Tripti R Kulkarni | Vasudeva G "CMOS Technology Across Technology Nodes: A Comprehensive Review of Device Scaling, Architectural Evolution, and Future Directions" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-10 | Issue-4, August 2026, pp.939-944, URL: www.ijtsrd.com/papers/ijtsrd142126.pdf

IJTSRD142126

1. INTRODUCTION

The relentless miniaturization of the metal-oxide-semiconductor field-effect transistor (MOSFET) has been the single most important driver of the digital revolution. Since Gordon Moore first observed that the number of components on an integrated circuit was doubling on a regular cadence, the semiconductor industry has organized itself around sustaining that trend, commonly referred to as Moore's Law (Moore, 1965). What began as an economic observation about component density evolved into a self-reinforcing industrial roadmap that dictated investment, research priorities, and competitive strategy across the entire electronics value chain.

CMOS technology, in which paired n-type and p-type MOSFETs are combined to realize logic functions with near-zero static power dissipation, became the dominant implementation technology because of this scalability. For approximately three decades, transistor scaling followed the constant-field rules articulated by Dennard, Gaensslen, Yu, Rideout, Bassous, and LeBlanc (1974), under which shrinking the physical dimensions of a transistor by a factor S while simultaneously reducing voltage by the same factor kept power density constant and delivered simultaneous gains in speed, density, and energy efficiency. This review traces CMOS evolution across all major technology nodes, from early micron-scale

planar devices through the introduction of strained silicon and high-k/metal-gate stacks, the transition to three-dimensional FinFET architectures, and the current generation of gate-all-around (GAA) nanosheet transistors, before examining the material and architectural innovations positioned to extend scaling beyond the 1 nm node.

2. Historical Foundations: Moore's Law and Dennard Scaling

Moore's original 1965 paper extrapolated five data points to predict that the number of components per integrated circuit would double annually, a rate he later revised to roughly every two years (Moore, 1965). The economic logic underpinning this prediction was that a shrinking minimum feature size reduced the cost per transistor even as it increased the number of devices that could be placed on a single die, creating a durable incentive for the entire industry to invest in lithographic and process advances aligned to a shared cadence.

Dennard et al. (1974) provided the physical scaling theory that made Moore's Law technologically realizable for CMOS logic. By simultaneously scaling gate length, gate oxide thickness, junction depth, and supply voltage by a common factor while increasing substrate doping, the electric field within the transistor channel remained constant across generations, allowing switching speed to increase, power density to remain flat, and static leakage to stay negligible. As documented in retrospective analyses of the scaling paradigm, Dennard's own paper anticipated a fundamental limitation: subthreshold slope does not scale with the same factor as other device dimensions, meaning that below a certain gate length, further voltage reduction causes off-state leakage current to grow exponentially rather than remaining negligible (Bohr, 2007). This limitation, largely dormant while feature sizes remained above roughly 130 nm, became the central constraint governing CMOS design from the 90 nm node onward and is discussed further in Section 6.

3. Planar CMOS Scaling and Its Physical Limits

Through the micron and deep-submicron eras (10 µm down to approximately 130 nm), CMOS scaling proceeded largely as classical geometric shrinkage of the bulk planar MOSFET: shorter gate lengths, thinner silicon dioxide gate dielectrics, and shallower source/drain junctions. As gate lengths fell below roughly 100 nm, short-channel effects (SCEs) — including drain-induced barrier lowering (DIBL), threshold-voltage roll-off, and degraded subthreshold swing — began to erode the electrostatic integrity that planar single-gate control could provide (Radamson et al., 2024). Gate oxide thickness scaling, essential for maintaining channel control, simultaneously drove direct quantum-mechanical tunneling current through the gate dielectric to unacceptable levels, motivating the search for alternative dielectric materials with higher permittivity than SiO_2.

4. High-k/Metal-Gate Integration and Strain Engineering (90-32 nm)

The 45 nm node marked a watershed in CMOS process technology with the first high-volume introduction of a hafnium-based high-k gate dielectric paired with dual metal gates, replacing the polysilicon-gate/SiO_2 stack that had been used since the invention of the MOSFET (Auth et al., 2008). The high-k/metal-gate (HKMG) approach allowed a physically thicker dielectric layer to be used while maintaining an equivalent oxide thickness compatible with continued scaling, reducing gate leakage by roughly three orders of magnitude for PMOS and 25-fold for NMOS relative to a SiO_2 baseline, while dual band-edge work-function metals eliminated the polysilicon depletion effect that had limited drive current in earlier generations (Robertson & Wallace, 2015).

