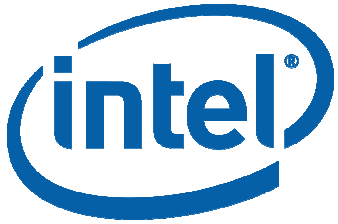


Low Power/Low Voltage Computing

Shih-Lien Lu

Intel Labs

October 13, 2010

Acknowledgement:
Alaa Alameldeen, Keith Bowman, Zeshan Chishti,
Dinesh Somasekhar, James Tschanz, Chris Wilkerson, Wei Wu

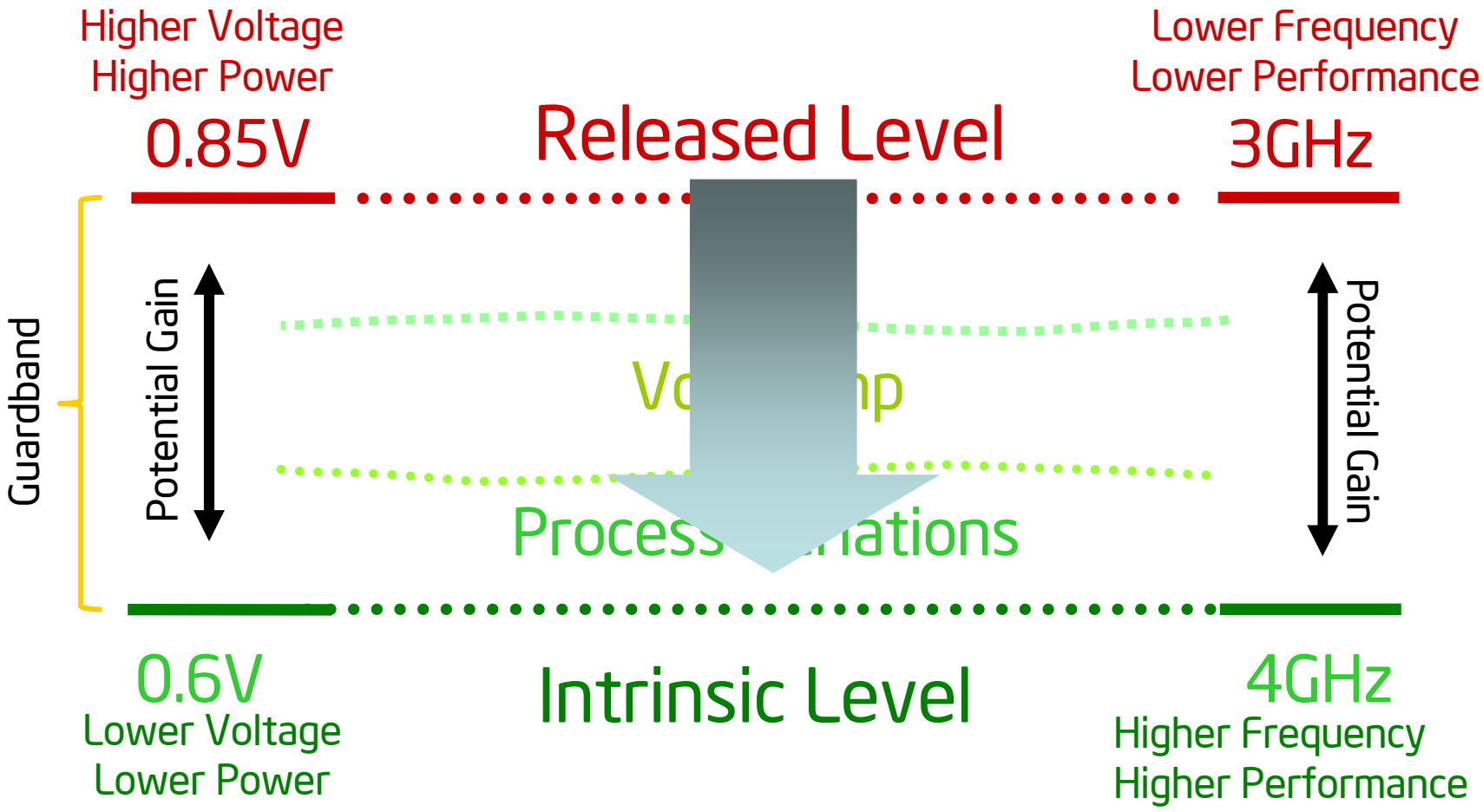
Power Limit

- Mark Seager
  - 2 Petaflop -> 6 MW (1.27 PF -> 2MW)
  - Linear scaling: 1 Exaflop -> 3GW (1.6GW)
  - With idea technology scaling @ constant die size and freq
    - 3 generation: ~380MW (200MW)
    - 4 generation: ~190MW (100MW)
  - This talk addresses challenges facing aggressive voltage scaling

Resiliency: A Familiar Topic

- Resilient design employs techniques that handle faults to give correct operation
- *Past Focus: Increase Reliability*
- **Resiliency as Part of Optimization**

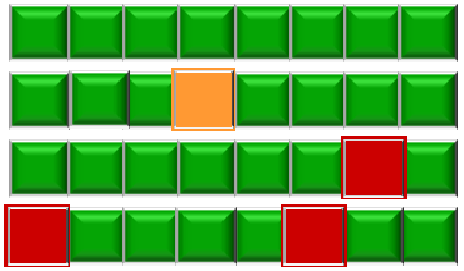
Reduce or Eliminate Guardbands



*Margins added for rare occurrences impact **Power** and **Performance***

Faults Caused by Vcc Reduction

- Decreases SRAM stability, more failing bits

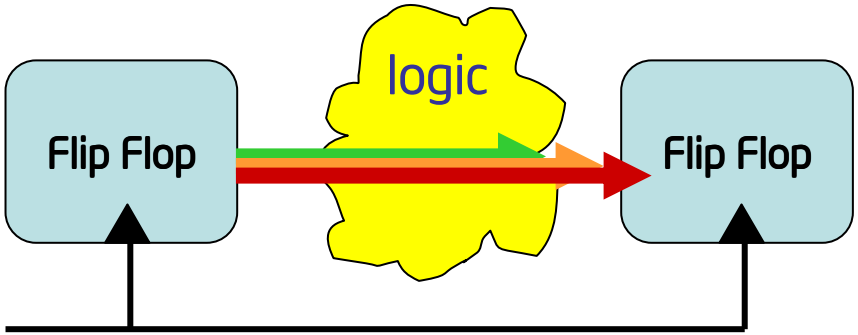


Vcc > Vmin : Memory fully functional

Vcc < Vmin : A few bits fail

Vcc << Vmin : Multiple bits fail

- More timing violations—reduces frequency



Vcc > Vmin : No timing violations

Vcc < Vmin : Some violations

Vcc << Vmin : More violations

Addressing On-die Memory Errors

- Cache line disabling
  - Coarse grain
  - Fine grain – Wilkerson et. al. ISCA 2008
- Multi-segmented ECC
  - Take part of the cache to store ECC bits
  - Segmented protection
  - OLSC: simple and modular encode/decode
- Variable Strength ECC
  - Current project

