Self-tuning of Clock Latencies Using Floating Gates

Kambiz Rahimi and Chris Diorio

University of Washington

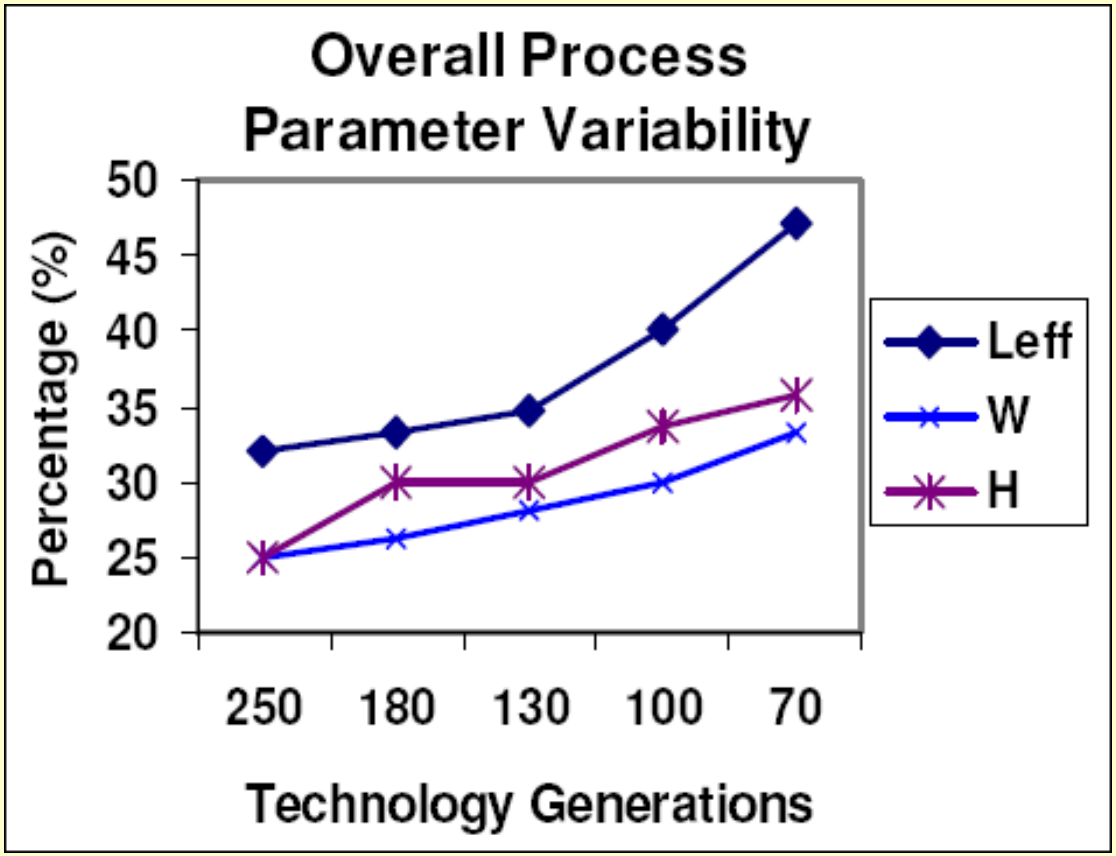
and

Impinj Inc.

Process Variability Grows

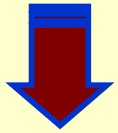
- Process variations cause delay variations
- Delay variations cause timing violations
- Delay sensitivity to operating conditions is also growing

- (S. Nassif, 2000)

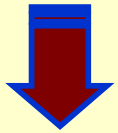


Cycle Time Wasted on Margins

- Larger design margins AND
- Deeper pipelines



Timing constraints of sequential elements have become a major yield and performance bottleneck



Self-tuning with floating gates solves the problem by tuning clock delays after fabrication

Outline

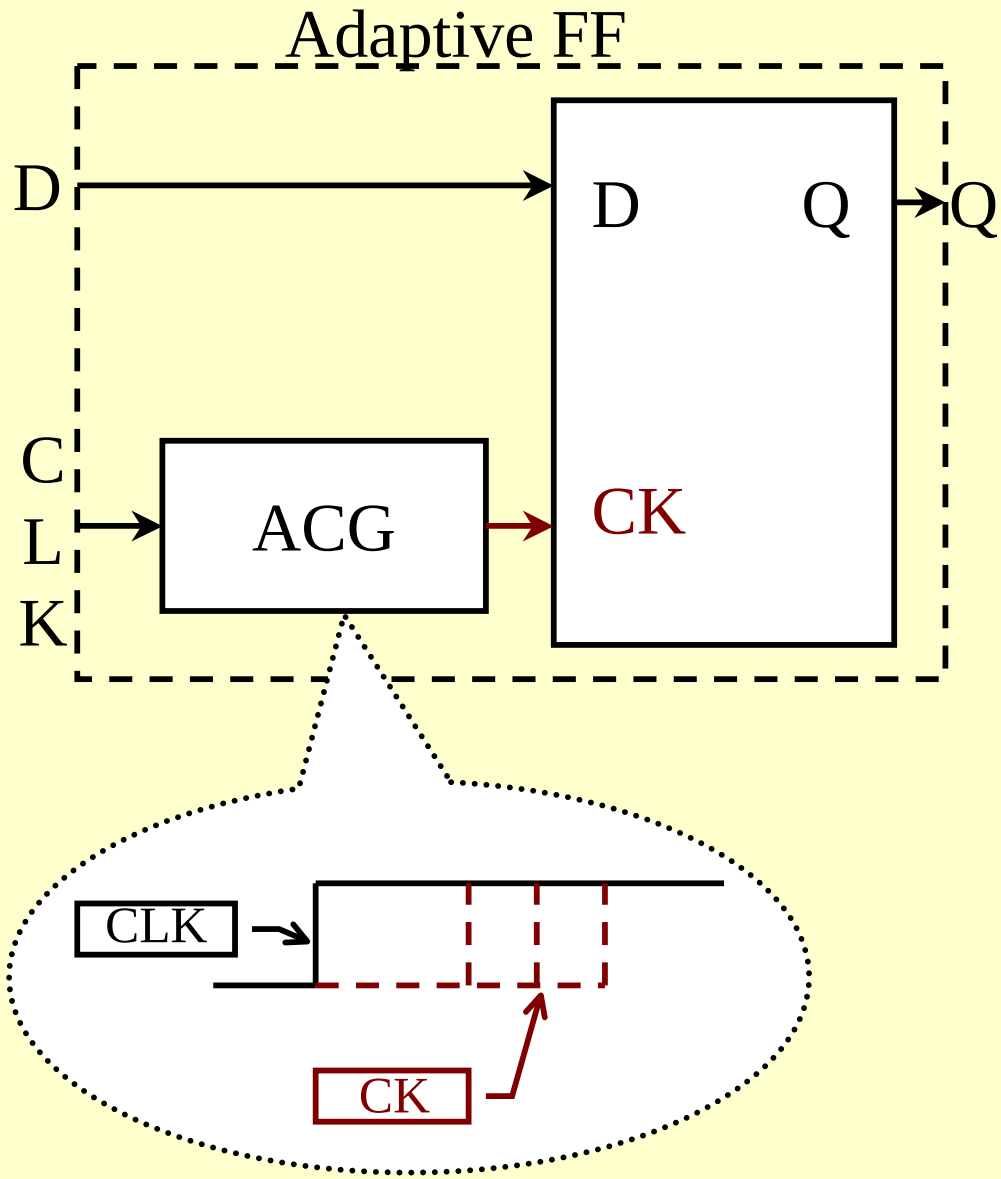
- Background
- Delay tuning with PMOS floating gates
- Self-tuning Adaptive-delay Sequential Elements (SASEs)
- Experimental results
- Tuning procedure for pipelines

