Facelift: Hiding and Slowing Down Aging in Multicores

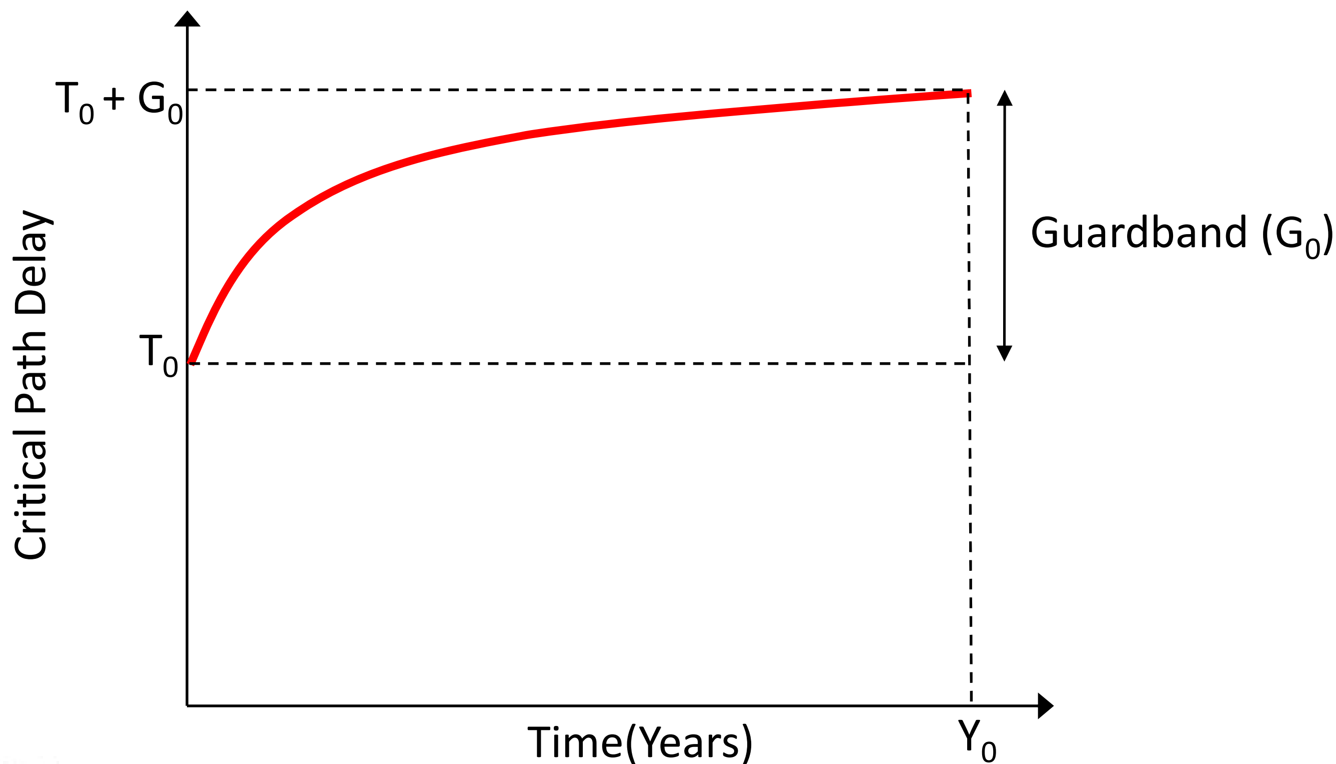
Abhishek Tiwari and Josep Torrellas
Department of Computer Science
University of Illinois at Urbana-Champaign
http://iacoma.cs.uiuc.edu

International Symposium on Microarchitecture (2008)

i-acoma group

Motivation: Aging or Wearout

- Transistor delay gradually increases with time under normal use
- Processor critical paths take longer
- Maximum attainable frequency decreases slowly with time