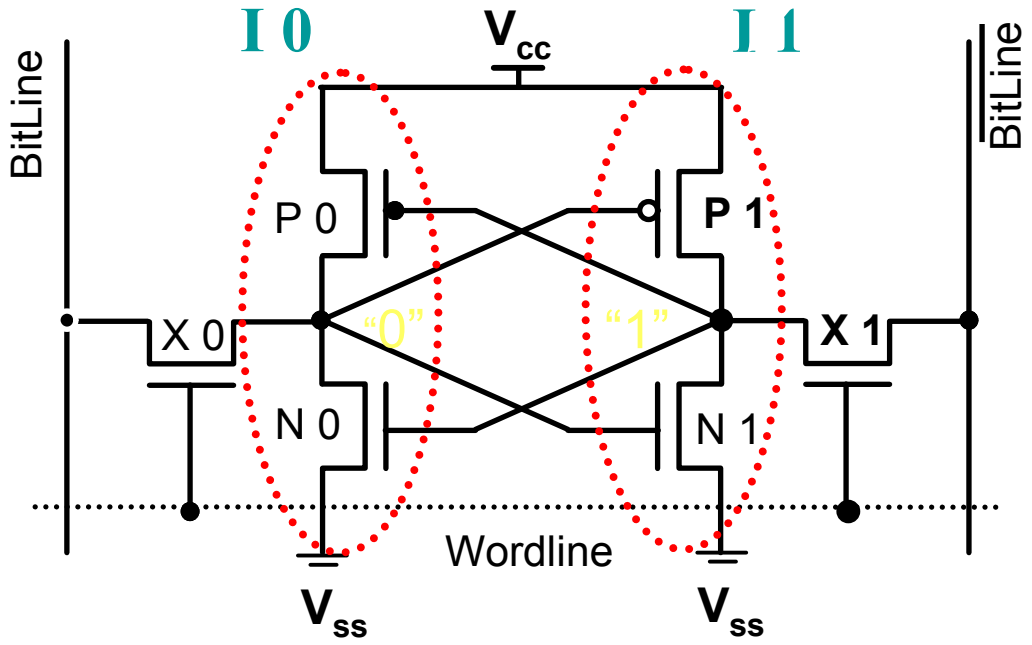
50% EPI reduction
With small performance loss

On-die Memory Fault Types

- Persistent
  - Permanent defect
  - Read/Write stability
  - Retention
- Transient
  - Particle strike
  - Proximity disturbance
- Erratic

SRAM Failures Result from Mismatched Devices in a Single Cell

Contention between READ and WRITE on device sizes
Ex: Weak pass device (X1) vs. a strong pull up device (P1) can cause a write failure



- **Random within-die variations primarily responsible**

4 Types of SRAM Failures

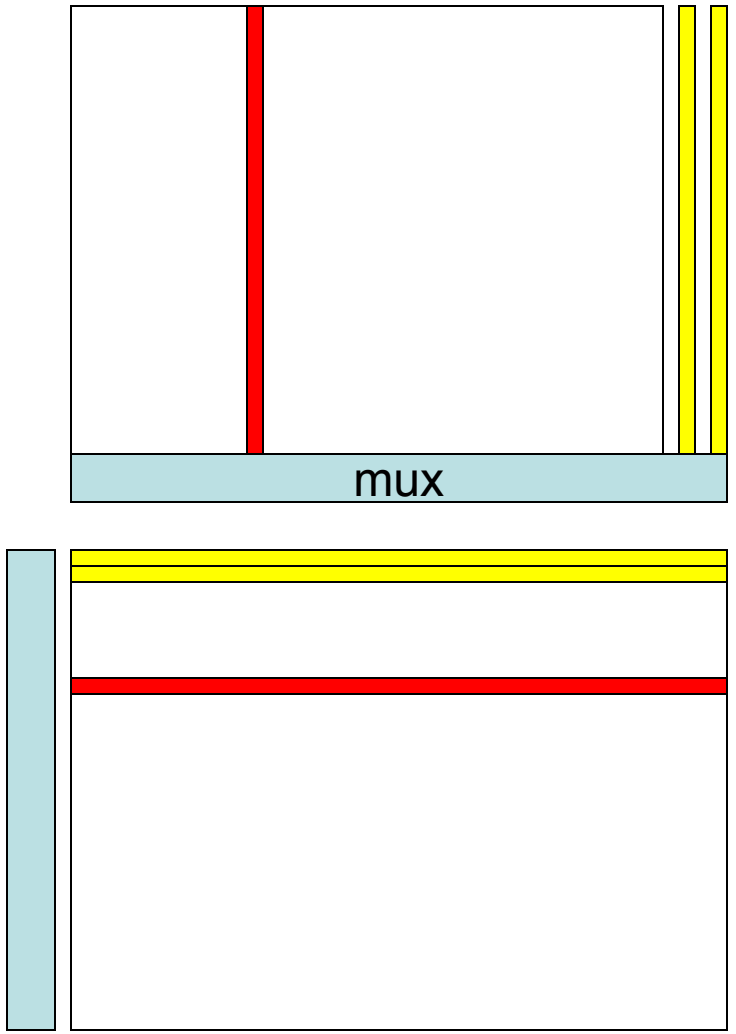
- Write failure
  - Device mismatch prevents cell from flipping
- Read failure
  - Cell flips during read
- Access failure
  - Insufficient differential increases latency.
- Retention failure
  - Reduced margin, failures occur due to noise

Resiliency Techniques

- Require testing (a priori info)
  - Advantages
    - More information means simpler (cheaper) remedy
  - Disadvantages
    - Test cost
    - Faults cannot be tested not covered
  - Techniques
    - Sparring – physical redundancy
    - Disabling – graceful degradation

Spares

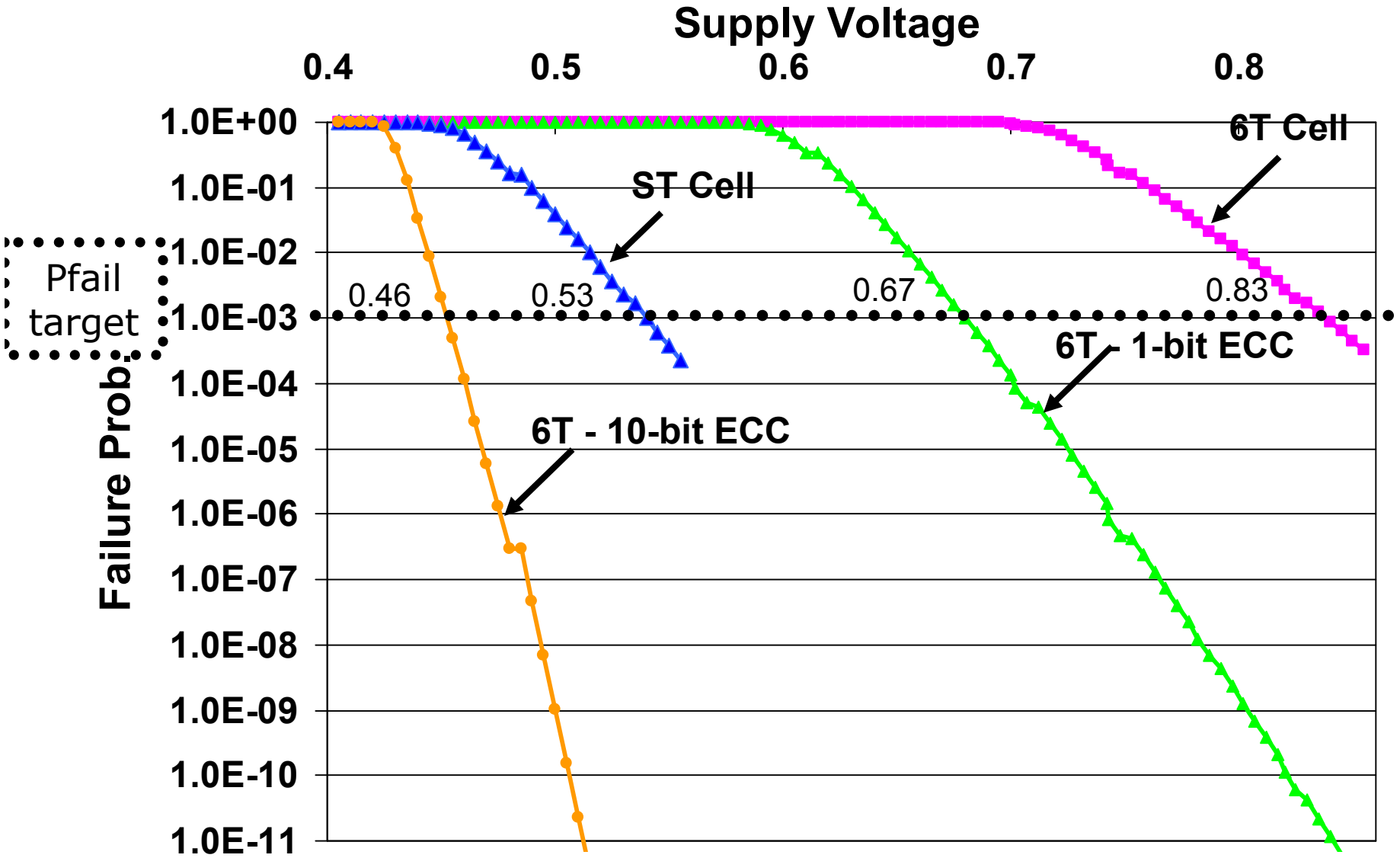
- Extra capacity
- Column redundancy
  - Random cell failure
  - Column mux
  - BIST & fuse
- Row redundancy
  - Word line failure
  - Multi-bit failure
  - Word-line segmentation
- Block redundancy



