Managing Standby and Active Mode Leakage Power in Deep Sub-micron Design

Lawrence T. Clark
Dept. of Electrical Engineering
Arizona State University

Rakesh Patel
Intel Corp.

Timothy S. Beatty
Intel Corp.

Outline

- Introduction and motivation
- Standby leakage management
  - Drowsy mode: Reverse body bias and supply collapse
    - Circuit design and operation
    - System level results
    - Limitations
  - Thick gate shadow latches
- Active leakage management
  - Multiple V_t and channel assignment
  - Drowsy memories
  - Thick gate SRAM
- Conclusions

Outline

- **Introduction and motivation**
- Standby leakage management
  - Drowsy mode: Reverse body bias and supply collapse
    - Circuit design and operation
    - System level results
    - Limitations
  - Thick gate shadow latches
- Active leakage management
  - Multiple V_t and channel assignment
  - Drowsy memories
  - Thick gate SRAM
- Conclusions

Power in Hand-held Electronics

- Battery capacity is limited
  - Batteries are heavy
    - Capacity is proportional to weight
- Problematic for hand-held devices
  - Cell phone batteries typically 600 to 1200 mA hrs
    - Power budget shared between analog, digital, and transmit
  - Digital IC budget decreasing while performance increases
- Two scenarios
  - Active operation—100's mW
    - Limits talk time (typically few hrs)
  - Standby—100 μW
    - Limits time waiting for calls (typically 100's hrs)
  - There is active power in standby mode
    - Each contact with the cell is an active transmit/receive operation—occurs on the order of once every 1-2 seconds

Voltage scaling

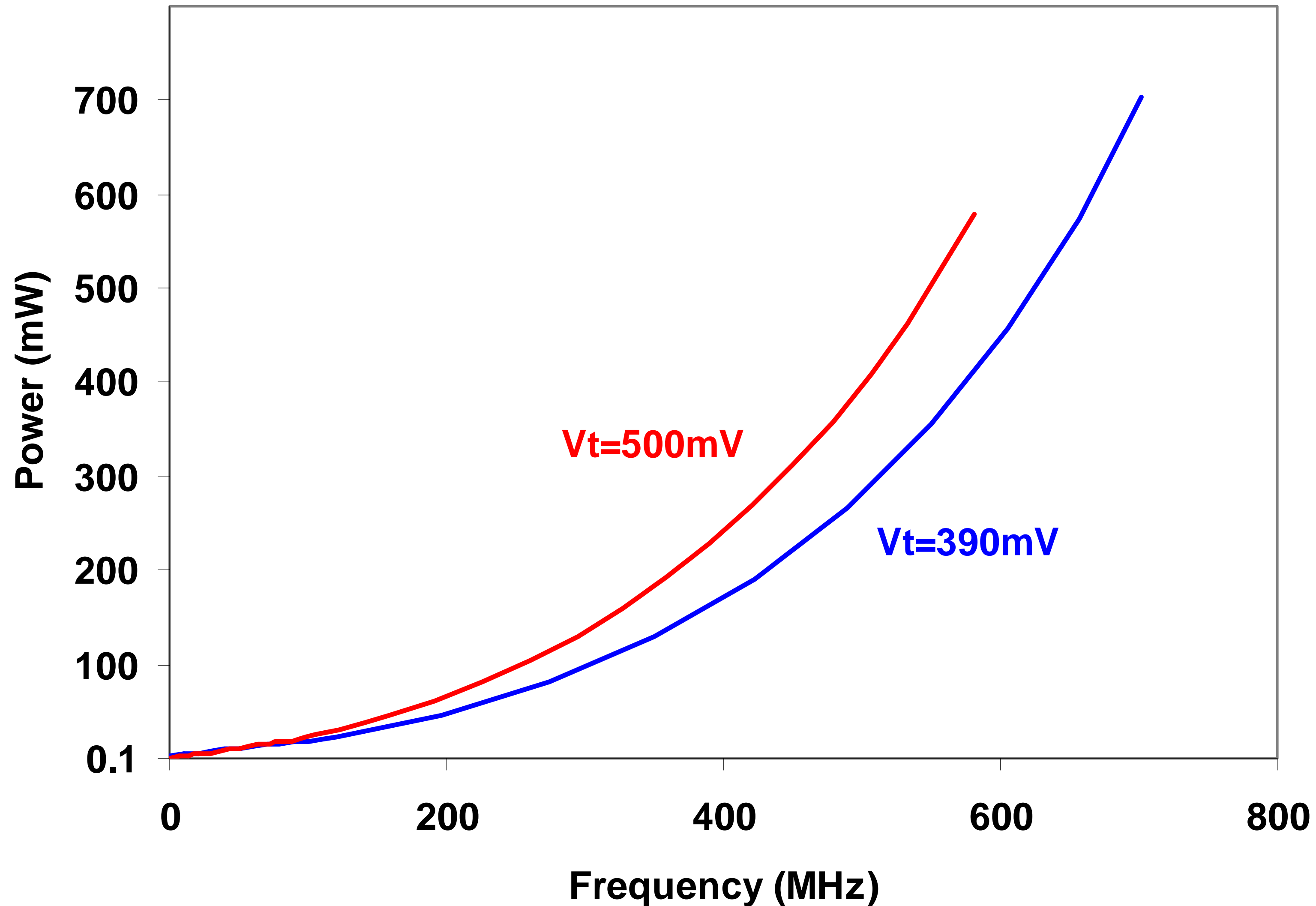
- Voltages must scale as transistors scale to avoid excessively high fields—this in turn requires V_t scaling→Higher leakage
- Total IC power

$$P_{TOTAL} = \alpha C V_{DD}^2 F + I_{LEAK} V_{DD}$$

- V_{DD}^2 active power dependence makes supply scaling the most effective lever for low power design
  - Makes low power and high performance design the same
    - Higher absolute (maximum V_{DD} performance) affords meeting lower application demand for performance at lower voltages
- High performance equals low power if done *efficiently*
  - Assumes that operating voltage is not a constraint
- Frequency proportional to voltage so dropping voltage derives roughly V_{DD}^3 change in power
  - Assumes that lower frequency is acceptable
- **Also greatly affects leakage components on advanced processes**

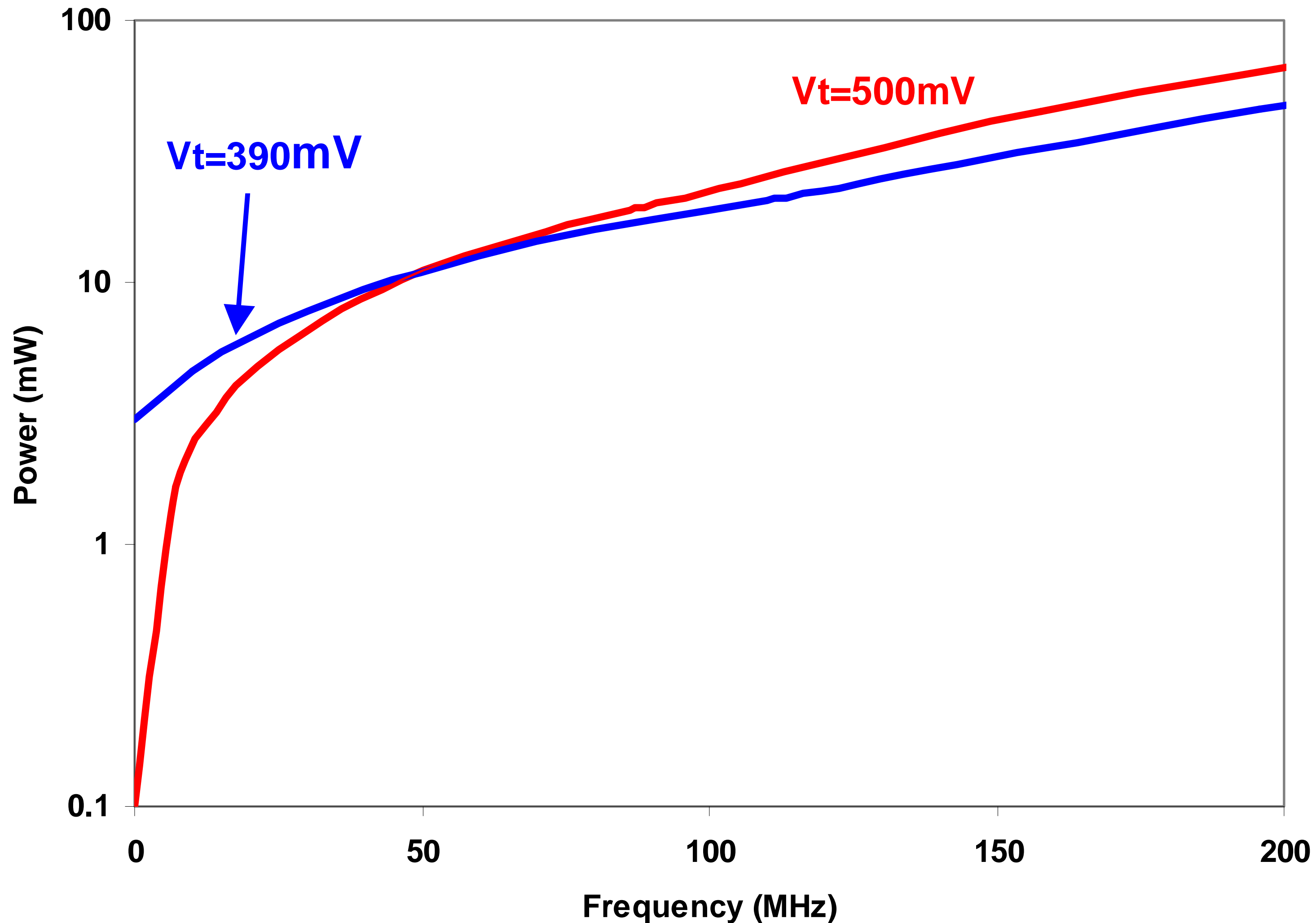
Voltage scaling: Effect of V_t

- Low V_t helps active power if V_{DD} scaled



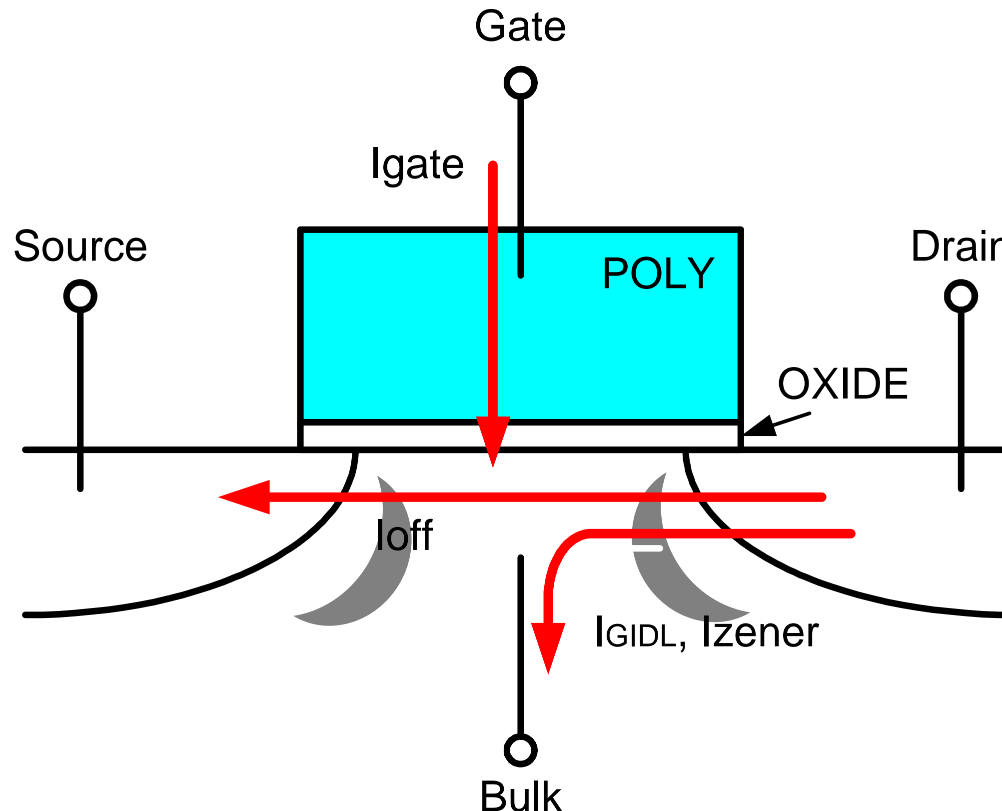
Voltage Scaling: Effect of V_t

- Low V_t problematic for standby power



Deep Sub-micron MOSFET Leakage

- Four Primary Components
  - Drain Source Leakage (I_{off})
  - Gate Leakage
  - Gate induced drain leakage (GIDL)
  - Junction band to band tunneling currents

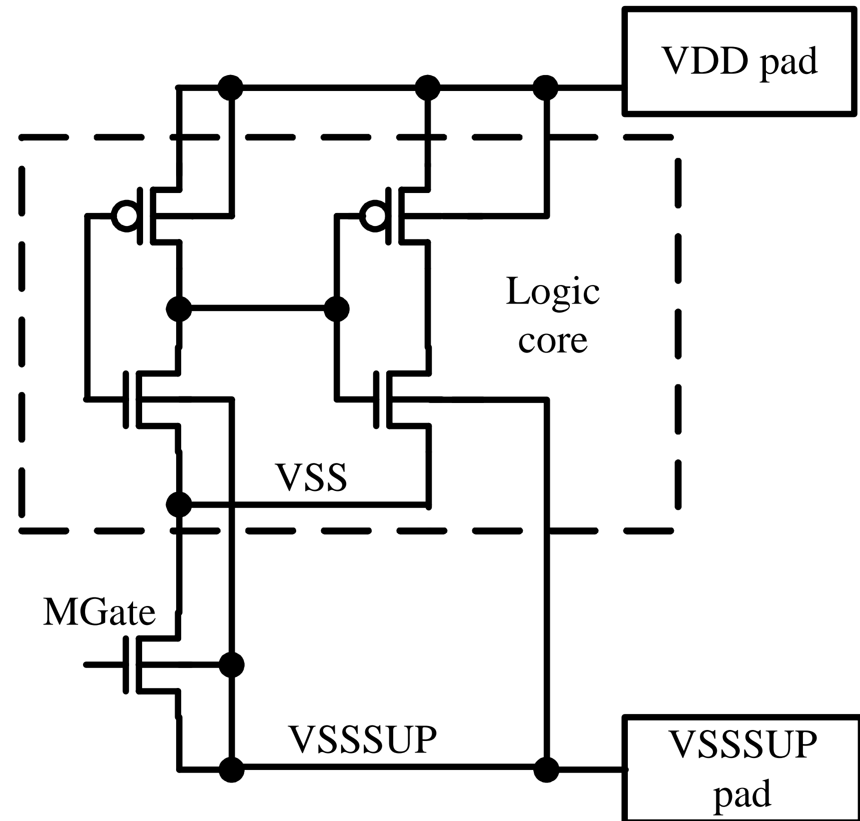


Circuit Methods for Leakage Control

- Body bias techniques
  - Reverse body bias (RBB) [1-2]
  - Forward body bias (FBB) [3]
  - These are the least invasive to the design, small area cost
- MTCMOS techniques [4]
  - State retentive
    - Balloon latches [5]
    - Multi-V_t design [6-7]
  - Non-state retentive
- Thick gate storage
  - Alleviates gate leakage [8-10]
    - Essentially store state in a generation N-x transistor
  - Highest area cost, most effective, potentially difficult design

Multi-threshold CMOS (MTCMOS)

- High V_t transistors gate power to low V_t circuits [4]
  - Leakage dominated by the high V_t gating transistors
- Not inherently state retentive
  - Power cost of moving state off chip is a penalty paid on entry and exit to low power state



The Need for State Retention

- Integrated circuits have increasing storage capacity
  - The storage constitutes the “state” of the machine
- Many commercially shipping low power standby modes have not been state retentive
  - Save in external memory
    - Incurs power penalty for the IO to save state
    - Still requires low power storage—somewhere
- Example: SA-1100 StrongARM microprocessor [11]
  - Write back cache state requires 16 μs at 66 MHz for 8 kB
  - 3.3V IO pins loaded with 35pF each gives 12.5 μJ
    - This creates a power floor of 1.25 mW if used 100 times/second
  - Still must account for the external storage power
  - Standby power and leakage of IO ring and real-time clock specified to be 165 μW

Outline

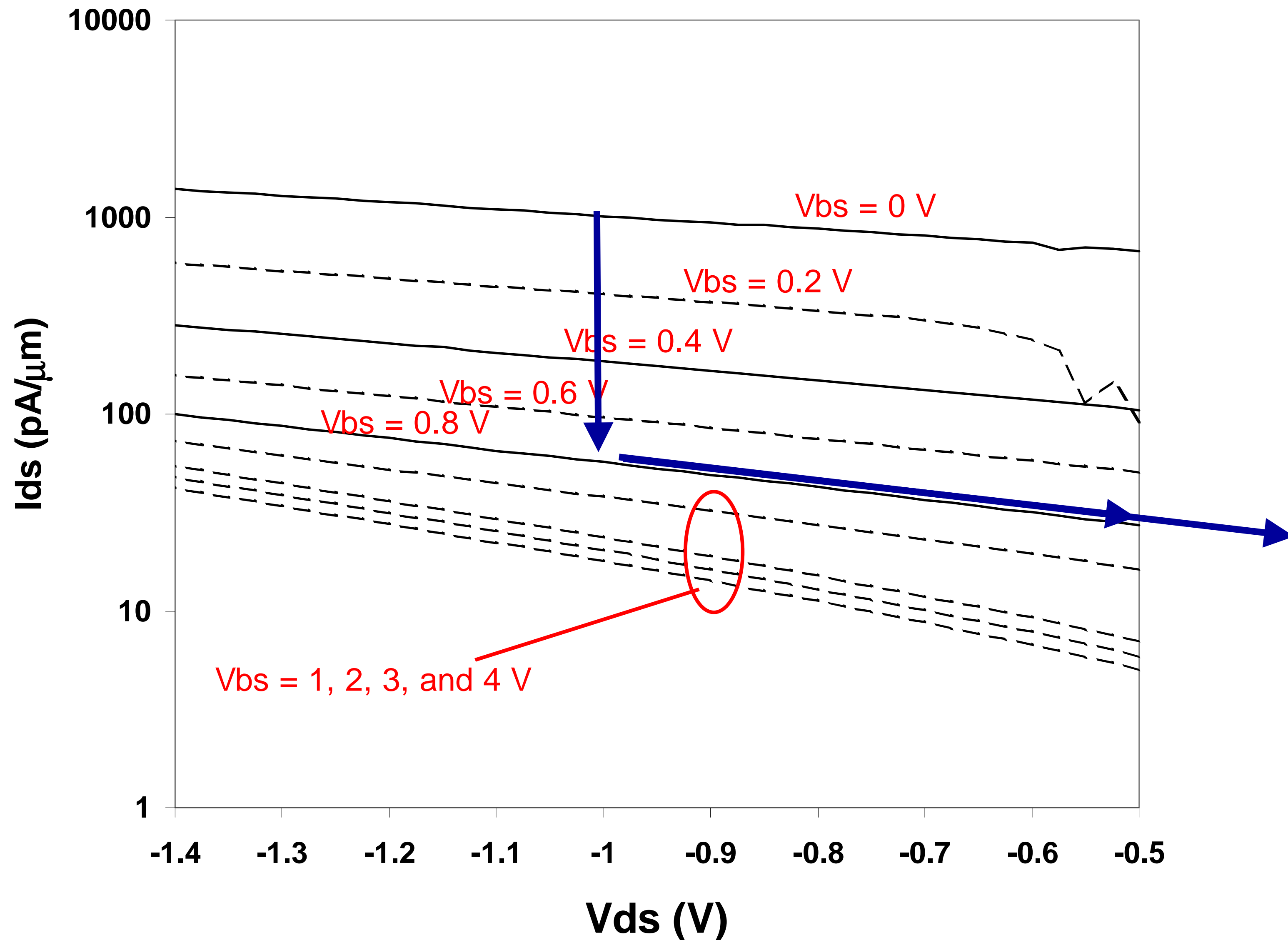
- Introduction and motivation
- **Standby leakage management**
  - Drowsy mode: Reverse body bias and supply collapse
    - Circuit design and operation
    - System level results
    - Limitations
  - Thick gate shadow latches
- Active leakage management
  - Multiple V_t and channel assignment
  - Drowsy memories
  - Thick gate SRAM
- Conclusions