In parallel, strain engineering emerged as the principal lever for boosting carrier mobility without further geometric shrinkage. Techniques such as embedded silicon-germanium (eSiGe) source/drain stressors for PMOS, tensile contact-etch-stop liners for NMOS, and stress-memorization techniques were combined to increase channel carrier mobility and drive current at a fixed gate length (Auth et al., 2008). Tri-gate transistor prototypes incorporating high-k dielectrics, metal gates, and strain engineering simultaneously demonstrated that three-dimensional channel geometries could further improve short-channel control while preserving the mobility benefits of strain (Kavalieros et al., 2006), foreshadowing the architectural transition that would follow at the 22 nm node.

5. The FinFET Era (22-7 nm)

Beginning around the 22 nm node, the industry transitioned from planar MOSFETs to the fin field-effect transistor (FinFET), a tri-gate architecture in which the gate wraps around three sides of a thin, vertical silicon fin that forms the channel. This geometry dramatically improves electrostatic control relative to a planar device of comparable footprint, suppressing DIBL and subthreshold leakage while permitting continued reduction of the contacted gate pitch (Wang et al., 2019, as reviewed in comparative CMOS scaling literature on FinFET and beyond architectures). FinFET technology proved sufficiently scalable to serve as the workhorse architecture from the 22 nm through the 7 nm nodes, with successive

generations refining fin height, fin pitch, and strain integration to sustain performance and density gains (Radamson et al., 2024).

Nonetheless, as contacted gate pitch continued to shrink, FinFETs began to encounter their own limits: fin-width variability, parasitic capacitance from tightly packed fins, and, critically, rising contact resistivity as the contact area between metal and silicon shrank in proportion to the pitch. Analyses of contact scaling indicate that sustaining FinFET performance below a 30 nm contacted gate pitch would require specific contact resistivities on the order of 1×10^{-10} Ω·cm², a target that pushed materials and interface engineering to its practical limits and motivated a transition to device architectures in which contact area is decoupled from fin or gate pitch (as documented in comparative reviews of FinFET scaling challenges).

6. Gate-All-Around Architectures: Nanosheet and Nanowire FETs (5-2 nm)

At the 5 nm and 3 nm nodes, the gate-all-around (GAA) transistor, in which the gate electrode fully encloses the channel on all sides, succeeded the FinFET as the leading architecture. Implemented predominantly as stacked silicon nanosheets or nanowires, GAA devices provide near-ideal electrostatic control of the channel, enabling further reduction of effective gate length while suppressing DIBL and improving subthreshold swing relative to FinFETs of comparable footprint (Zhang, Zhang, Luo, & Yin, 2024). Because nanosheet width can be tuned independently of fin height, GAA architectures also offer a drive-current-versus-density trade-off that is not available in fin-based designs, while remaining largely compatible with established FinFET process flows, which has eased industrial adoption (Zhang et al., 2024). Authors of paper[11,12] discuss on FINFET technology for 22nm.

Beyond the 2-3 nm regime, vertical transistor stacking has emerged as the next architectural lever. Complementary FETs (CFETs), in which an NMOS device is monolithically stacked atop a PMOS device to form a single logic cell, have been shown in device-level modeling to reduce standard-cell layout area by roughly 25 percent and to increase pin density correspondingly relative to FinFET-based layouts, while offering greater routing flexibility (Zhang et al., 2024; comparative CFET modeling literature). Because Moore's Law-style geometric shrinkage is not expected to remain the primary density lever beyond the 1 nm node, such vertical integration strategies — including 3D-stacked FETs and vertical-channel transistors — represent the industry's principal near-term response to the exhaustion of lateral scaling (Zhang et al., 2024).

7. Lithography as an Enabling Technology: Extreme Ultraviolet (EUV)

Continued reduction of minimum feature size has depended as much on lithographic capability as on device architecture. Extreme ultraviolet (EUV) lithography, operating at a 13.5 nm exposure wavelength compared with the 193 nm wavelength of deep ultraviolet (DUV) immersion systems, became the enabling patterning technology for the 7 nm node and below, first deployed in high-volume manufacturing by leading foundries in 2018. By replacing multiple DUV multi-patterning exposures with a single EUV exposure for critical layers, EUV reduces process complexity, mask count, and the associated overlay error accumulated across multiple patterning steps, while imposing new challenges related to limited source power, stochastic resist effects at the smallest feature sizes, and mask defectivity. EUV systems remain capital-intensive, with individual scanners costing well over one hundred million dollars, concentrating access to leading-edge lithography among a small number of foundries. [18] this says design of VCO with output peak to peak

