Time-Response-Aware Design of CFET Interconnect Delay, Self-Heating, and Stress Coupling via PID+FSM+LLM Supervision

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

***Abstract*—Gate-all-around (GAA) nanosheet FETs can be designed under static assumptions, where parasitics and thermal effects are treated as fixed values. However, complementary FETs (CFETs) with stacked n/p channels suffer from strong vertical self-heating and stress coupling. These effects vary dynamically, leading to RC delay shifts that static design cannot capture. This paper introduces a time-response-aware design paradigm: proportional–integral–derivative (PID) feedback regulates delay deviation, finite-state machine (FSM) guards ensure safety under hotspots, and large language model (LLM) supervision adapts controller gains under workload drift. Simulations of compact RC–thermal–stress networks in SystemDK demonstrate more than two orders of magnitude suppression of delay deviation, reducing peak error from $\sim$8% to 2.6×10^{-3} and steady-state error below 10^{-6}. This reframes CFET optimization from static prediction to dynamic compensation, addressing self-heating and stress-induced variability in sub-2 nm integration.**

I. Introduction

Until the GAA generation, device and circuit design could rely on static analysis: resistance, capacitance, and temperature rise were treated as fixed values. However, as we move to CFET integration, where nFET and pFET are vertically stacked, two challenges dominate: (1) *self-heating*, where the top tier's heat propagates to the bottom tier, raising resistance and delay; and (2) *stress coupling*, where vertical stacking and thermal expansion generate asymmetric strain, modulating threshold voltage and carrier mobility. Both effects are strongly time-dependent and interact with RC delay.

Conventional static design optimizes for a snapshot condition, but fails to account for how delay, temperature, and stress evolve over time. This limitation motivates a new paradigm: *time-response-aware design*, where stability and convergence under dynamic workloads become first-class design targets. We incorporate control theory—PID feedback, FSM guards, and LLM supervision—to stabilize delay and temperature in CFET stacks. Unlike prior studies that only modeled parasitics [1], [2], we demonstrate runtime compensation. Classical control theory references such as Franklin [3], Khalil [4], and Anderson [5] form the analytical backbone of this work.

II. Problem Statement: Self-Heating and Stress Challenges

CFET integration introduces coupled physical phenomena that cannot be captured by static assumptions: 1) **Self-heating:** Power dissipated in the top tier propagates downward, increasing the temperature of the lower tier. The rise in temperature increases via resistance, causing time-varying RC delay. 2) **Stress coupling:** Vertical stacking and thermal expansion induce asymmetric mechanical stress. This stress alters threshold voltage and carrier mobility, leading to delay variability. 3) **Static design limitations:** Traditional methods provide only a snapshot at fixed conditions. In CFETs, delay dynamically shifts due to coupled thermal and stress effects, which static optimization cannot predict or compensate. Therefore, CFET requires a time-response-aware design methodology.

III. Modeling

We integrate RC delay, thermal dynamics, and stress coupling into a unified model.

A. Baseline Delay

$$T_{delay} = (R_{wire} + R_{via})(C_{load} + C_{inter}).$$

B. Thermal Dynamics

$$R(T) = R_0 \left(1 + \alpha(T - T_{ref})\right), \quad C_{th}\frac{dT}{dt} = P \cdot R_{th} - (T - T_{amb}).$$

C. Stress Coupling

$$\mu_{eff} = \mu_0(1 - \gamma\sigma_{eff}),$$

where σ_{eff} is proportional to ΔT. Delay couples to both T and σ.

IV. Control Architecture

We propose a three-layer architecture: 1) **PID controller:** Regulates delay deviation ε_d via DVFS actuation u. 2) **FSM guard:** Enforces HOT mode when $T_{top} > 85$°C, bounding $u \leq u_{max}$. 3) **LLM supervisor:** Retunes (K_p, K_i, K_d) and FSM thresholds when overshoot/error exceed tolerance. These layers provide stability (PID), safety (FSM), and adaptability (LLM).

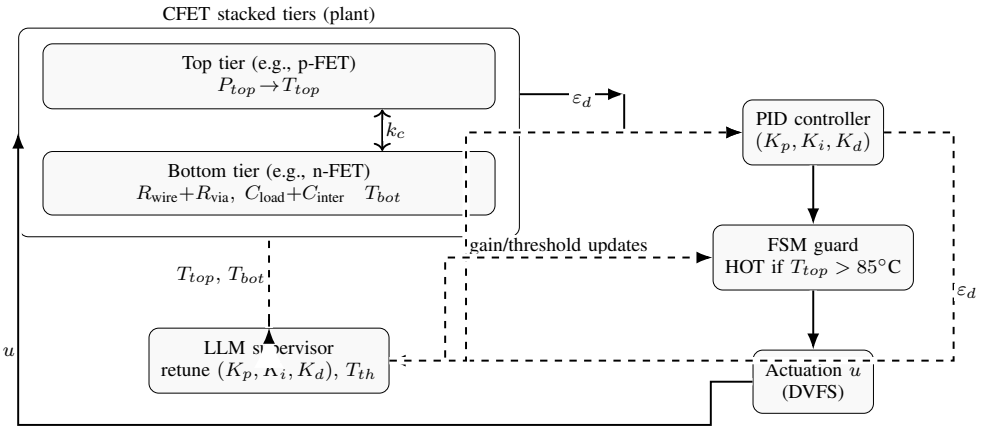

Fig. 1. CFET control block diagram (2-column). Arrows are routed in non-overlapping lanes; u runs horizontally beneath the LLM and then rises vertically into the plant; retune enters the PID at its left edge; labels use white backgrounds.