Reverse Body Bias

- State retentive
- Design only
  - Can be used on any process
  - Allows use of a leakier, faster process at same I_{SB}
- Increase V_{SB} during standby
  - Raises V_t due to "body effect"
    - Electrical control allows this only during standby
  - Use $V_{SB} = 0$ during active operation
- Done first commercially on 0.25 μm microprocessor [1]
  - $V_{DD} = 1.8$ V
  - N-well driven to IO voltage (3.3 V)
  - Charge pump drives P type substrate to –2 V
  - Fine granularity power supply grids
    - 1000's of local supply switches

RBB and Power Supply Collapse

- 0.18 μm PMOS transistor measurements

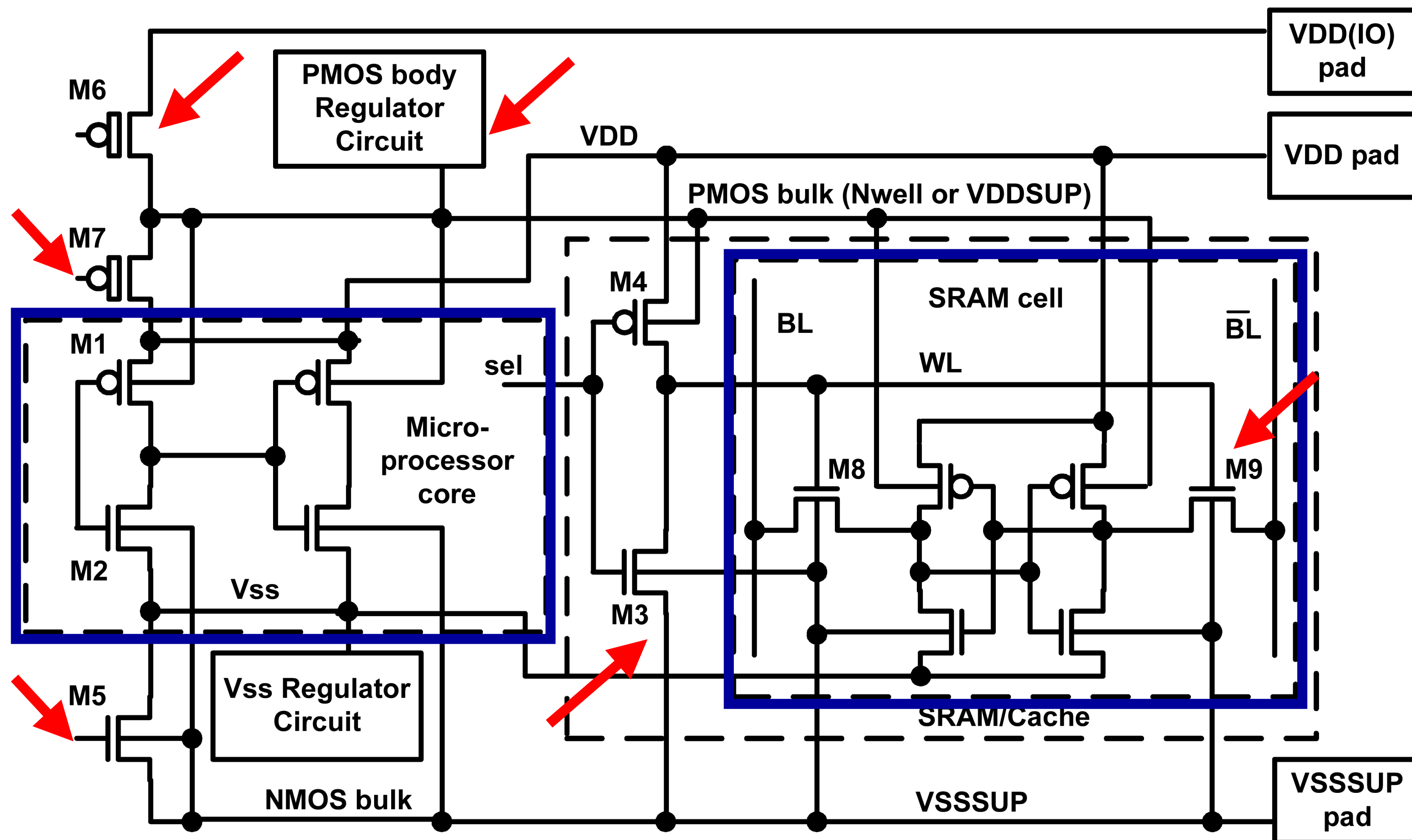


Drowsy: RBB and Supply Collapse

- Drain Source Leakage (I_{off})
  - Decreases at linear or better rate with V_{DD} collapse
    - Depends on process DIBL
  - Decreases with a square root V_{SB} dependency
- Gate Leakage
  - Decreases faster than V_{DG}^2 (can give V^4 power impact)
    - Sensitive to physical oxide thickness
- Gate induced drain leakage (GIDL)
  - Lower voltage has a very large effect
    - Essentially eliminated with supply collapse
- Junction band to band tunneling currents
  - Unaffected!
    - This requires careful transistor design and circuit design interaction
    - Otherwise likely to be *the limiting factor* in future usage

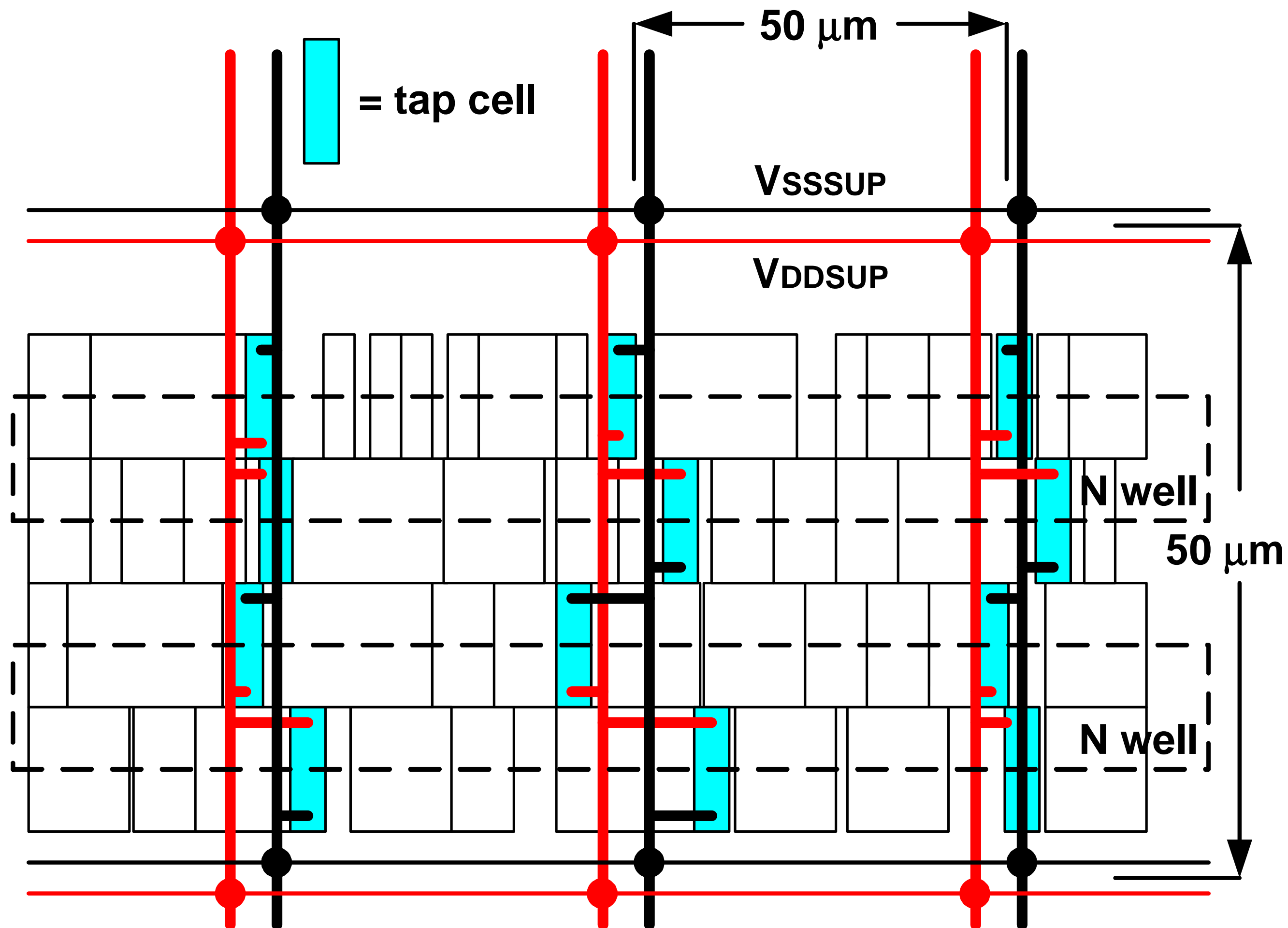
RBB Circuit Design

- Apply body bias by raising the source
  - Naturally applies supply collapse



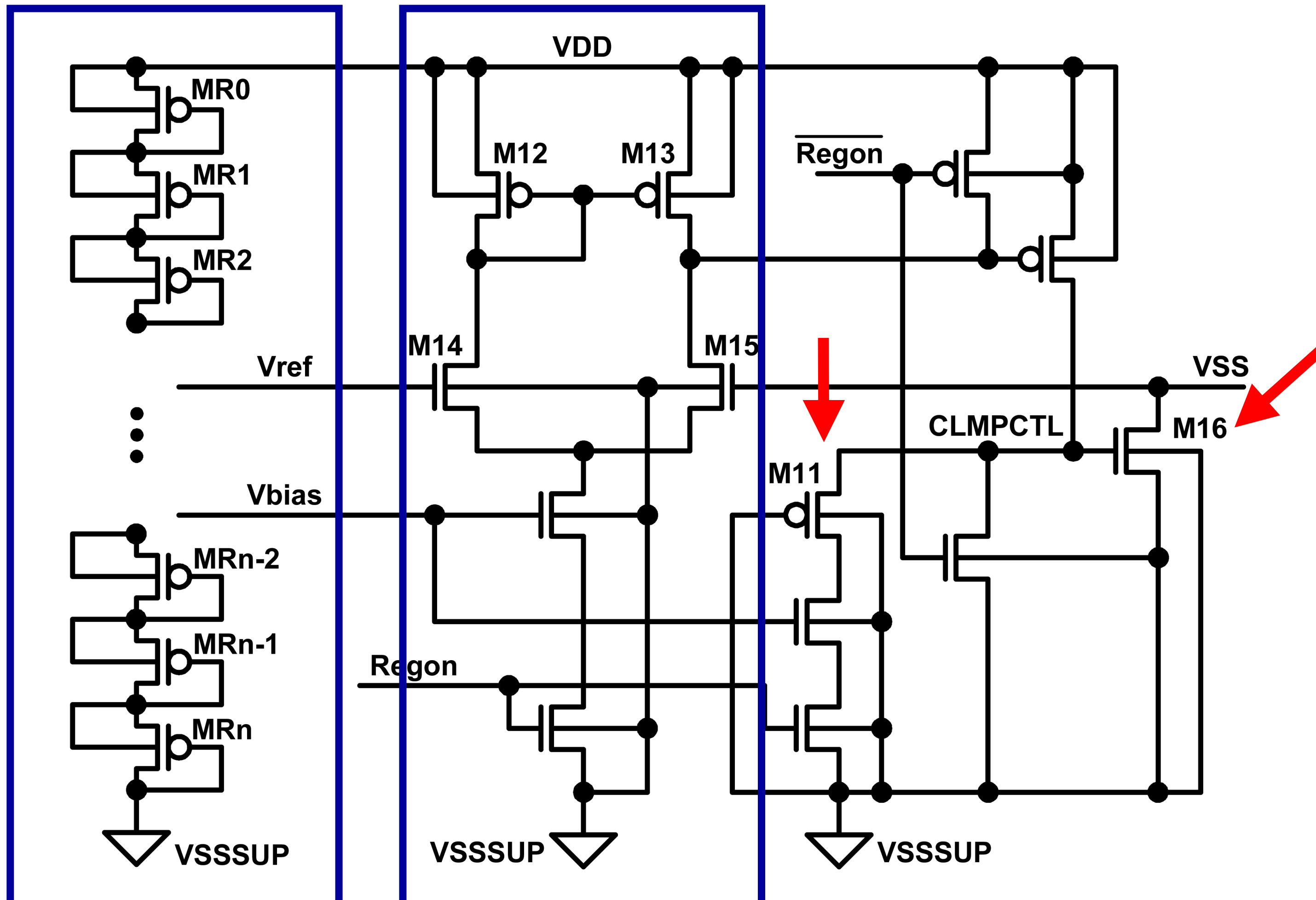
Power Supply Routing

- Substrate and well taps on a 50 μm grid [11]
  - Highly doped epi substrate for low V_{SSSUP} impedance
  - N-wells contiguous to grid in substrate for V_{DDSUP}



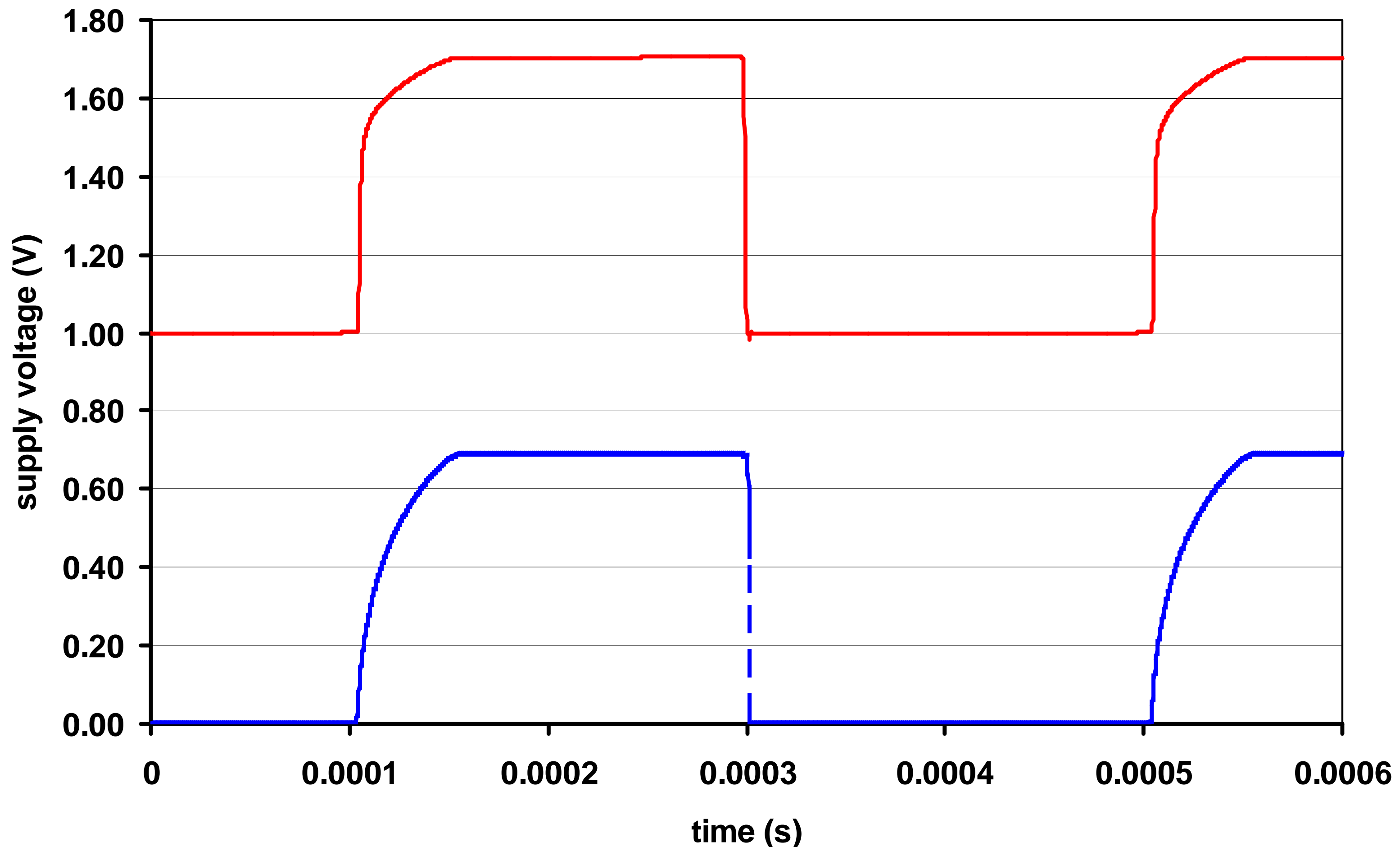
RBB Circuit Design

- Amplifier & reference voltage based V_{SS} regulator
  - Reference tracks with V_{DD}
    - Allows larger V_{DS} and V_{SB} if needed



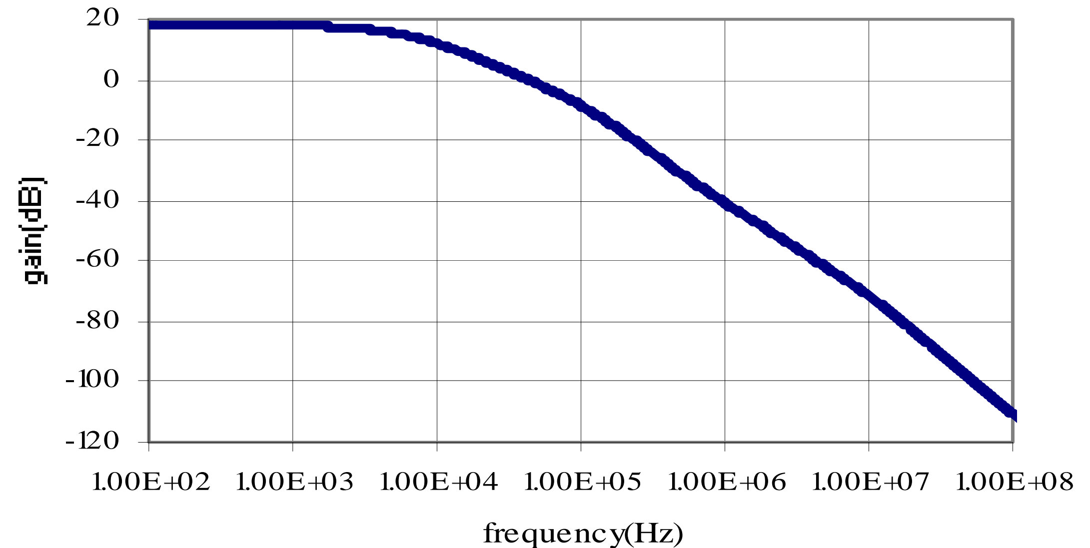
Drowsy Operation

- Well is actively pulled up for RBB
  - Logic circuit leakage passively pulls V_{SS} up to produce RBB and power supply collapse