8. Power, Leakage, and the End of Dennard Scaling

As transistor dimensions shrank below approximately 90-65 nm, the assumptions underlying Dennard's scaling theory began to break down. Threshold voltage could no longer be scaled proportionally with supply voltage without causing subthreshold leakage current to grow exponentially, since subthreshold swing does not improve with geometric scaling in the way channel length or oxide thickness does (Bohr, 2007). [19,20,21]This paper presents FGMSVM with one-against-one and maximum voting, using K-means clustering for noninvasive cardiovascular disease screening. It also explores a deep learning method for detecting five arrhythmia types via PPG and surveys approaches for noninvasive fetal oxygen saturation measurement using PPG Sub-threshold leakage, which was negligible relative to dynamic switching power in devices with structures larger than roughly 130 nm, has since become a substantial fraction of total chip power, with contemporary analyses attributing on the order of 40 percent of total power dissipation in deep-submicron circuits to various leakage mechanisms (IEEE conference literature on low-power and low-leakage CMOS design).

The practical consequence, often termed the end of Dennard scaling, is that power density is no longer constant across process generations: shrinking a

design at fixed supply voltage now increases power density, forcing designers either to reduce clock frequency, disable portions of the chip during operation (the so-called "dark silicon" phenomenon), or adopt heterogeneous and specialized architectures to control energy consumption. Multiple circuit-level mitigation strategies — including power gating, multi-threshold CMOS, body biasing in fully depleted silicon-on-insulator (FDSOI) technology, and transistor-stacking techniques for leakage suppression — have been developed to partially compensate for this loss of a free scaling dimension, but none fully restores the power-density constancy that characterized the classical Dennard regime.[13,14,16] papers tell us about speech processing ,cancer diseases and how to sanitize the place

9. Beyond Silicon: 2D Materials and Heterogeneous Integration

As silicon channel thickness in nanosheet devices approaches a few nanometers, carrier mobility degradation from surface roughness scattering and short-channel tunneling currents becomes increasingly difficult to manage, motivating research into channel materials with intrinsically superior electrostatic scalability. Two-dimensional transition metal dichalcogenides (TMDs) such as molybdenum disulfide (MoS_2) and tungsten diselenide (WSe_2) have attracted particular attention because their atomically thin, dangling-bond-free structure suppresses short-channel effects even at extreme gate-length scaling, while offering a sizeable bandgap and favorable on/off current ratios not available in graphene (Zheng, Meng, & Li, 2025).

Recent demonstrations have integrated large-area n-type MoS_2 and p-type WSe_2 field-effect transistors into complementary logic circuits, achieving tailored threshold voltages and reduced subthreshold leakage through high-κ gate dielectric integration and optimized channel-length scaling, illustrating a credible pathway toward 2D-channel CMOS logic (as reported in recent 2D CMOS integration studies). Nonetheless, TMD-based CMOS still faces substantial integration barriers, including wafer-scale single-crystal growth, low-resistance contact formation, gate-oxide interface quality, and device-to-device variability, all of which must be resolved before 2D-channel transistors can be considered a manufacturable alternative to silicon nanosheets (Zheng et al., 2025). [17] speaks how subthreshold leakage happens in MOSFETs.

10. Challenges and Future Outlook

Several interlocking constraints will shape CMOS development beyond the 2 nm node. First, contact and interconnect resistance are increasingly dominant contributors to overall delay as cross-sectional dimensions shrink, requiring new low-resistivity metals, liner-free contacts, or backside power delivery schemes to prevent parasitic resistance from negating the benefits of smaller transistors. Second, the economic cost of each successive node — driven by EUV capital expenditure, multi-patterning complexity, and design-for-manufacturability overhead — has grown substantially, concentrating leading-edge fabrication among very few foundries and raising questions about the long-term sustainability of Moore's Law as a purely economic proposition. Third, architectural innovation, rather than pure geometric shrinkage, is increasingly the primary source of density and performance gains, as reflected in the industry's transition toward CFETs, monolithic 3D stacking, and backside power delivery networks (Zhang et al., 2024). [15] speaks of LC oscillator using machine learning.