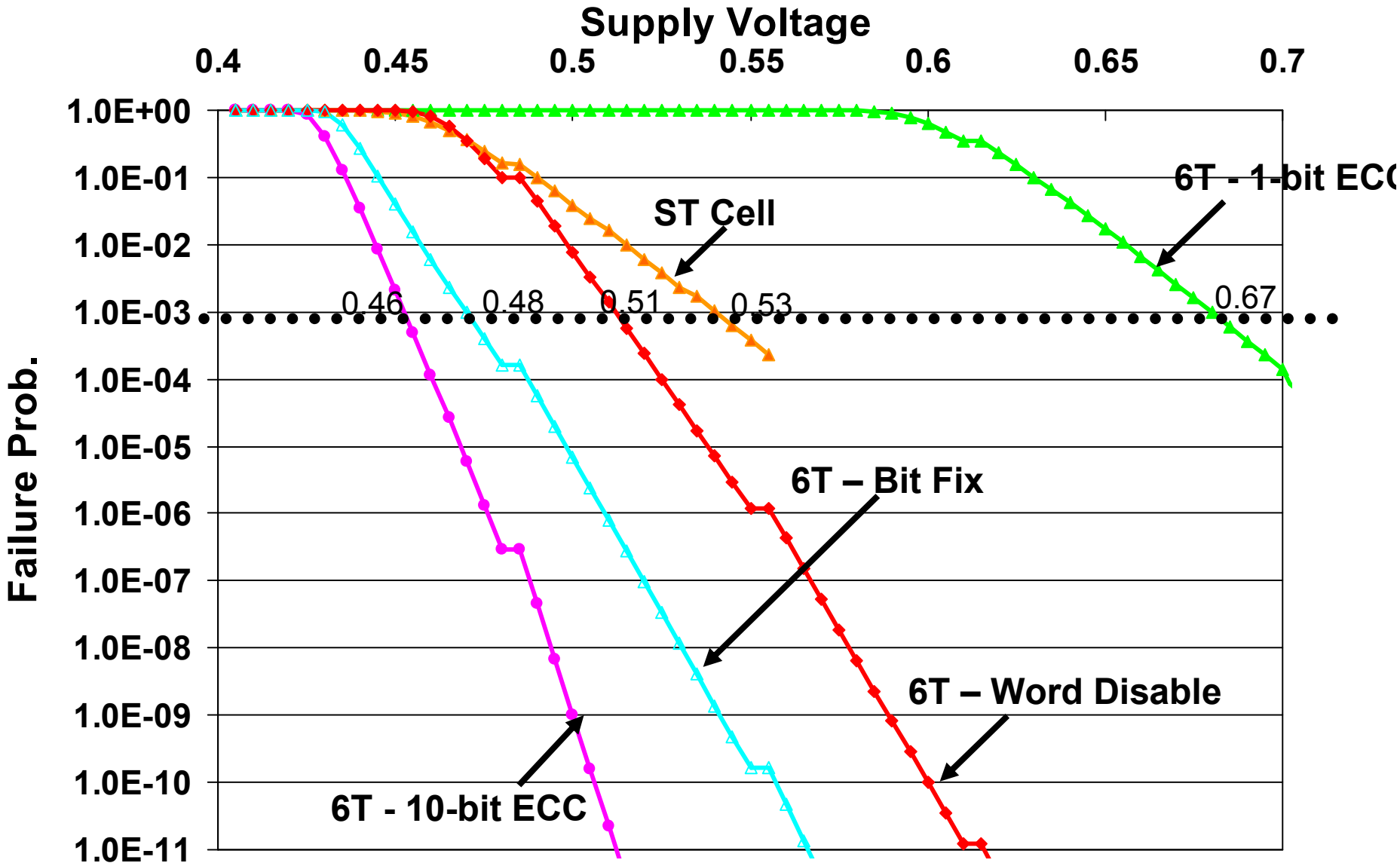
Disabling

- Wide dynamic range
- Graceful degradation
  - Static and dynamic sizing
  - Bank disabling
  - Fine grain disabling
- Example
  - Cache design

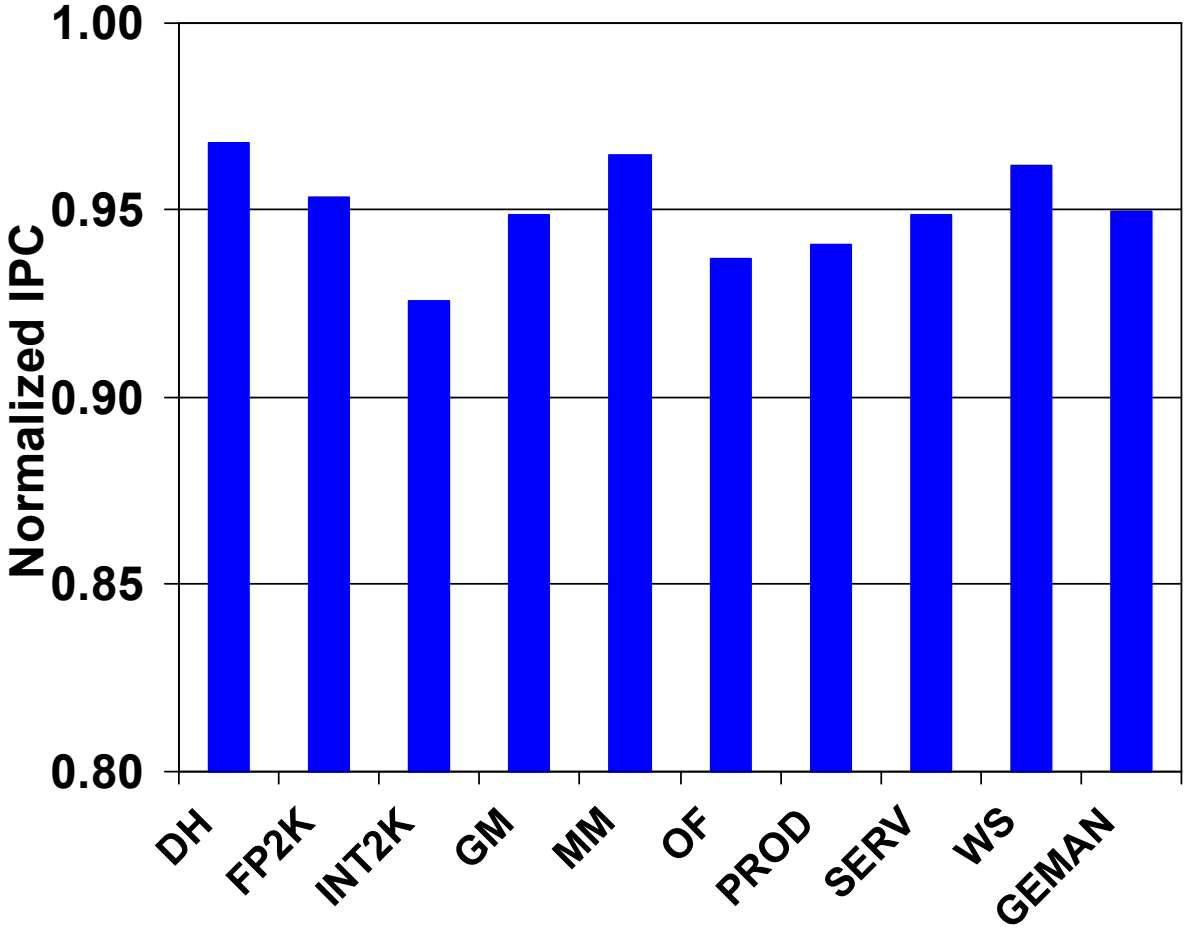
Known Methods (2MB Cache)