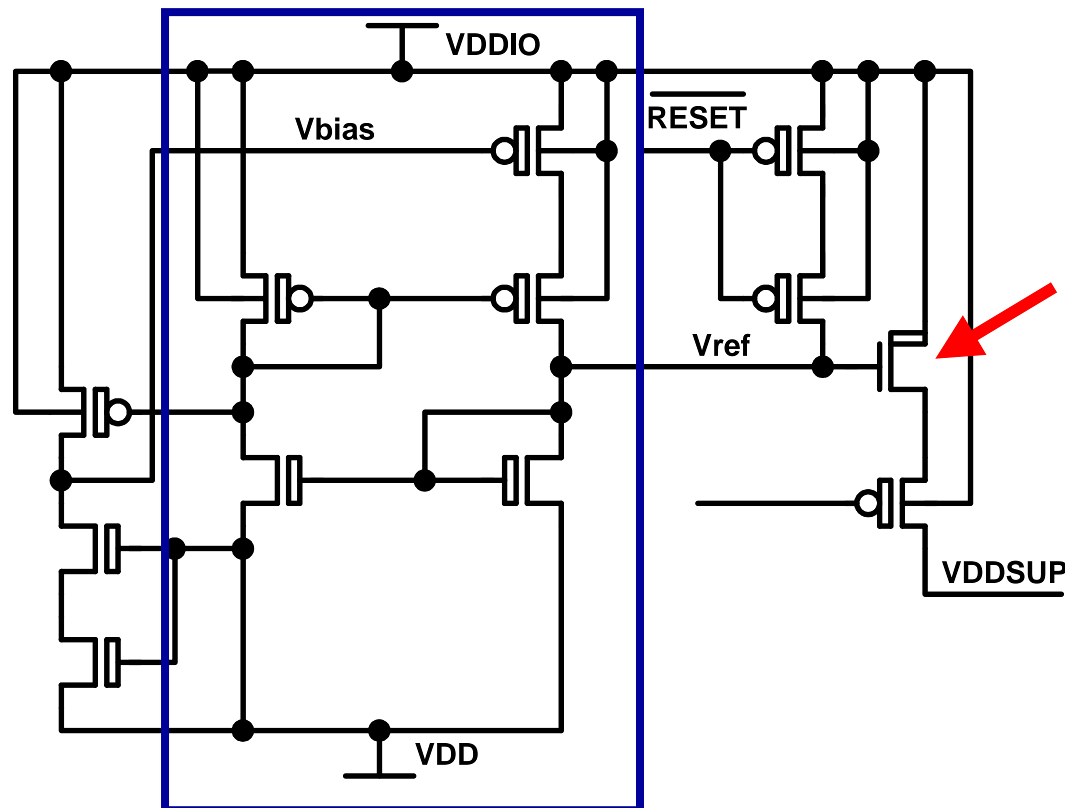
V_{SS} Regulator Stability and Power

- Design achieves 60° phase margin at all process corners
- Amplifier operates in subthreshold
  - Low gain
- V_{SS} regulator consumes less than 4 μA
  - Key since it contributes to total power consumption in Drowsy mode



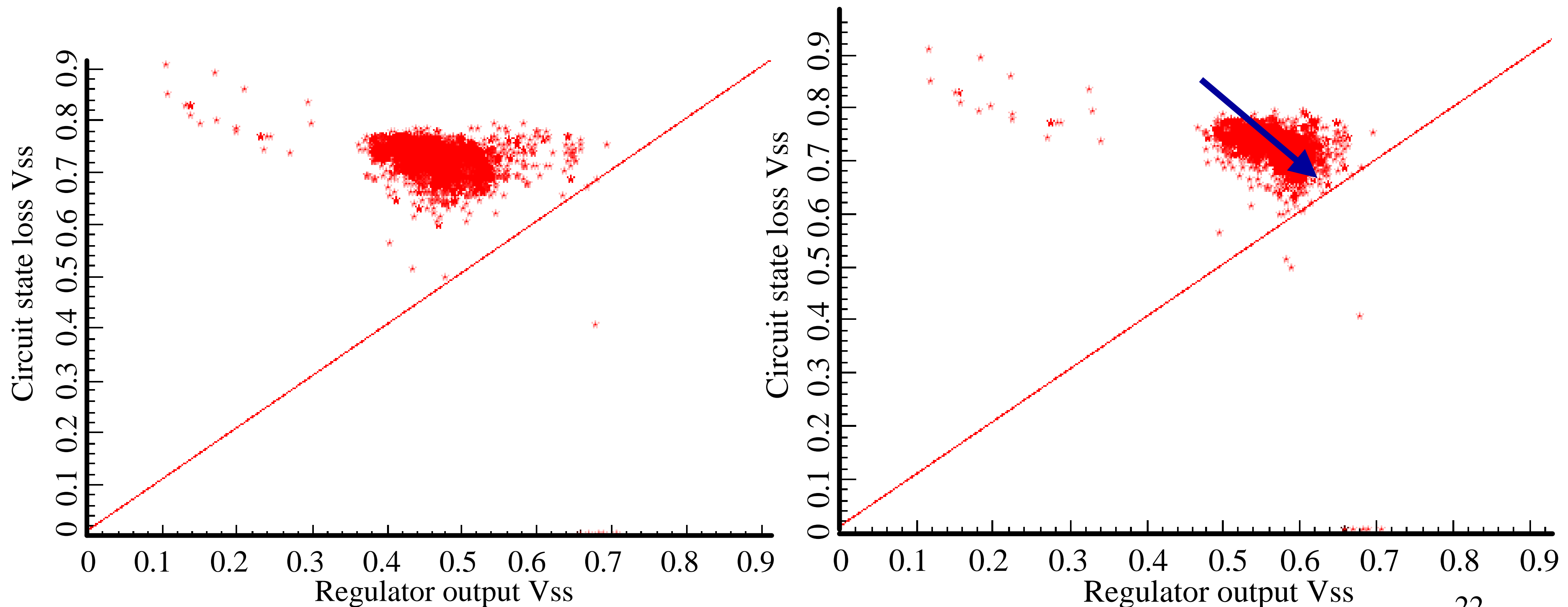
N-well (V_{DDSUP}) Regulator Design

- Low value and high cost to using high voltage
- Use a textbook bootstrapped voltage reference
  - Low V_t VDNMOS source follower
    - Provides very low dropout even with high body bias
    - Note startup circuit



Testing with Guardband

- External access to the internal supply nodes essential
  - Allows observability [12]
  - Allows controllability
    - Find point of fail independent of regulator
    - Drive current into core to provide guardband



Outline

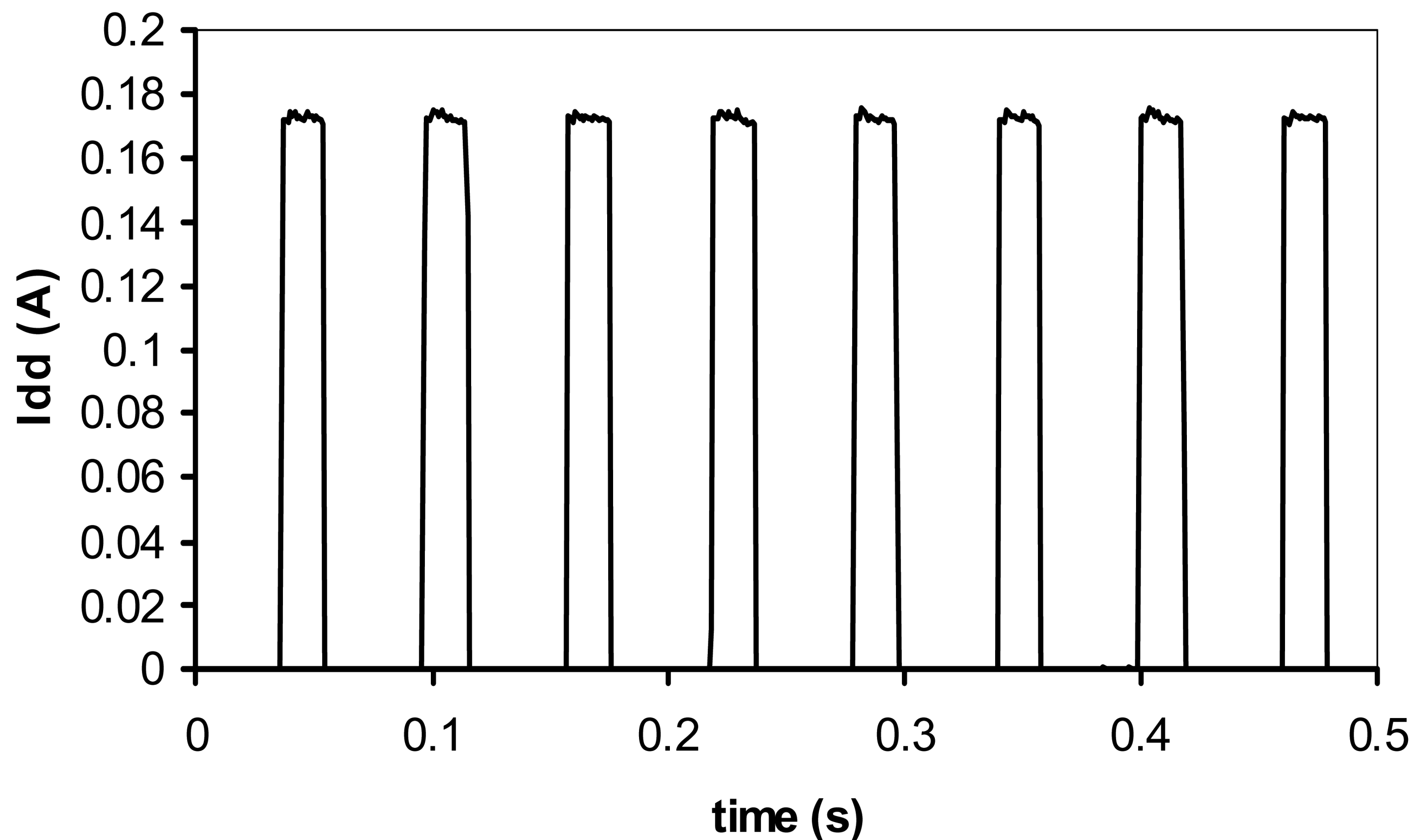
- Introduction and motivation
- Standby leakage management
  - Drowsy mode: Reverse body bias and supply collapse
    - Circuit design and operation
    - **System level results**
    - Limitations
  - Thick gate shadow latches
- Active leakage management
  - Multiple V_t and channel assignment
  - Drowsy memories
  - Thick gate SRAM
- Conclusions

Time Division Multiplexed Drowsy

- Time Multiplex between active operation and Drowsy mode to simulate a low leakage process [11]
  - At low *effective frequency* (F_{EFF}) burst operate at a high frequency to make time for low standby power mode
  - E.g., 300 MHz operation, 30 bursts per second, 100k instructions per burst achieves 3 MHz F_{EFF}
    - 99% of the time is spent in the low standby power mode
- Energy cost of entry and exit must be small
  - Need to amortize this penalty with leakage savings
  - Can't know *a-priori* duration of standby state
- Applicable to cellular communications
  - 1-2 seconds between contact with cells in standby
- Applicable to hand-held devices, e.g., PDAs
  - Between keystrokes or pen-strokes

Experimental Operation

- Board using an Intel XScale 80200 microprocessor
  - Power supplies brought external to measure power consumption using Agilent ammeter and PC
  - Separate measurements accounted for IR drop



TDM Drowsy: Code and Behavior

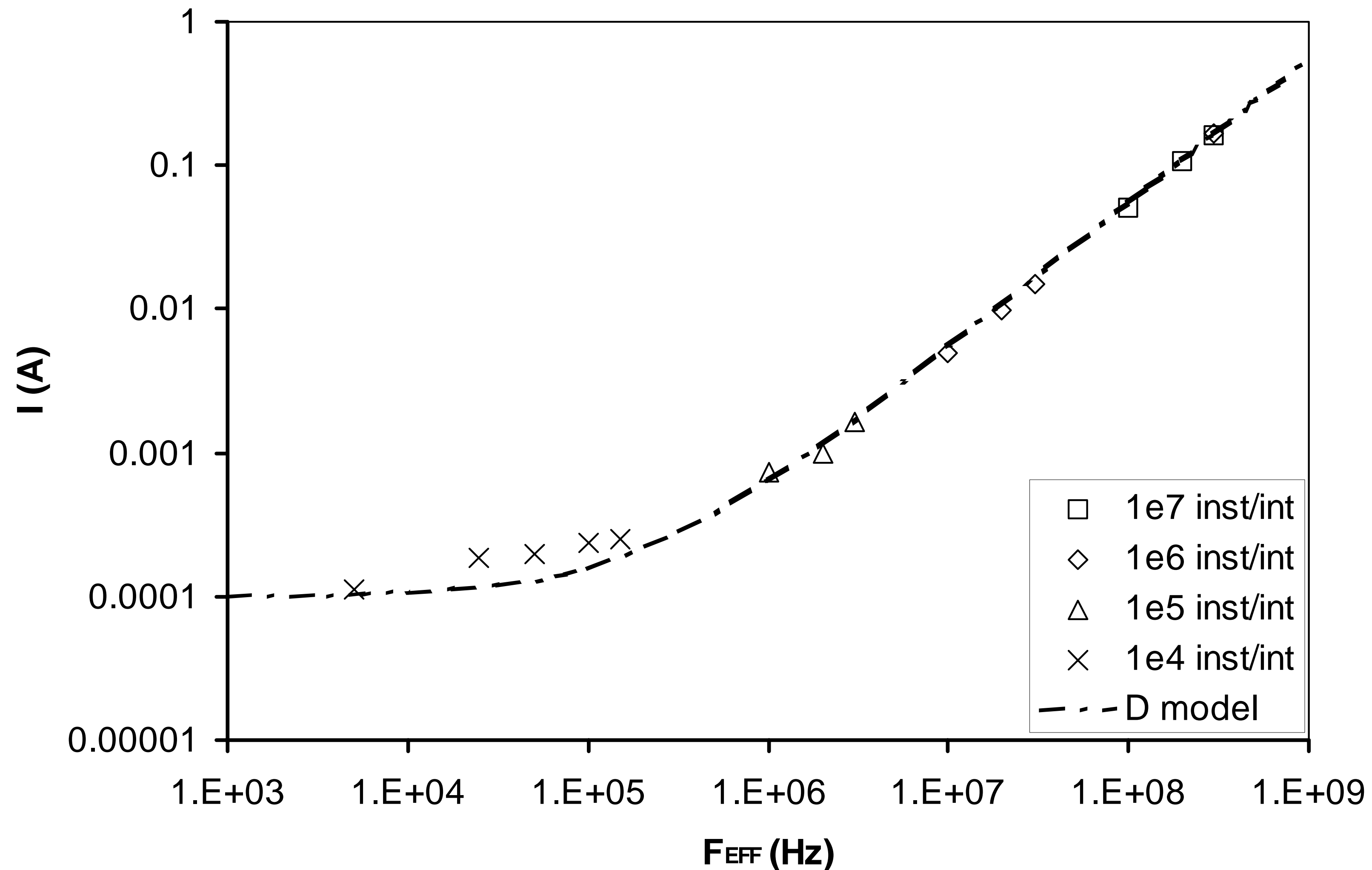
- Code was programmable to run a loop of instructions as the interrupt handler
  - Loop counter determined the interrupt instruction count
- At the end of the loop, the microprocessor re-entered Drowsy mode
  - Drowsy mode is exited by interrupts
- Code:

```
outerLoop:
        MOV     R0, #instructions_per_interrupt
work:   SUBS    R0, R0, #1                      ; decrement count
        BNE     work                            ; loop while count != 0
        DROWSE                                  ; wait for interrupt
        B       outerLoop
```

- BTB holds state, no cache misses

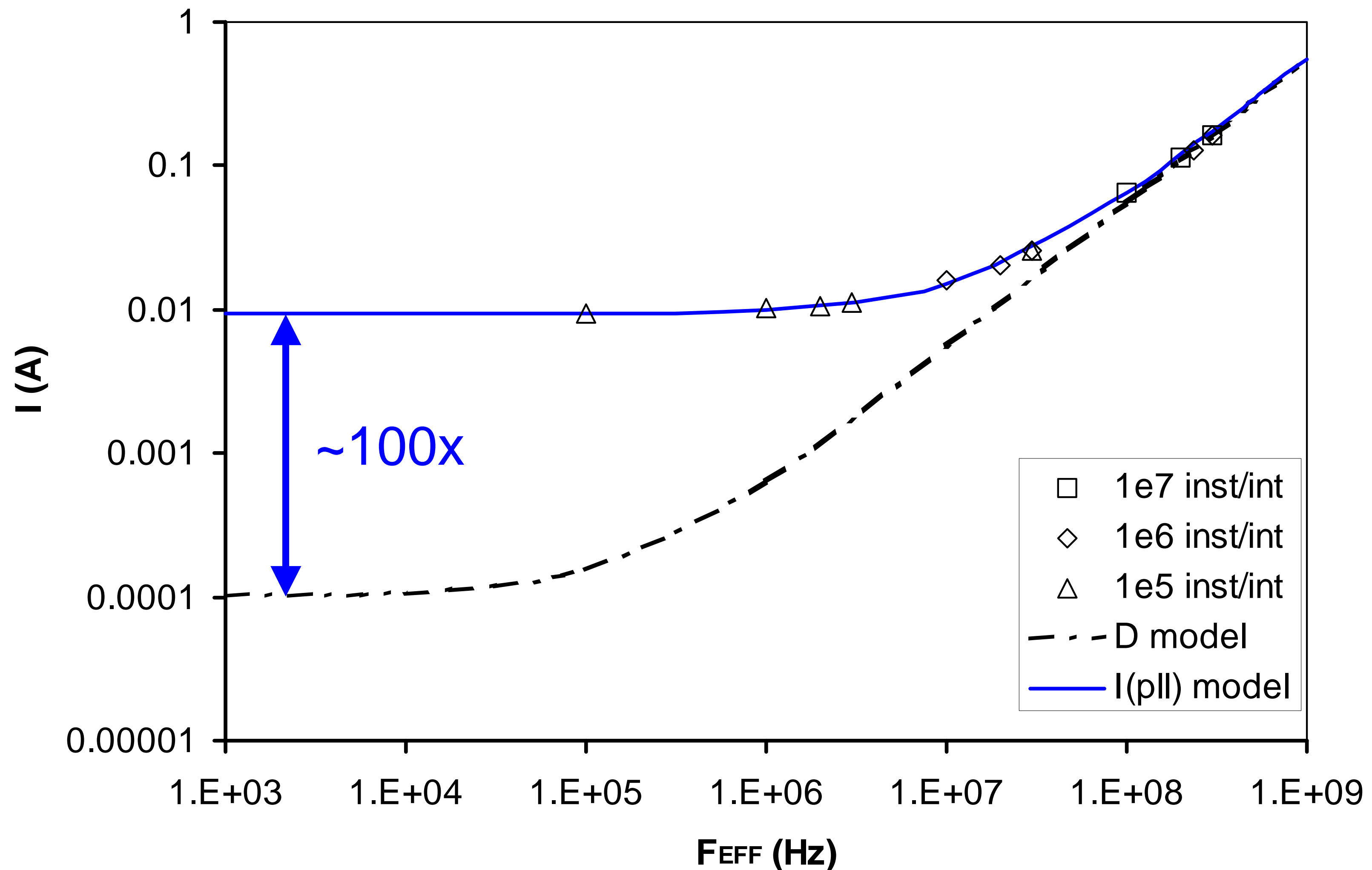
TDM Drowsy: Results

- Measured system results
  - 300 MHz, V_{DD} = 1V



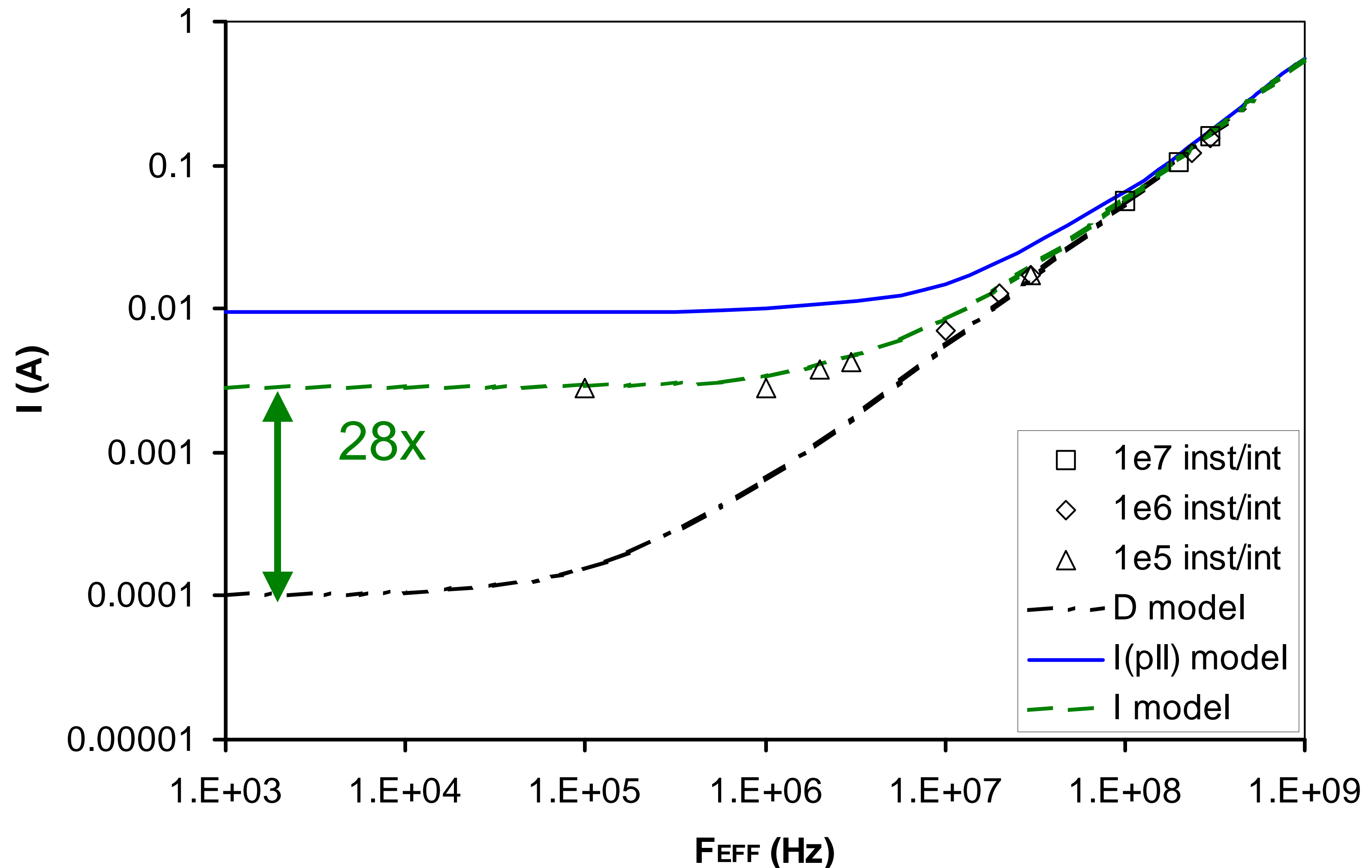
TDM Drowsy: Results

- Comparison with "Standby" mode
  - Standby has single IO clock interrupt latency