Background

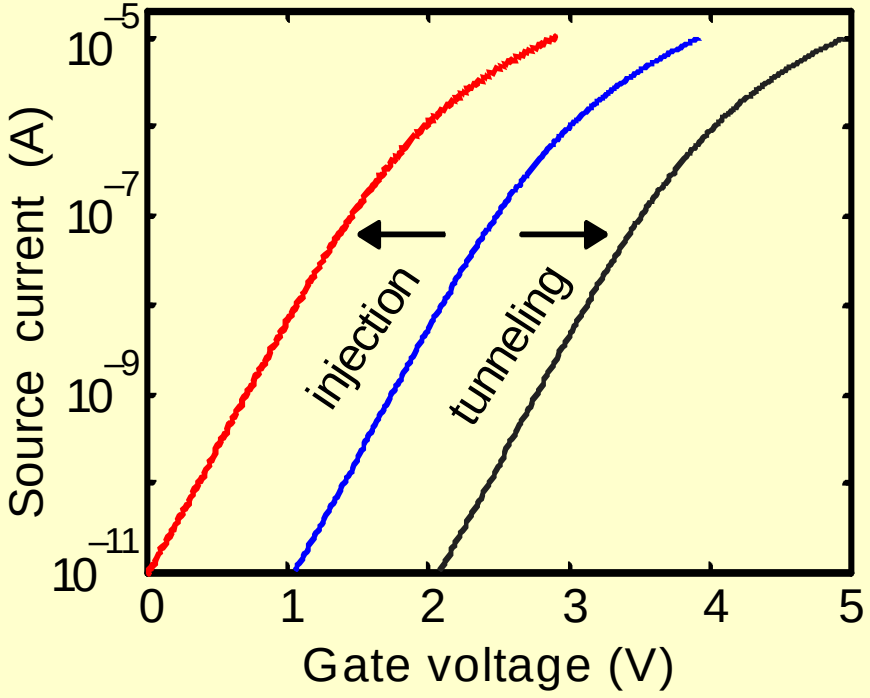
- Path duplication methods:
  - Correlating VCO - Gutnik et al. (1996), Burd et al. (2000)
  - Delay Line Speed Detector - Dhar et al. (2002), Uht (2003)
- In-situ methods:
  - Triple-Latch Monitor - Kehl (1993)
  - Razor - Ernst et al. (2003)
- Previous tuning methods monitored delays to adjust voltage or frequency

Adapting Delays with Floating Gates

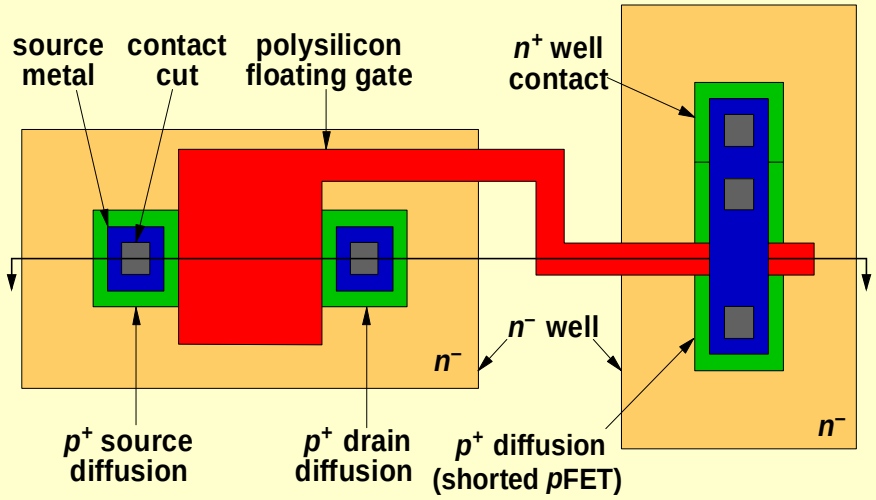
- Adjust the delay between input and internal clocks
- PMOS Synapse Transistor inside ACG
- ACG provides nonvolatile, electrically programmable delays