Vmin for Proposed Techniques



L1WDis_L2BFix Normalized to ST Cell



Performance loss ~5%

Voltage/Area/Energy Comparison

	Vccmin (mV)	Norm Area	Norm EPI
6T Cell	825	1.0	1.0
ST Cell (circuit sol.)	530	2.0	0.45
L1WDis_L2BFix	500	1.08 (L1) 1.00 (L2)	0.47

- Lower Voltage & EPI (Energy Per Inst) vs 6T
- Much less area overhead than ST Cell

Summary for this Example

- Vccmin limits energy scaling
  - Ability to reduce voltage critical but limited by memory reliability
- The ability to detect/avoid failures allows low voltage operation and reduces energy
- Configurable approach that “trades off cache capacity”
  - Maximizes performance at high voltage
  - Enables ~50% improvement in energy per instruction when operating at low voltage

Resiliency Techniques (2)

- No a priori information (through testing)
  - Adv
    - Lower test cost
  - Disadv
    - Overhead
- Techniques
  - ECC
    - Random error
    - Correlated error

Our Observations with ECC

- Use of systematic H' matrix instead of non-systematic
- Separate error detection from correction
- Codes with different strength can share H matrix
- Trade code density with logic complexity and modularity

Non-Persistent Failures

- Exhibit sporadic failing behavior
- Examples: soft errors, erratic failures
- Also exhibit supply voltage dependence
- Cannot be detected by apriori memory testing
- Both persistent and non-persistent failures affect Vccmin

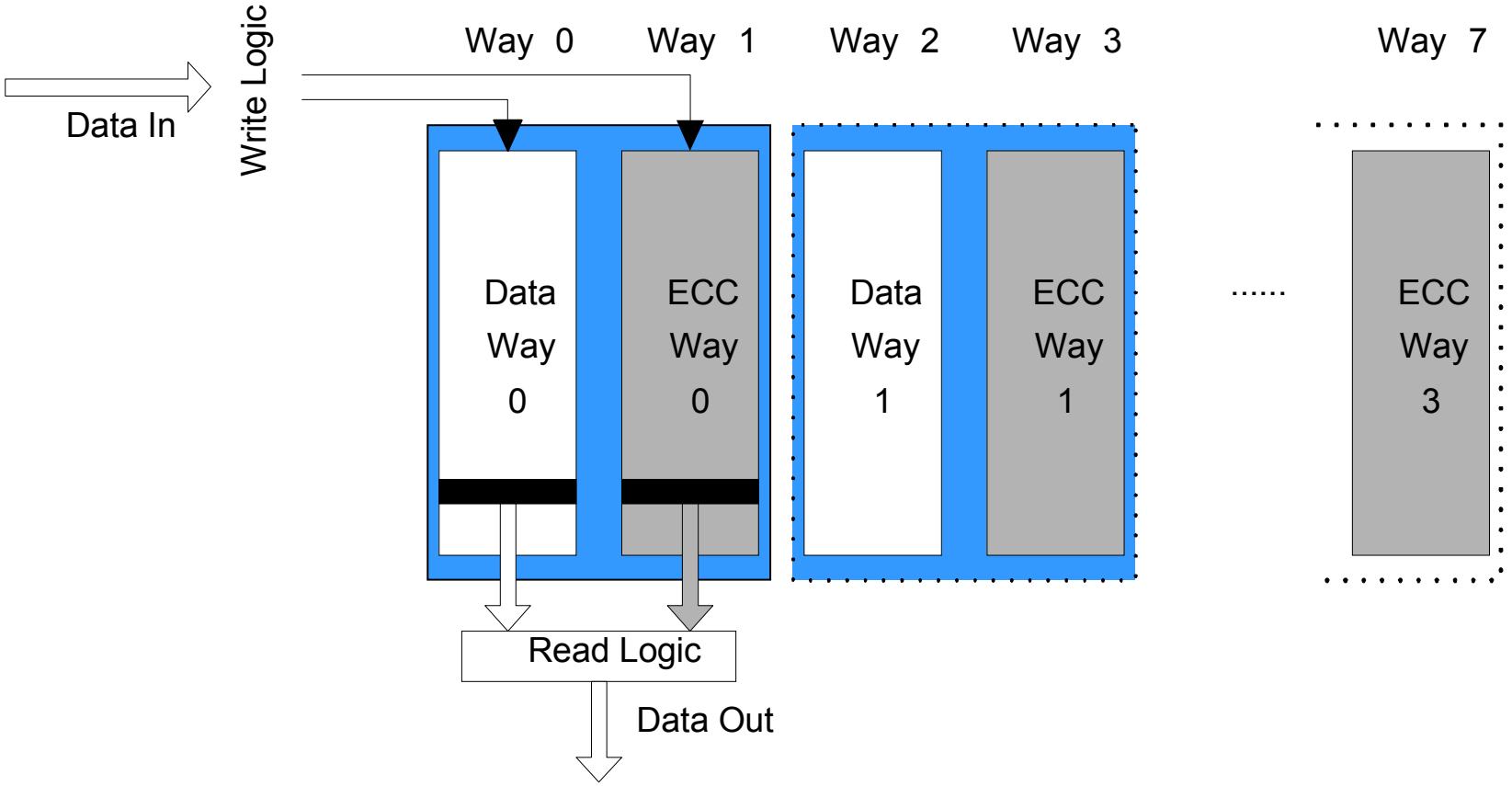
Approach

- Key Idea:
- Adaptive cache that works at both high and low voltages
  - As big as possible when performance is important (high voltages)
  - Sacrifice capacity when power is most important (low voltages)
  - In low voltage mode, use a portion of cache to store ECC
  - Enough check bits to correct both persistent and non-persistent errors
  - No additional testing to isolate defective bits