TDM Drowsy: Results

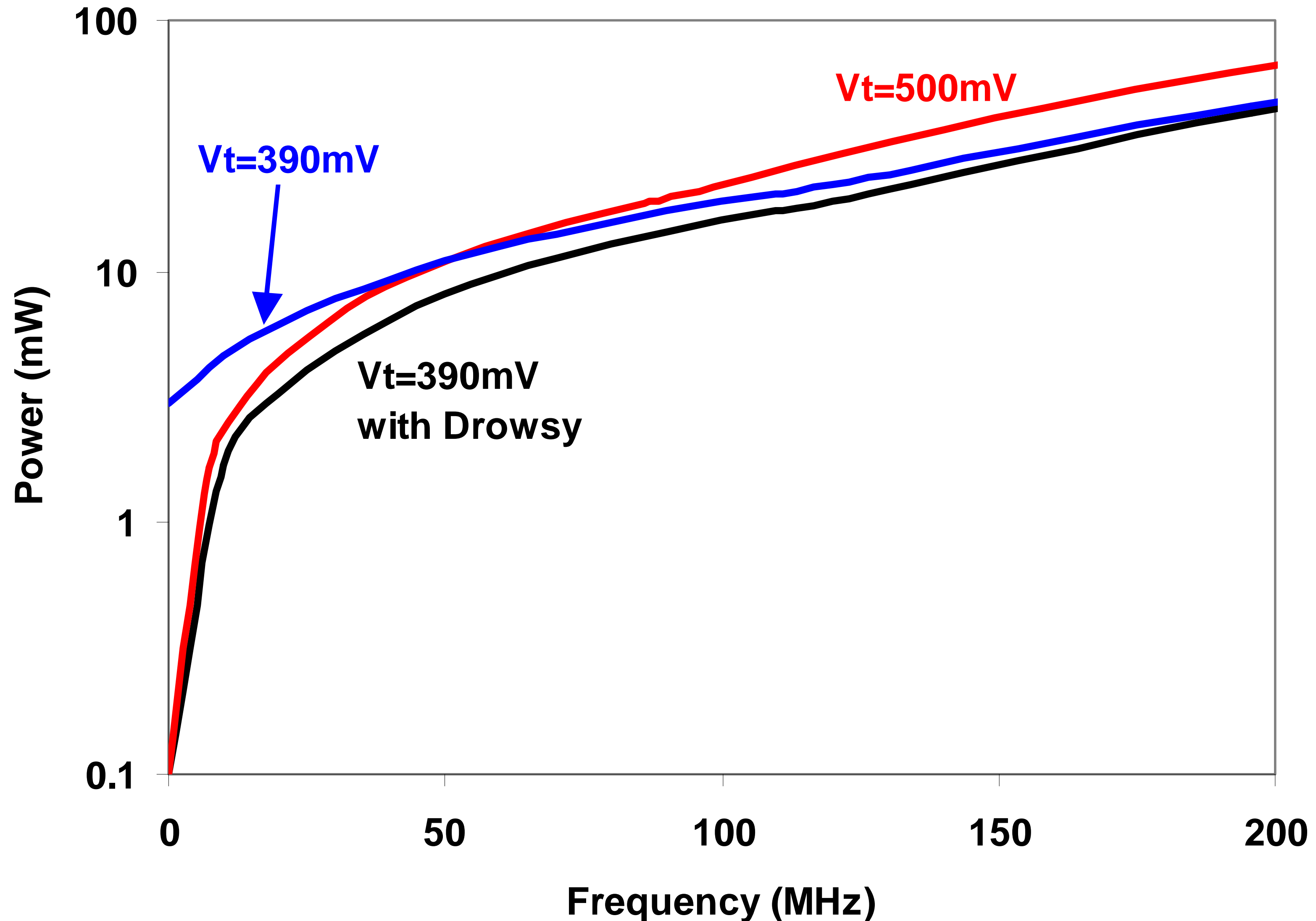
- "Standby" with PLL disabled
  - Leakage reduction limited by regulator resolution



Energy Cost of Entry and Exit

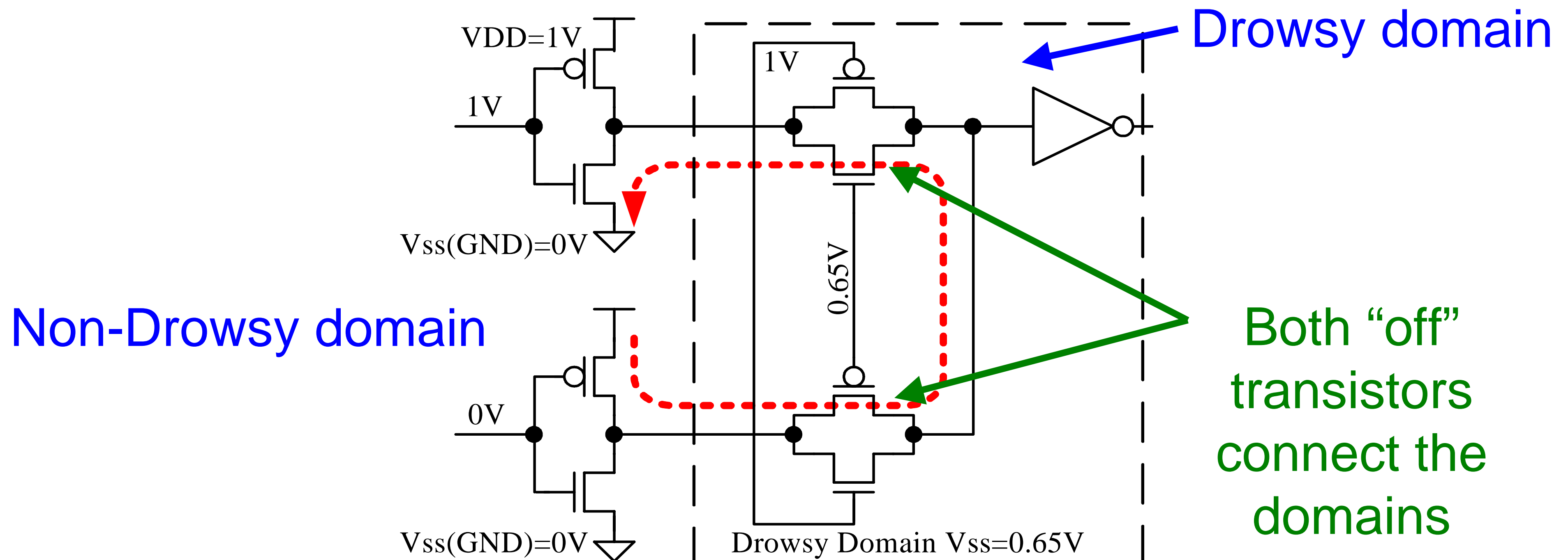
- Power Components
  - Active power
    - Active mode leakage—lumped with the above
  - PLL power
    - During operation and for 20 μs before active operation to lock
  - Drowsy mode leakage
  - Power supply movement power
    - Passive entry saves ½ of the V_{SS} component
    - Also saves power if resume soon after entering Drowsy
    - Both V_{SS} and V_{SSSUP} are small swing
    - C_{VSS} = 55 nF
    - C_{VSSSUP} = 5 nF
- Total energy overhead equivalent to approximately 60 clock cycles of active power at V_{DD} = 1V

Voltage scaling: Effect of Drowsy



Interfacing Domians

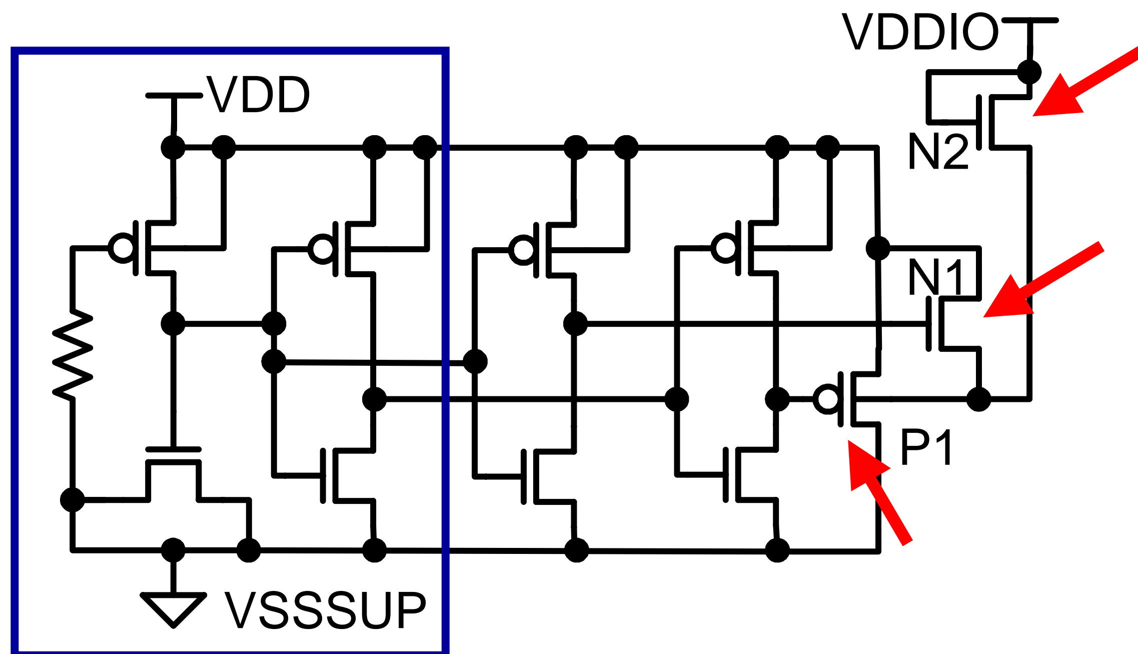
- It is easy to create sneak paths! [12]
  - Signals between voltage domains must be driven full rail
  - Avoid pass gate interfaces
    - Latches suffice to isolate domains



See also [13] for a set of rules for MTCMOS designs

Low Leakage ESD Clamping

- Drowsy mode current small enough to make otherwise negligible contributors significant
  - Large PMOS transistors in ESD clamps important
  - Fixed by RBB on clamp devices [14]
    - Also provides FBB during ESD transients
    - Equivalent or better performance when tested using HBM, MM



Other Implementations

- Same scheme used in [15] for 0.13 μm SRAM
  - No PMOS body bias
    - We can speculate that the PMOS leakage was substantially lower than NMOS, so no value in PMOS RBB
    - We have seen this on other 0.13 μm processes
- This approach will be used for 65 nm handheld devices [16]
  - 0.5 V V_{DD}-V_{SS}

Outline

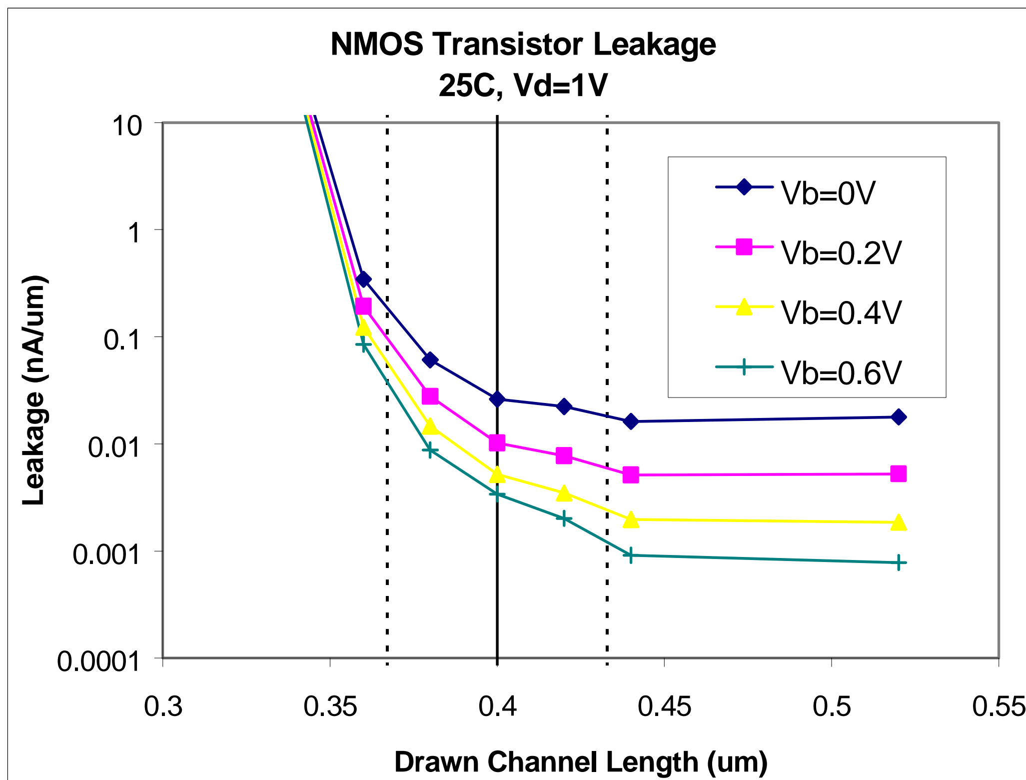
- Introduction and motivation
- Standby leakage management
  - Drowsy mode: Reverse body bias and supply collapse
    - Circuit design and operation
    - System level results
    - **Limitations**
  - Thick gate shadow latches
- Active leakage management
  - Multiple V_t and channel assignment
  - Drowsy memories
  - Thick gate SRAM
- Conclusions

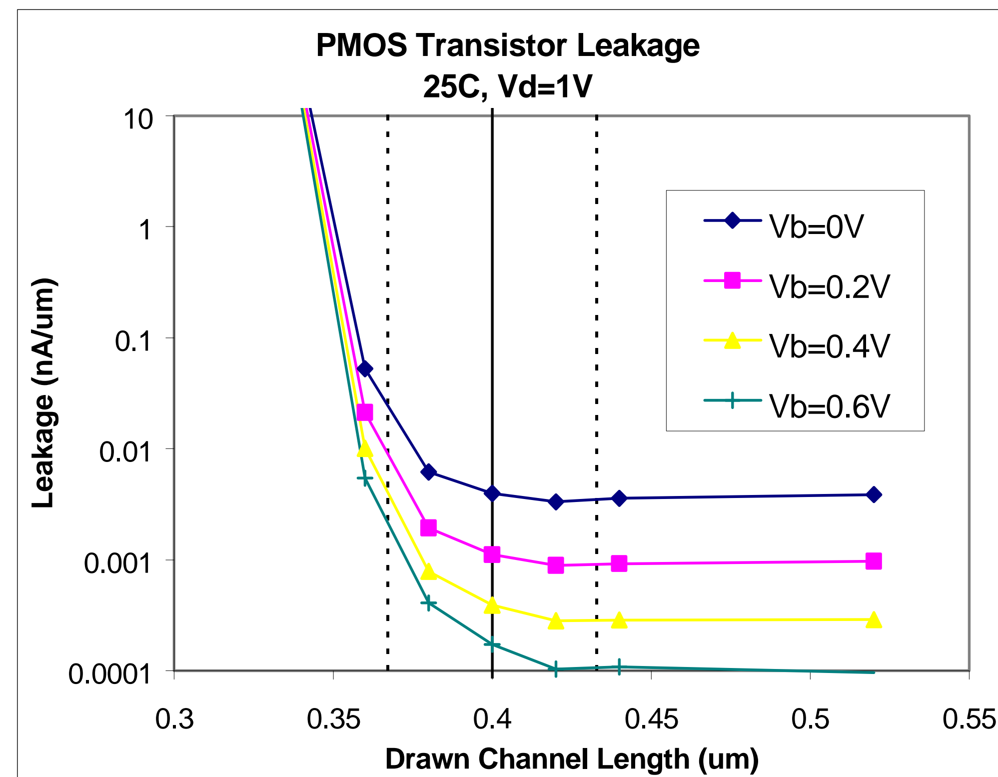
Limitations of Drowsy Modes

- State stability [2]
  - High fan-in domino circuits can have N to P ratios of 100’s to 1
    - Need highly balanced storage, MTCMOS logic
- Channel length
  - Aggressively scaled transistors have poor body transconductance g_{MB}
    - Need to back off from the highest performance possible
- Drain to bulk tunneling currents
  - Requires less steep halo doping gradient at drain
    - No halo is best—this will limit transistor scaling
- Defects
  - Stacking faults generate nearly 10 μA of leakage
    - Turns Drowsy cells into defect detectors
    - May be problematic for strained silicon