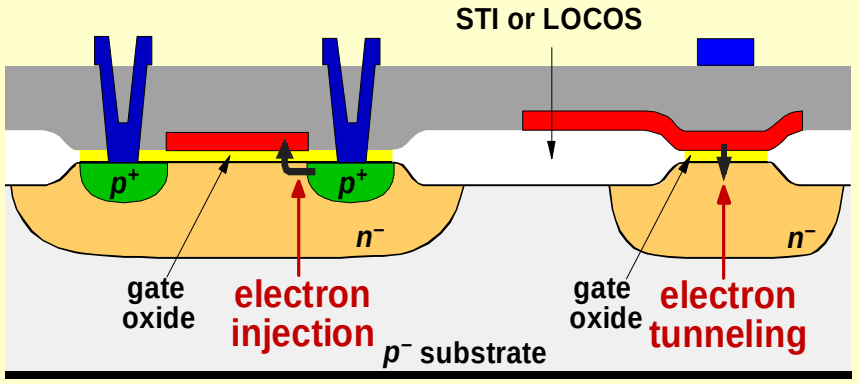
Floating-Gate Synapse Transistor



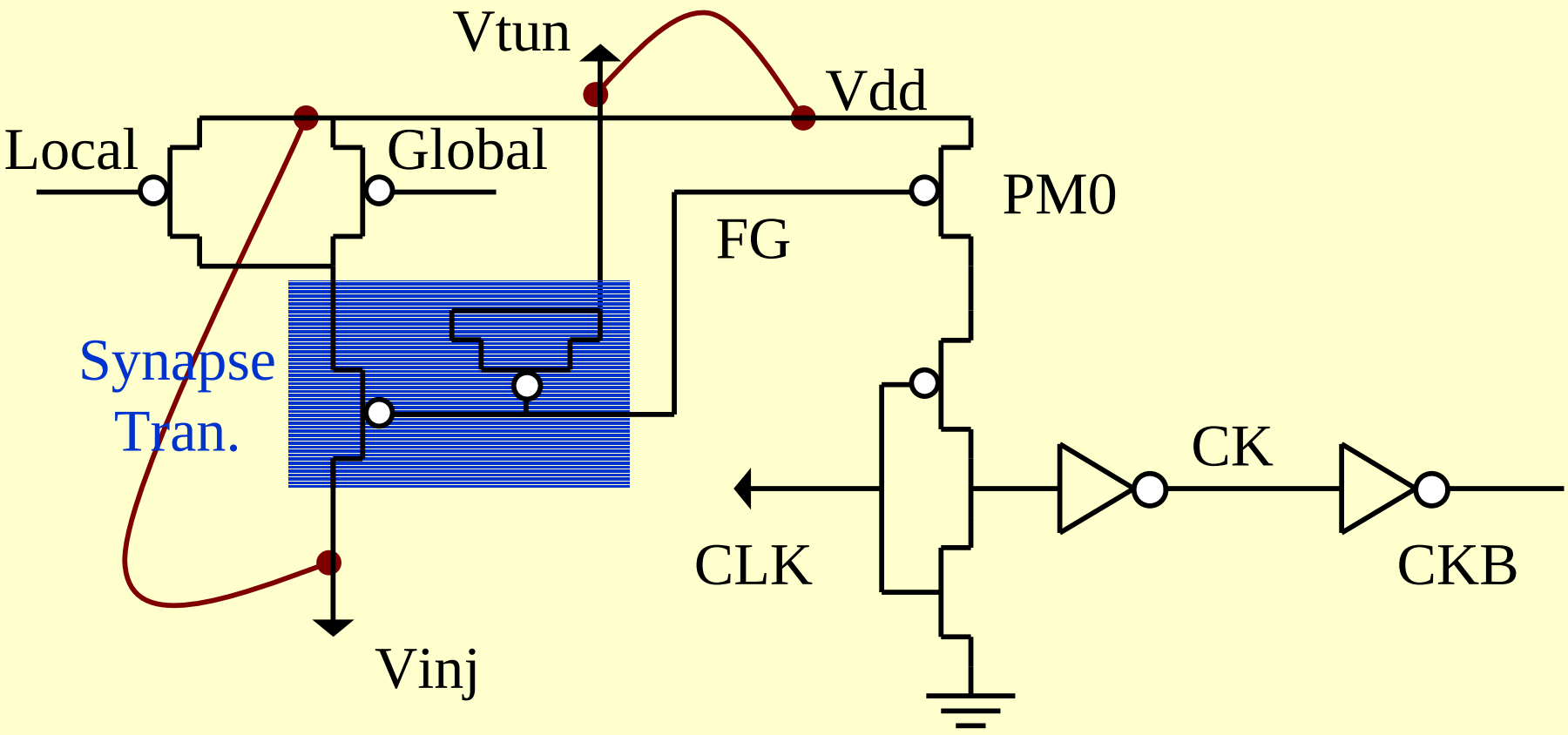
Top View



Side View