A. Tiwari and J. Torrellas, Facelift ...

Contributions

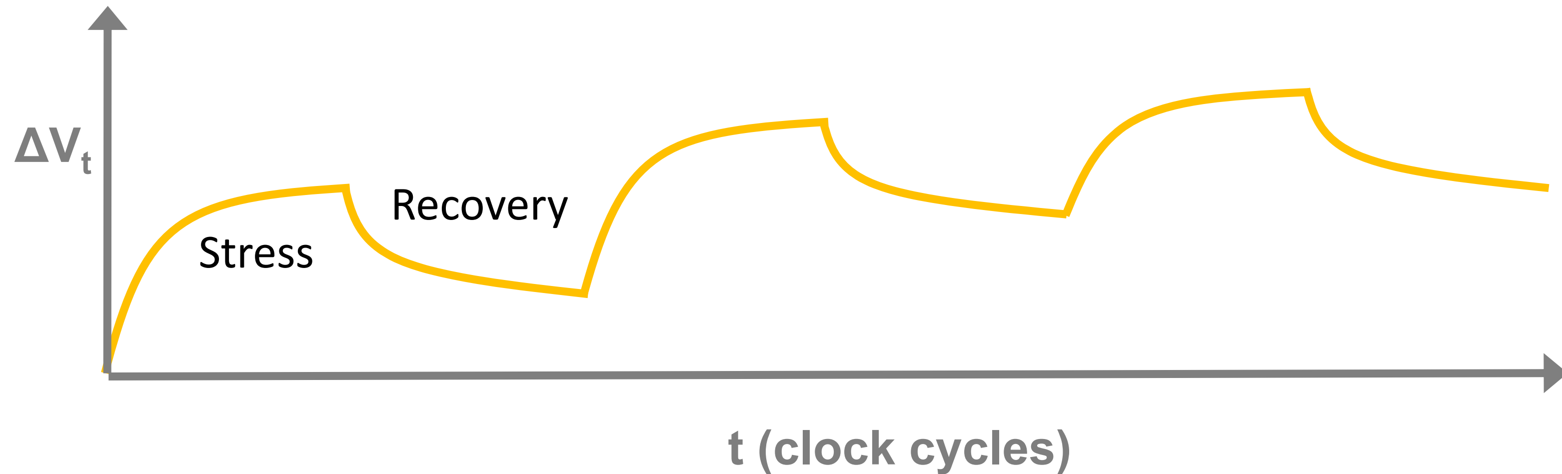
- A framework of techniques, called Facelift
  - Hide the effect of aging in a multicore
  - Slow down aging
  - Configure chip for a short service life
- The shorter guardband enabled by Facelift can be used to:
  - Design a less refined version of the processor
  - Increase clock frequency
- Results:
  - Multicore can be run on avg. at a 14-15% higher frequency
  - Alternatively: it can be designed for 5-7 month service life rather than for 7 years
  - Low hardware overhead

Aging Mechanisms

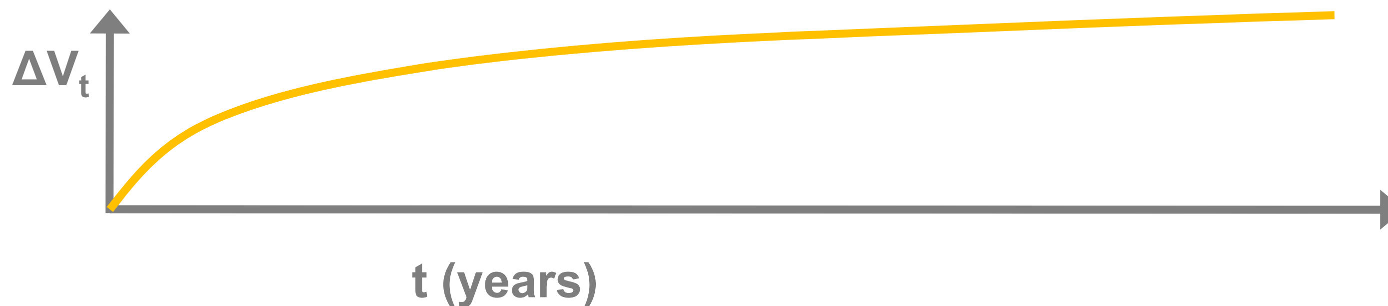
- Two key mechanisms:
  - Negative Bias Temperature Instability (NBTI)
  - Hot Carrier Injection (HCI)
- Gradual elevation of transistor threshold voltage (V_t) with time
- Increases transistor switching delay through the alpha power law
  - $T_s \propto 1/(V_{dd}-V_t)^{\alpha}$

Aging Mechanisms

- PMOS: Negative bias temperature instability (NBTI)