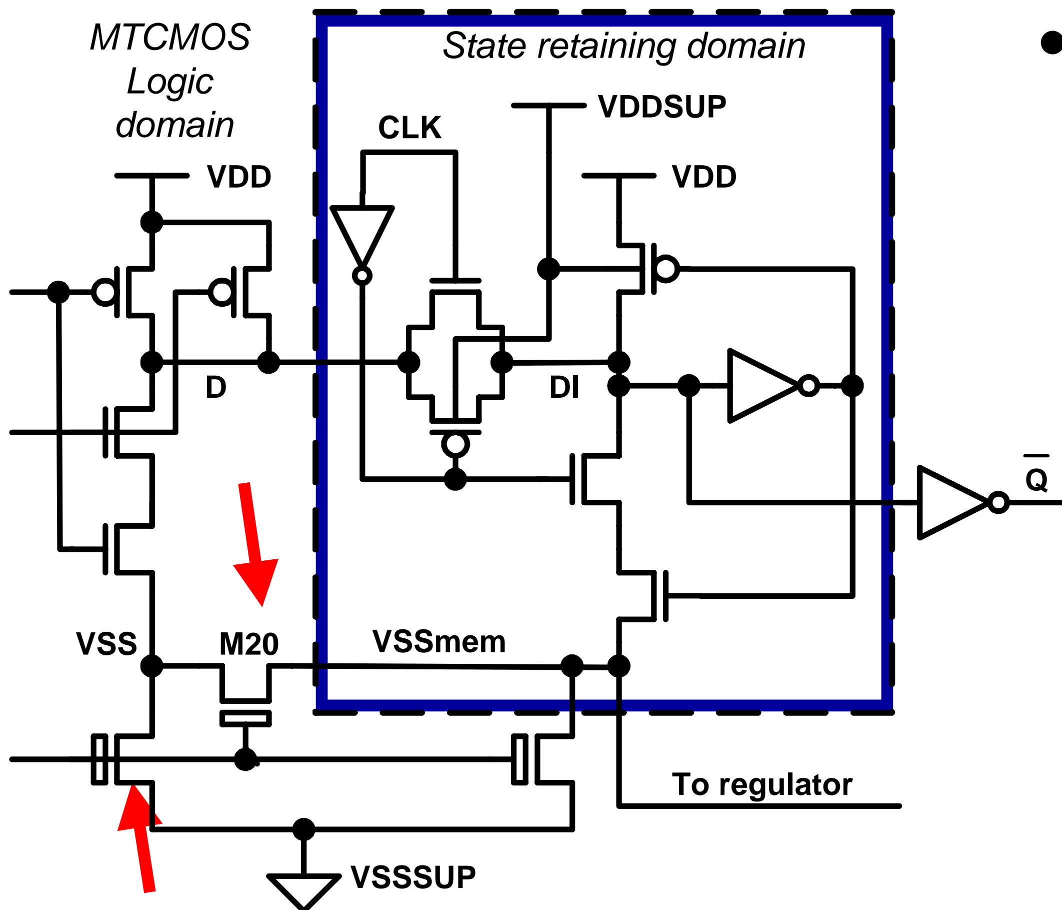
Leakage Control Limitations

- Body bias vs. channel length
  - Bulk control lost as transistors approach punchthrough
    - This is where high performance processes are often targeted



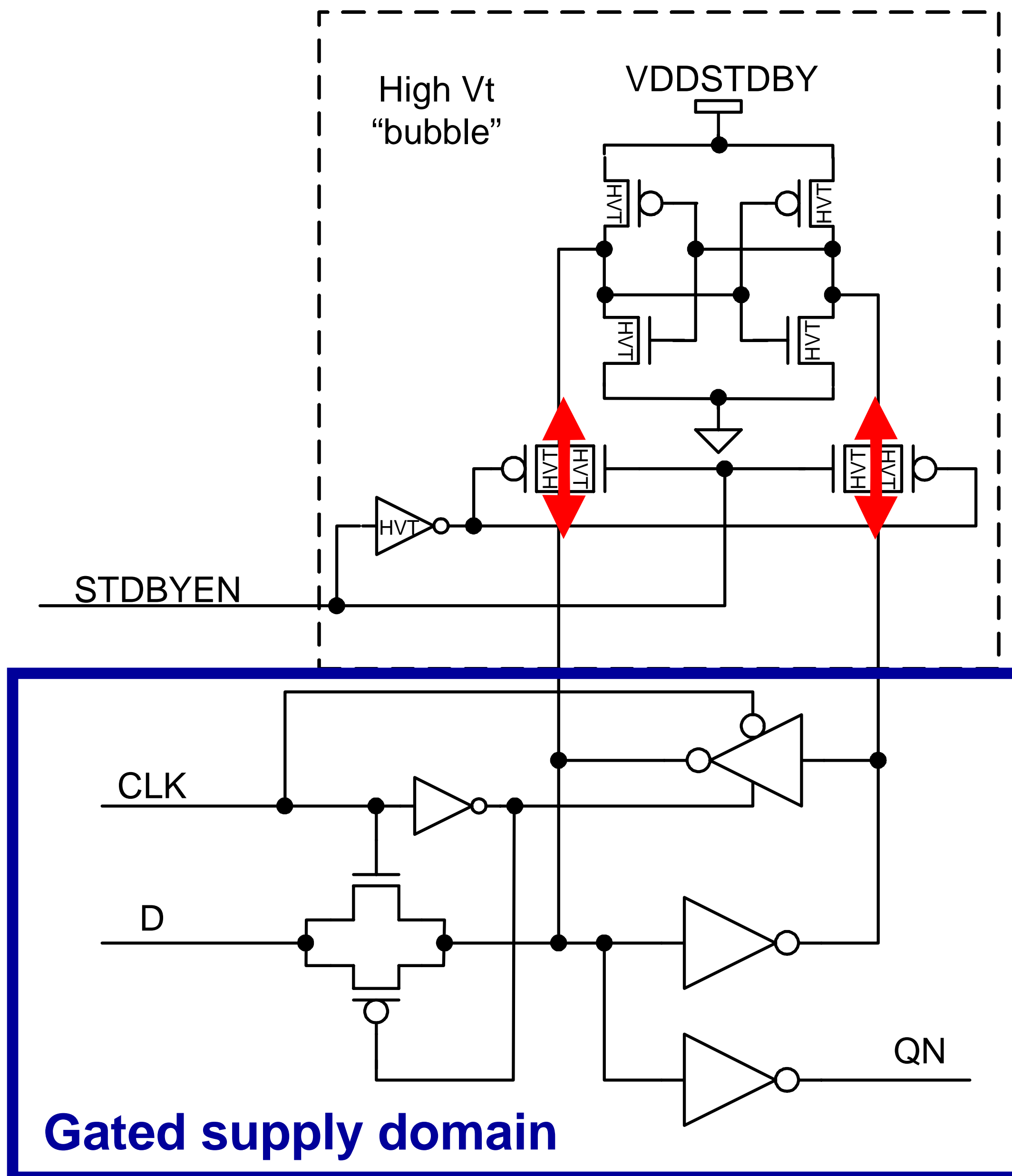


Combining MTCMOS and Drowsy



- Only state elements have RBB applied [11]
  - The rest of the circuits are "slept" using MTCMOS
    - This eliminates about 2/3 of the total leakage
  - Allows highly balanced state elements
    - Drowsy can be pushed to even lower V_{DS}
    - Leverage high I_{gate} V_{GS} dependency

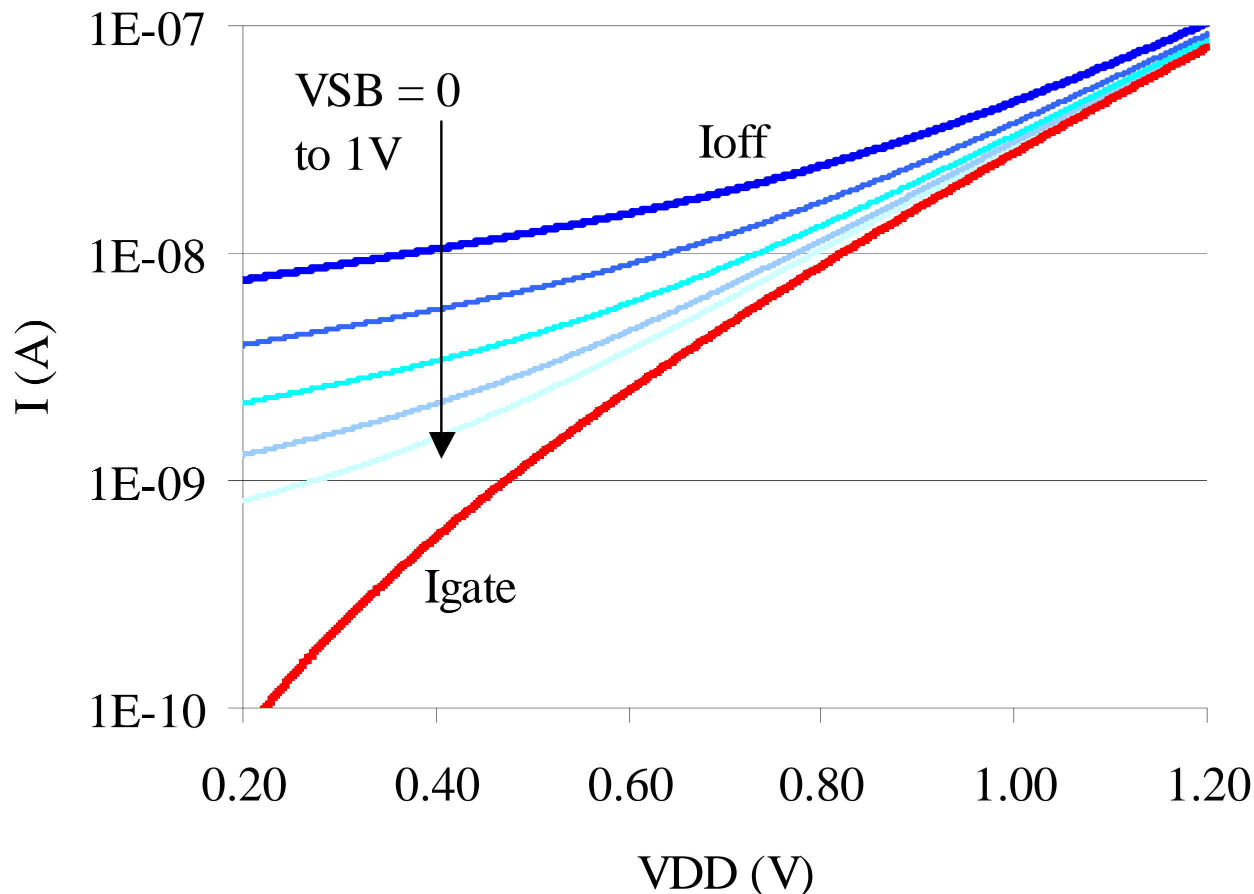
“Balloon” Latches



- A state retentive MTCMOS scheme [5]
  - High V_t transistors gate low V_t circuits
  - State retained in high V_t balloons
    - Circuit speed remains a function of low V_t
- Does not address I_{gate} leakage component

Leakage on Advanced Processes

- Many of the old techniques will still be applicable
  - RBB and supply collapse still works
    - Supply collapse (**VOLTAGE SCALING**) is key
  - 10 µm wide 65 nm NMOS characteristics using BPTM [17]



Leakage on Advanced Processes

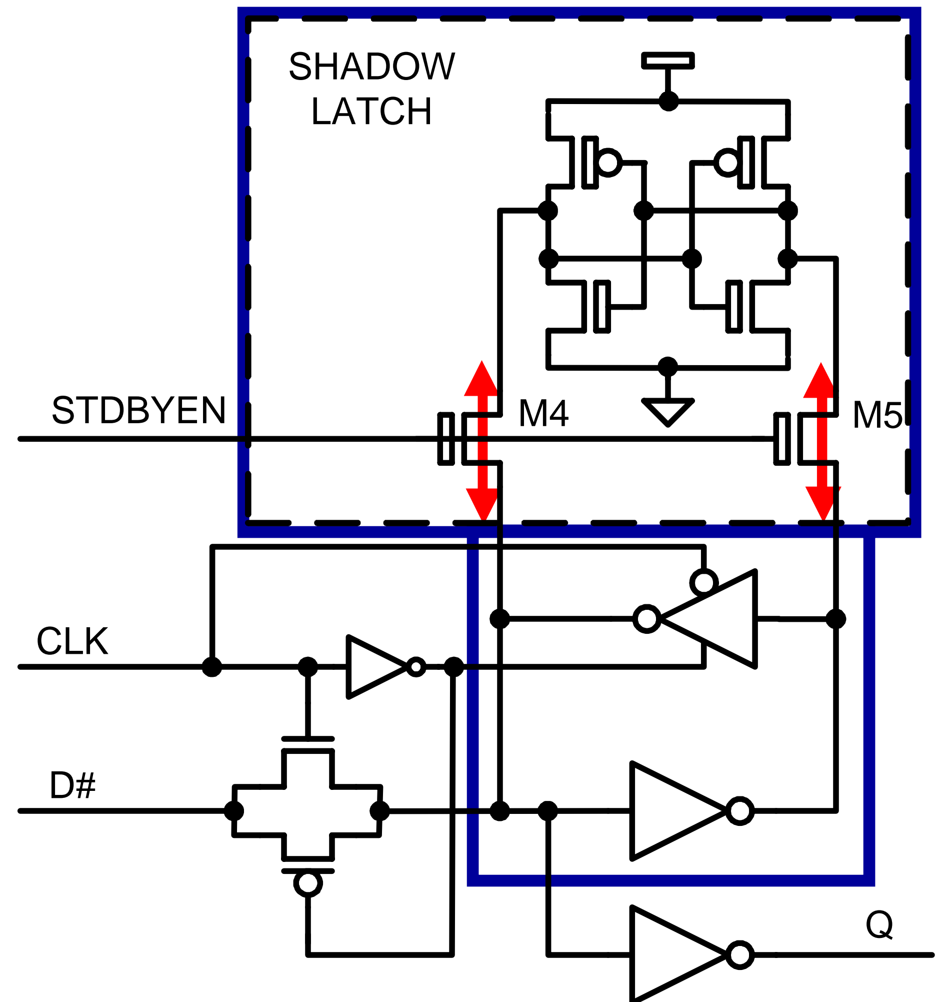
- Gate leakage suppressed 2 orders of magnitude
  - Consistent with previous results [18]
- RBB and supply collapse still works
  - Cutting the voltage is critical
    - Lowers I_{off} by the DIBL coefficient
    - Pulls the transistor away from punchthrough and gives control back to bulk
    - Longer channel suppresses DIBL the same way
- Key point not obvious is drain to bulk band to band tunneling
  - This is becoming the dominant component and is helped by lower voltage
  - Transistor design is also important
- Future devices will apply “Drowsy” style RBB and supply collapse for SRAM’s on 90 and 65 nm [16]

Outline

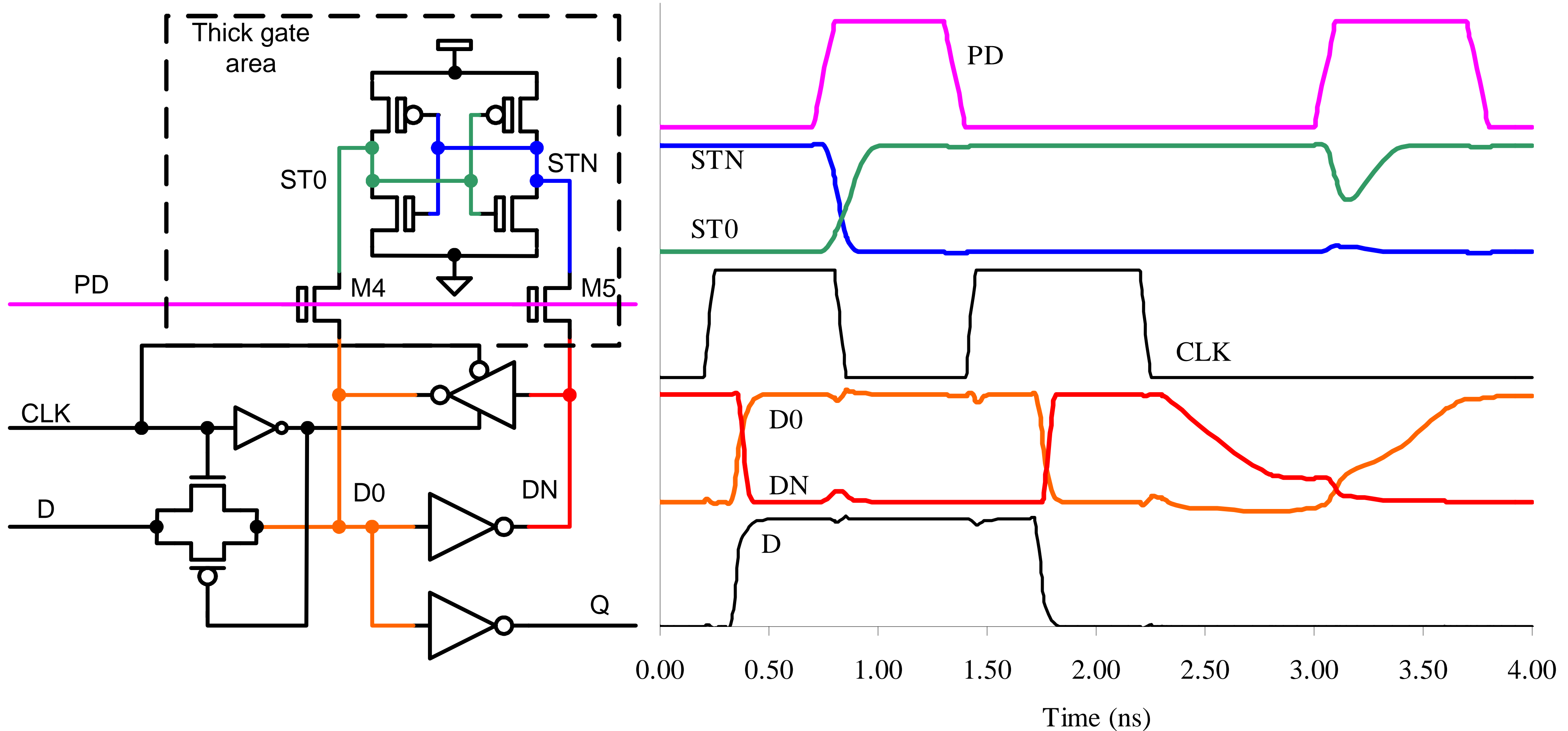
- Introduction and motivation
- Standby leakage management
  - Drowsy mode: Reverse body bias and supply collapse
    - Circuit design and operation
    - System level results
    - Limitations
  - **Thick gate shadow latches**
- Active leakage management
  - Multiple V_t and channel assignment
  - Drowsy memories
  - Thick gate SRAM
- Conclusions

Addressing I_{gate}: Thick Gate State Retention

- Add a “shadow” thick gate (and high V_t) latch to retain state during standby [8]
  - Using the thick gate IO transistors implies higher V_t
    - The gate length can be pushed--not exposed to high drain voltages
- Eliminate thin gate latch if speed is unimportant