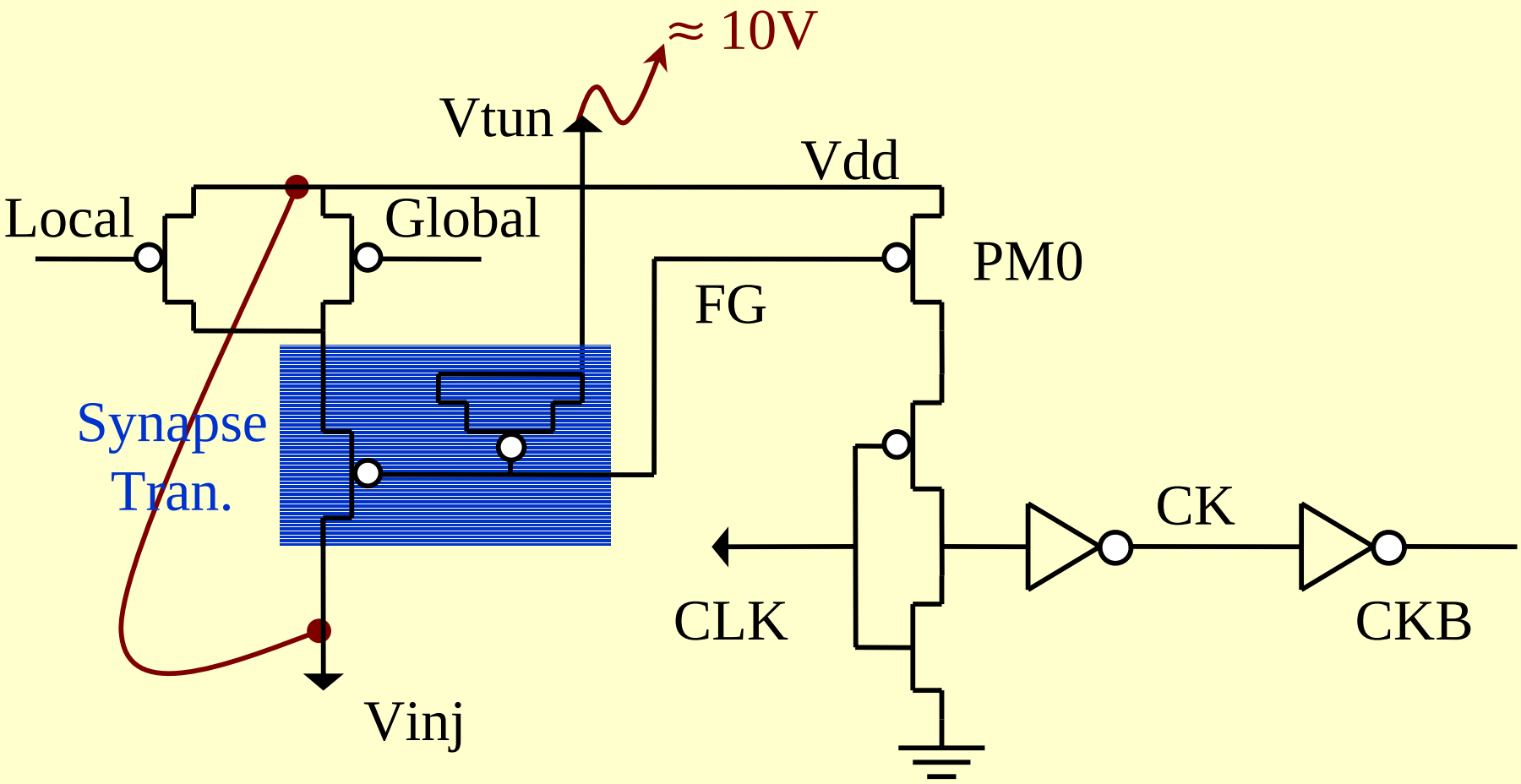


Adaptive Clock Generation



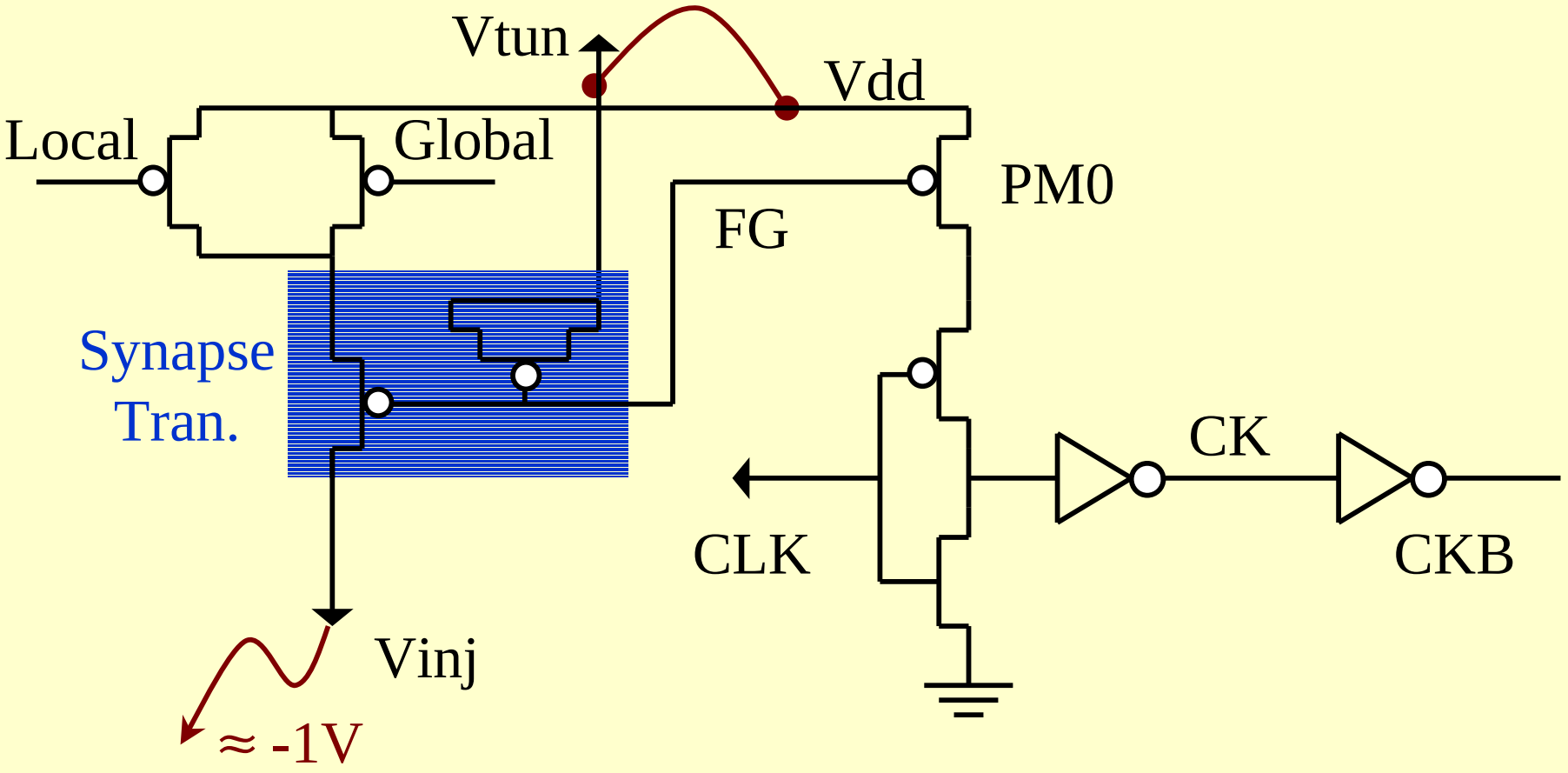
- Normal operation:
  - Vinj = VDD, Vtun = VDD

