Post-CFET Device Architectures: Materials, Integration, and Design Perspectives

Shinichi Samizo
Independent Semiconductor Researcher
Project Design Hub, Samizo-AITL
Email: shin3t72@gmail.com *GitHub:* Samizo-AITL

***Abstract*—CMOS scaling has evolved from planar MOSFETs to FinFETs, Gate-All-Around (GAA) nanosheets and Complementary FETs (CFETs). While CFET improves electrostatic control and mitigates local wiring bottlenecks, further gains are constrained by material, thermal and reliability limits of silicon. This paper surveys post-CFET device options—two-dimensional (2D) material FETs, monolithic 3D (M3D) integration, spintronics/quantum devices, and heterogeneous atomic-scale integration. We discuss their physical principles, process and integration challenges, demonstrated performance, reliability concerns, design/EDA implications, and educational aspects. A comparison matrix and a 2030–2045 roadmap are provided.**

I. Introduction

Over five decades, the industry advanced by shrinking devices and re-architecting transistors: Planar → FinFET → GAA → CFET. Dennard scaling has long broken; beyond ∼3 nm nodes, *interconnect delay and power density* dominate, and device self-heating and variability constrain further gains. Wiring RC delay already exceeds intrinsic device delay for many paths; thermal design power densities approach $1\,\mathrm{W/mm^2}$ in high-performance logic, stressing reliability (BTI, EM) and packaging. Consequently, progress increasingly relies on **materials innovation**, **vertical integration**, and **alternative state variables** (spin, photon, charge+heat co-design) rather than lateral device scaling alone. This work positions "post-CFET" as that transition and compares four concrete vectors.

II. Evolution from CMOS to Post-CFET

Fig. 1 sketches the pathway from CMOS to post-CFET candidates.

A. Planar MOSFET to FinFET

Below 45 nm gate lengths, short-channel effects (SCE) and junction leakage degraded SS and I_{off}. Tri-gate FinFETs provided stronger gate coupling and reduced SCE, enabling continued voltage scaling and variability control.

B. FinFET to GAA

To further compress cell height and suppress corner leakage, stacked nanosheets fully surrounded by the gate were introduced. GAA improves electrostatics and V_T tunability, at the cost of increased process complexity (sacrificial layers, release uniformity) and tighter variability control (line-edge roughness, work-function granularity).

C. GAA to CFET

CFET vertically stacks complementary devices (nFET over pFET or vice versa), improving cell density and reducing local interconnect delay by gate-level vertical proximity. Key challenges include *thermal budget management*, *alignment of stacked devices*, and *contact/via resistance* between tiers.

D. Beyond CFET

As wiring dominance and heat density increase, simply stacking silicon devices hits diminishing returns. Post-CFET options expand along three axes: (i) introduce new channel or functional materials (2D/TMDs), (ii) pursue true *monolithic* 3D integration to shorten interconnects and co-locate memory/logic, and (iii) adopt *non-charge* state variables (spin) and cross-domain fusion (photonics/MEMS/bio).

III. Post-CFET Candidate Technologies

A. 2D Material FETs

Principle: Atomically thin semiconductors (e.g., MoS_2, WSe_2) provide excellent electrostatic control and suppressed short-channel leakage owing to sub-nm body thickness. Van der Waals interfaces relax lattice matching constraints, enabling heterogeneous stacks.

State-of-the-art: Lab-scale devices with $L_g \sim$ 10–20 nm show $I_{on}/I_{off} \sim 10^7$ and $SS \sim$ 60–70 mV/dec. Contact resistivity remains high ($R_c \sim 0.5$–$1\,\mathrm{k\Omega \cdot \mu m}$); wafer-scale CVD growth shows 5 %–10 % non-uniformity and grain-boundary-induced variability.

Integration: Low-temperature BEOL-compatible processes (<450 °C) are desired for sequential stacking or CFET+2D hybrids. Key knobs are phase/defect control, substitutional/charge-transfer doping, metal/TMD interface engineering, and clean transfer (or direct growth) on dielectrics and Cu-compatible liners.

