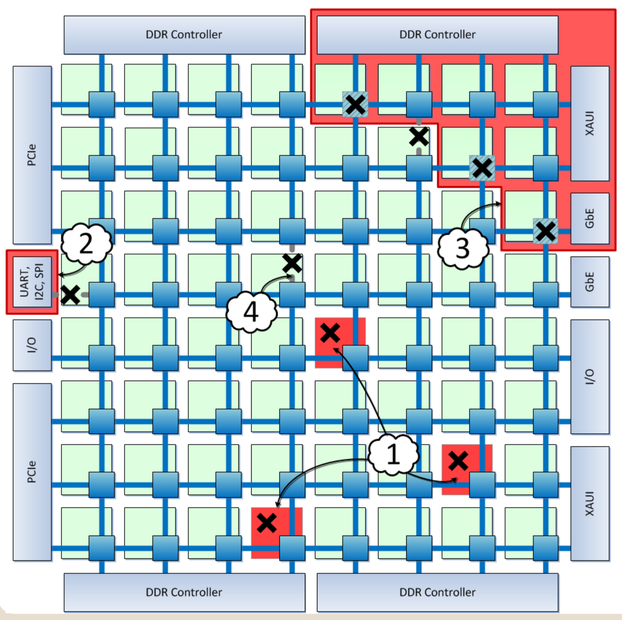
USE IT OR LOSE IT :
WEAR-OUT AND LIFETIME IN FUTURE CHIP MULTIPROCESSORS

Hyungjun Kim, Paul V. Gratz (*Texas A&M University*)

Arseniy Vitkovskiy, Vassos Soteriou (*Cyprus University of Technology*)

MOTIVATION

Single failure in the interconnect could incapacitate the entire CMP



ITRS: 10-fold decrease in wear-rate required to maintain current design lifetimes without dramatically increasing timing margins.

- **Recoverable faults:**
  1. Failures of cores
- **Non-Recoverable faults:**
  2. Failure of certain links leading to disconnection of important peripherals
  3. or even a network segment
  4. Failure of other links may lead to communication deadlock.

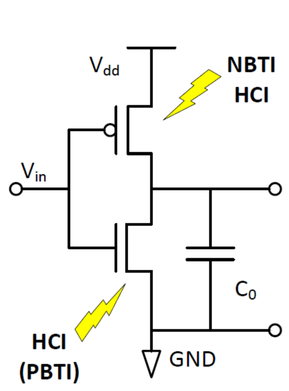
Transistor failure mechanisms:
- *Negative-bias temperature instability (NBTI)*
- *Hot carrier injection (HCI)*

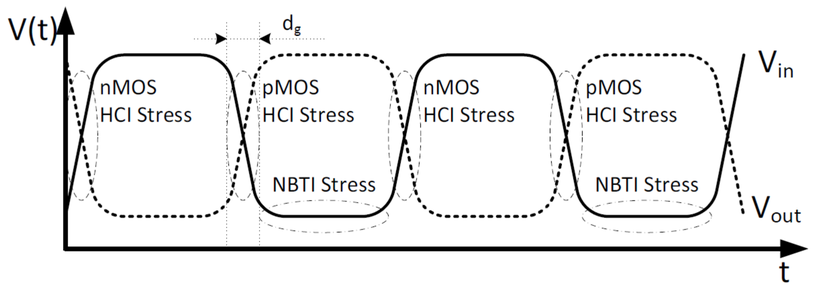
CONTRIBUTIONS

- **Theory**: microarchitecture-level HCI and NBTI wearout models
- **Analysis**: Router/link wearout characterization under real NoC traffic (PARSEC benchmarks)
  - *Lack* of load leads to interconnect wear
- **Design**: Novel wear-resistant router microarchitecture, 22X lifetime improvement

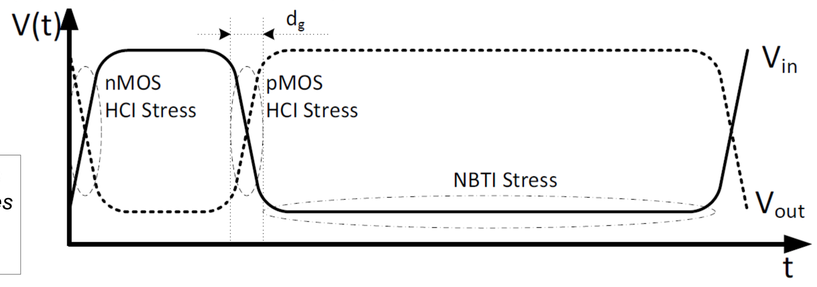
FAILURE MECHANISMS

HCI occurs when input toggles. *Related to **activity factor**.*
NBTI occurs when input holds value "0". *Related to **duty cycle**.*





(a) Higher switching activity, lower duty cycle.