Adaptive Clock Generation



- Increase delay:
  - Vinj = VDD, Vtun ≈ 10V

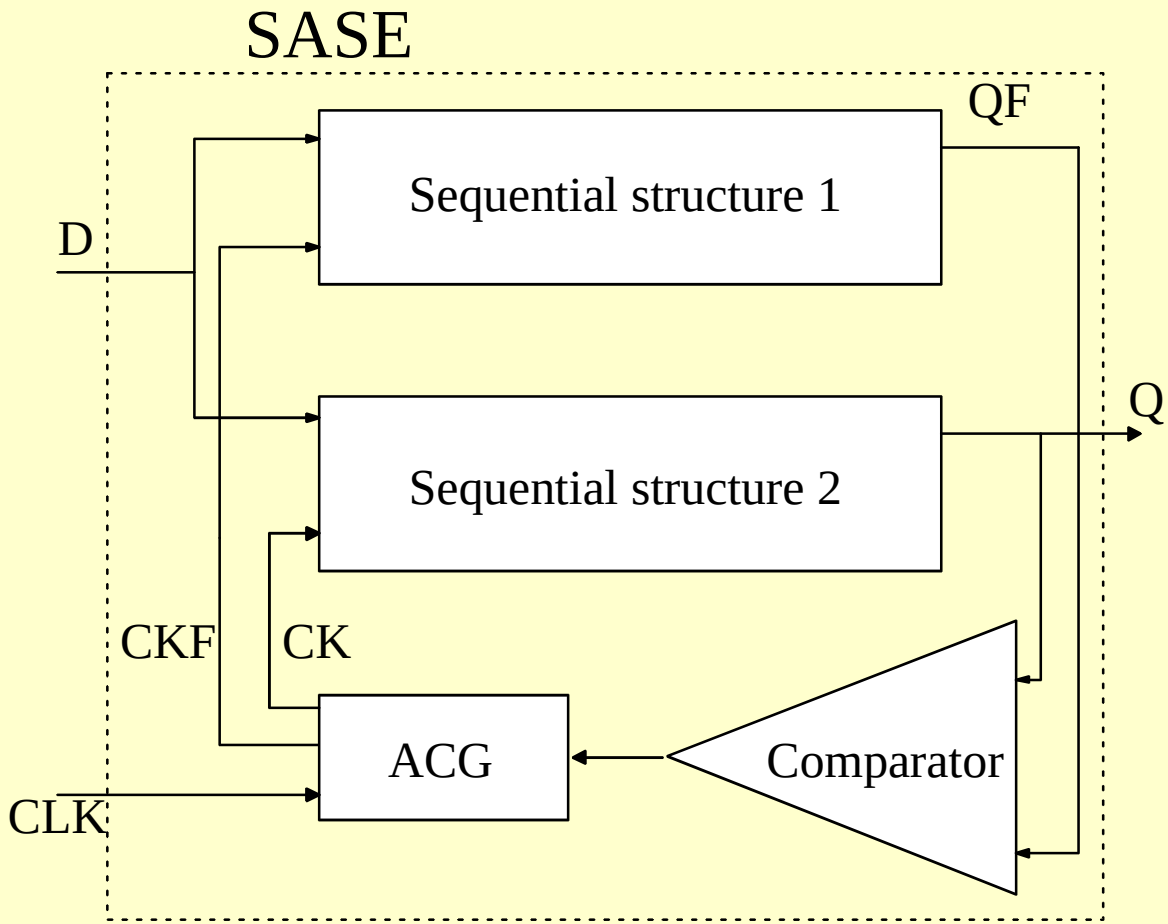
Adaptive Clock Generation



- Decrease delay:
  - Vinj < 0V, Vtun = Vdd

Self-Tuning Principles

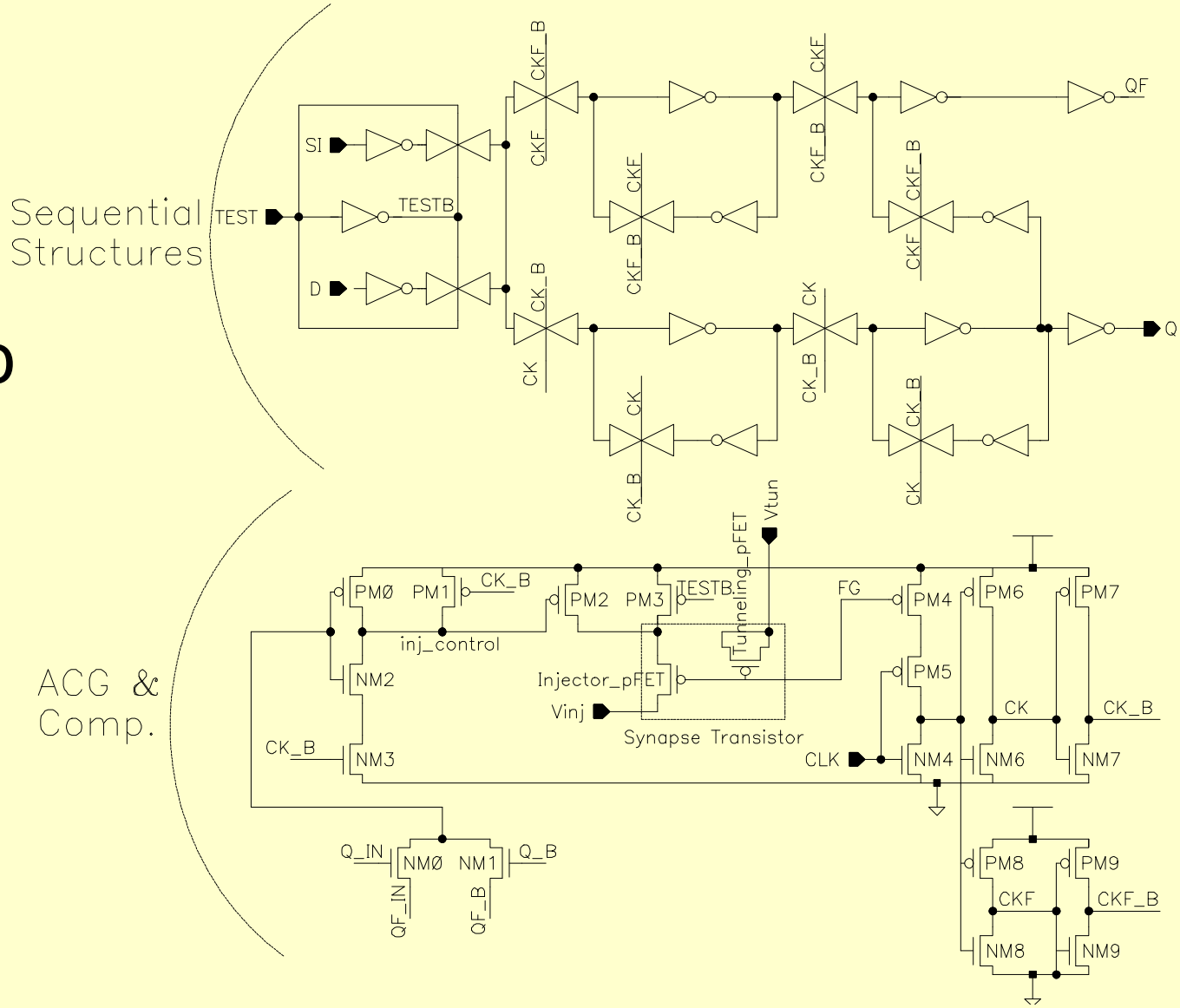
- Double sample the D input
- Comparator provides local injection control
- Failure Margin: $FM = |D_{CK} - D_{CKF}|$