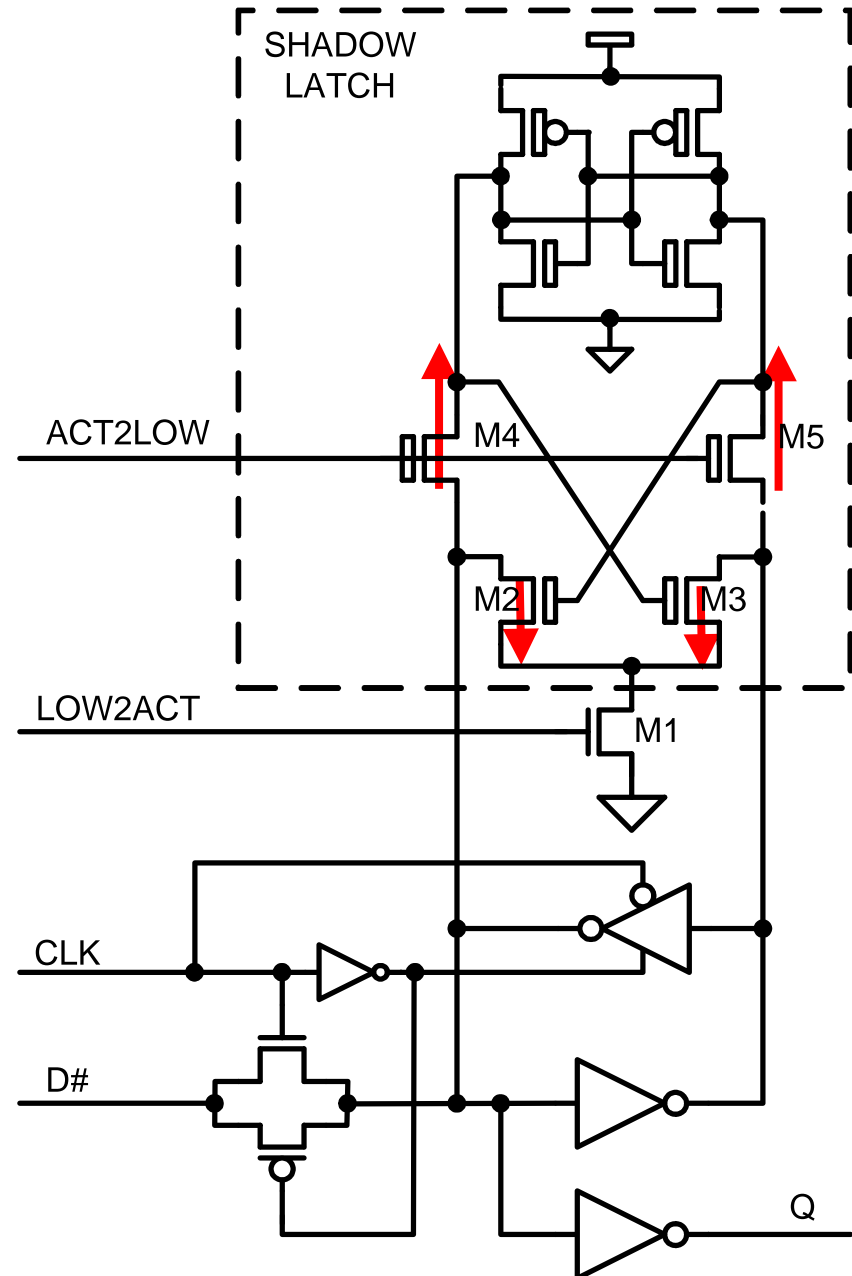
Thick Gate State Retention Operation



– TSMC 0.18 µm thick gate and BPTM 65 nm thin gate

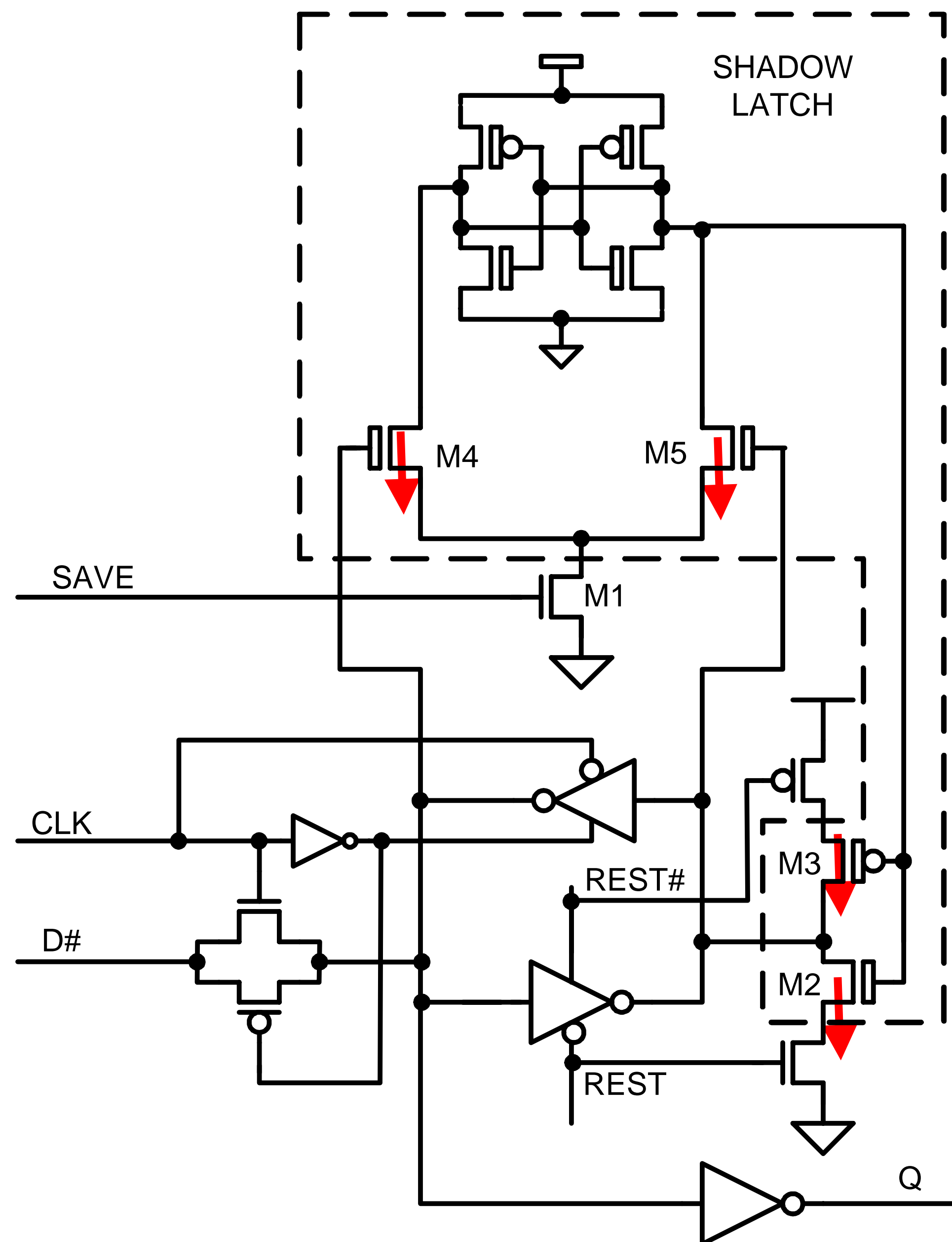
Safer Thick Gate State Retention

- The V_{DD} supply needn't be completely discharged if a uni-directional path is provided from the thick to thin gate circuitry [9]
  - Recall that it is best to disable supplies, allowing movement via leakage rather than driving them low



Thick Gate State Retention: Another Approach

- This approach uses uni-directional circuits in both directions
  - More transistors
  - May ensure thick gate write-ability over a wider voltage range
- Adding transistors to slave in MSFF does not incur a speed penalty
  - From USPTO website [10]



Thick Gate State Retention: Results

- Master-slave flip-flop
  - Using projected 65 nm thin gate transistors
    - I_{off} = 10.1 nA/μm
    - I_{gate} = 8.0 nA/μm
  - Using TSMC 0.18 μm thick gate
    - I_{off} = 10 pA/μm
    - 40 angstrom electrical t_{ox} provides negligable gate leakage
  - Result is over 7200x leakage savings for just MSFF!
    - Does not include thin gate logic between the flip-flops
- So we've eliminated the I_{gate} standby contribution
- **But… The real limiter will be drain edge band to band tunneling**
  - Not modeled in this analysis
  - Will require process work
    - Cost?

Outline

- Introduction and motivation
- Standby leakage management
  - Drowsy mode: Reverse body bias and supply collapse
    - Circuit design and operation
    - System level results
    - Limitations
  - Thick gate shadow latches
- **Active leakage management**
  - Multiple V_t and channel assignment
  - Drowsy memories
  - Thick gate SRAM
- Conclusions

Managing Active Leakage

- Leakage is becoming a large part of overall power
  - Up to 40% of total power at the worst-case corners on high performance processes
  - Not as problematic on low power processes…yet
- Transistor scaling will increasingly force "low power," really low-leakage processes to lower V_t's or some of the scaling value will be lost
  - Designers that deal with the leakage will provide a competitive advantage compared to those that don't
- But this is hard
  - All schemes create some kind of cost
  - Cost must be optimized to balance active/standby power
    - Including the energy cost of moving between the states

Shutting Down Units

- Can work well for low activity factor blocks
  - Floating point, small granularity cache banks on μP
  - Some blocks may be unused for applications on SOC
- Key factor is the cost of discharging and charging the block power supply
  - The power cost must be amortized by the leakage savings
  - Also must be careful about IR drop through the switches
    - Very difficult
- Switch overhead is very low
  - Beware of inductive effects on supplies [18]
  - My solution is under-driving the switch transistor gates [2]
    - [18] used staged turn-on of the switches

Shutting down units: Example

- Multiplier total transistor width is 60 mm
- Leakage power savings is

  $P = VI = ½ (0.8)\ V\ I_{leak} = 6.5\ mW$

  (essentially 0.065 nJ per cycle at 1 GHz)

  at $V_{DD} = 1.3\ V$ and $I_{leak} \sim I_{off} = 100\ nA/\mu m$ @ 100^oC
- Power supply capacitance of 0.5 nF
  - One on gating is 0.33 nJ
    - That's 5 clock cycles at 1 GHz
  - Results will be very sensitive to decoupling capacitance
    - More is better for performance and noise
    - Less is better for gating
- It can be difficult to do this effectively and easy for particular behaviors to become higher power

Outline

- Introduction and motivation
- Standby leakage management
  - Drowsy mode: Reverse body bias and supply collapse
    - Circuit design and operation
    - System level results
    - Limitations
  - Thick gate shadow latches
- Active leakage management
  - **Multiple V_t and channel assignment**
  - Drowsy memories
  - Thick gate SRAM
- Conclusions

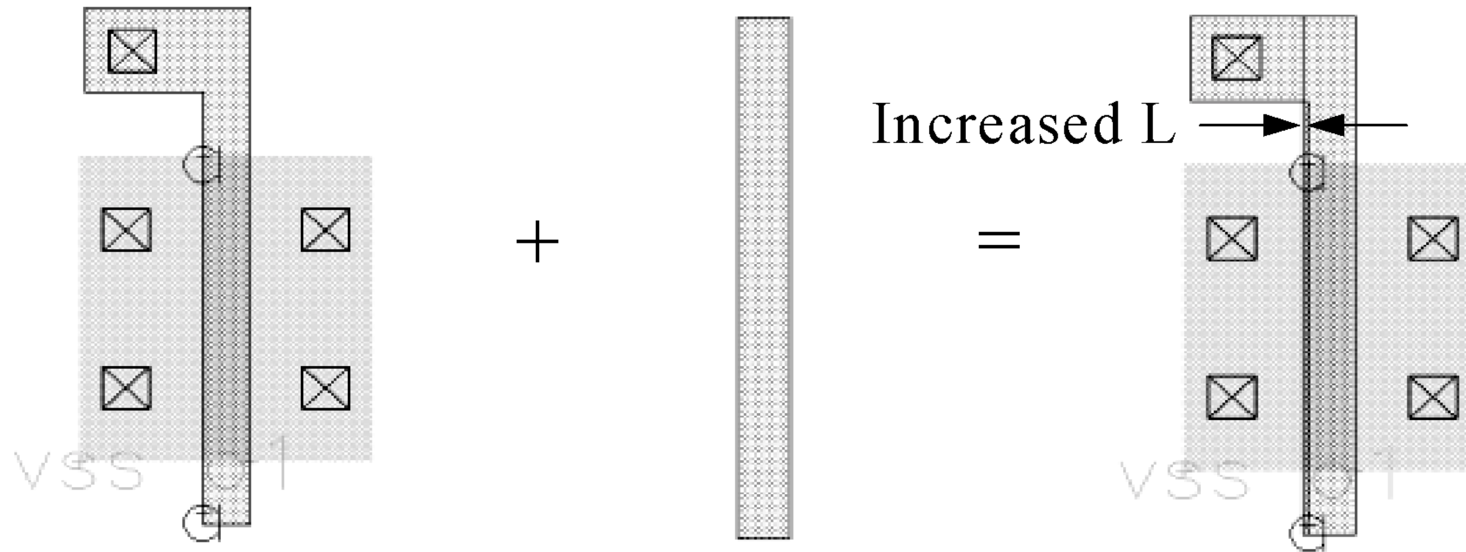
Dual Threshold Voltages

- About 5% fabrication cost adder
  - No active power adder
  - Leakage cost is over 10x per low V_t compared to high V_t
- Minimize low V_t transistors for low power
  - Design with high V_t
    - Don't forget this increases active (switching) power
    - Higher supply voltage for the same speed
  - Insert low V_t only on difficult speed paths
  - Tendency for widely separated high and low V_t targets
- Can be very difficult for high performance design
  - Tendency to over-insert
  - Difficulty with noise on dynamic circuits
  - Less separated high and low V_t targets
  - Timing accuaracy effects
    - High and low V_t needn't track each other

Multiple Channel Lengths

- Longer channel decreases DIBL rapidly
  - About 7x I_{off} decrease on aggressive 90 nm process
- No process cost
  - But a small size adder (1-2%)
- Active power cost
  - Up to 15%, depending on activity factor
    - Caution required with insertion on high speed circuits
  - Slow circuits dominated by leakage can be all long L
- High V_t and low V_t track each other at process corners
  - Essentially, both get faster and slower together

Multiple Channel Lengths: Physical Design



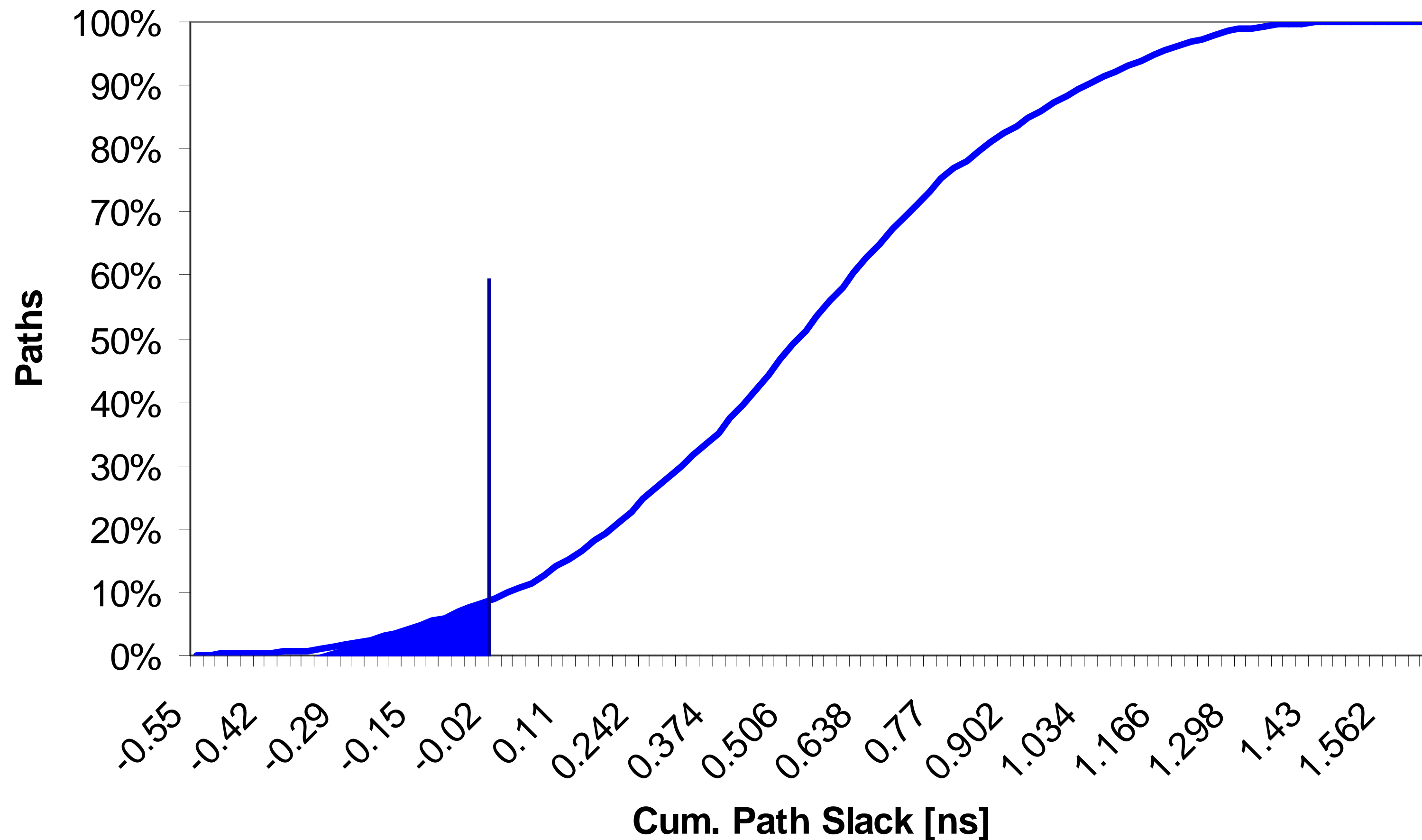
- Add one grid
  - Layout must have one grid of space to add the gate length
  - Small overall area cost
- Can also be added at mask synthesis
  - Better resolution, but harder to check

Multiple Channel Lengths: Timing

- Priority for insertion
  - Low activity factor
    - Leave the clocks alone! Maximize leakage savings
  - Wide transistors
    - Maximize leakage savings
  - Low activity factor
- Work on post-layout data
  - Otherwise you don't really know your timing margin
- Fix hold time violations after long channel insertion
  - Slower (long L) gates fix these for free
- Logic block long L insertion must be automated
  - Timing must be re-calculated after each insertion
  - We used an insertion tool using Langrangian Relaxation

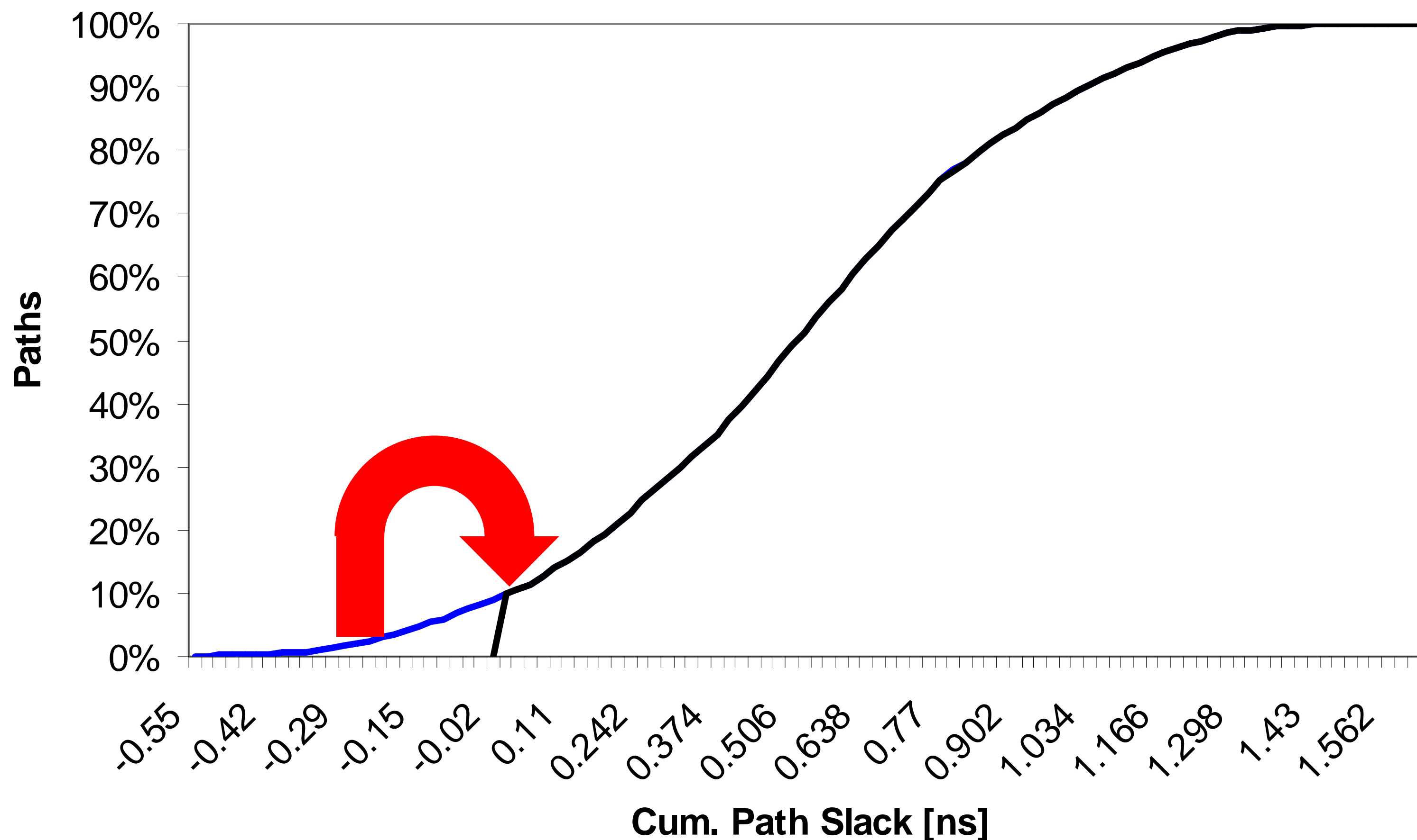
Timing Margin

- Cumulative block timing shows timing slack
  - Negative path fixes comprise the work to meet timing