Trading off Code Density for Simplicity

- Traditional codes optimized for check bit overhead
- But, complexity grows rapidly with no. of corrections
- We need large number of corrections
  - E.g., up to 10 corrections in each cache line for 500 mV operation
- Traditional BCH-based code too complex for such corrections
- Solution: Orthogonal Latin Square Codes (OLSC)
- Less complexity at the cost of more check bits

Multi-bit Segmented ECC (MS-ECC)



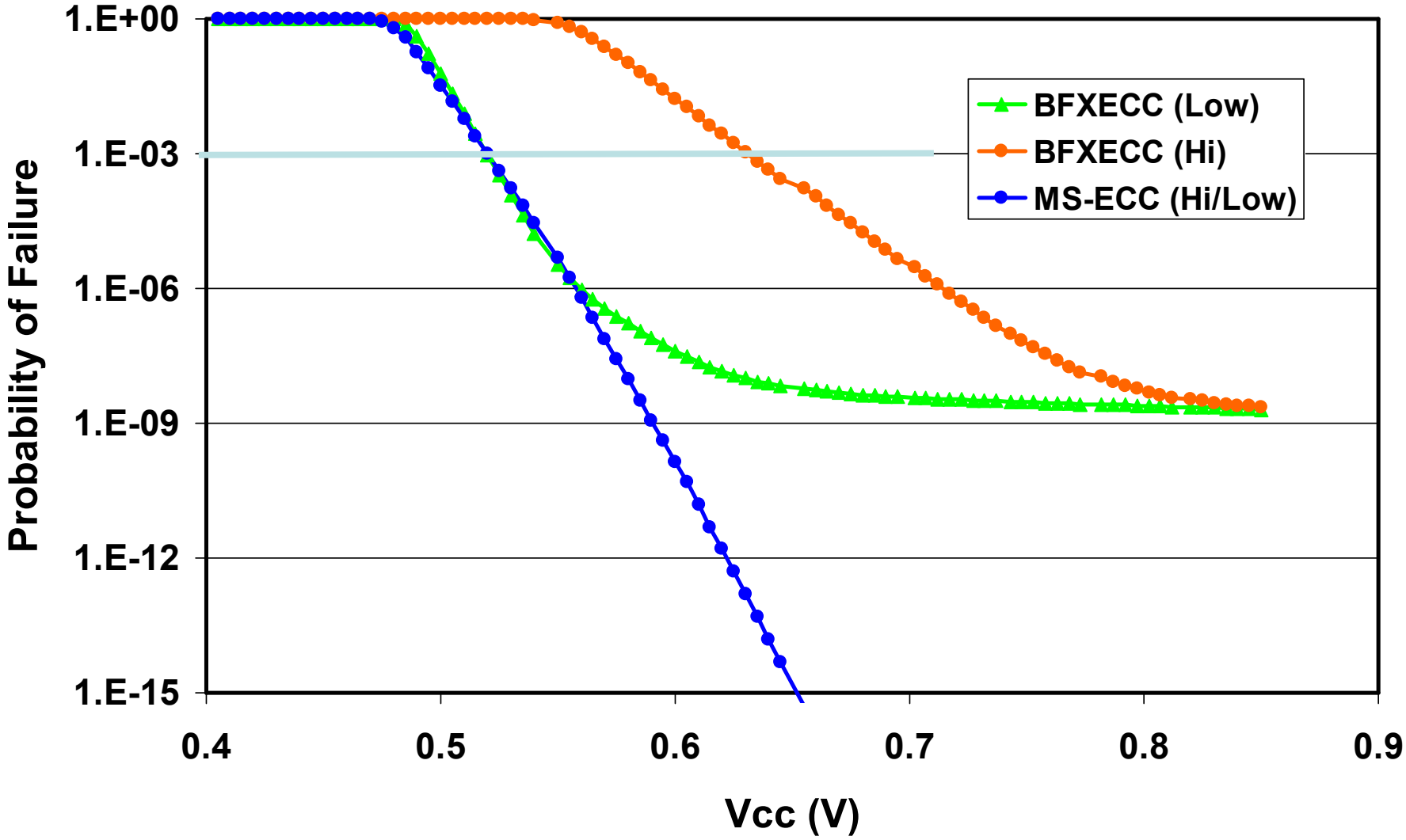
Orthogonal Latin Square Codes (OLSC)

- Modular error correction hardware
  - More regular implementation than BCH
- Based on majority voting
  - Example: TMR triplicates data and uses majority function
- Instead of keeping multiple copies of data bits,
  - Encode orthogonal groups of data bits to form check bits
  - For t-corrections in m2 data bits, need 2tm check bits