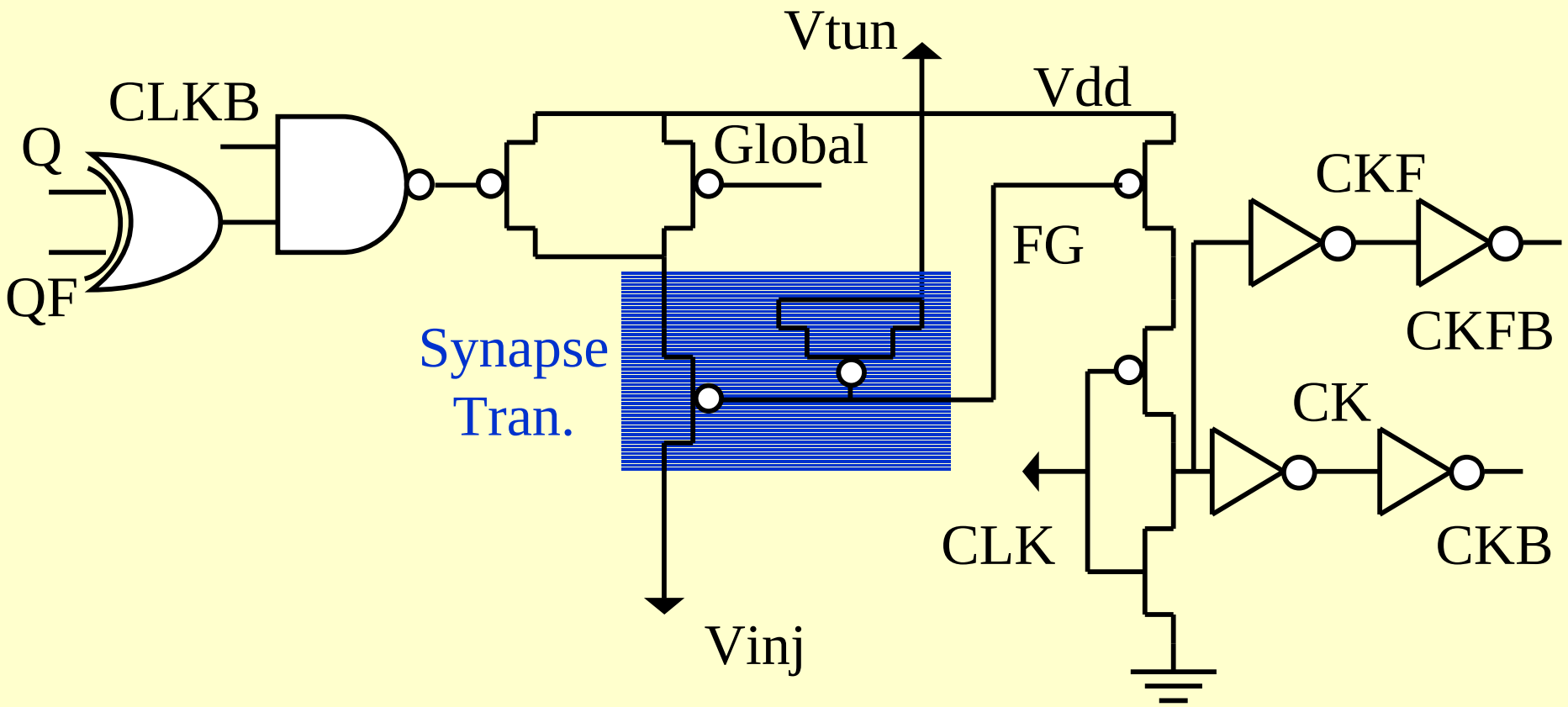
Starting with a large ACG delay decrement delay by ΔD if Q and QF match

Self-tuning Adaptive-delay Implicit-pulsed Master-slave Static FF (SIMS)

- QF is set to Q in every clock cycle



Self-tuning ACG

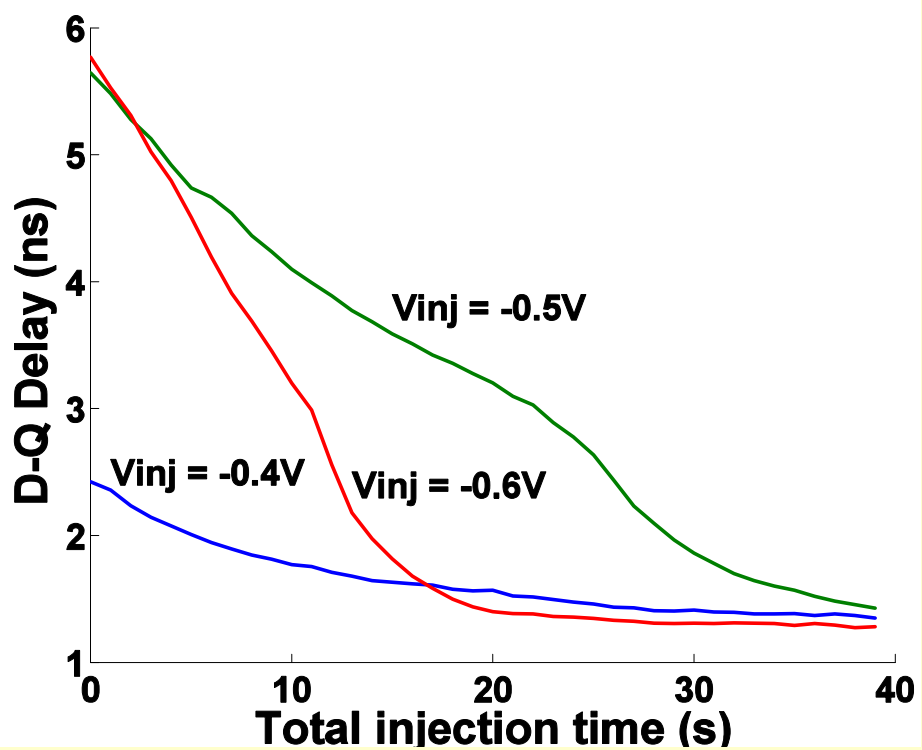


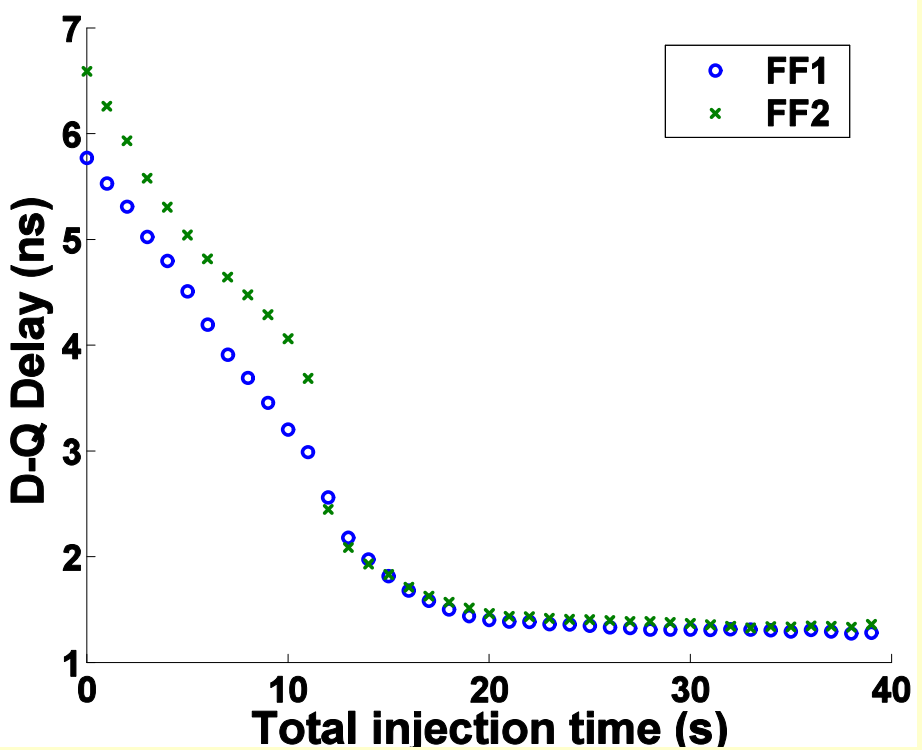
- Local injection control generated from Q, QF

