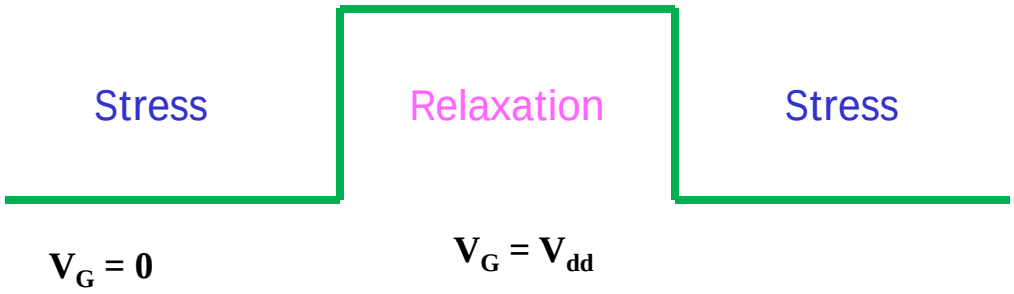
Adaptive Techniques for Overcoming Performance Degradation due to Aging in Digital Circuits

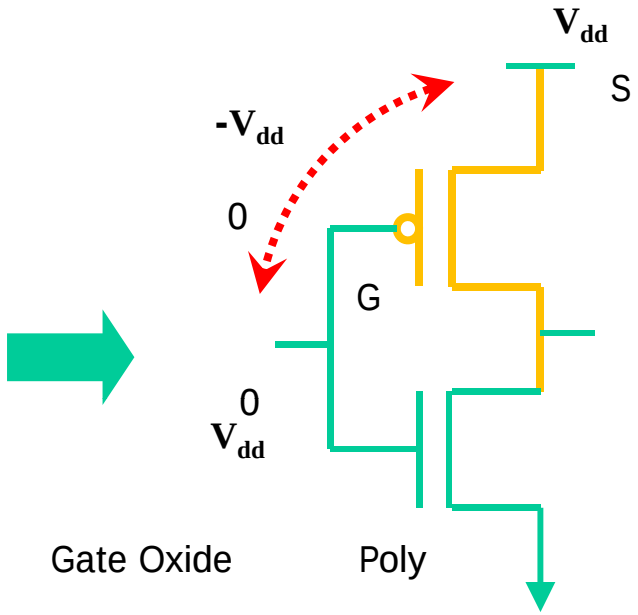
Sanjay Kumar

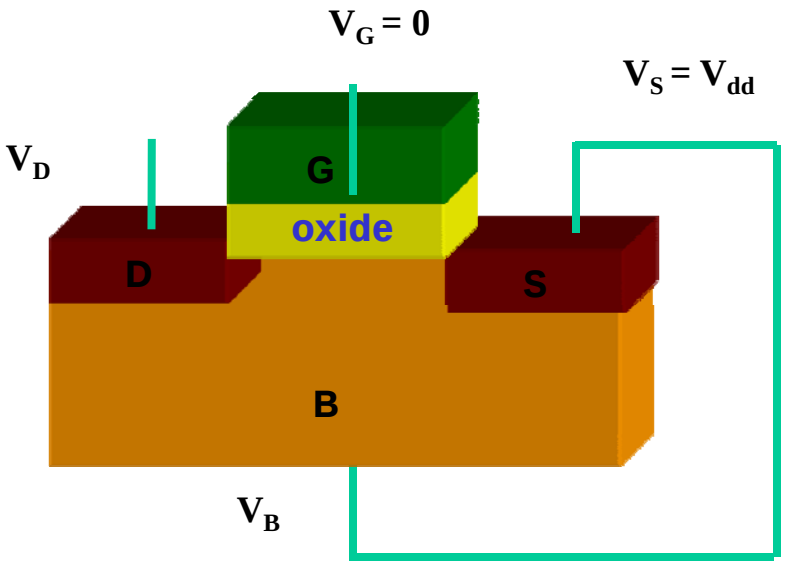
Chris Kim

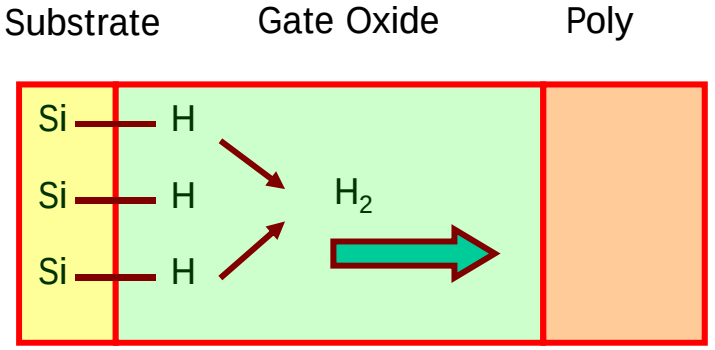
Sachin Sapatnekar

Negative Bias Temperature Instability (NBTI)