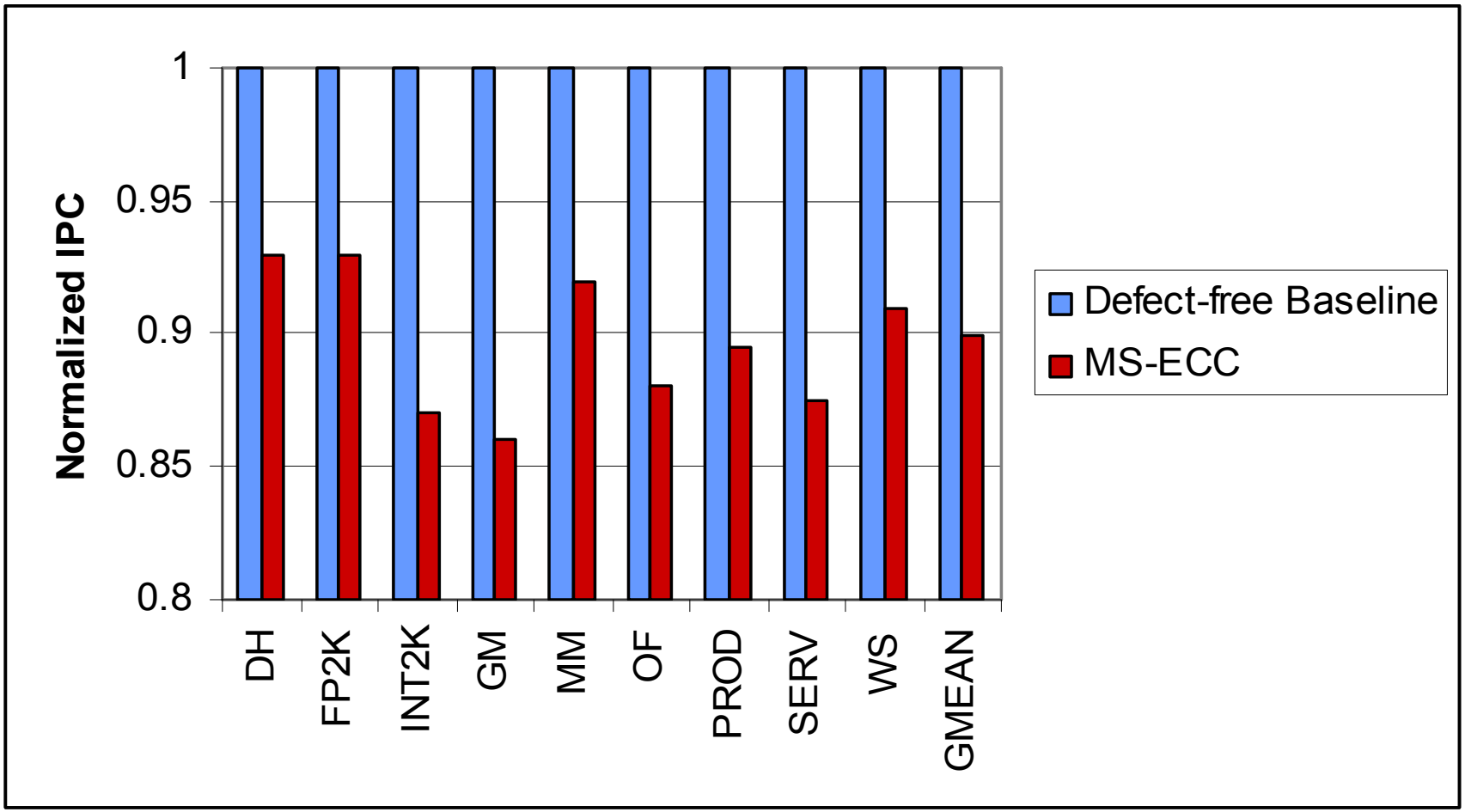
Methodology

- Two modes of operation:
  - High voltage: 1.3V, 3 GHz
  - Low voltage: 0.5V, 500 MHz
- 32K 8-way L1 caches, 2M 8-way L2 cache
- Compare
  - Baseline: SECDED ECC
  - MS-ECC: 64-bit segments, 4 corrections per segment
    - 50% capacity, 1-cycle added latency overhead
    - Compare against: Bit-fix with SECDED ECC (BFXECC)
    - Can correct 10-bit persistent and 1-bit non-persistent errors

Reliability (2MB Cache)



Performance Overhead



10% IPC degradation relative to unrealistic defect-free baseline

Energy

Scheme	Vccmin (mV)	Frequency (MHz)	Norm. Power	Norm. EPI
Baseline	725	1400	1	1
BFXECC	630	1000	0.57	0.8
MS-ECC	520	700	0.29	**0.58**

MS-ECC reduces energy-per-instruction by 42% relative to baseline SECDED ECC

Summary for This Example

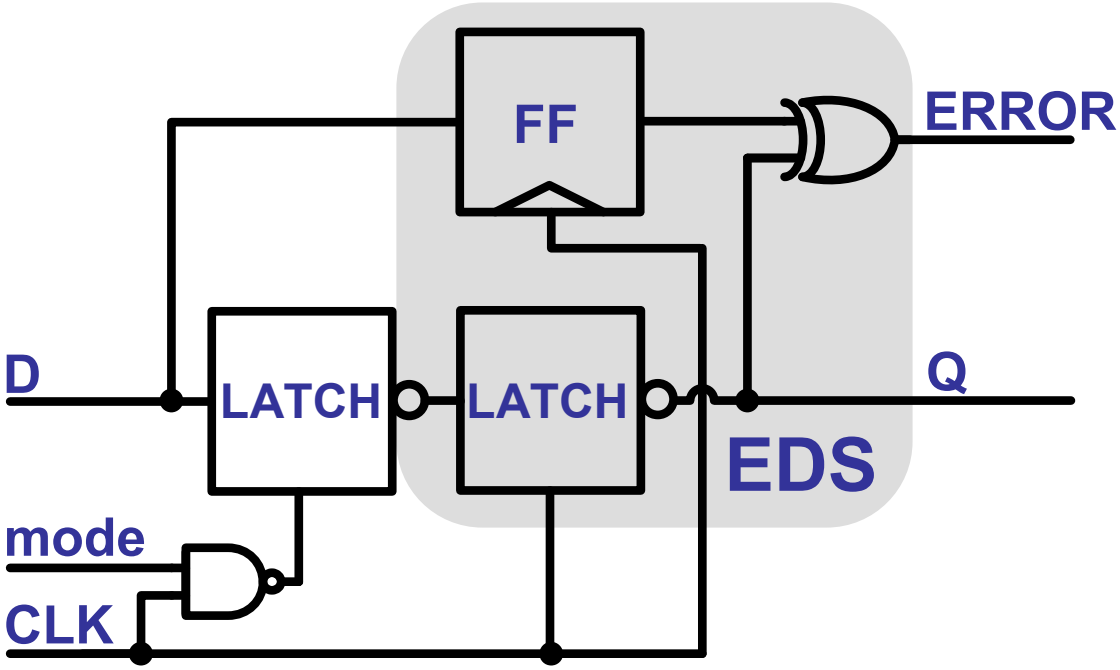
- Reducing supply voltage key to higher energy efficiency
- Supply voltage reduction limited by memory reliability
- MS-ECC: novel technique to mitigate bit failures
  - Leverages error correction codes based on OLSC
  - Does not rely on testing to isolate defects
  - Reduces Vccmin by ~ 200 mV, EPI by 42%

Addressing Logic Errors

- Timing faults
- Error detection sequential
- Detect timing faults at the circuit level
- Replay pipeline at the microarchitecture level
- A research processor in 45nm
  - "A 45nm resilient and adaptive microprocessor core for dynamic variation tolerance," ISSCC 2010

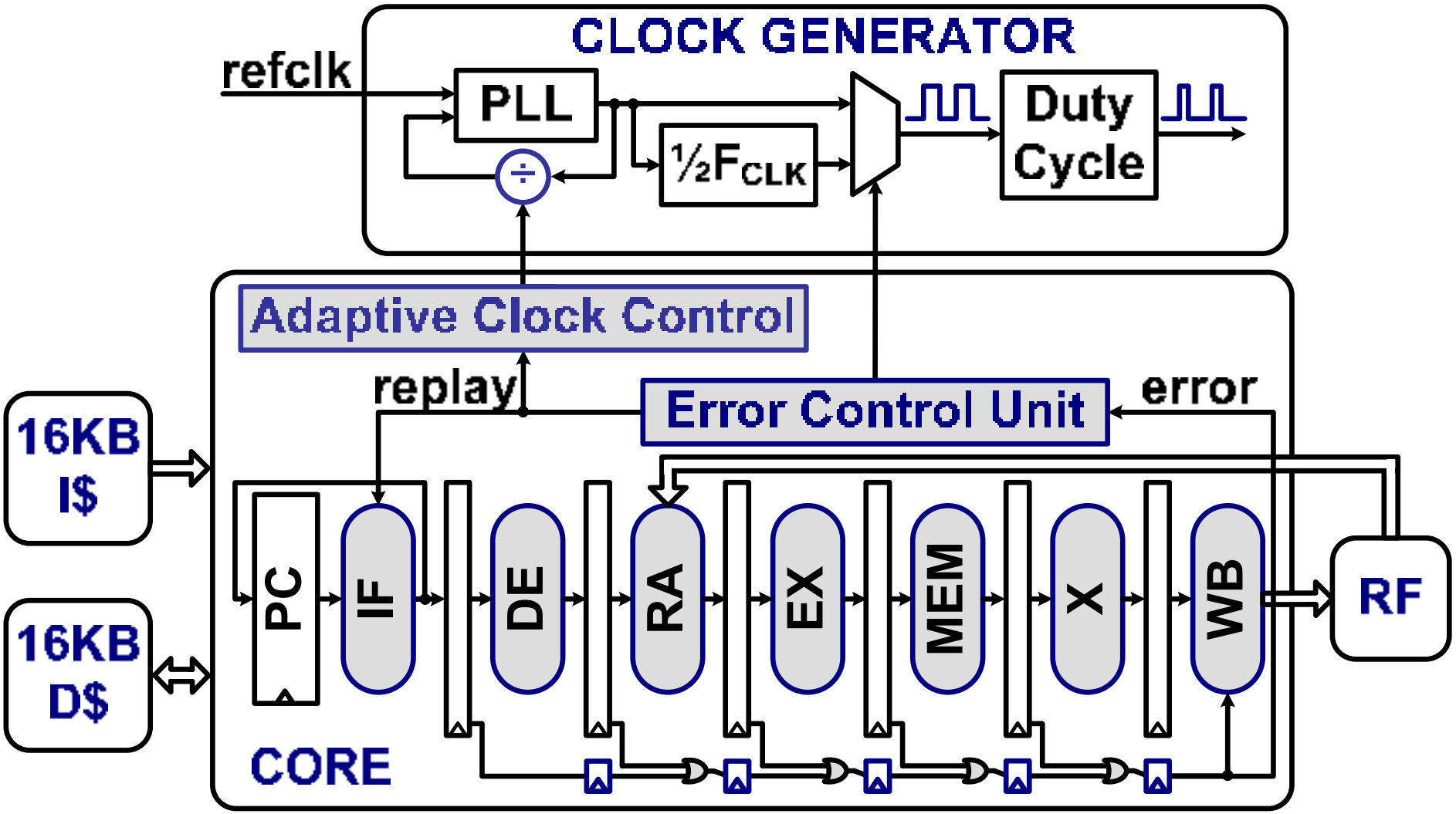
Error-Detection Sequential (EDS)