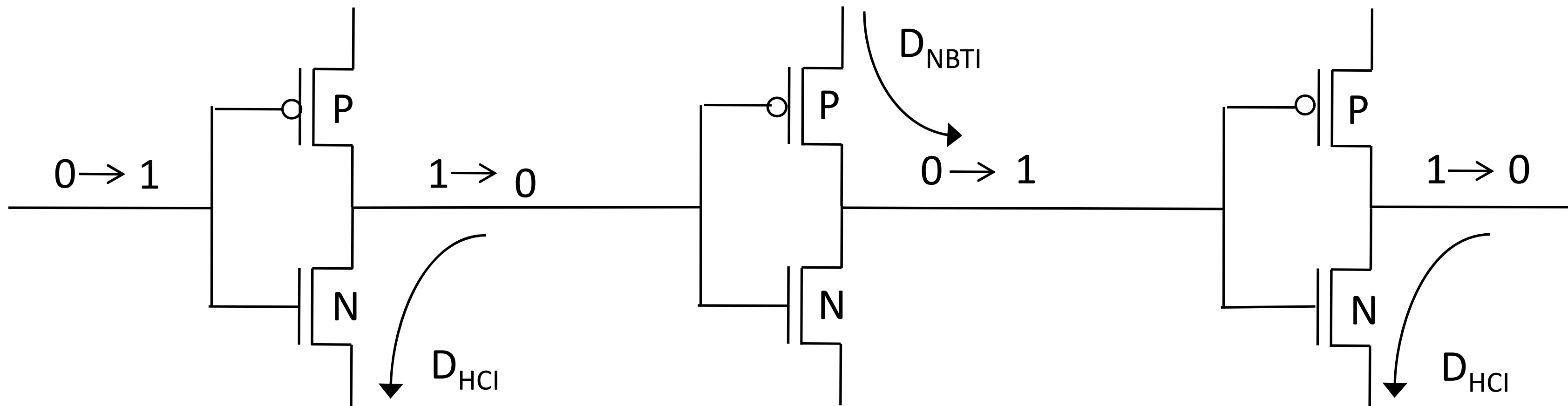


- NMOS: Hot Carrier Injection (HCI)



i-acoma group

Impact of Aging on Inverter Chain



$$\Delta \text{ (Inverter chain delay)} = (\Delta NBTI + \Delta HCI) * N/2$$

What Affects Aging Rate?

Factor	NBTI	HCI
V_{dd} - V_t	exponential	exponential
T	exponential	linear
f, activity (α)	---	linear

- Job scheduling:
  - Temperature, activity (α)
- Voltage changes at key times of processor service
  - Changes V_{dd} , V_t
  - High impact on temperature

i-acoma group

Proposed Approaches

- Hide aging
  - Aging-driven job scheduling in multicores

- Slow down aging:

  Apply at key times of service life:
  - ASV (Adaptive Supply Voltage): ASV- or ASV+
  - ABB (Adaptive Body Bias): FBB or RBB

Aging-Driven Job Scheduling

- Different policies are possible
- A possible one:
  - Send aging-intensive applications to the faster cores
    - High-T, high activity apps
    - Fast aging of cores that have more room to age
  - Send less aging-intensive applications to slower cores
    - Low-T, low activity, memory bound apps
    - Slow aging of cores that have less room to age