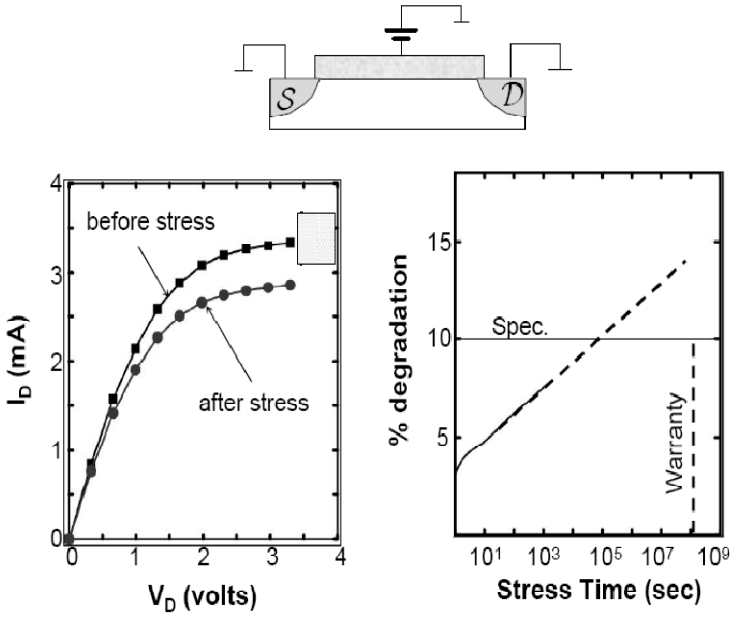


$SiH + h^+ \rightarrow Si^+ + H$

$H + H \rightarrow H_2$

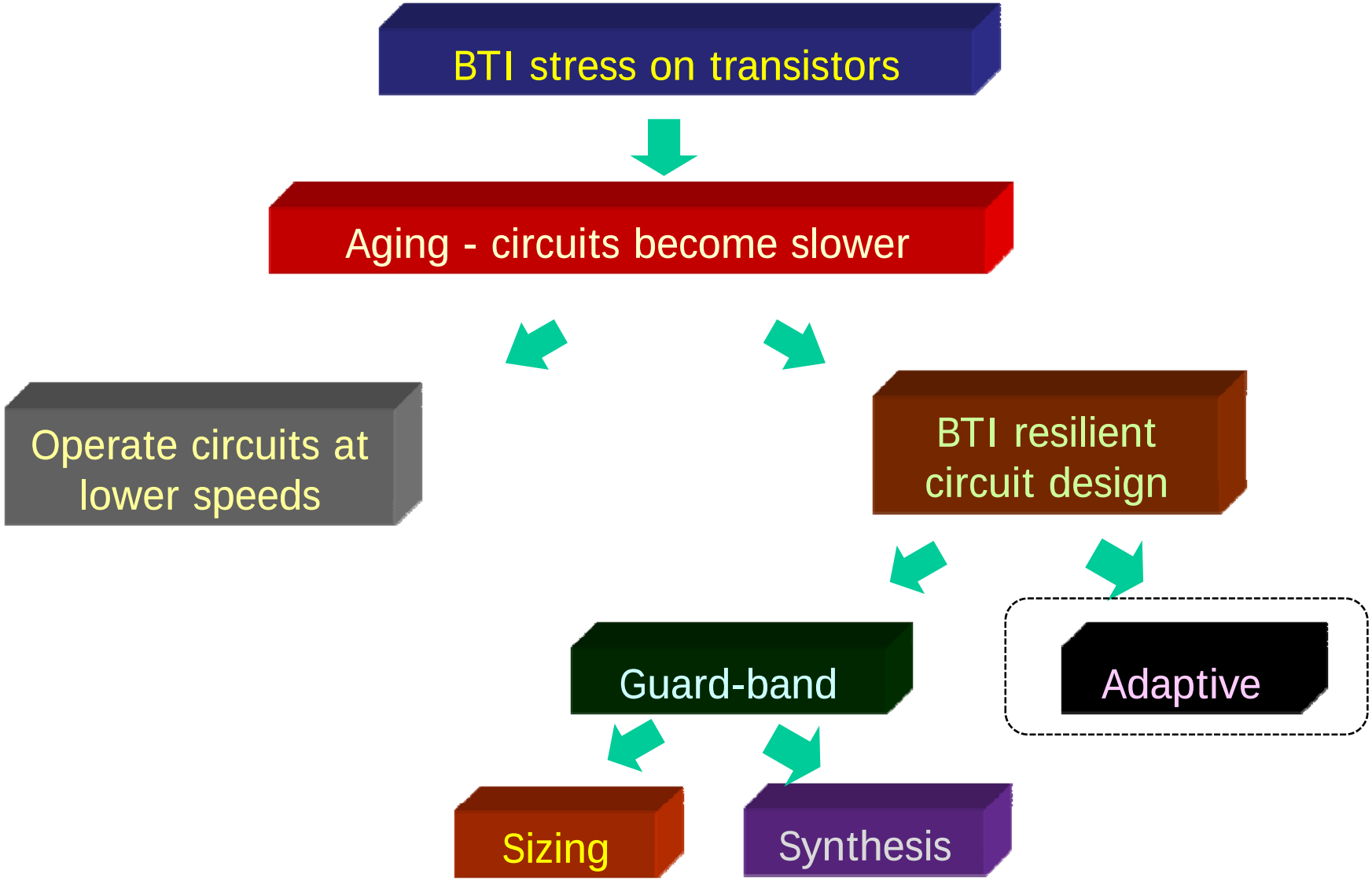
Impact of BTI

- 25-30% degradation in PMOS V_{th} – drain current reduces
- Positive Bias Temperature Instability (PBTI)
  - In NMOS devices when $V_{gs} = V_{dd}$
  - Lower impact reported as compared with PMOS NBTI
  - Increasing impact with Hf-based high-k dielectrics
- Challenges in nanometer design
  - Quantify the impact of BTI on circuit performance
  - **Design robust circuits**

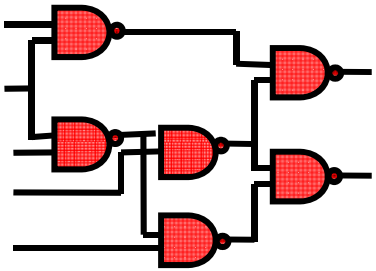


[Alam, IRPS05]

Overcoming BTI in Digital Circuits

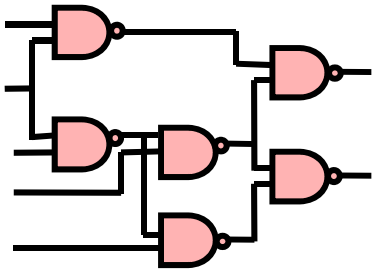


Sizing for Reliability [DATE06, ICCD06]



Original design

After 3 years

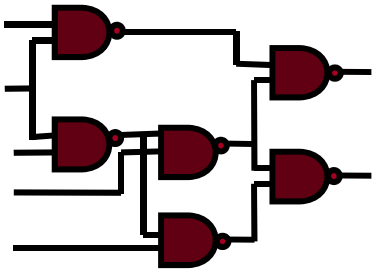


Gates become weak, target freq not met

Are

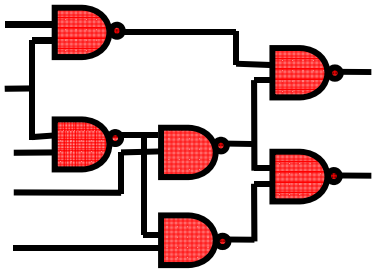
Can change gate sizes only

ble design



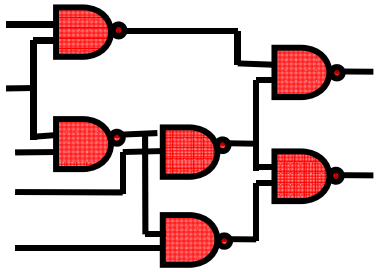
Design sized taking into account aging

After 3 years



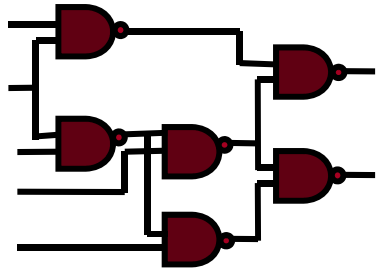
Still meets specs

BTI-Aware Synthesis [DAC07]