Reliability: Interface traps, charge trapping/de-trapping, and self-heating through low-k dielectrics impact BTI-like shifts and mobility. Environmental stability (oxidation, moisture) requires encapsulation (e.g., h-BN, ALD Al_2O_3).

Design/EDA implications: Compact models must capture quantum confinement, Schottky contacts, and variability (grain boundaries, thickness steps). PDKs need parameter corners for R_c, mobility, contact asymmetry, and layout rules addressing transfer seams and 2D-specific proximity effects.

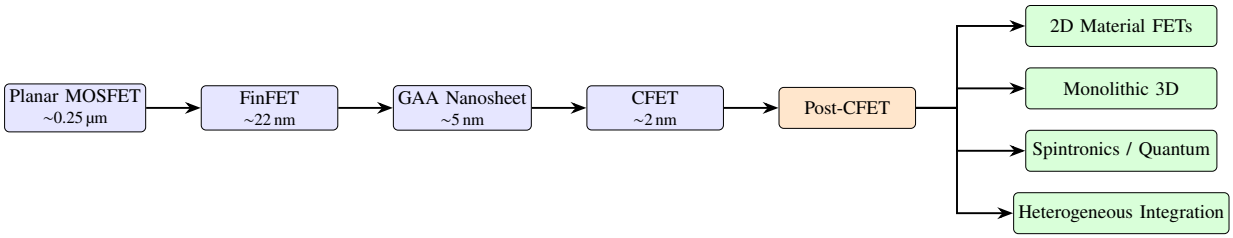

Fig. 1. Evolution tree: CMOS → CFET → post-CFET candidates.

Applications: Ultra-low-power edge nodes, flexible/bio electronics, steep-slope concepts and sensor-logic co-integration. Near term, 2D devices likely *complement* CFETs (e.g., analog sensing layers atop logic).

B. Monolithic 3D Integration (M3D)

Principle: Sequential device fabrication forms multiple active layers on a single wafer with fine-pitch inter-layer vias (ILVs), shortening interconnects by $\gg 100\times$ versus 2.5D/TSV stacks.

State-of-the-art: Demonstrations include 3D SRAM/logic with delay reductions (~30%) and area savings (~40%), as well as memory-centric AI tiles showing $>1.5\times$ system energy efficiency.

Integration: Main constraints are **thermal budget** (<450 °C after the first tier), dopant activation at low-T, and **alignment/yield**. Laser/flash anneal, solid-phase epitaxy, and low-T materials (oxide semiconductors, 2D) are active areas. ILV pitch < 200 nm with acceptable resistance and electromigration margins is a key milestone.

Thermal/mechanical: Vertical heat flow and local hotspots require thermal-aware placement, TSV/ILV planning, and stress-aware signoff (Coffin–Manson fatigue, wafer bow).

Design/EDA implications: 3D placement & routing with layer assignment, ILV budgeting, thermal and stress co-simulation, and timing closure that includes vertical parasitics. Power delivery network (PDN) partitioning across tiers and integrated thermal throttling policies are needed.

Applications: AI accelerators (near-/in-memory compute), memory-wall mitigation for graph/transformer workloads, and low-latency sensor fusion.

C. Spintronics / Quantum Devices

Principle: Magnetic tunnel junctions (MTJs) exploit spin-dependent tunneling (CoFeB/MgO stacks), enabling non-volatile memory (STT/SOT-MRAM) and logic-in-memory concepts. Quantum/topological devices explore spin–orbit coupling and edge states.

State-of-the-art: STT-MRAM endurance of 10^{12} cycles at room temperature; SOT-MRAM reduces write latency and improves endurance by separating read/write paths, at the expense of area/driver complexity.

Integration: BEOL-friendly temperatures (<400 °C) are feasible; challenges include write current reduction (mA → μA), variability of H_k/TMR, and read disturb. Hybrid stacks (e.g., MRAM over logic tiers in M3D) enable ultra-fast, non-volatile buffers and checkpointing.

Design/EDA implications: Compact models must include stochastic switching, retention distributions, and temperature dependence. For compute-in-memory (CIM), EDA flows need array-level analog behavior, error-rate aware mapping, and ECC co-design.