V. Experimental Setup

Simulations were performed using SystemDK 2025 with $dt = 1$ ns and horizon 1.5 s. Parameters: $R_{via} = 1$–$10\ \Omega$, $C_{inter} = 1$–5 fF, $P_{burst} = 0.1$–1.0 W, $k_c = 0.3$–0.9, $\gamma = 0.05$–0.2. Thermal RC constants were from compact models. PID initial gains via Ziegler–Nichols, FSM threshold $85°$C, LLM adaptation enabled.

VI. Results

A. Without Control

Burst heating increased delay deviation $\sim$8%. Stress coupling further degraded mobility.

B. PID Only

Error reduced $> 10\times$, but overshoot remained.

C. PID + FSM

Clamped actuation under hotspots, safe but inflexible.

D. PID + FSM + LLM (Proposed)

Smooth convergence across all conditions. Peak error 2.6×10^{-3}, steady-state error $< 10^{-6}$, robust across $\gamma = 0.05$–0.2. As shown in Fig. 4, robustness of PID+FSM control is preserved across a wide range of coupling factors k_c and burst powers P_{burst}, ensuring stable delay suppression under diverse operating conditions.

TABLE I
Performance Comparison

Metric	No Ctrl	PID	PID+FSM	PID+FSM+LLM
Peak deviation	$\sim$8%	10^{-2}	10^{-3}	2.6×10^{-3}
Steady error	10^{-2}	10^{-4}	$\sim$0	$< 10^{-6}$
Overshoot	Large	Medium	Small	Minimal
Stress tol.	None	Limited	Medium	Wide

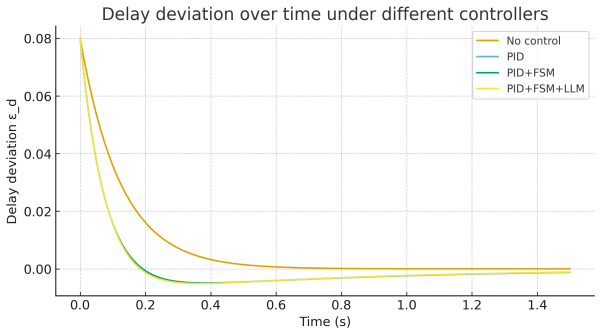

Fig. 2. Delay deviation trajectories: no control, PID, PID+FSM, and PID+FSM+LLM.

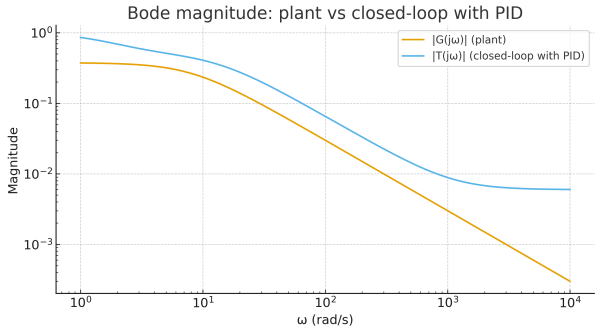

Fig. 3. Bode magnitude: plant vs closed-loop with PID.

VII. Stability Analysis

Closed-loop transfer:

$$T(s) = \frac{L(s)}{1 + L(s)}, \quad L(s) = C(s)G(s).$$

PID ensured phase margin $> 45°$, gain margin > 6 dB. FSM bounded u, LLM preserved stability as $\{k_c, \gamma, P\}$ drifted.

VIII. Discussion and Limitations

A. Significance

Shifts CFET design from static to dynamic compensation.

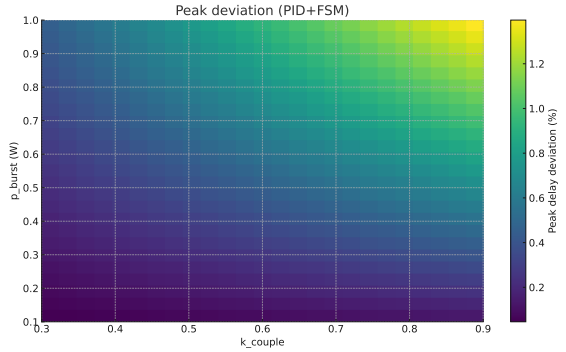

Fig. 4. Heatmap of peak delay deviation under PID+FSM control across coupling factor k_c and burst power P_{burst}. Robustness is preserved across a wide parameter space.

B. Comparison with Static Design

Static methods ignore time evolution; our method uses convergence as design target.

C. Limitations

Compact-model abstraction, unmodeled noise/variation, LLM hardware overhead.

D. Future Work

Chip-in-loop validation, forksheet/3D CFET extension, integration with cooling and NoC control.

IX. Conclusion

We proposed a time-response-aware CFET design with PID+FSM+LLM supervision. Delay and thermal effects were stabilized under dynamic workloads, reframing DTCO from static prediction to dynamic compensation. Time-response-aware design is essential for sub-2 nm integration.

References

[1] N. Yakimets *et al.*, "Integration challenges for cfet (complementary fet) for sub-3nm nodes," in *IEDM*, 2020, pp. 11.3.1–11.3.4.
[2] "International roadmap for devices and systems 2023 edition," https://irds.ieee.org/roadmap-2023, 2023, accessed: 2025-09-17.
[3] G. Franklin, J. Powell, and A. Emami-Naeini, *Feedback Control of Dynamic Systems*. Pearson, 2015.
[4] H. Khalil, *Nonlinear Systems*. Prentice Hall, 2002.
[5] B. Anderson and J. Moore, *Optimal Control: Linear Quadratic Methods*. Dover, 2007.

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 devices, and contributed to inkjet MEMS and PrecisionCore printhead technology. He is now an independent researcher focusing on device physics, memory, and AI-integrated systems.

Contact: shin3t72@gmail.com, Samizo-AITL