i-acoma group

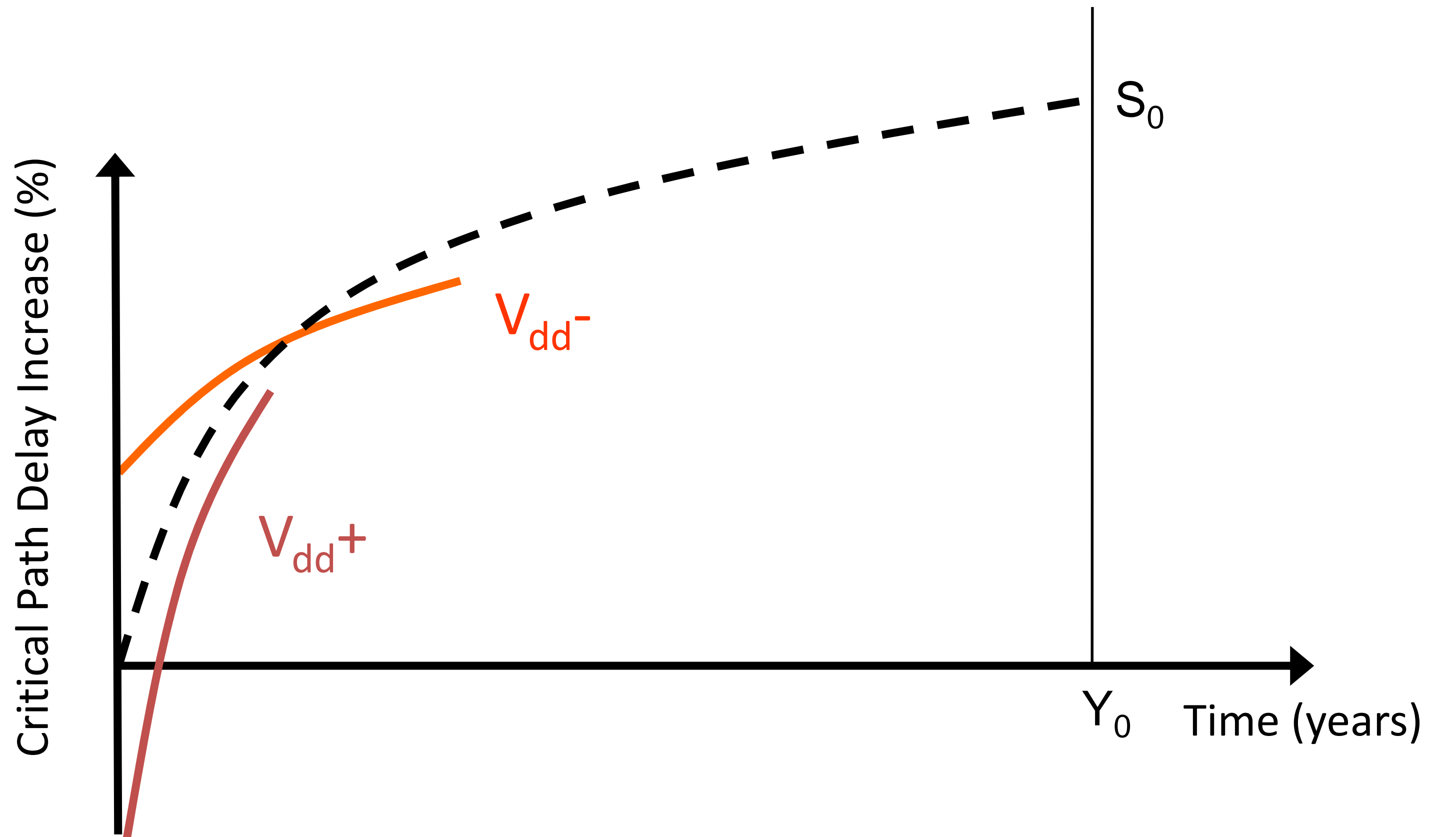
Proposed Approaches

- Hide aging
  - Aging-driven job scheduling in multicores

- Slow down aging:

  Apply at key times of service life:
  - ASV (Adaptive Supply Voltage): ASV- or ASV+
  - ABB (Adaptive Body Bias): FBB or RBB

Voltage Changes (Vdd)