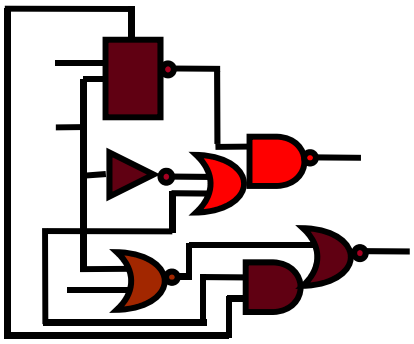
Original design

Sizing solution



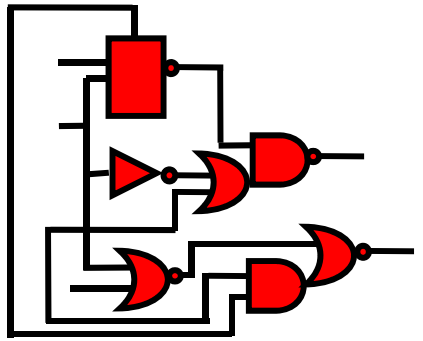
Gate sizes increased

Can change circuit topology



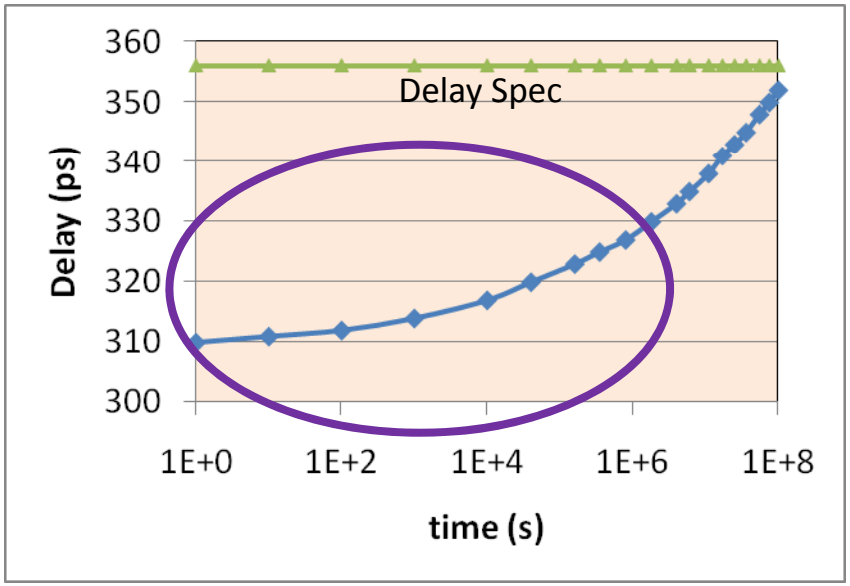
Design remapped taking into account aging

After 3 years

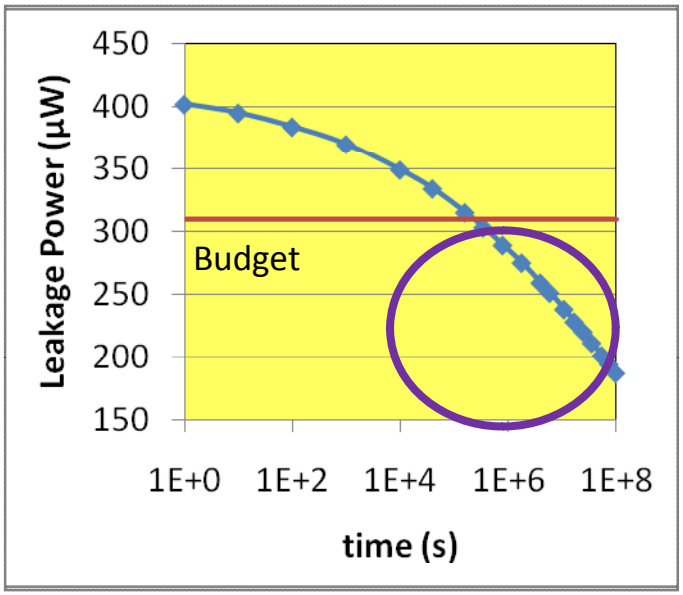


Still meets spec

Limitations of "One-time" Fixes



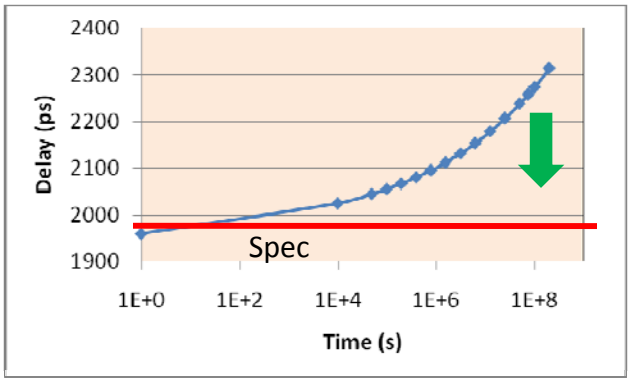
Circuit synthesized s.t. $D(t_{life}) \leq D_{spec}$



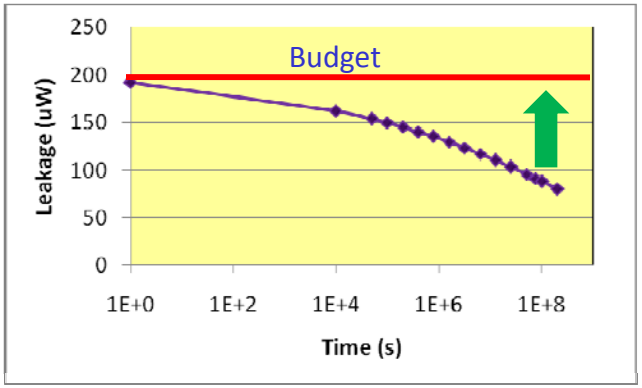
Temporal leakage of BTI-aware synthesized circuit

- Circuit runs faster than spec for $t < t_{life}$
- Burns additional power in comparison with nominal design due to design-time (one-time) fix
- Leakage decreases below budget for t closer to t_{life}
- Potential for leakage-performance tradeoff not utilized