Implementation



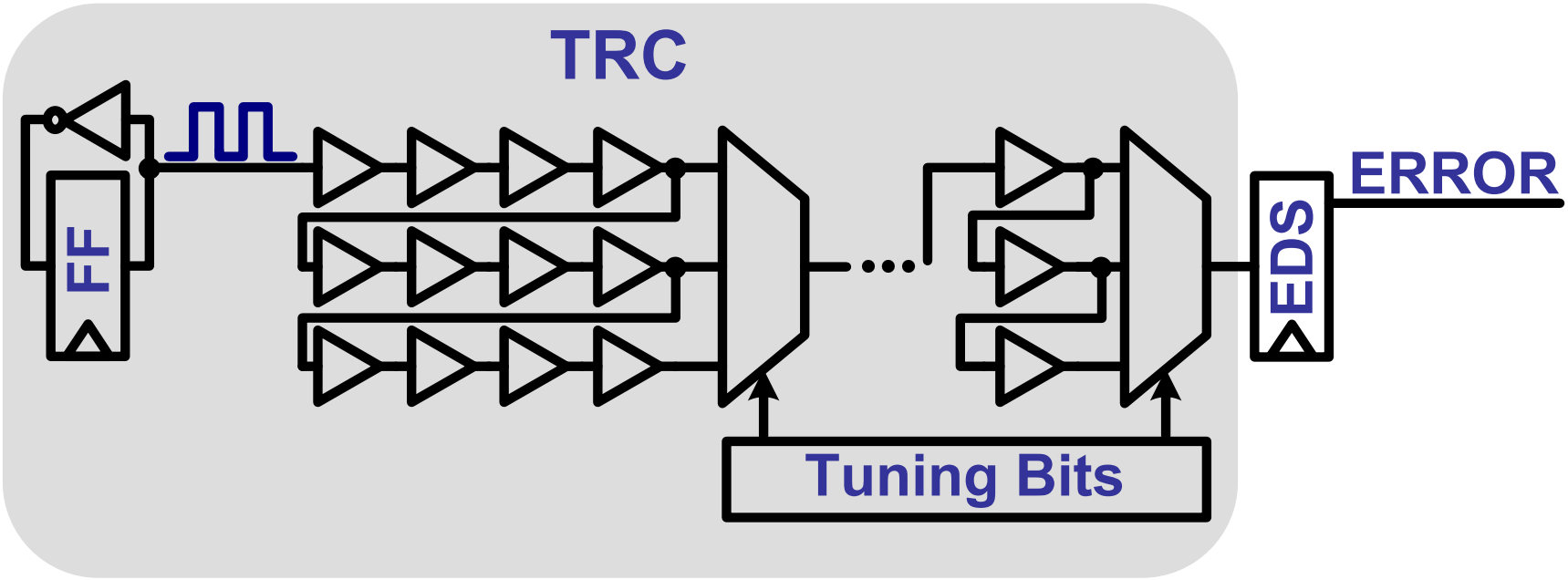
- Contains additional scan-enabled latch for testing
  - mode=0: EDS
  - mode=1: FF

Microprocessor Core Overview



- Adaptive clock control enables dynamic F_{CLK} change

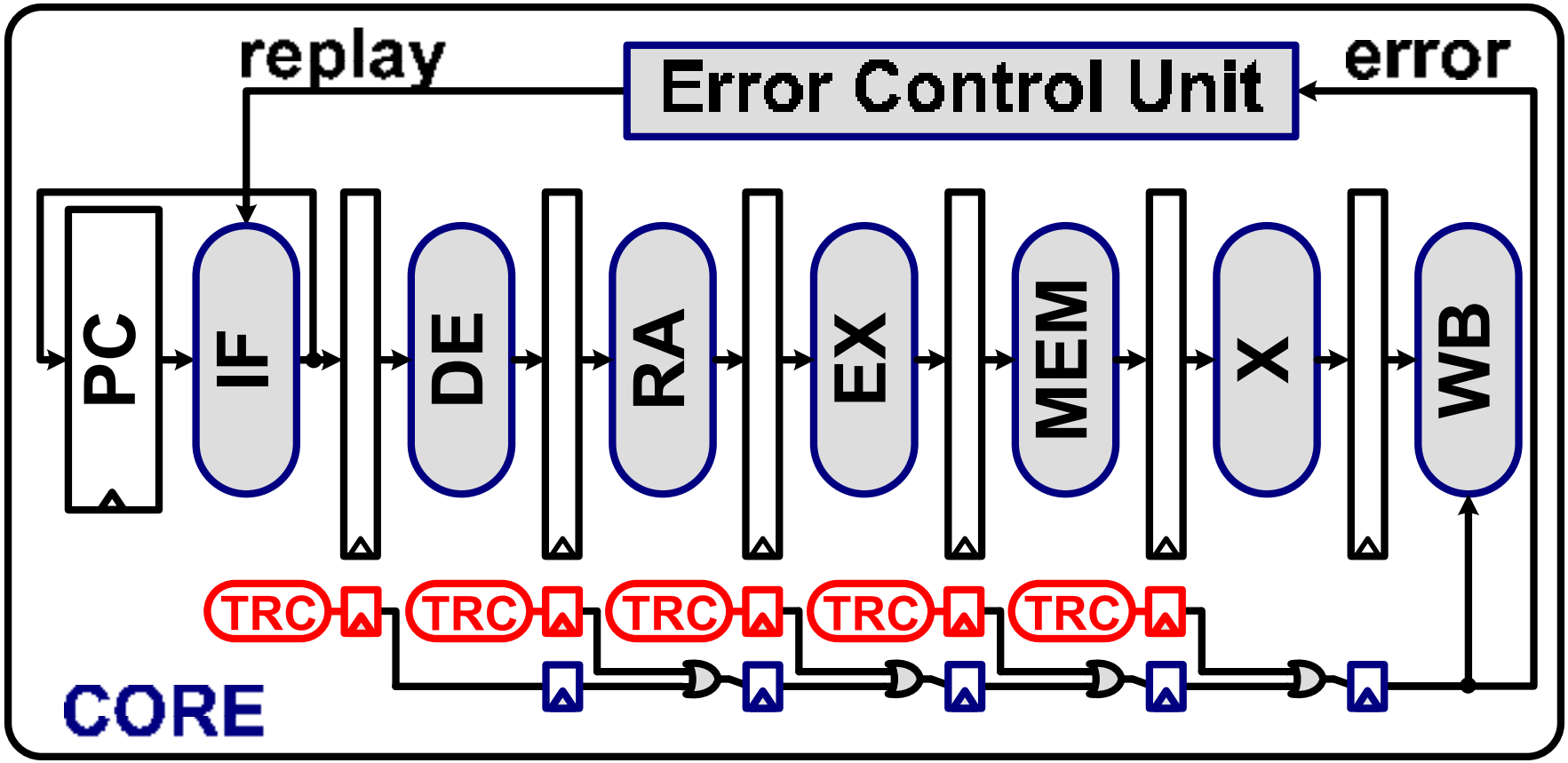
Tunable Replica Circuit (TRC)