SASE Impact

- What architectural changes are needed?
  - No additional addressing and support circuitry
  - Two global signals for tunneling and injection
  - Two I/O pads
- What is the area overhead?
  - FFs constitute a fraction of total area
  - Use SASEs only on near-critical paths
  - Chip area overhead is 10-20%
- How often do you self-tune?
  - Changes in operating conditions
  - Leakage in floating gates

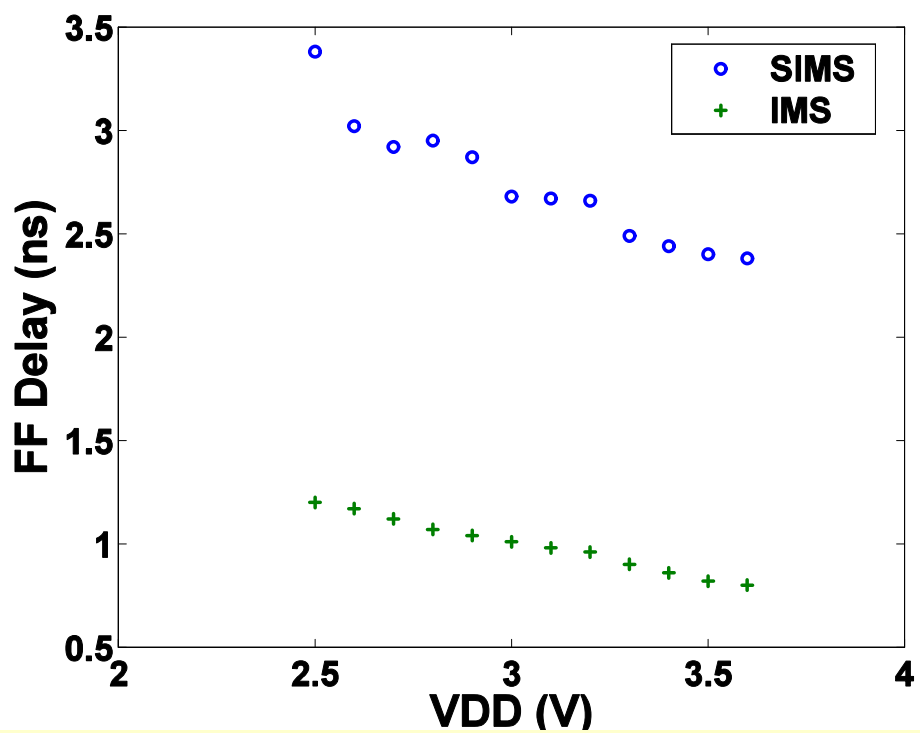
SIMS Tuning Experiments

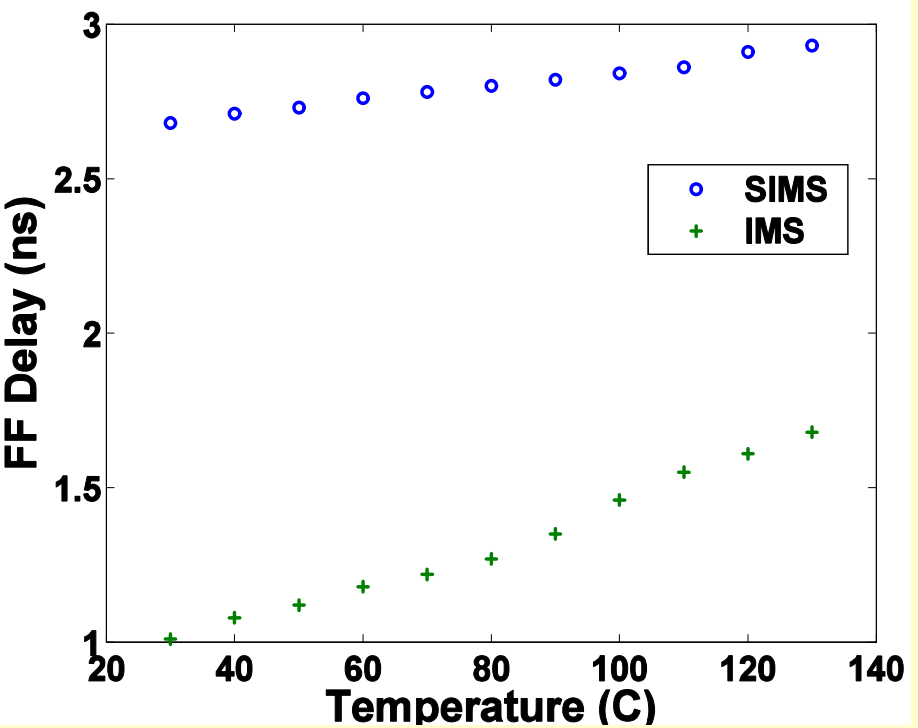




- Failure Margin was 400ps
- Average delay difference after self-tuning is 54ps, std. dev. is 15ps.