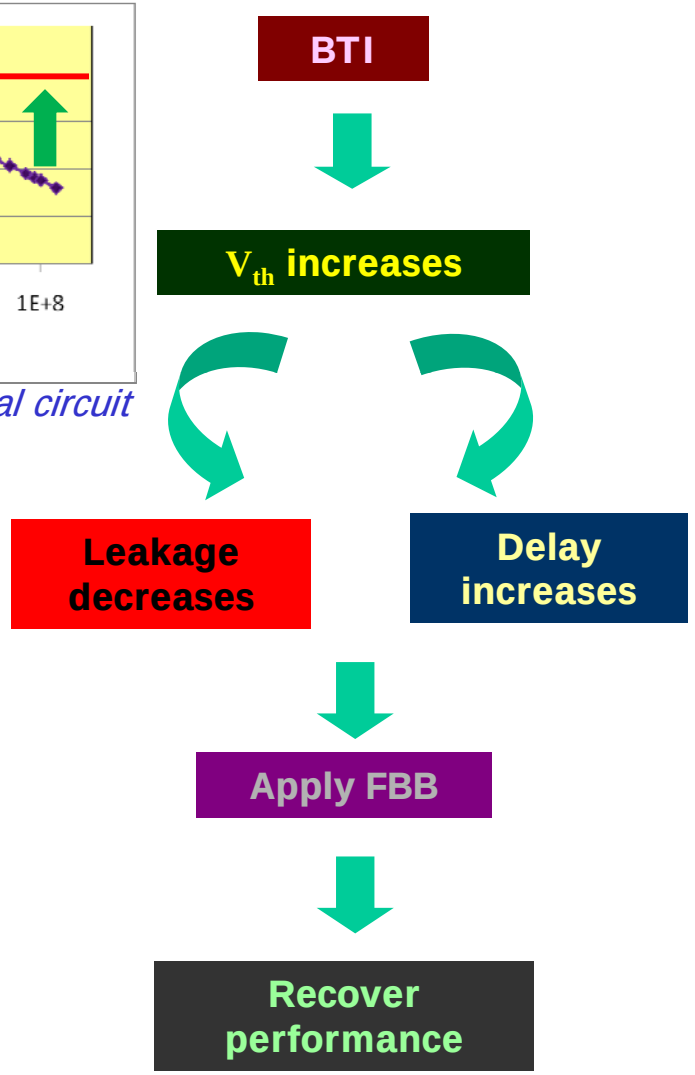
Adaptive Techniques



Increase in delay of nominal circuit

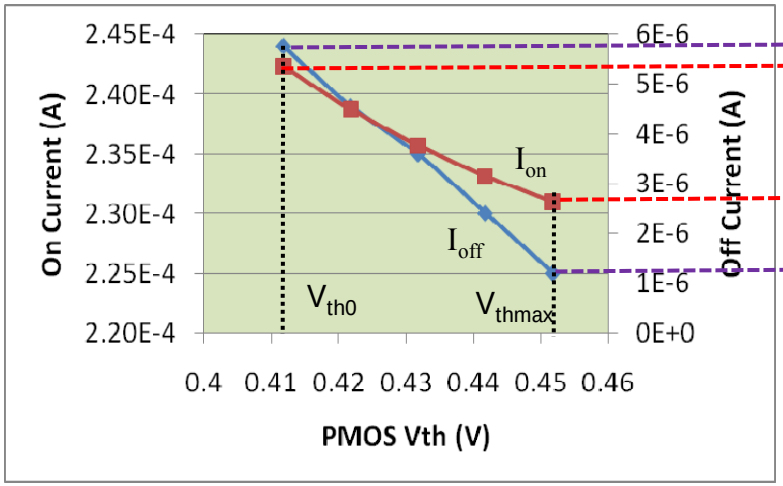


Decrease in leakage of nominal circuit

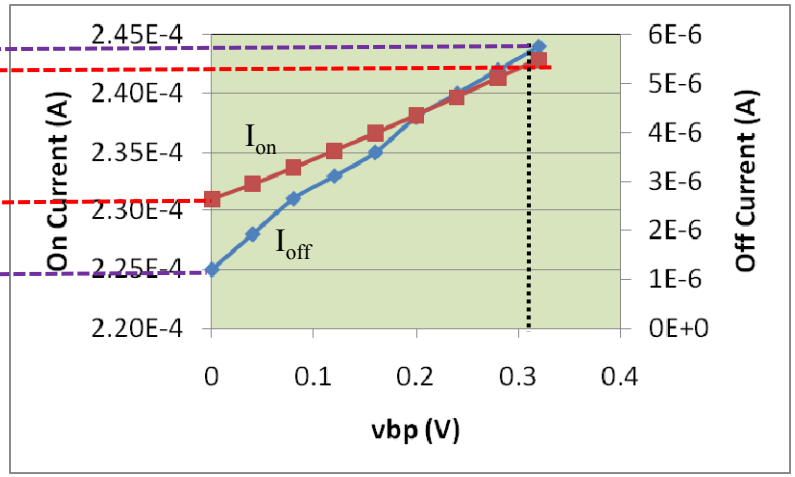


- Use available slack in leakage budget
- Dynamically apply FBB to recover performance
- Can we recover performance completely without leakage overhead?

Ideal Case



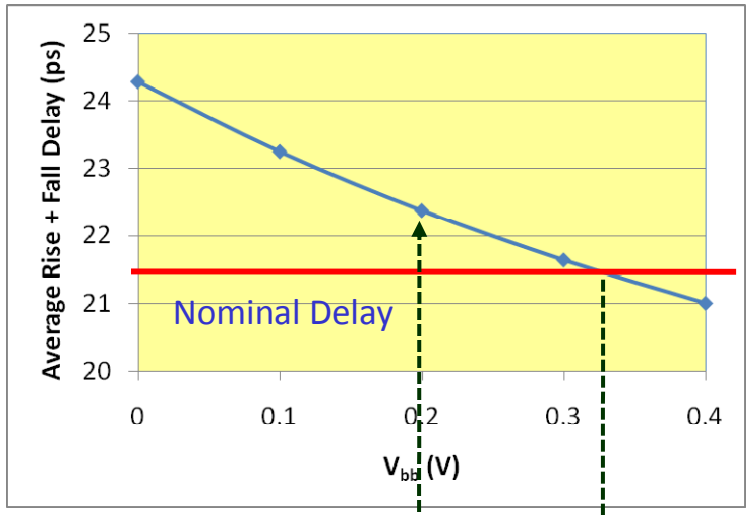
Decrease in currents with increase in V_{th}



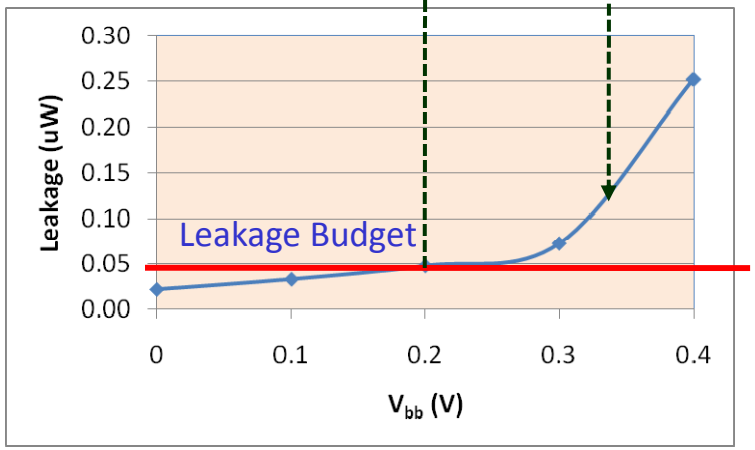
Increase in I_{on} and I_{off} (measured at t_{life}) with PMOS body bias

- NBTI causes on and subthreshold currents to decrease
- FBB (of around 0.3V) to the PMOS device sets V_{th} back to V_{th0}
- On and subthreshold currents back to their nominal values
- Effectively, device reset to its original state?