Looking forward, near-term roadmaps (through the 1.4 nm and 1 nm nodes) are expected to rely on further refinement of GAA nanosheet transistors, the introduction of CFET stacking, and backside power rails, while longer-term research continues to explore 2D channel materials, negative-capacitance and ferroelectric gate stacks for sub-thermionic switching, and tunneling FETs as potential successors once conventional MOSFET electrostatics reach their physical limits (Radamson et al., 2024; Zhang et al., 2024).

11. Conclusion

CMOS technology has sustained more than five decades of exponential transistor scaling through a sequence of distinct innovation phases: classical geometric shrinkage of the planar MOSFET, the introduction of high-k/metal-gate stacks and strain engineering, the transition to three-dimensional FinFET architectures, and, most recently, gate-all-around nanosheet transistors. Each transition was driven by the exhaustion of the electrostatic or material margins of the preceding architecture, and each was enabled by parallel advances in lithography, materials science, and process integration. With lateral geometric scaling approaching fundamental limits, the industry's continued progress increasingly depends on vertical integration strategies such as CFETs and monolithic 3D stacking, alongside longer-horizon material innovations including 2D channel semiconductors. Sustaining the pace of improvement that has defined the CMOS era will require coordinated advances across device physics, lithography, materials engineering, and circuit design, reinforcing that the future of Moore's Law is now an

architectural and materials challenge as much as a purely dimensional one.

References

[1] Auth, C., Cappellani, A., Chun, J.-S., Dalis, A., Davis, A., Ghani, T., Glass, G., Glassman, T., Harper, M., Hattendorf, M., Hentges, P., Jaloviar, S., Joshi, S., Klaus, J., Kuhn, K., Lavric, D., Lu, M., Mariappan, H., Mistry, K., ... Bai, P. (2008). 45nm high-k + metal gate strain-enhanced transistors. In 2008 Symposium on VLSI Technology, Digest of Technical Papers (pp. 128–129). IEEE.

[2] Bae, G., Bae, D.-I., Kang, M., Hwang, S. M., Kim, S. S., Seo, B., Kwon, T. Y., Lee, T. J., Moon, C., Choi, Y. M., Oikawa, K., Masuoka, S., Chun, K. Y., Park, S. H., Shin, H. J., Kim, J., Bae, S., Park, K.-M., Yoon, J. S., ... Jang, S. (2018). 3nm GAA technology featuring multi-bridge-channel FET for low power and high performance applications. In 2018 IEEE International Electron Devices Meeting (IEDM) (pp. 28.7.1–28.7.4). IEEE.

[3] Bohr, M. (2007). A 30-year retrospective on Dennard's MOSFET scaling paper. IEEE Solid-State Circuits Society Newsletter, 12(1), 11–13.

[4] Colinge, J.-P. (2008). FinFETs and other multi-gate transistors. Springer.

[5] Dennard, R. H., Gaensslen, F. H., Yu, H.-N., Rideout, V. L., Bassous, E., & LeBlanc, A. R. (1974). Design of ion-implanted MOSFET's with very small physical dimensions. IEEE Journal of Solid-State Circuits, 9(5), 256–268.

[6] Kavalieros, J., Doyle, B., Datta, S., Dewey, G., Doczy, M., Jin, B., Lionberger, D., Metz, M., Rachmady, W., Radosavljevic, M., Shah, U., Zelick, N., & Chau, R. (2006). Tri-gate transistor architecture with high-k gate dielectrics, metal gates and strain engineering. In 2006 Symposium on VLSI Technology, Digest of Technical Papers (pp. 50–51). IEEE.

[7] Loubet, N., Hook, T., Montanini, P., Yeung, C.-W., Kanakasabapathy, S., Guillom, M., Yamashita, T., Zhang, J., Miao, X., Wang, J., Young, A., Metz, A., Katragadda, R., Kim, Y., Brand, A., Radens, C., Boyd, D., Waskiewicz, S., Rankin, J., ... Khare, M. (2017). Stacked nanosheet gate-all-around transistor to enable scaling beyond FinFET. In 2017 Symposium on VLSI Technology, Digest of Technical Papers (pp. T230–T231). IEEE.

[8] Moore, G. E. (1965). Cramming more components onto integrated circuits. Electronics, 38(8), 114–117.

[9] Natarajan, S., Agostinelli, M., Akbar, S., Bost, M., Bowonder, A., Chikarmane, V., Chouksey, S., Dasgupta, A., Fischer, K., Fu, Q., Ghani, T., Giles, M., Govindaraju, S., Grover, R., Han, W., Hanken, D., Haralson, E., Haran, M., Heussner, M., ... Bai, P. (2014). A 14nm logic technology featuring 2nd-generation FinFET, air-gapped interconnects, self-aligned double patterning and a 0.0588 µm² SRAM cell size. In 2014 IEEE International Electron Devices Meeting (IEDM) (pp. 3.7.1–3.7.3). IEEE.