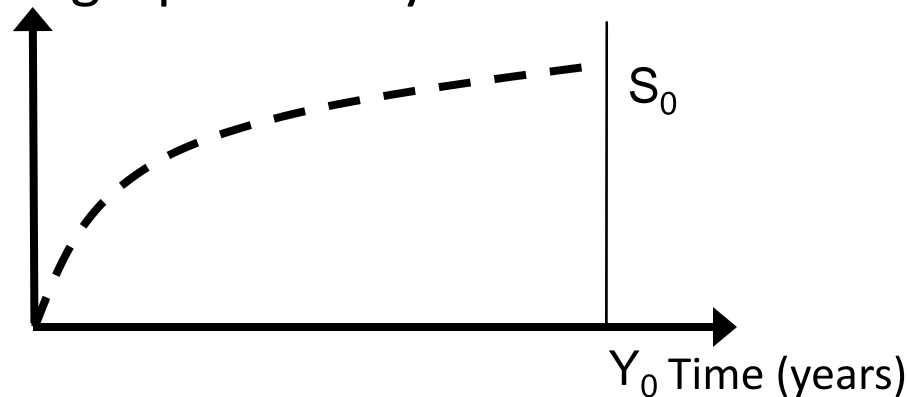
No frequency change

Two groups of Techniques

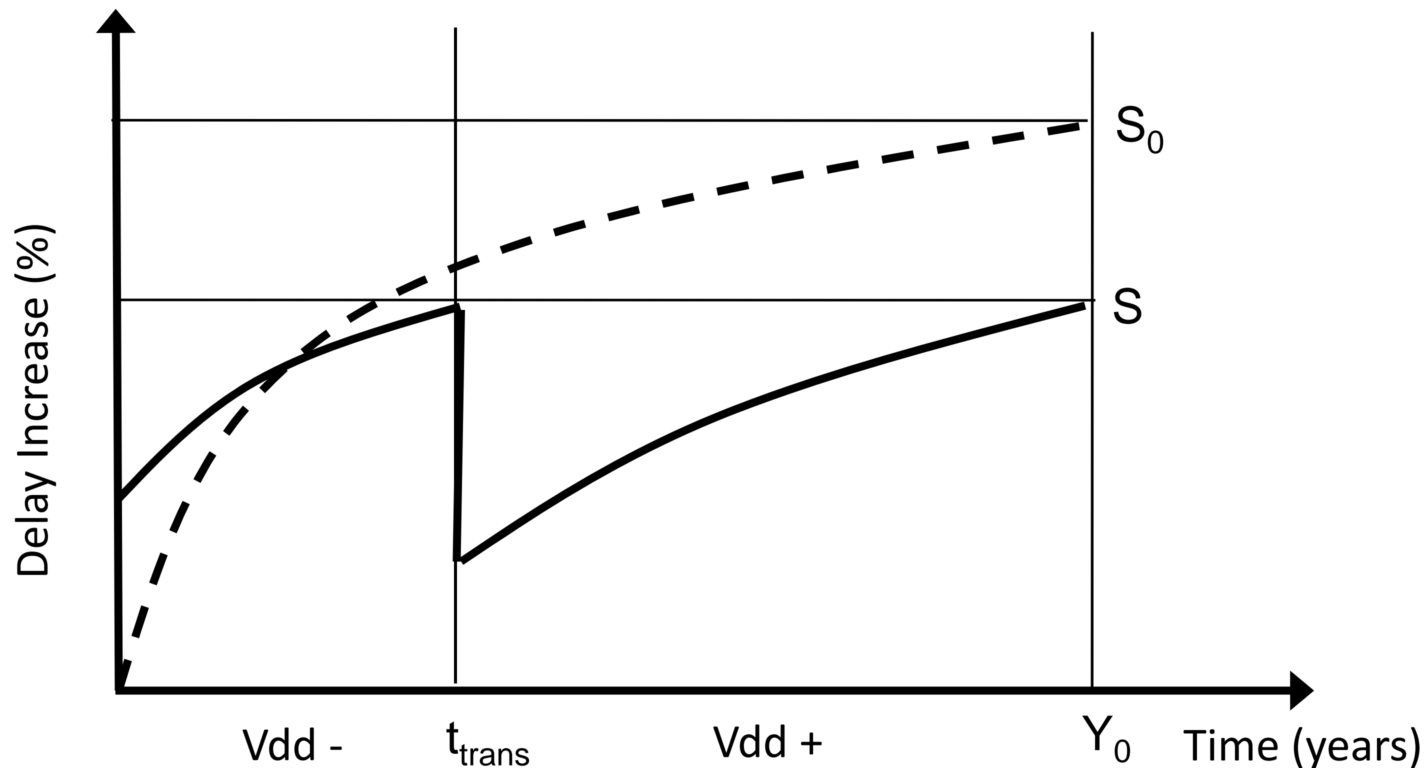
- SlowAge: RBB or ASV-
  - Reduces T and/or Vdd
  - Slows down aging rate
  - Increases delay of logic paths
- HighSpeed: FBB or ASV+
  - Increases T and/or Vdd
  - Increases aging rate
  - Reduces delay of logic paths

When to Apply These Techniques?

- Aging rate is higher in beginning and lower at end
- Apply V_{dd}- in beginning
  - Slow down aging the most
  - Have room to slow down paths
- Apply V_{dd}+ at end
  - Little impact on aging rate anyway
  - Reduce logic path delays