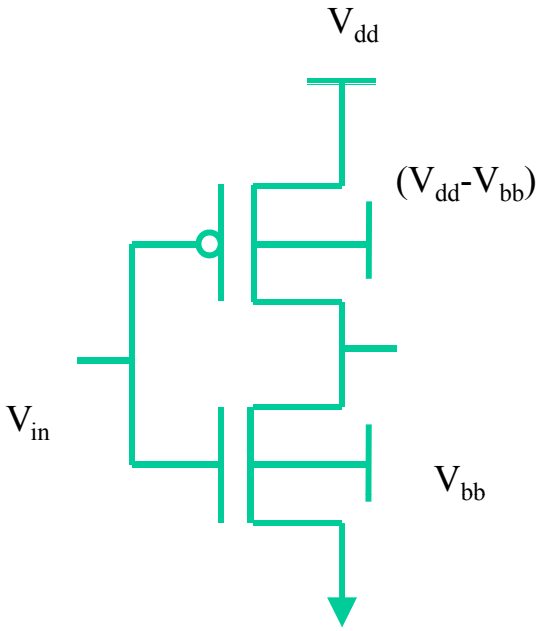
Leakage Components

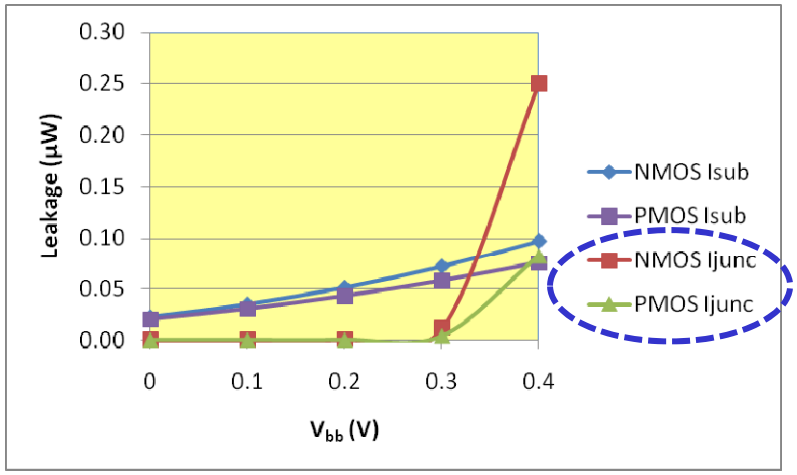


Delay (Rise + Fall)/2 for an inverter with FBB



Leakage power of an inverter with FBB





Components of leakage power

Problem Formulation

- Exponential increase in junction leakage with FBB
- Complete recovery in performance without leakage overhead not possible with ABB
- Use ASV (Adaptive Supply Voltage) as an additional knob

- ASV (Adaptive Supply Voltage)
  - Better control over performance (delay) with V_{dd} than V_{bb}
  - Leakage still increases (exponentially) with V_{dd}
  - Active power increases quadratically with V_{dd}
  - Minimize overall power consumed subject to delay constraints