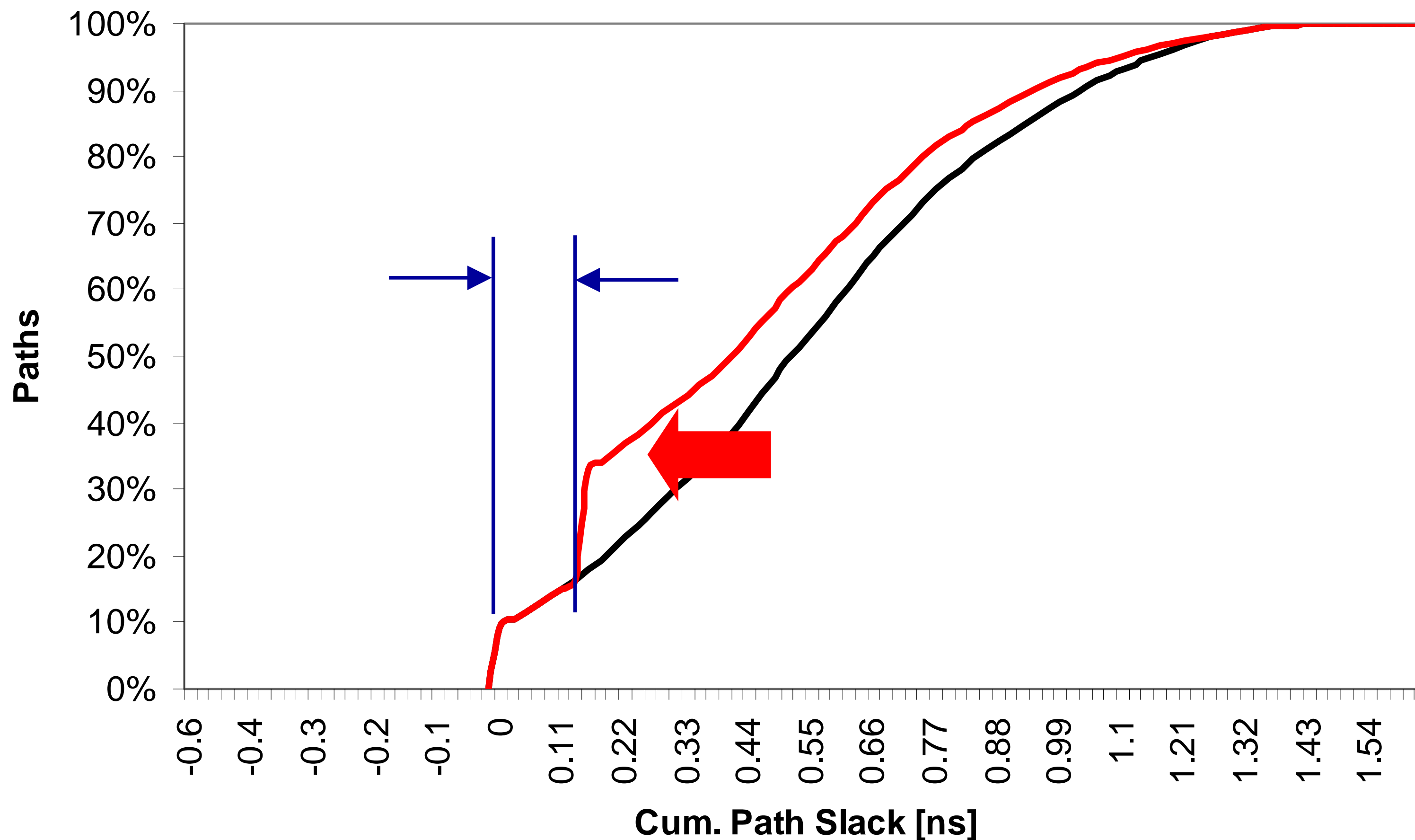
Timing Margin

- Low V_t insertion moves paths from negative slack to zero slack
  - Without Low V_t, this requires logic, sizing changes



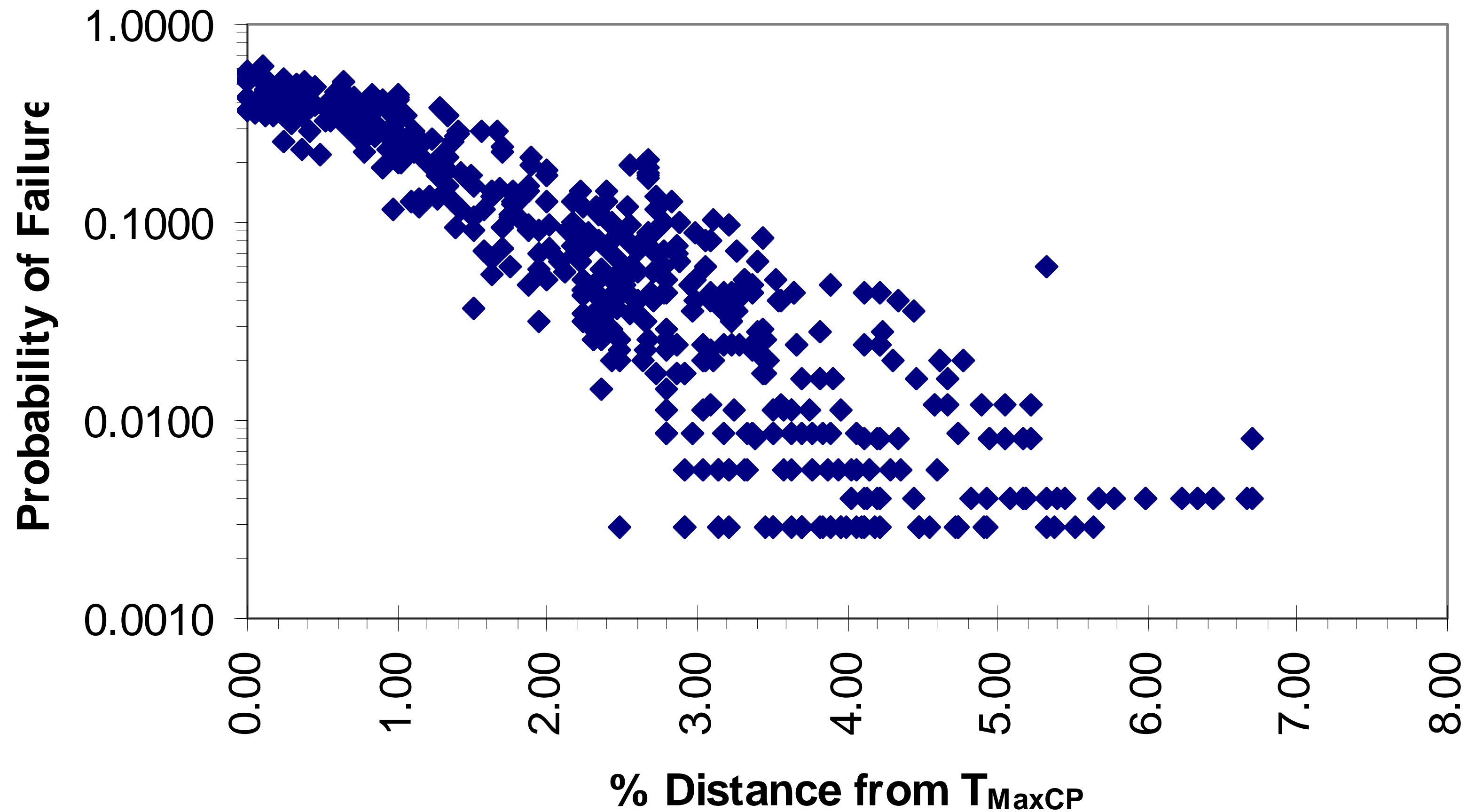
Timing Margin: Long L insertion

- Don't fix every path to zero timing margin
  - Statistical variation will impact the yield
  - Timing tools are not perfect



Timing Margin: Effect of Variation

- Variation modeled as channel length
  - Likelihood of a path becoming worse than 0 ns slack with variation (creating a timing failure) vs. original path distance from critical [19]



Results: Long L insertion

- Automatic insertion on microprocessor logic blocks
  - 90 nm process

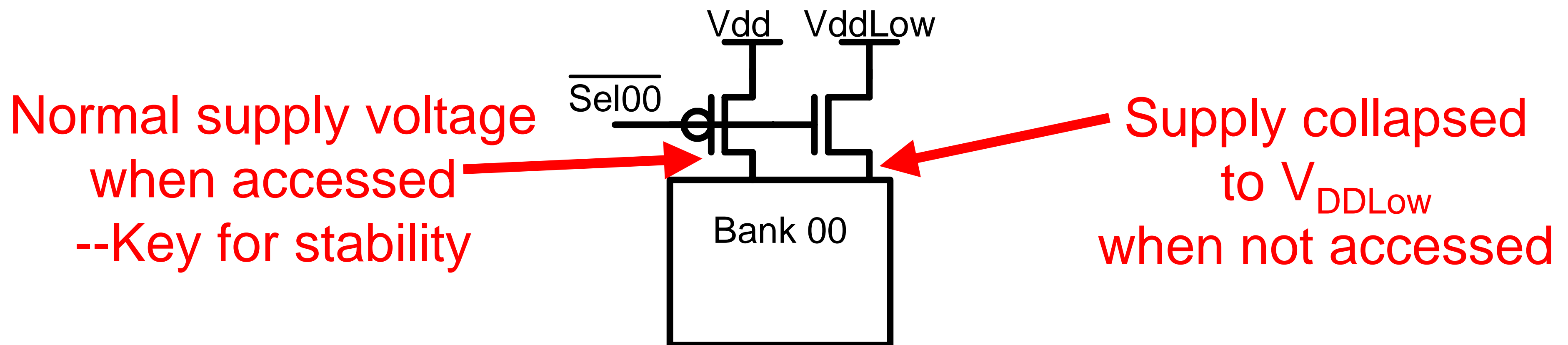
Block	1	2	3	4	5	6	7	8	9	10	11	12	13
I_{off} reduction (%)	39	43	44	48	45	38	49	49	48	47	39	42	39

Outline

- Introduction and motivation
- Standby leakage management
  - Drowsy mode: Reverse body bias and supply collapse
    - Circuit design and operation
    - System level results
    - Limitations
  - Thick gate shadow latches
- Active leakage management
  - Multiple V_t and channel assignment
  - **Drowsy memories**
  - Thick gate SRAM
- Conclusions

Drowsy Memories

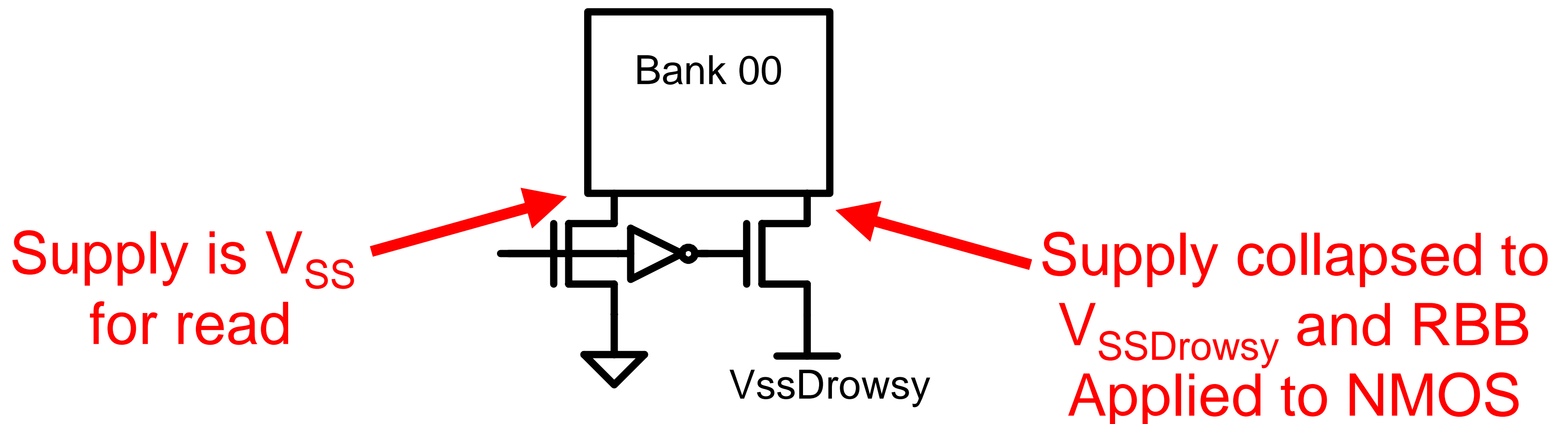
- Leakage dominates memory power on deep submicron processes
  - Low activity factor leads to low active power
- Supply collapse suggested for limiting cache memory leakage [20]
  - Can be done by bank or row



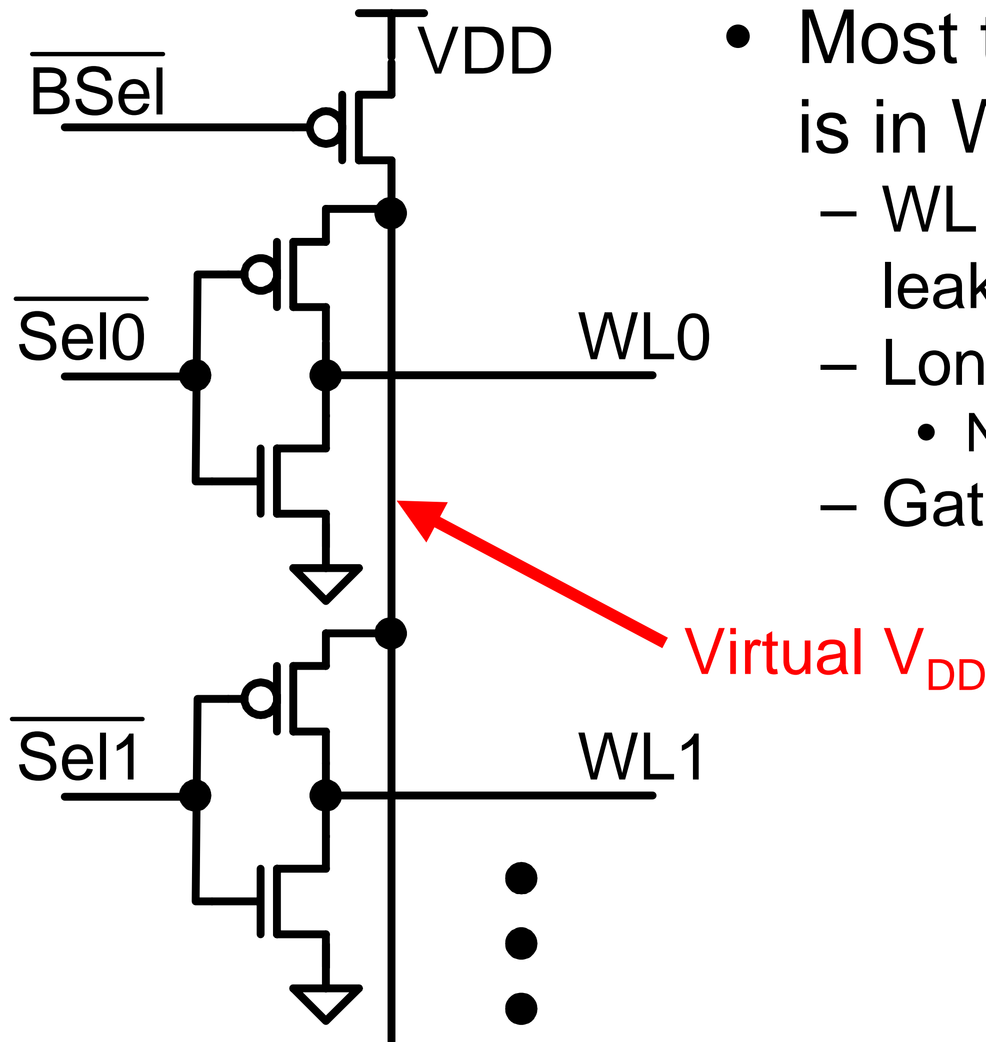
  - Very aggressive processes have low RBB impact
    - Backing off the gate length fixes this—usually needed anyways

Drowsy Memories: Driving V_{SS}

- Driving V_{SS} allows greater leakage reduction
  - Can still be done on a row by row basis
    - Helps write speed, Ref. [3] used FBB to improve read speed and stability
  - Used effectively on a 0.13 µm process [21]
    - Note this is the same as the full-chip Drowsy mode described previously
    - No PMOS RBB
    - NMOS and PMOS leakage not always balanced



Memory Decode



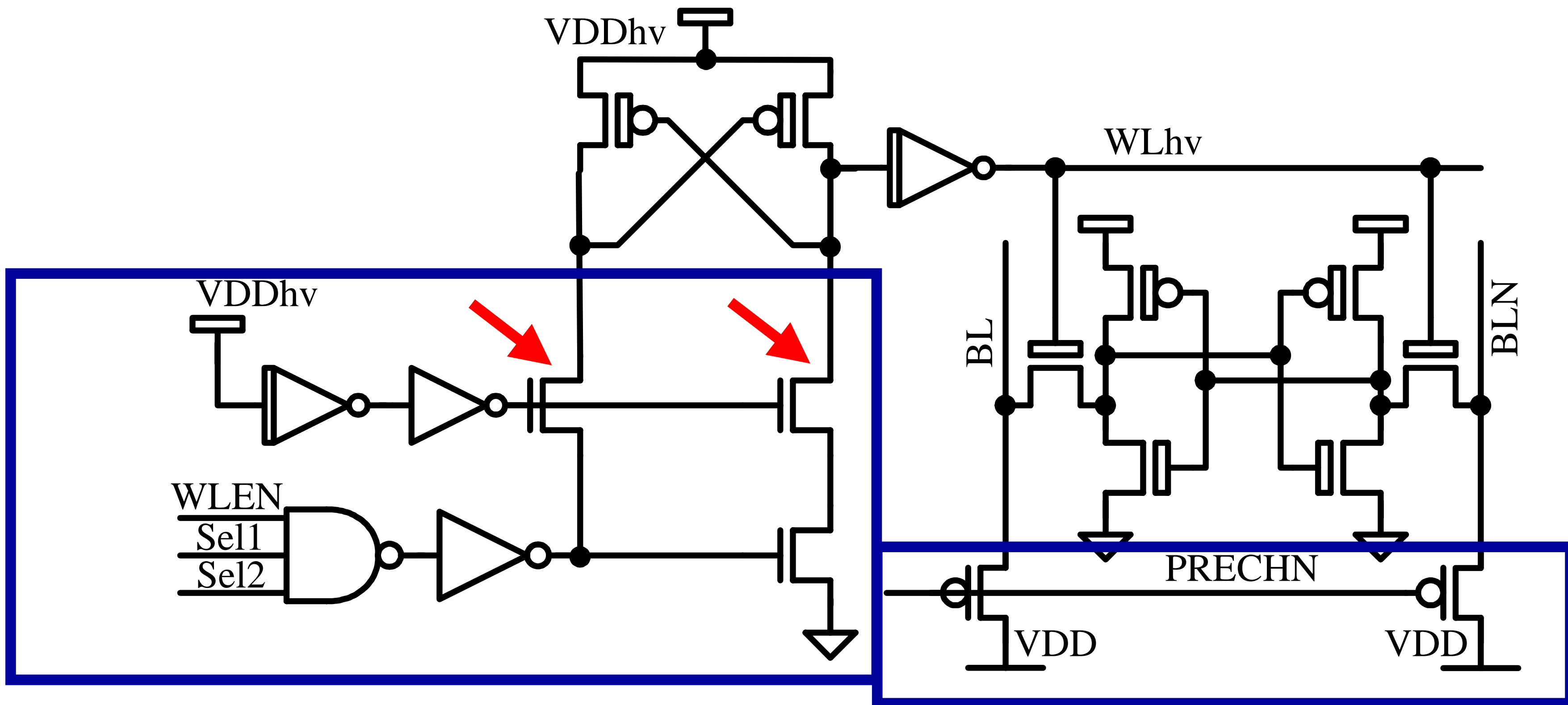
- Most transistor width in decode is in WL drivers
  - WL low when inactive, PMOS leakage predominates
  - Long channel can be used
    - No logic change
  - Gating V_{DD} more effective [22]