Applications: Radiation-tolerant systems (space/avionics), last-level cache augmentation, always-on inference, neuromorphic primitives.

D. Heterogeneous Atomic-Scale Integration

Principle: Fuse dissimilar domains (Si CMOS + photonics + MEMS + sensors + 2D) by leveraging van der Waals interfaces, direct wafer bonding, or hybrid bonding.

State-of-the-art: Examples include Si+MoS_2 photodetectors (responsivity ~ 200 mA/W at 1.55 µm), CMOS+MEMS inertial sensors, and on-chip photonics for low-latency I/O.

Integration: Interface cleanliness/planarity, CTE/lattice mismatch, and bonding yield dominate. Cross-domain verification (optical S-parameters, mechanical eigenmodes) must be integrated with electrical timing and power.

Design/EDA implications: Compact models spanning domains, co-simulation backplanes (SPICE + FDTD/RCWA + FEM), and PDKs that provide optical/mechanical layer stacks, DRC, and reliability rules.

Applications: Optical interconnect, medical/industrial sensing, aerospace instruments, and co-packaged optics.

TABLE I
COMPARISON OF POST-CFET CANDIDATES (REPRESENTATIVE, NOT EXHAUSTIVE)

Tech.	Demo (rep.)	Bottlenecks	Reliability	Design impact
2D-FET	$I_{on}/I_{off} \sim 10^7$, $SS \sim 65$ mV/dec	R_c, film uniform.	Traps, self-heat	New compact models, var. kits
M3D	Delay -30%, Area -40%	$<450^\circ$C, ILV R	Hotspots, stress	3D P&R, thermal/stress co-sim
Spin	MRAM 10^{12} cyc	I_{write}, TMR var.	Retention, disturb	Stochastic models, ECC/CIM
Hetero	Si+2D PD 200 mA/W	Bonding, CTE mismatch	Interface aging	Cross-domain EDA/PDK

IV. Comparison and Positioning

Table I contrasts each vector from *device*, *integration*, and *design* angles. A pragmatic path is hybrid: CFET for density, M3D for distance reduction, spintronics for non-volatility, and 2D materials or photonics for sensing/I/O. The *system* wins when interconnect energy is minimized and memory/compute locality is maximized.

V. Design and Educational Perspectives

A. EDA/PDK Requirements

- **Compact models:** Quantum confinement, Schottky contacts (2D), stochastic switching (spin), thermo–electro–mechanical coupling (M3D/hetero).
- **3D physical design:** Tier-aware floorplanning, ILV budgeting, thermal-aware placement, vertical timing/parasitics extraction.
- **Multi-physics co-simulation:** Electrical–thermal–mechanical–optical closed loops with sign-off criteria (hotspot temp., stress limits, optical loss budgets).
- **Reliability sign-off:** BTI/HCI/EM with temperature gradients, MRAM retention/disturb models, interface aging for bonded stacks.
- **Variability/DFM:** Film thickness, grain boundary density, R_c spread, bonding voids; provide statistical corners and Monte Carlo kits.

B. Curriculum and Workforce Development

Graduate curricula should include: (i) device-to-system co-design, (ii) variability-aware modeling labs, (iii) 3D PDN/thermal projects, (iv) cross-domain capstone (e.g., CMOS+photonics+MEMS). Industry–academia pilot lines (multi-project wafers, shuttle services) can accelerate translational education.

VI. Future Scenarios (2030–2045)

2030–2035 (TRL 3–4): 2D-FET/CFET hybrid cells and M3D SRAM/logic pilots; ILV pitch $\lesssim$200 nm; MRAM pervasive in caches; first cross-domain PDKs.
2035–2040 (TRL 5–6): Sequential 3D logic+memory tiles in AI accelerators; reliable $< 450\,^\circ$C flows; SOT-MRAM adoption; wafer-level bonding for photonics/2D stacks.
2040–2045 (TRL 7–8): HPC/aerospace systems using non-volatile 3D logic–memory, radiation-hard 2D stacks, and co-packaged optics; cross-domain EDA becomes mainstream.