[10] Radamson, H. H., Miao, Y., Zhou, Z., Wu, Z., Kong, Z., Gao, J., Yang, H., Ren, Y., Zhang, Y., Shi, J., Xiang, J., Cui, H., Lu, B., Li, J., Liu, J., Lin, H., Xu, H., Li, M., Cao, J., He, C., Duan, X., Zhao, X., Su, J., Du, Y., Yu, J., Wu, Y., Jiang, M., Liang, D., Li, B., Dong, Y., & Wang, G. (2024). CMOS scaling for the 5 nm node and beyond: Device, process and technology. Nanomaterials, 14(10), 837. https://doi.org/10.3390/nano14100837

[11] Vasudeva, G. et al. (2025). Design of OPAMP Based on 22 nm FinFET Technology. In: Kaiser, M.S., Singh, R., Bandyopadhyay, A., Mahmud, M., Ray, K. (eds) Proceedings of the Fifth International Conference on Trends in Computational and Cognitive Engineering. TCCE 2023. Lecture Notes in Networks and Systems, vol 1208. Springer, Singapore. https://doi.org/10.1007/978-981-96-0185-1_32

[12] Vasudeva G et.al "Design of FinFET Based Op-Amp Using High-K Device 22 nm Technology", Applied Mathematics, Modeling and Computer Simulation, 2023, doi:10.3233/ATDE231033

[13] Tripti R, Kulkarni and Bharathi, Gururaj and Aditi, Jaiswal (2025) Comparison of Implementation in Blood Cancer Causes and Diseases. International Journal of Trend in Scientific Research and Development, 9 (1). pp. 503-512. ISSN 2456-6470

[14] Tripti R, Kulkarni and Arun Vikas, Singh and Vasudeva, G and Bharathi, Gururaj (2025) *Sanitization and Sterilization Robot Ultra Violet Based.* International Journal of Trend in Scientific Research and Development, 9 (2). pp. 1004-1010. ISSN 2456-6470

[15] Jatkar, M., G, V., Kulkarni, T.R. et al. LC oscillator frequency prediction using machine

learning linear regression algorithm. Discov Electron 2, 52 (2025). https://doi.org/10.1007/s44291-025-00092-9

[16] Kulkarni, T.R., Vasudeva, G., Renukaprasad, G., Roopa, K.R., Srividya, B.V., Jatkar, M. (2026). Speech Detection of Pathological Voice Disorders Non-Invasively Using Machine Learning Algorithms. In: Joshi, A., Ragel, R.G., S., K. (eds) ICT: Applications and Social Interfaces. ICTCS 2025. Lecture Notes in Networks and Systems, vol 1874. Springer, Cham. https://doi.org/10.1007/978-3-032-19681-1_2

[17] Kulkarni, T.R. et al. (2026). Modeling Subthreshold Leakage in MOSFETs: Supervised Regression with Noise-Aware Synthetic Datasets. In: Senjyu, T., Mahmud, M., Joshi, A. (eds) Smart Trends in Computing and Communications. SmartCom 2026. Lecture Notes in Networks and Systems, vol 1990. Springer, Cham. https://doi.org/10.1007/978-3-032-26908-9_42

[18] Srivathsava N. L & Tripti Kulkarni," Novel Design of VCO with Output Peak to Peak Control", International Journal of Instrumentation, Control and Automation (IJICA) ISSN : 2231-1890 Volume-1, Issue-2, 2011

[19] (Kulkarni, 2021)Tripti R Kulkarni, N.D. Dushyant, Early and noninvasive screening of common cardio vascular related diseases such as diabetes and cerebral infarction using photoplethysmography signals, Results in Optics, Volume 3,2021,100062,ISSN 2666 9501,https://doi.org/10.1016/j.rio.2021.100062.

[20] Performance evaluation of deep learning models in detection of different types of arrhythmia using photo plethysmography signals TR Kulkarni, ND Dushyanth - International Journal of Information Technology, 2021,https://doi.org/10.1007/s41870-021-00795-8

[21] T. Kulkarni and B. N. Krupa, "Noninvasive method to find oxygen saturation of the fetus using photoplethysmogram," 2016 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT), Mysuru, India, 2016, pp. 40-42, doi: 10.1109/ICEECCOT.2016.7955182