Outline

- Introduction and motivation
- Standby leakage management
  - Drowsy mode: Reverse body bias and supply collapse
    - Circuit design and operation
    - System level results
    - Limitations
  - Thick gate shadow latches
- Active leakage management
  - Multiple V_t and channel assignment
  - Drowsy memories
  - **Thick gate SRAM**
- Conclusions

Thick Gate SRAM

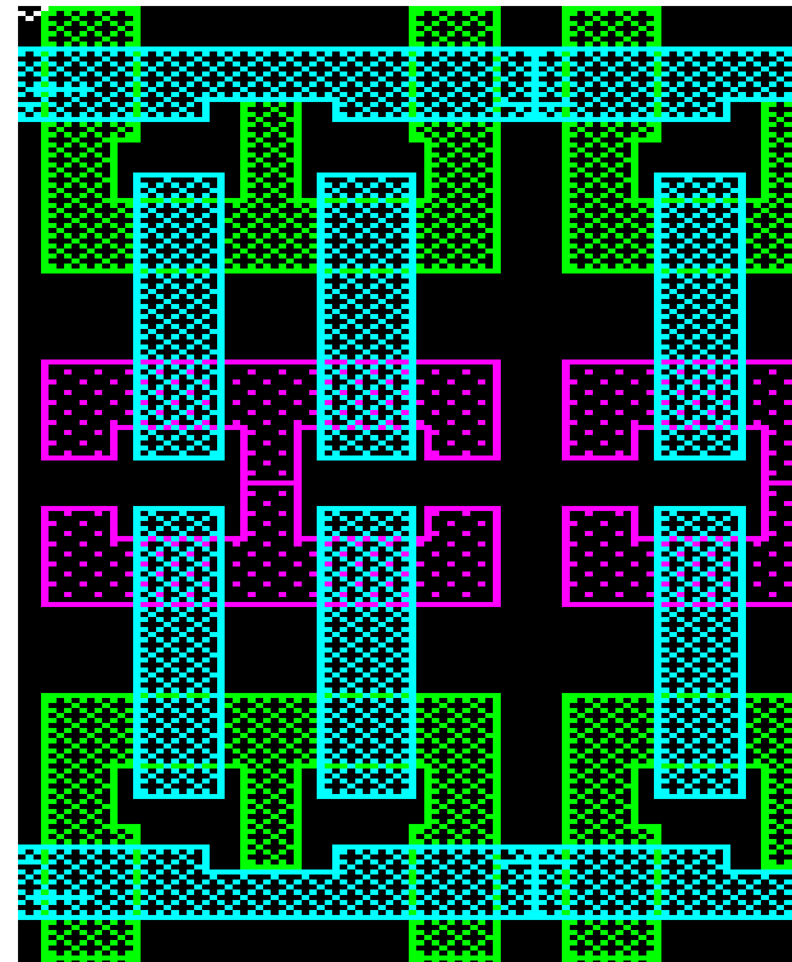
- Use thick gate transistors for SRAM
  - High V_t, no appreciable I_{off} or I_{gate} currents



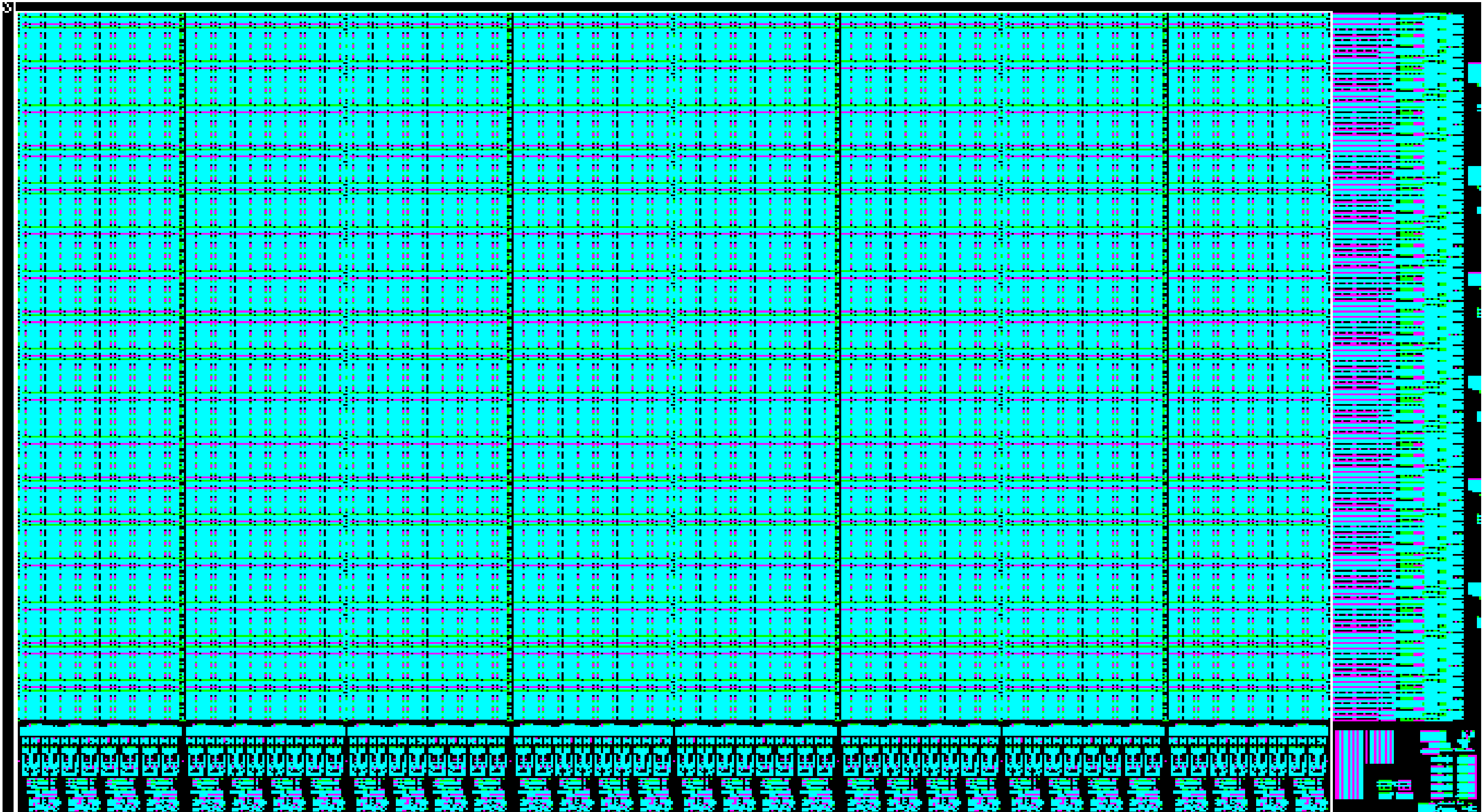
- Bitlines precharged to the core V_{DD}
  - SRAM cells operate from V_{DDhv} –ensures stability
  - Level shift at the WL driver--keeps decoder low power

Thick Gate SRAM: Layout and Size

- Gate length cannot scale with thicker t_{OX}
- Transistors must be re-targeted from the I/O transistors
  - Avoid punchthrough with high V_t
  - Eliminate halo for low drain to bulk band to band tunneling
  - Longer gate to keep gate control (but short as possible)
- Cell about 20-40% larger than thin gate cells

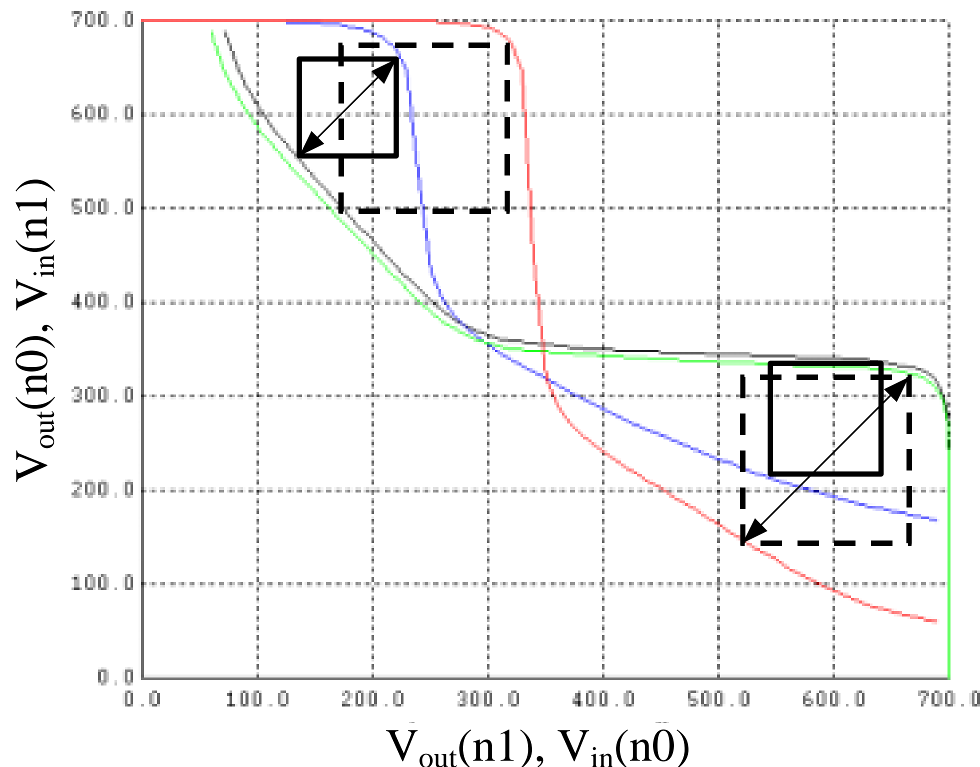


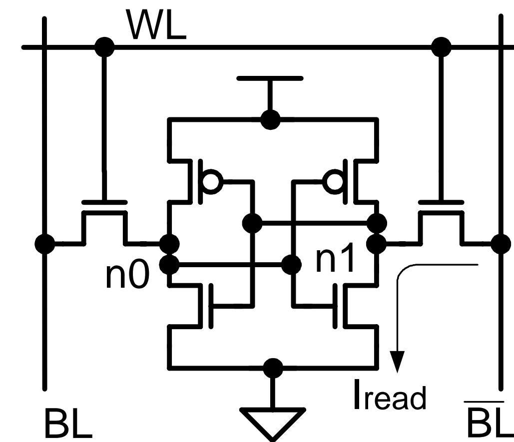
Thick Gate SRAM: Array Layout



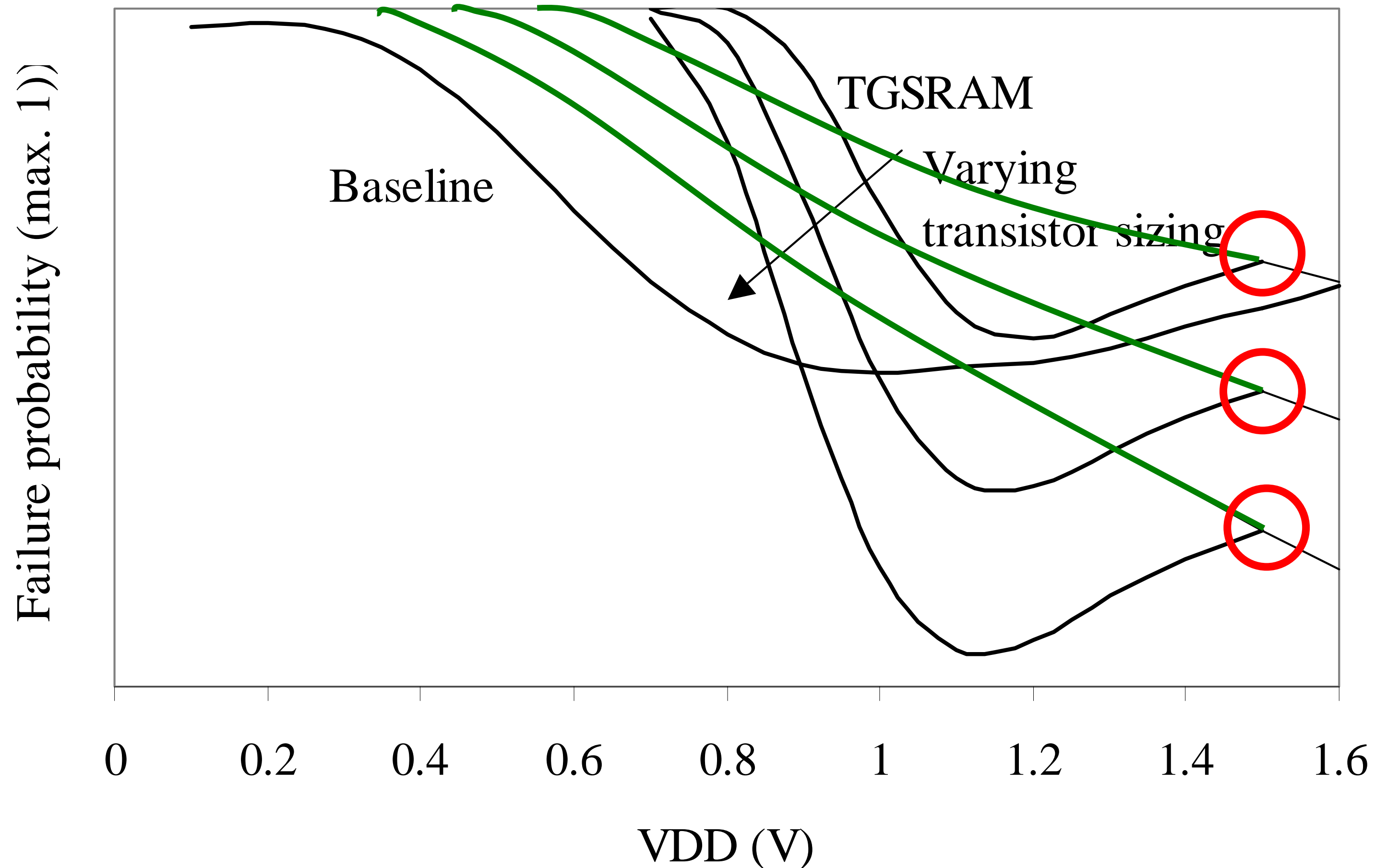
SRAM Cell Read Stability

- Current through inverter pulldown raises cell logic low level during read—particularly at low voltage
  - Due to mis-match (even RDF [23]) the static noise margin can be much smaller than expected
  - Some cells flip when read





Thick Gate SRAM Cell Stability



- Stability improved in thick gate SRAM design
  - Size matters! Better matching
  - Also greatly helped by lower precharge voltage…
    - Until V_{DD} – V_{tTG} reached—then looks like a write

Conclusions

- Standby power is moving from a process to a design problem
  - Process scaling increases leakage
- There is a lot of room for improvement
  - Huge and growing hand-held, wearable, and medical markets will stimulate creative solutions
- Design solutions can limit many components
  - Low V_{GS} limits, thick gate eliminates I_{gate}
  - High V_t or body bias limits I_{off}
  - Drowsy limits GIDL
    - Combines low V_{GS}, simulates high V_t
- Transistor design will also matter
  - Drain to bulk tunneling current will be limiting
    - Requires limited or no halo implants

Questions?

Acknowledgment

- Many people contributed to this work:
  Kim Velarde
  Shay Demmons
  Franco Ricci
  Manish Biyani
  Ed Bawolek
  Neil Deutscher
  Mike Morrow
  Bill Brown
  Dave McCarroll
  Eric Hoffman
  Alfredo Barrenechea

References

[1] H. Mizuno, et al., An 18-μA standby current 1.8-V, 200-MHz microprocessor with self-substrate-biased data-retention mode, *IEEE JSSC*, 34, Nov., 1999, pp. 1492-1500.

[2] L. Clark, N. Deutscher, F. Ricci, and S. Demmons, Standby power management for a 0.18 μm microprocessor, *Proc. ISLPED*, 2002, pp. 7-12.

[3] H. Mizuno and T. Nagano, Driving source-line cell architecture for low-V high-speed low-power applications, *IEEE JSSC*, 31, April, 1996, pp. 552-558.

[4] S. Mutoh, et al., 1V power supply high-speed digital circuit technology with multithreshold-voltage CMOS, *IEEE JSSC*, 30, Aug., 1995, pp. 847-854.

[5] S. Shigematsu, et al., A 1-V high-speed MTCMOS circuit scheme for power-down application circuits, *IEEE JSSC*, 32, June, 1997, pp. 861-870.

[6] J. Kao and A. Chandrakasan, "Dual-threshold voltage techniques for low-power digital circuits," *IEEE JSSC*, 25, July, 2000, pp. 1009-1018.

[7] Q. Wang and S. Vruhula, "Algorithms for minimizing standby power in deep submicrometer, dual-Vt CMOS circuits," *IEEE Trans. On Computer-aided Design of Int. Circuits and Systems*, pp. 306-318.

[8] L. Clark and F. Ricci, "Low standby power using shadow storage," US Patent #6,639,827

[9] L. Clark, F. Ricci, and M. Biyani, "Low standby power state retention for sub-130 nm processes," *accepted to IEEE JSSC*.

[10] U. Ko, D. Scott, S. Gururajarao, H. Mair, "Retention register for system-transparent state retention," US patent application 2004000871.

[11] L. Clark, M. Morrow, and W. Brown, Reverse Body Bias and Supply Collapse for Low Effective Standby Power, *to appear in IEEE Trans. VLSI*, Sept., 2004.

[12] L. Clark, D. McCarroll, and E. Bawolek, Characterization and debug of reverse body bias low power modes, *Electronic Device Failure Analysis*, 6, Feb., 2004, pp. 13-21.

[13] B. Calhoun, F. Honore, and A. Chandrakasan, "A leakage reduction methodology for distributed MTCMOS," *IEEE JSSC*, 39, May, 2004, pp. 818-827.

[14] T. Maloney, S. Poon, and L. Clark, "Methods for Designing Low-leakage Power Supply Clamps," *to appear in Journal of Electrostatics*, October, 2004.

[15] K. Osada, Y. Saitoh, E. Ibe, K. Ishibashi, "16.7pA/cell tunnel-leakage-suppressed 16Mb SRAM for handling cosmic-ray-induced multi-errors," *ISSCC Proc.*, 2003.

[16] S. Zhao, et al., "Transistor optimization for leakage power management in a 65 nm CMOS technology for wireless and mobile applications," *VLSI Symp. Tech. Dig.*, 2004, pp. 14-15.

[17] UC Berkeley Device Group. Berkeley predictive technology model [online]. http://www-device.eecs.berkeley.edu/~ptm.

[18] S. Kim, S. Kosonocky, and D Knebel, "Understanding and minimizing ground bounce during mode transition of power gating structures," *Proc. ISLPED*, 2003, pp. 22 – 25.

[19] A. Barrenechea, "Design impact of process variation," M.S. Thesis, University of New Mexico, 2004.

[20] K. Flautner, et al., Drowsy caches: Simple techniques for reducing leakage power, *Proc. ISCA'02*, p. 148, 2002.

[21] K. Min, K. Kanda, and T. Sakurai, "Row-by-row dynamic source-line voltage control (RRDSV) scheme for two orders of magnitude leakage current reduction of sub-1-V-VDD SRAM's," *Proc. ISLPED*, 2003, pp. 66-71.

[22] K. Itoh, "Trends in low-voltage embedded-RAM technology," *Proc. 23rd Int. Conf. on Microelectronics*, 2002, p. 497.

[23] A. Bhavnagarwala, T. Xinghai, J. Meindl, The impact of intrinsic device fluctuations on CMOS SRAM cell stability, *IEEE JSSC*, 36, no. 4, 2001, pp. 658-665.