VII. Conclusion

Post-CFET shifts progress from lateral device scaling to a co-optimized space of materials, vertical integration, and new state variables. A hybrid trajectory—CFET density, M3D distance reduction, spintronic non-volatility, and heterogeneous fusion—offers the best path to energy-efficient systems. Success hinges on multi-physics EDA/PDKs, reliability- and variability-aware models, and curricula that train device-to-system engineers.

References

[1] IRDS, *International Roadmap for Devices and Systems*, 2024.
[2] S. Takagi, *et al.*, IEDM Tech Digest, 2023.
[3] Z. Liu, *et al.*, *Nature Electronics*, 2022.
[4] A. Fert, *et al.*, *Rev. Mod. Phys.*, 2019.
[5] H. S. P. Wong, *Nat. Rev. Mater.*, 2020.
[6] P. Batude, *et al.*, IEDM, 2019.

Author Biography

Shinichi Samizo received the M.S. degree in Electrical and Electronic Engineering from Shinshu University, Japan. He worked at Seiko Epson Corporation on semiconductor memory and mixed-signal device development and contributed to inkjet MEMS actuators and PrecisionCore printhead technology. He is currently an independent semiconductor researcher focusing on process/device education, memory architecture, and AI system integration.
Contact: shin3t72@gmail.com *GitHub:* Samizo-AITL.

Fig. 2. Conceptual block diagram of candidate device/integration options.

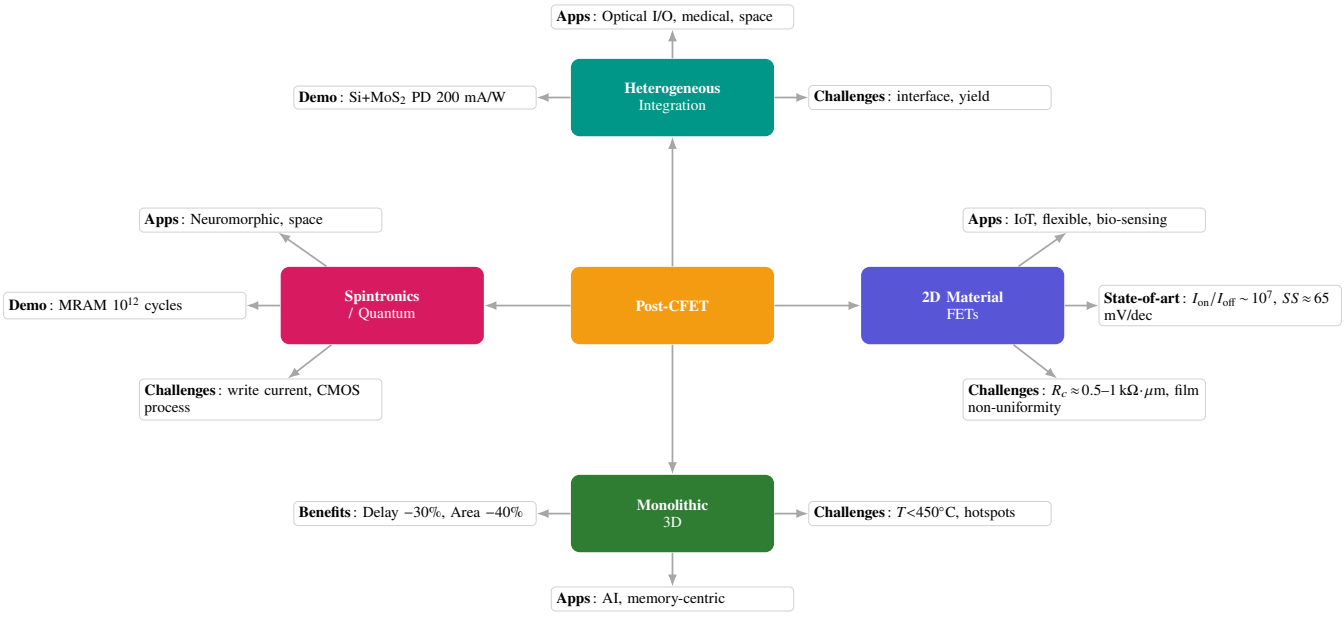

Fig. 3. Post-CFET technology mind map.