Problem Formulation

Total Power $P = f(P_{active}, P_{leakage})$

Active power: $P_{active}(t, V_{dd})$

Leakage power: $P_{leakage}(t, v_{bn}, v_{bp}, V_{dd})$

Minimize P

s.t. $D(t, v_{bn}, v_{bp}, V_{dd}) \leq D_{spec}$

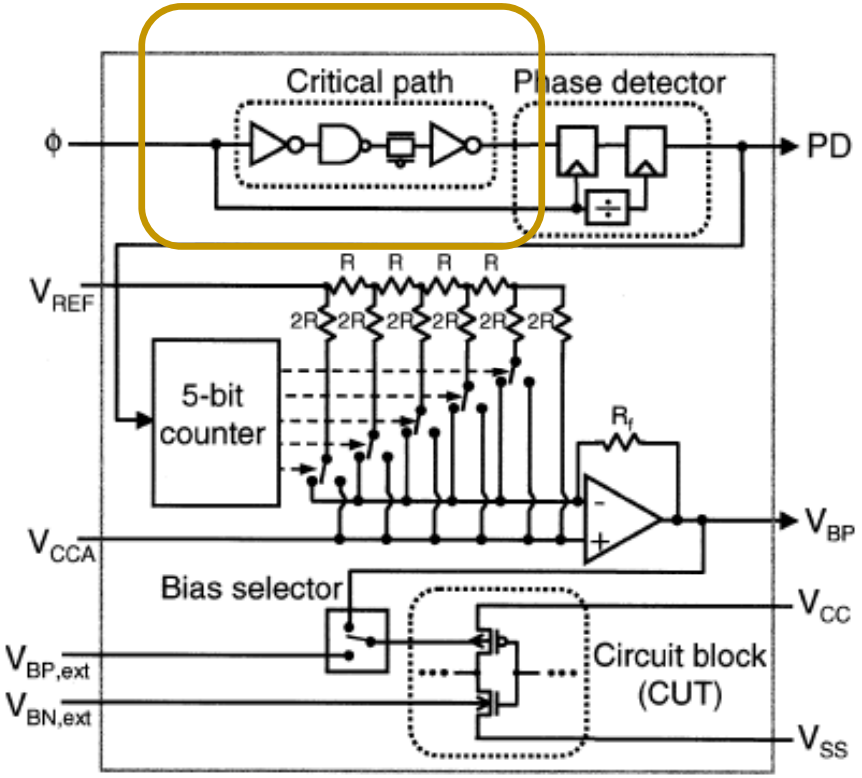
$0 \leq v_{bn}(t) \leq v_{bnmax}$

$0 \leq v_{bp}(t) \leq v_{bpmax}$

$0 \leq t \leq t_{life}$

Control System

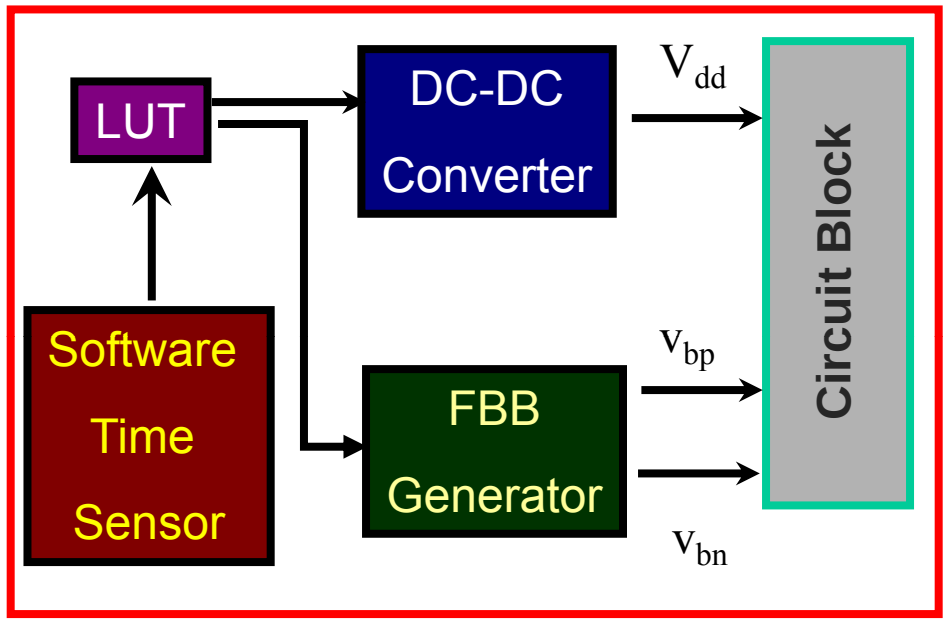
- Critical path replica based
  - Large number of critical paths required for an identical f_{max} distribution as CUT
  - Aging of critical path replicas depend on signal probabilities, is usage specific – cannot be predicted *a priori*
  - Critical paths can change temporally based on relative aging of paths



[Intel, JSSC2002]

Control system

- Lookup table based
  - Stores optimal values referenced by time
  - Software routine (assumed) to track time of usage
  - On chip local body bias and V_{dd} generators
  - Optimal (v_{bn}, v_{bp}, V_{dd}) precomputed at design time by estimating degradation in delay considering BTI



Time (s)	v_{bn} (V)	v_{bp} (V)	V_{dd} (V)

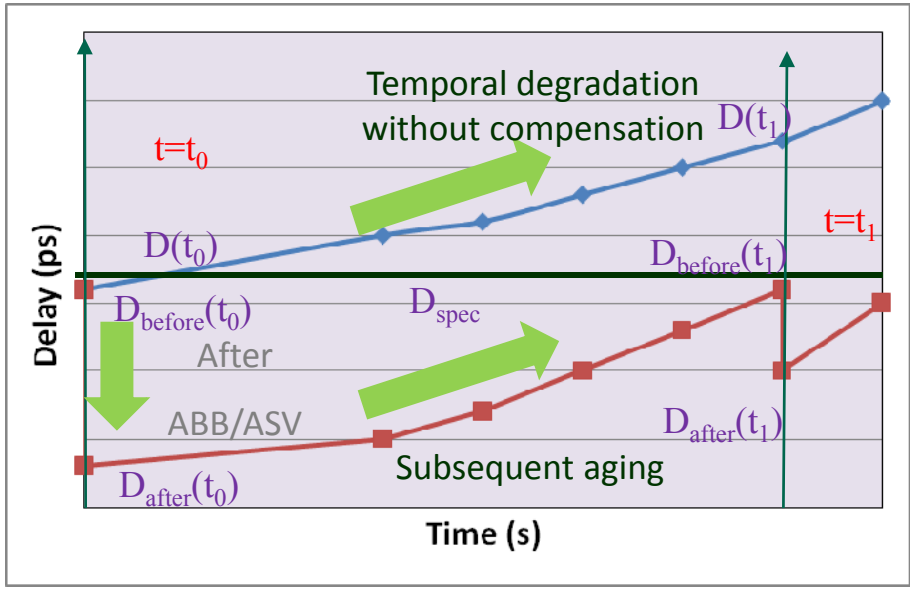
How to precompute delay degradation

- Signal activity based model
  - Cannot predict signal probabilities *a priori*
  - Circuit must work under all conditions
- Worst-case method
  - Assume maximal degradation of all NMOS and PMOS devices
  - Compute delay of the circuit at different times
  - Upper bound over the temporal delay of the circuit

Optimal ABB/ASV Computation

- Delay must be less than D_{spec} at all times

$$D_{after}(t_i - 1) < D_{before}(t_i - 1) \leq D_{spec}$$
$$D_{after}(t_i - 1) < D_{before}(t_i) \leq D_{spec}$$