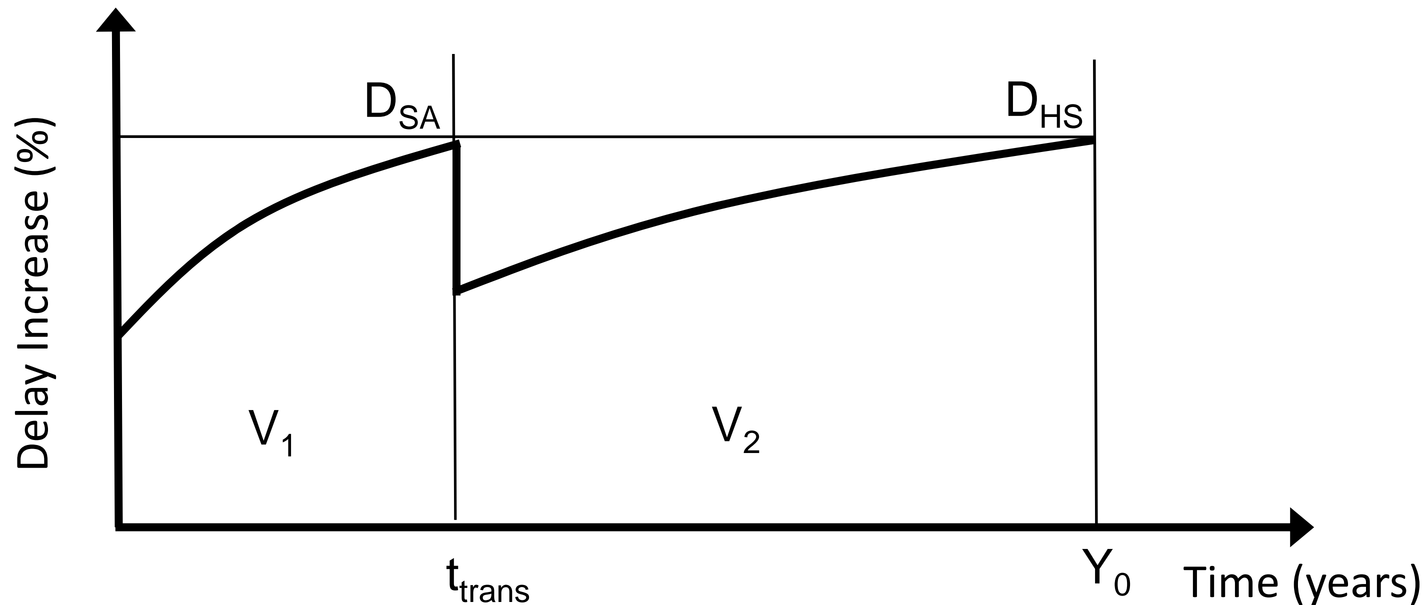


Impact on the Aging Curve



- Need less guardband (S% rather than S_0%)
- Can cycle it at higher frequency from the beginning

Non-linear Optimization



Minimize D_{HS}, such that

$D_{SA} = D_{HS}$

$D_{SA} = f(V_1, t_{trans})$

$D_{HS} = f(V_2, Y_0)$

Aging Setup

- Assume processors designed for 7 year service life
- ΔNBTI : ΔHCI = 3 : 1 [Bernstein'06]
- 2 different aging rates [Abella'07,Mitra'07]
  - Low Wearout: 10% delay increase in service life
  - High Wearout: 25% delay increase in service life
- 16-core CMP
- 26 SPEC workloads in random order

Facelift Frequency Improvements

Frequency Increase (%)

25
20
15
10
5
0

Sched

Low Wearout

High Wearout

Frequency increase corresponds to reduction in guardband

i-acoma group

Facelift Frequency Improvements

Frequency Increase (%)

25
20
15
10
5
0

Low Wearout
High Wearout

Sched
Sched+ASV

Facelift Frequency Improvements

Frequency Increase (%)

25
20
15
10
5
0

Low Wearout
High Wearout

Sched
Sched+ASV
Sched+ABB

Facelift Frequency Improvements

Frequency Increase (%)

25
20
15
10
5
0

Low Wearout
High Wearout

Sched
Sched+ASV
Sched+ABB
No Aging

Facelift is effective across a range of aging rates

Conclusions

- Facelift
  - Enables multicore to run on avg. at a 14-15% higher freq
  - Regains > 50% of performance losses due to aging
  - Alternatively, design for 5-7 months of service life and use for 7 years
- Simple implementation and low hardware overhead
- Architecture-level techniques can be effective at mitigating the effect of aging on performance

Facelift: Hiding and Slowing Down Aging in Multicores

International Symposium on Microarchitecture (2008)

Abhishek Tiwari and Josep Torrellas
Department of Computer Science
University of Illinois at Urbana-Champaign
http://iacoma.cs.uiuc.edu

i-acoma group

Related Work

- Lifetime reliability
- Detecting processor aging
- Reducing aging at circuit-level

Slowing Down Aging

Variable	NBTI	HCI
V_{dd}	exponential	exponential
T	exponential	Linear
V_t	exponential	exponential

- Change aging rate with
  - Adaptive Body Bias (ABB):
    - changes V_t
    - Exponential impact on T
  - Adaptive Supply Voltage (ASV):
    - changes V_{dd}
    - Exponential impact on T

i-acoma group

Lifetime Reliability

- J. Srinivasan, S.V. Adve, P. Bose, J. A. Rivers, ISCA04, ISCA05
- Similarities:
  - Notion of failure: timing error
  - Direct tradeoff between timing errors and performance
- Key difference
  - We leverage the performance degradation with time to improve reliability and performance

Detecting Aging

- Blome et al. MICRO 2007
  - Predict wearout faults based on circuit delay using exponential moving average
- M. Agarwal, B. Paul and S. Mitra VLSI Test Symposium 2007
  - Measuring timing margin violations using special latches
- Smolens et al. SELSE 2007
  - BIST