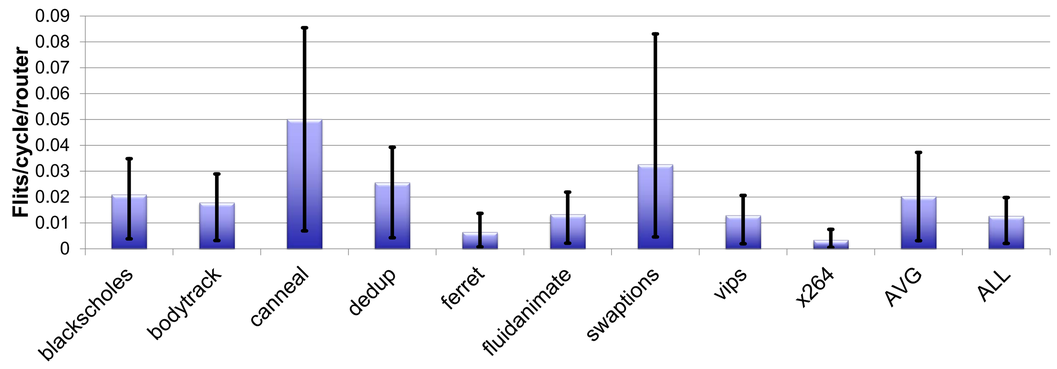
(b) Lower switching activity, higher duty cycle.

- **Activity factor:** *% of cycles where input of transistor toggles*
- **Duty cycle** *: % of cycles where NMOS gate value is 0*

Both mechanisms slow transistor switching speed

- Delay of each gate increases due to HCI and NBTI
  - HCI: $\Delta d_g = A(I_{sub})^m \left(d_g \alpha f t\right)^{n'}$, α = activity factor
  - NBTI: $\Delta d_g = \dot{A}\left(\frac{\beta}{1-\beta}\right)^n t^n$, β = duty cycle
- When the sum of increased delay on a pipeline stage exceeds a given margin, it will violate timing constraints.
- Thus, ***strong impact on timing critical path.***

REAL WORKLOAD CHARACTERIZATION



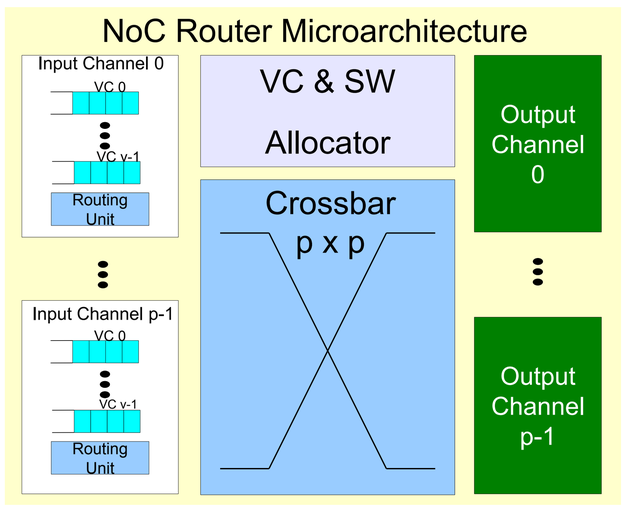
Load seen by each router under PARSEC suite benchmarks (incoming rate)
Average (bar), minimum and maximum (whiskers)

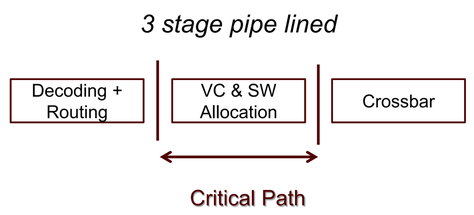
- Incoming rate very low generally, particularly for minimum-load routers
  - Network lifetime is determined by the router with the lowest incoming rate
- Activity factor ∝ incoming rate ∝HCI – Only minor impact
- Duty cycle ∝ 1/incoming rate ∝NBTI – ***Significant impact!***

- ***Use it or Lose it:***
  - Balance duty cycle by periodically by "exercise" or induced load when there is no flit in the router
  - Change slowly so it won't significantly activity factor

ROUTER MICROARCHITECTURE

Goal:
Find the critical path, which is the most vulnerable to NBTI aging.
Balance duty cycle along those paths to minimize NBTI stress.