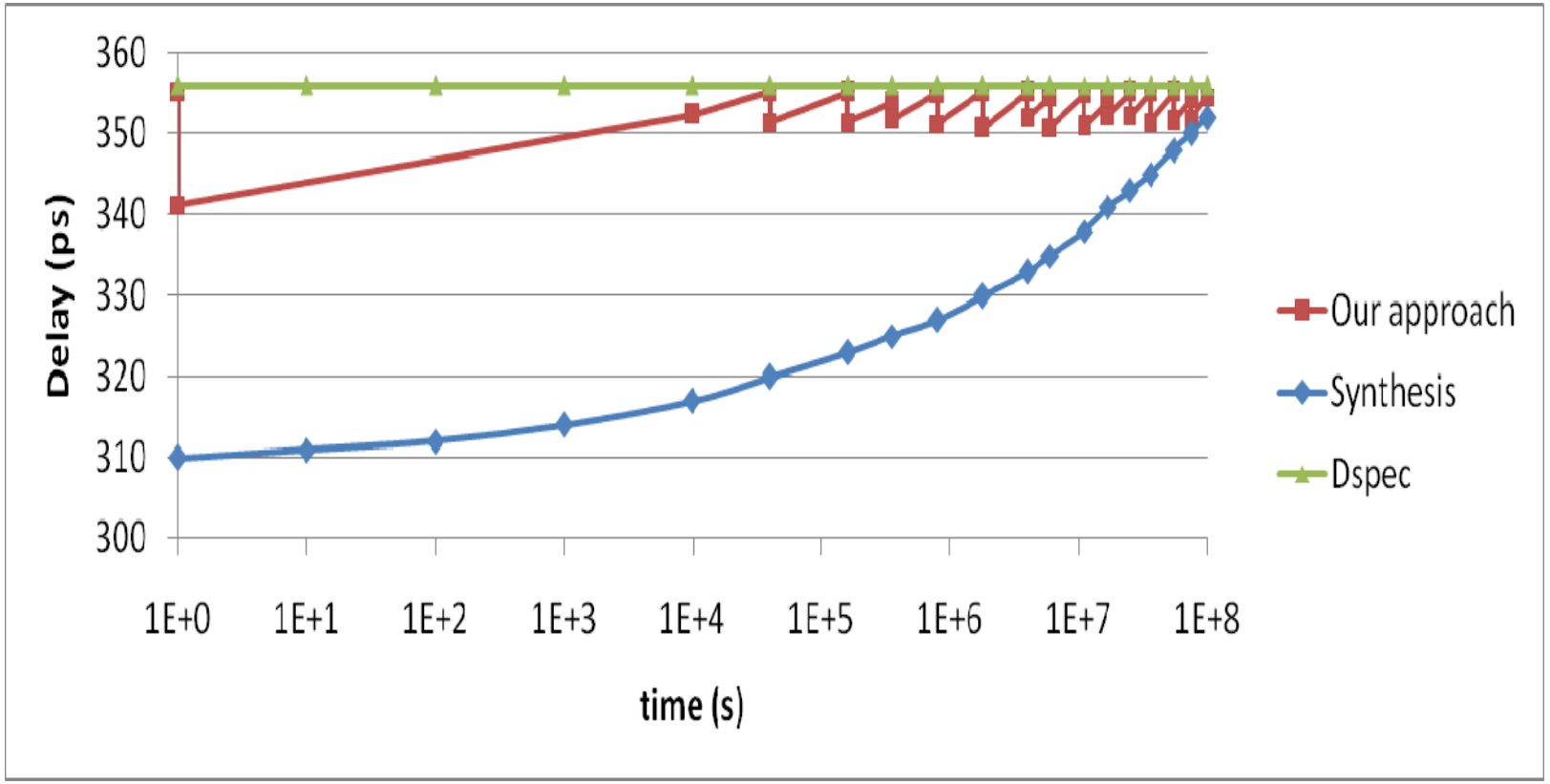


- Amount of compensation at t_0 depends on degradation in $[t_0,t_1]$
- Compute degradation assuming worst-case aging
- Second order dependence of the extent of trap generation on V_{dd}
- Determine optimal (V_{dd}, v_{bn}, v_{bp}) such that delay is met and power is minimized using an enumeration based algorithm [KumarTVLSI08]

Lookup Table (LGSYNTH93 Circuit "des")

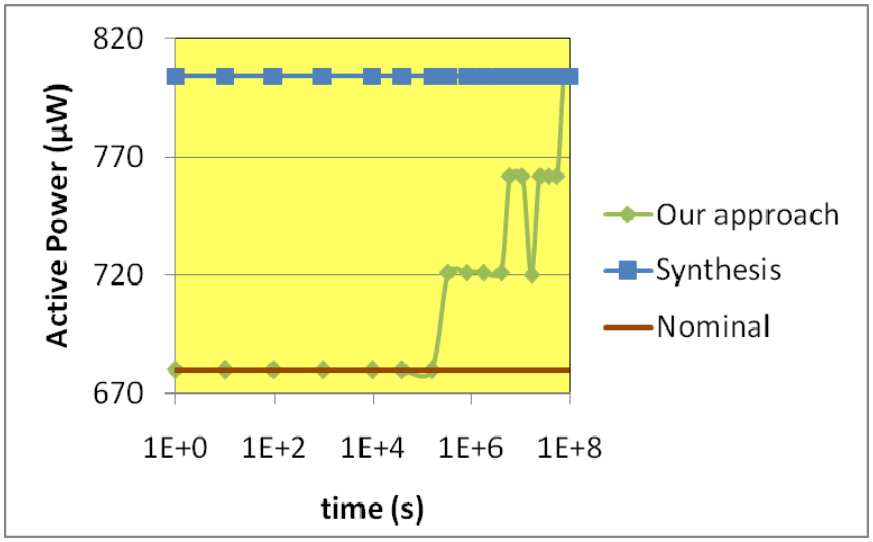
Time ($x10^8$s)	v_{bn} (V)	v_{bp} (V)	V_{dd} (V)	Delay (ps)	P_{act} (µW)	P_{lkg} (µW)	% Increase
Nominal	0.00	0.00	1.00	355	641	327	
0.0000	**0.00**	**0.05**	**1.03**	341	680	416	16%
0.0001	**0.00**	**0.05**	**1.03**	351	680	346	6%
0.0004	**0.00**	**0.10**	**1.03**	351	680	362	8%
0.0016	**0.05**	**0.10**	**1.03**	351	680	369	9%
0.0035	**0.00**	**0.05**	**1.06**	352	721	344	9%
0.0080	**0.05**	**0.05**	**1.06**	351	721	357	11%
0.0180	**0.05**	**0.10**	**1.06**	351	721	368	12%
0.0400	**0.10**	**0.10**	**1.06**	352	721	377	14%
0.0600	**0.00**	**0.10**	**1.09**	351	762	353	13%
0.1100	**0.05**	**0.10**	**1.09**	351	762	360	14%
0.1700	**0.10**	**0.20**	**1.06**	352	721	398	17%
0.2500	**0.05**	**0.15**	**1.09**	352	762	362	15%
0.3600	**0.05**	**0.20**	**1.09**	351	762	388	19%
0.5500	**0.10**	**0.20**	**1.09**	351	762	396	20%
0.7500	**0.05**	**0.15**	**1.12**	352	804	359	17%
1.0000				355	804	350	16%

Temporal Delay

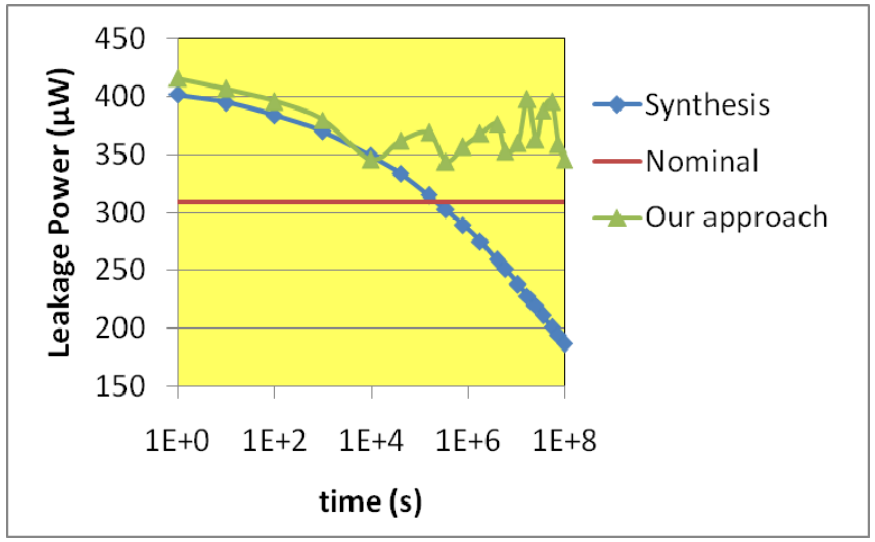


Temporal delay of a benchmark with worst-case synthesis and our ABB/ASV based adaptive approach