Circuit-level Techniques

- Bose ISCA 2008, Abella et al. MICRO 2007
- Improved design of memory cells resilient to NBTI-induced degradation
- Reduce stress time of PMOS
- Complementary to our architecture-level techniques

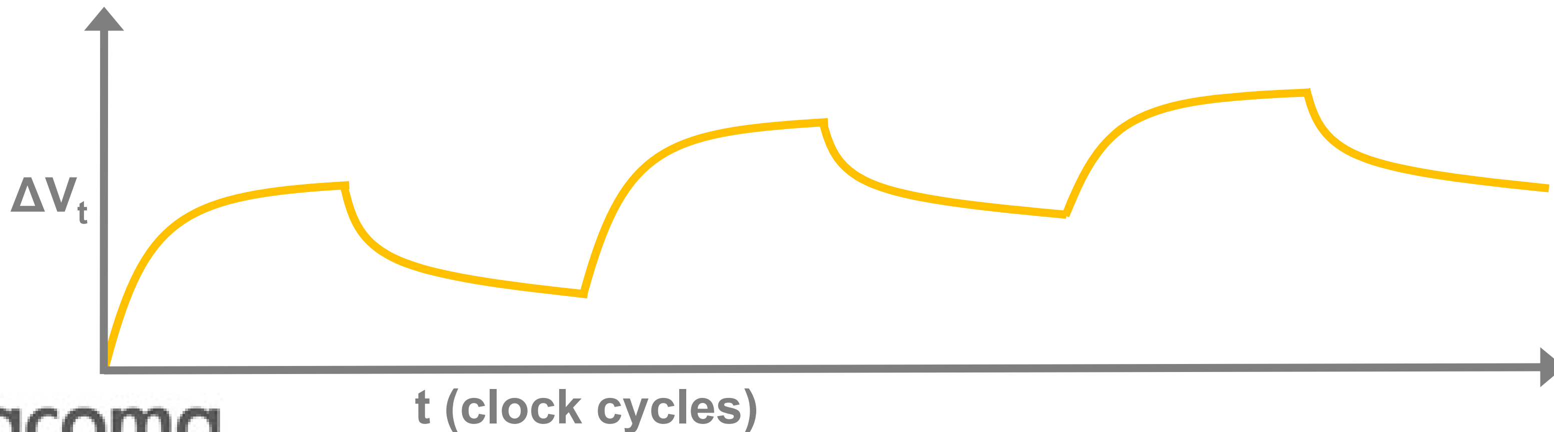
NBTI in PMOS

- Vattikonda et al DAC06 (Reaction-Diffusion Model)
- Stress phase (input=0)

$$\Delta V_{t_stress} \propto e^{(Vdd-Vt)} * e^{-Ea/kT} * (t_{stress})^{0.25}$$

- Recovery phase (input=1)

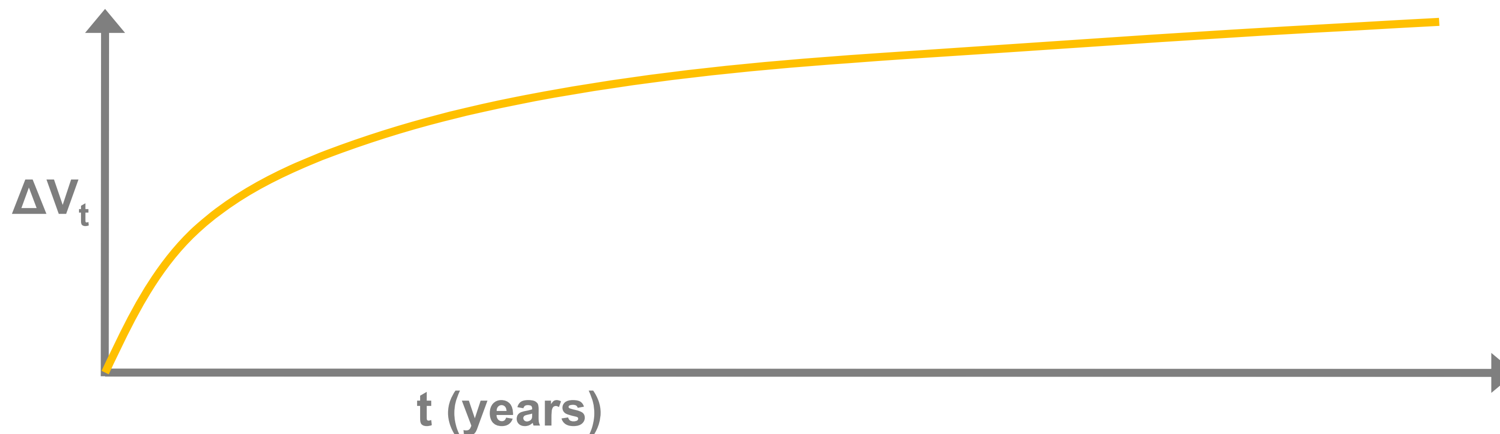
$$\Delta V_t = \Delta V_{t_stress}(1 - \sqrt{\eta t_{rec} / (t_{stress} + t_{rec})})$$