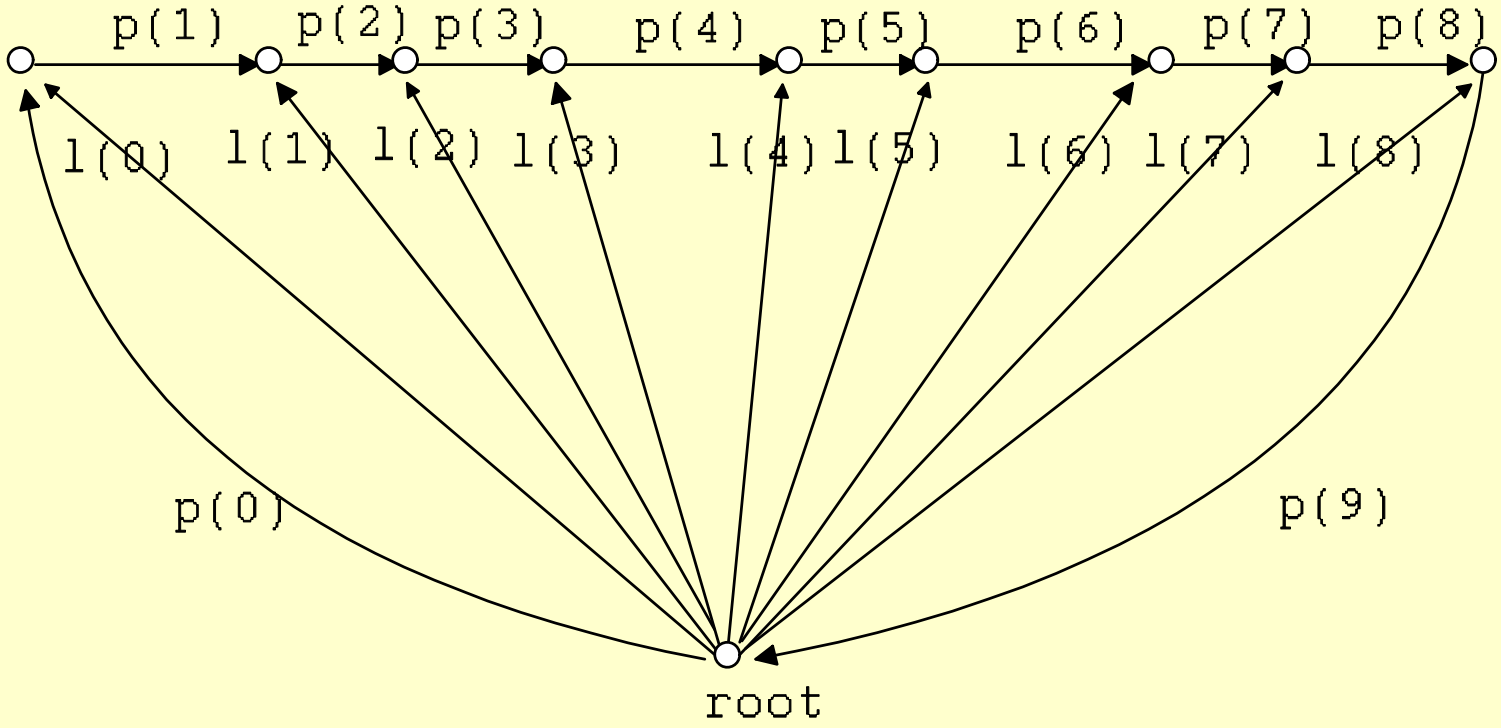
Voltage & Temperature Sensitivity





- Vdd sensitivity: SIMS = 1ps/mV, IMS = 0.4ps/mV
- Temp. sensitivity: SIMS = 2.7ps/C, IMS = 6.6ps/C

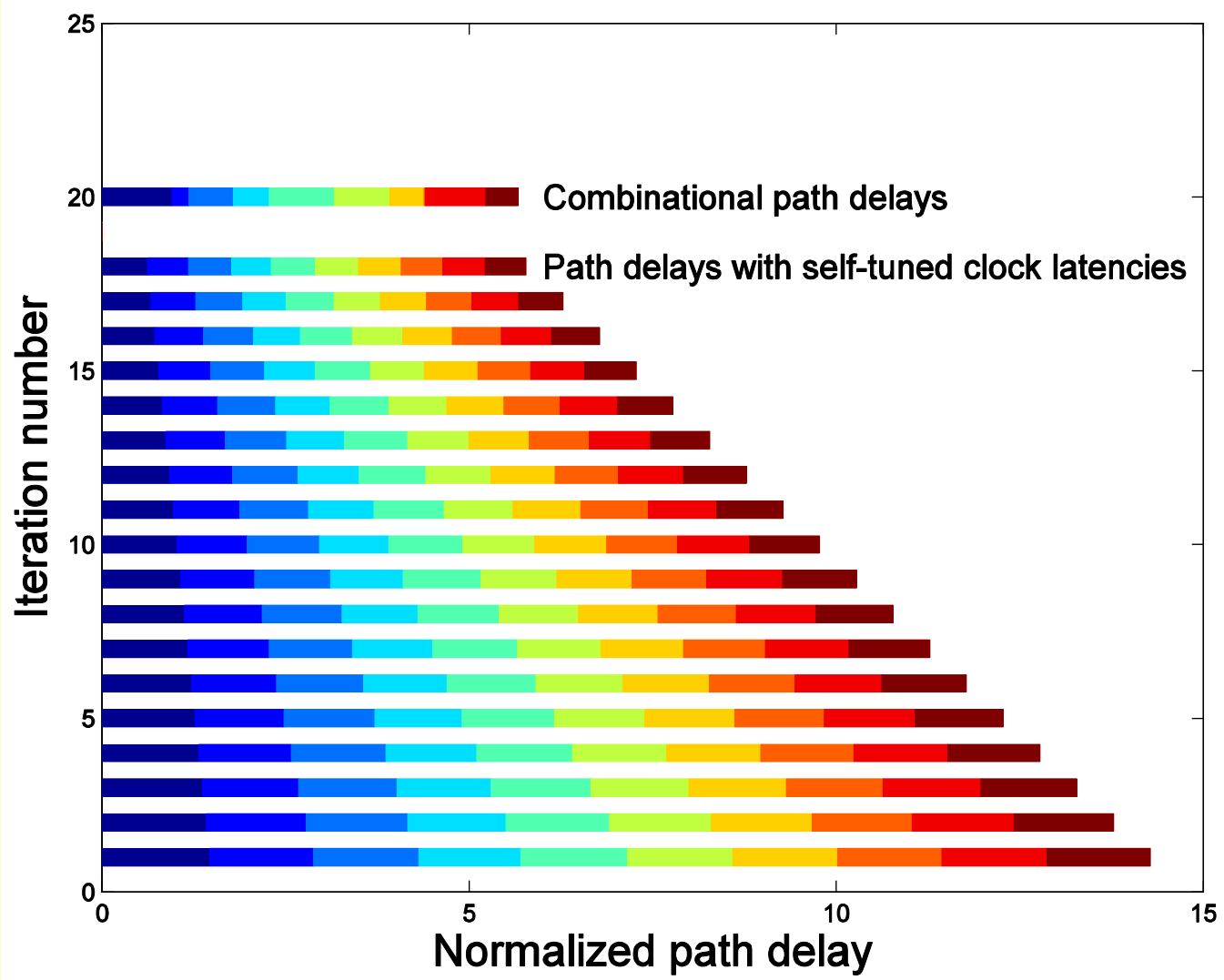
SIMS Tuning Procedure



- Choose $T = 1.5 \times \max(\bar{d}(i, j))$, $delay\_step > FM$
- While $T > mean(\bar{d}(i, j))$
  - Toggle clock and data $T / delay\_step$ times
  - $T = T - delay\_step$

Convergence of SIMS Tuning

- Normalized path delays
- delay_step = 0.05



Summary

- In-situ self-tuning of clocks using floating gates
- SASE prototype and tuning operations
- Tuning experiments demonstrate the range and precision of SASE delays
- SASE delay sensitivity are comparable to regular flip-flops
- Procedure for tuning SASEs in pipeline circuits