Temporal Power



Temporal active power with different approaches



Leakage power versus time using different approaches

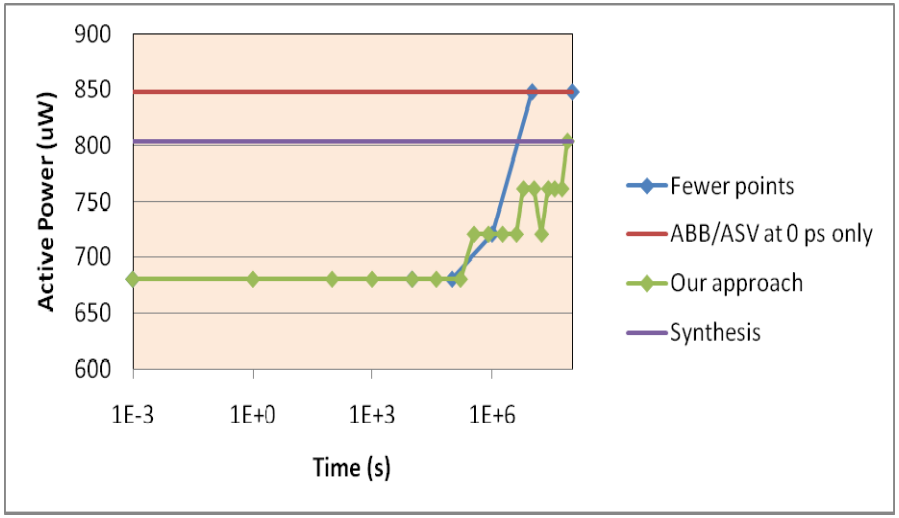
Power	Synthesis	Our (Adaptive) Approach
Active	Constant, large overhead	Increases in steps temporally with V_{dd}
Leakage	Highest at t=0s when there is no BTI Decreases below nominal value temporally	Varies with time – always greater than nominal value since ABB/ASV is applied to compensate for aging Max leakage (at t=0s) comparable with synthesis

Optimal Adaptation Times Selection

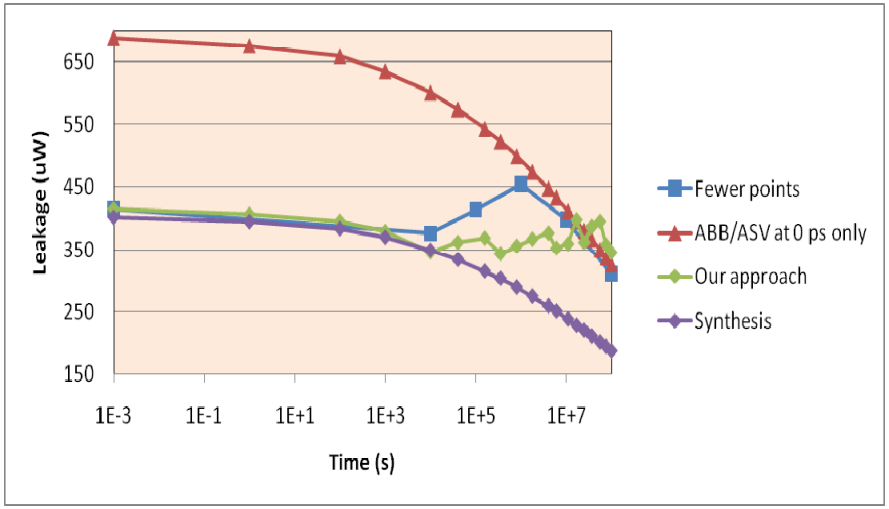
- Number of points chosen depends on
  - Ability of software routine to track time with desired accuracy
  - Discreteness in generating ABB/ASV voltages (50mV for v_{bn}, v_{bp}, and 30mV for V_{dd} in our case)
  - Minimum change in delay over $[t_i, t_{i+1}]$ – subject to modeling errors (assumed to be 1% in our case)
  - Resolution of mapping each delay to a unique optimal (v_{bn}, v_{bp}, V_{dd}) using our enumeration algorithm
  - BTI model accuracy particularly for very small values of $t << t_{life}$ (model asymptotically accurate beyond 10^4s with $t_{life}=10^8$s)

Optimal Adaptation Time Selection

- Want to compensate as much as is required only – keep delay closest to D_{spec}
- Larger number of points leads to
  - Lower degradation in each time interval
  - Minimal ABB/ASV to compensate for increase in delay in each $[t_i, t_{i+1}]$
  - Less overall temporal power overhead
- Compensating at $t=0$ps only is overkill (as compared with synthesis)



Active power versus time for different cases



Leakage power versus time for different cases

Area and Power Overhead

	Nominal Circuit			Our approach		Worst-case Synthesis		
Bench-mark	D_{spec} (ps)	Increase in delay (BTI for 3 years)	Area (μm)	Maximal Increase in active power	Maximal Increase in leakage power	Area Overhead	Maximal Increase in active Power	Maximal Increase in leakage Power
b14	1078	14%	95626	19%	26%	16%	17%	16%
b15	902	13%	179096	19%	26%	16%	18%	15%
C3540	769	14%	18692	25%	30%	32%	37%	38%
C5315	729	15%	29951	19%	29%	14%	18%	25%
C7552	616	15%	42261	19%	29%	18%	19%	15%
des	355	15%	81777	25%	27%	35%	38%	28%
i8	840	17%	55128	25%	26%	18%	71%	44%
i10	830	14%	4063	25%	32%	21%	26%	28%
Avg		15%		23%	27%	24%	30%	26%

Summary

- BTI causes delay to increase and leakage to reduce
- Existing “one-time” fix techniques (sizing, synthesis) lead to large area **and** power overhead
- Attempt to recover performance through available slack in leakage, adaptively
- ABB + ASV to combat increase in delay
- Lookup table based control system indexed by time of stress
- Similar power overhead as compared with synthesis with large area savings