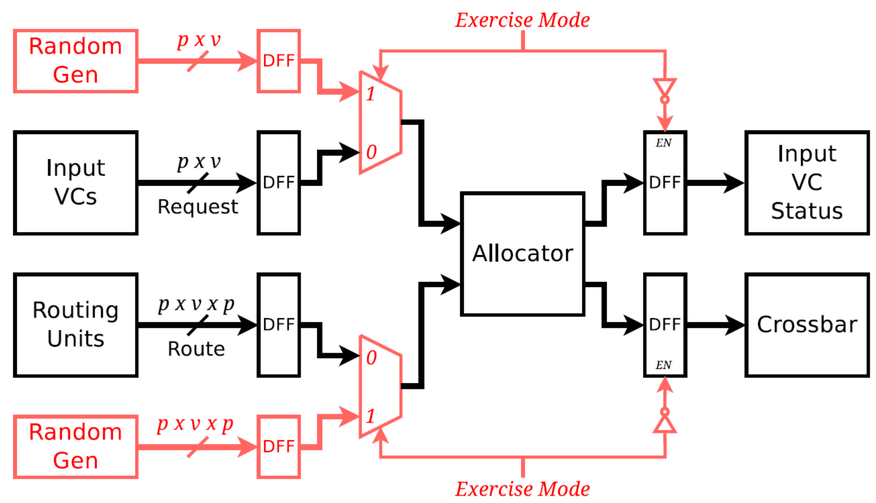




- Collected timing critical paths
  - Slack of less than 10% clock time
  - Mostly VC & SW allocation stage
- Duty cycle along critical path found to be extremely biased under low incoming rate
  - Allocation corner cases which occur infrequently except under high load
- The router ages due to NBTI because these paths rarely activate

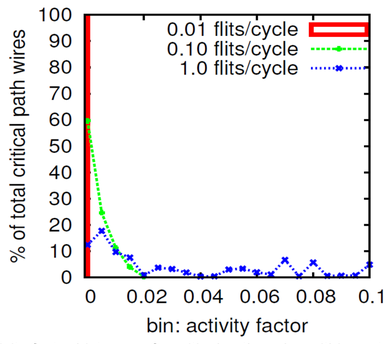
Exercise Mode:
Exercise the critical path when there are no flits to process

Inject random values to allocators while output pipeline latches are disabled
Change random values slowly to reduce impact on high activity factor

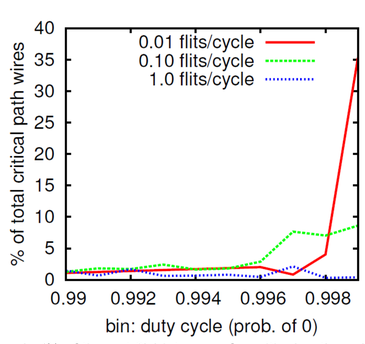


WORKLOAD AND CRITICAL PATH AGING

Wearout stress depends on workload :



Activity factor histogram for critical path nodes within router, under random injection of varying rate. Critical path nodes defined as nodes on all paths with latency within 10% of clock period.



Duty cycle (% of time at 0) histogram for critical path nodes within router, under random injection of varying rate. Critical path nodes defined as nodes on all paths with latency within 10% of clock period.

- Activity factor generally low
- Some sensitivity to router incoming rate

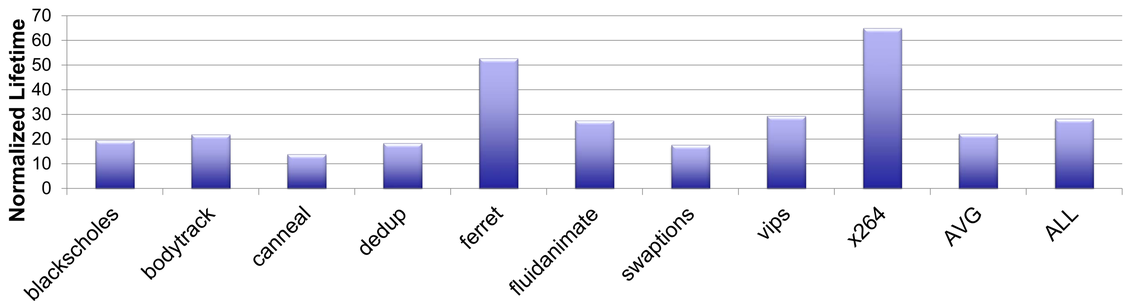
HCI stress negligible when incoming rate is low

- High fraction of wires on the critical path have "poor" duty cycle when load is low
- Highly sensitive to load

NBTI stress inversely proportional to incoming rate

Wear (due to NBTI) expected to occur when routers are underutilized

LIFETIME IMPROVEMENT



Normalized lifetime (acceleration factor) of CMP system with the *Use it or Lose it* router microarchitecture under loads from the PARSEC benchmarks. AVG is the geo-mean across benchmarks, ALL is the product of cycling through each benchmark sequentially.

DISCUSSION

- Proposed design yields a 13.8~65 x increase in CMP lifetime.
- ~1% increase in router power consumption.
- <1% increase in router area.
- No performance impact.

Contact : Hyungjun Kim (hyungjun@tamu.edu) Arseniy Vitkovskiy (arseniy.vitkovskiy@cut.ac.cy) Paul V. Gratz (pgratz@tamu.edu) Vassos Soteriou (vassos.soteriou@cut.ac.cy)