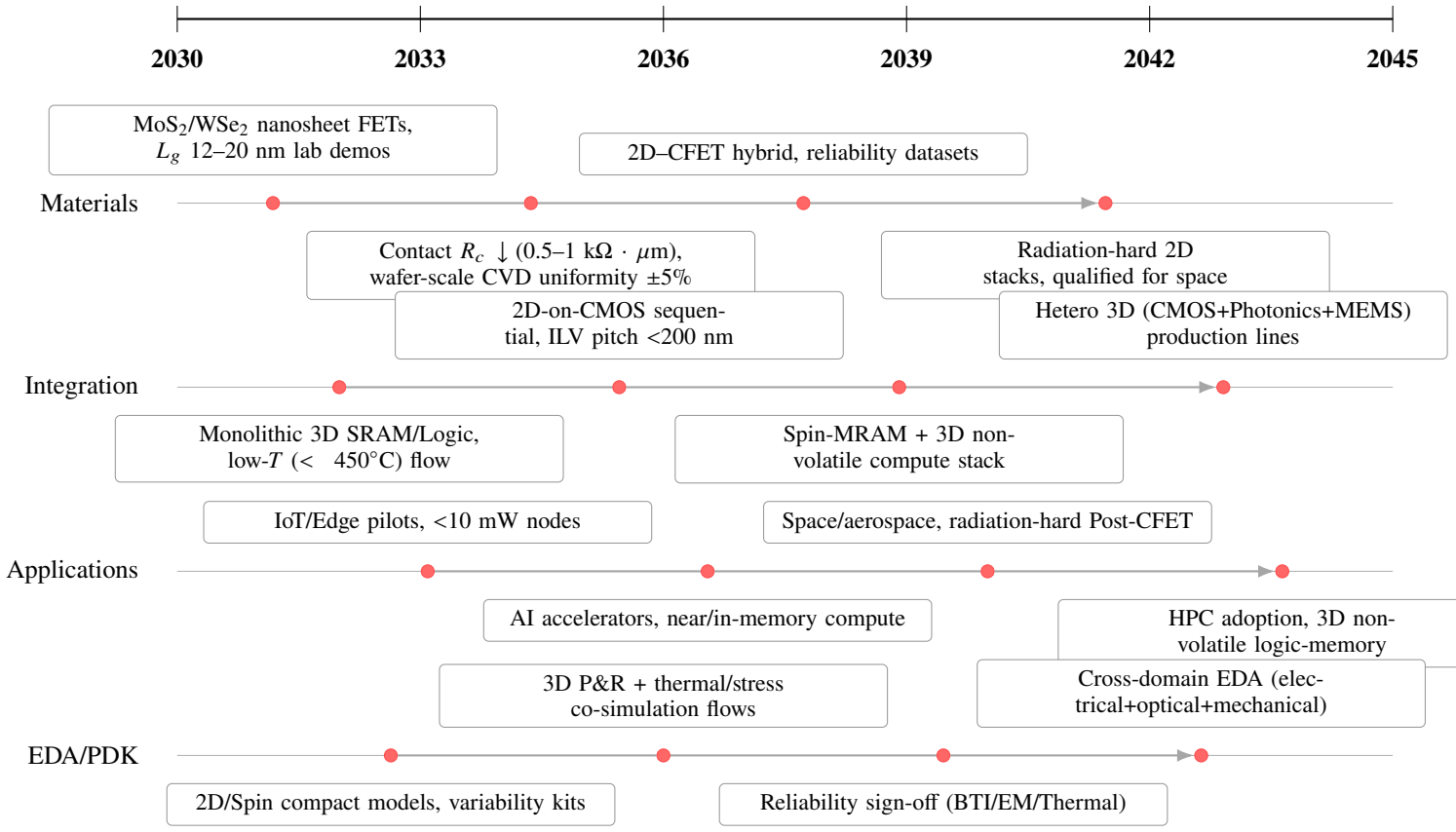

Fig. 4. 2030–2045 roadmap (materials, integration, applications, EDA).