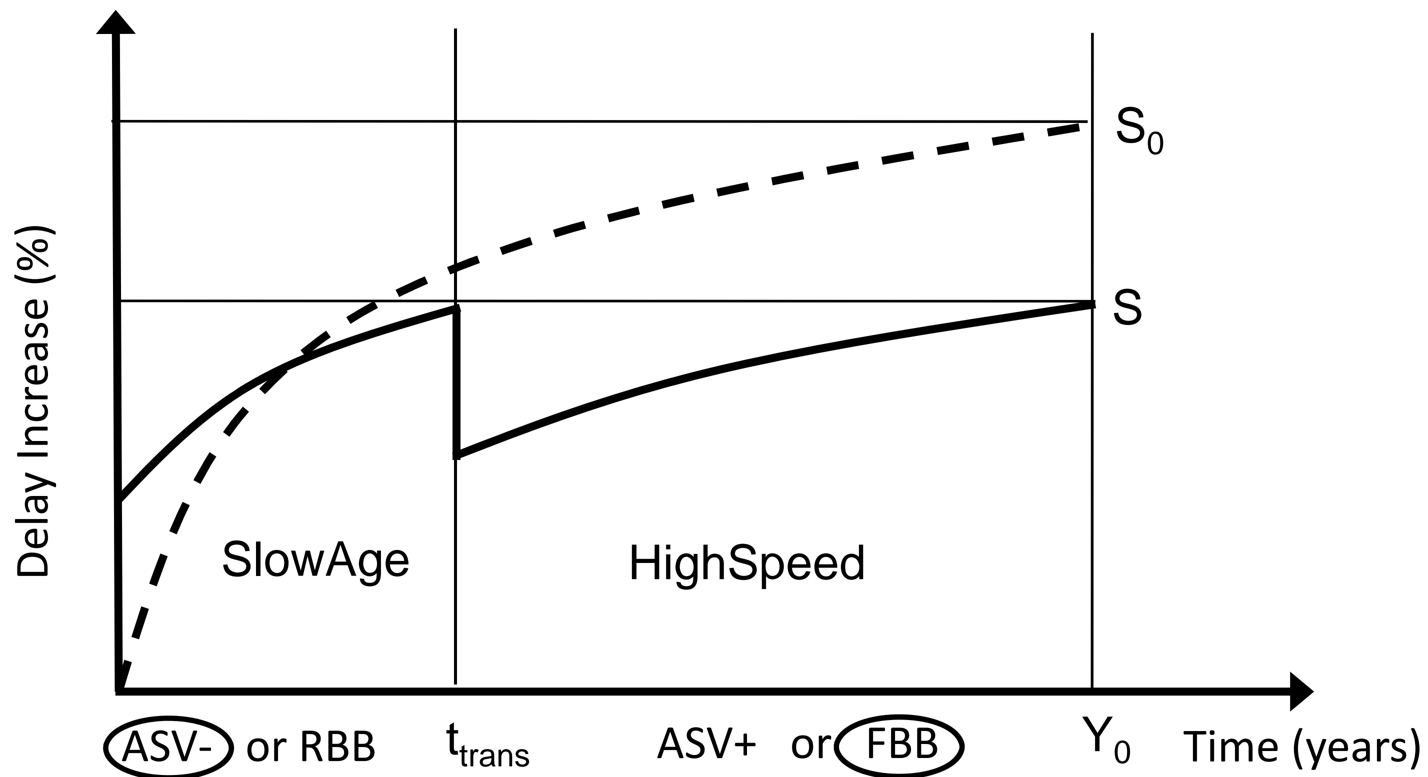
i-acoma group

HCl in NMOS

- Wang et al. TMDR07 and Takeda et al. *Hot Carrier Effects in MOS Devices* 1995
- $\Delta V_t \propto f * \alpha * e^{(Vdd-Vt)/toxE1} * t^{0.5}$



Impact on the Aging Curve



Need less guardband => can cycle it at higher frequency from the beginning