- TRC monitors critical path delays
- Non-intrusive design

J. Tschanz, et al., *Symp. VLSI Circuits*, 2009.

Tunable Replica Circuit (TRC)



- TRC tuned to track critical paths per pipeline stage
- TRC must always fail if any critical path fails
- TRC error initiates pipeline error recovery

EDS & TRC Overheads

Circuit Blocks	EDS	TRC
Error Detection & Accumulation Area Overhead	2.2%	0.8%
ECU & Clock Control Area Overhead	1.4%	1.4%
Min-Delay Buffer Insertion Area Overhead	0.2%	–
Total Area Overhead	3.8%	2.2%
Total Power Overhead (iso-F_{CLK}, iso-V_{CC})	0.9%	0.6%

Error-Recovery Circuits

1) Instruction Replay at ½F_{CLK}

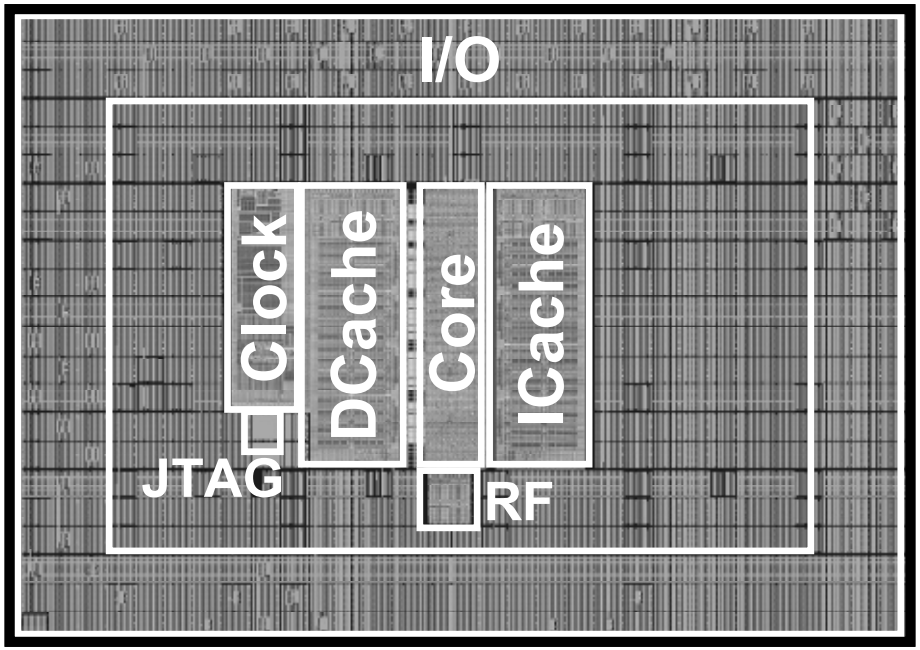
- Clock divider generates ½F_{CLK} without PLL re-lock
- Clock high-phase delay remains unchanged

2) Multiple Issue Instruction Replay at F_{CLK}

- Does not require clock control
- Issue ***replica instructions*** to setup pipeline registers
- Last issue is a valid instruction

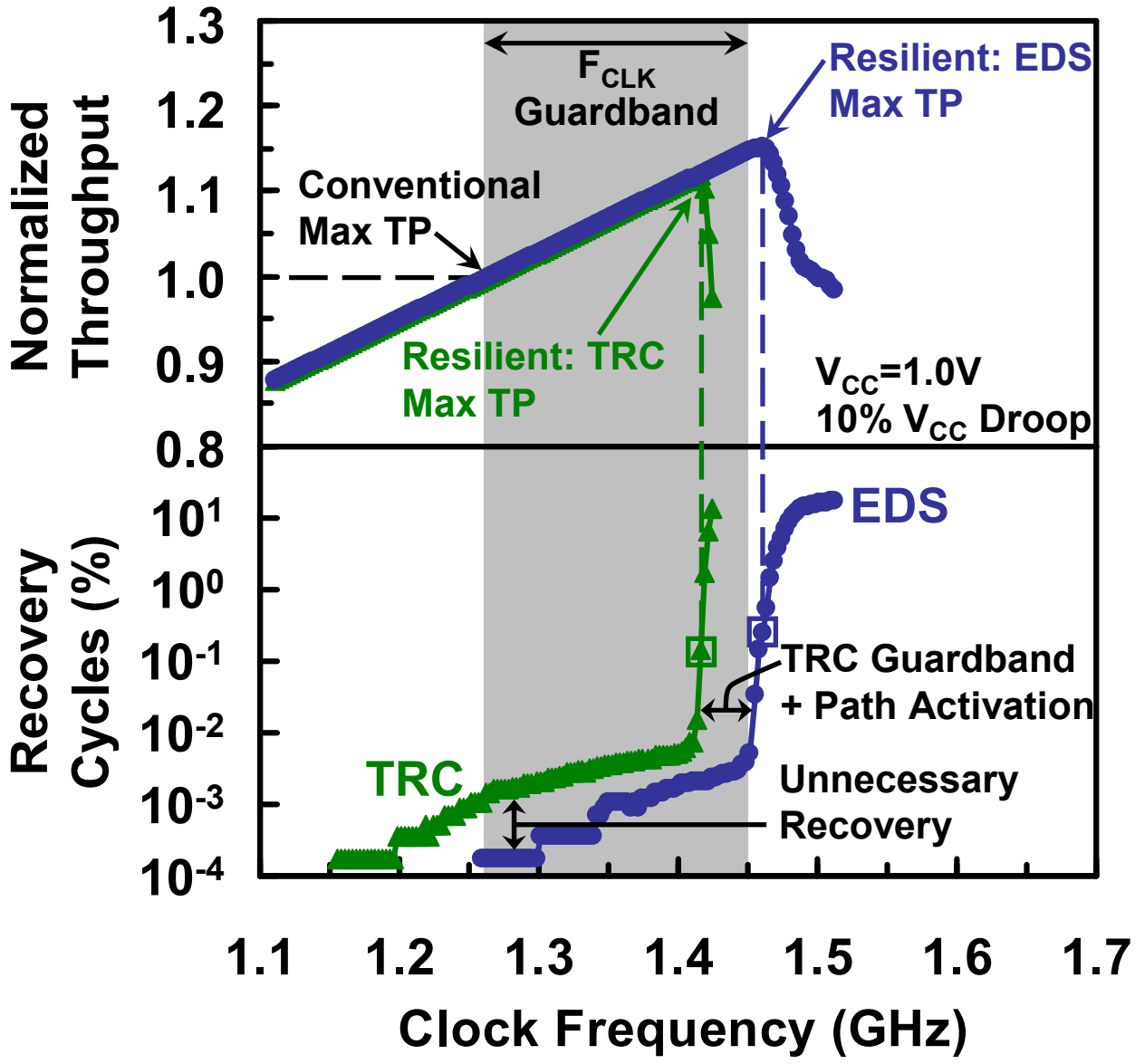
Characteristics & Measurements

Technology	45nm CMOS
Die Area	13.64 mm^2
Core Area	0.39 mm^2
Core F_{MAX}	1.45GHz at 1.0V
Core Power	135mW at 1.0V



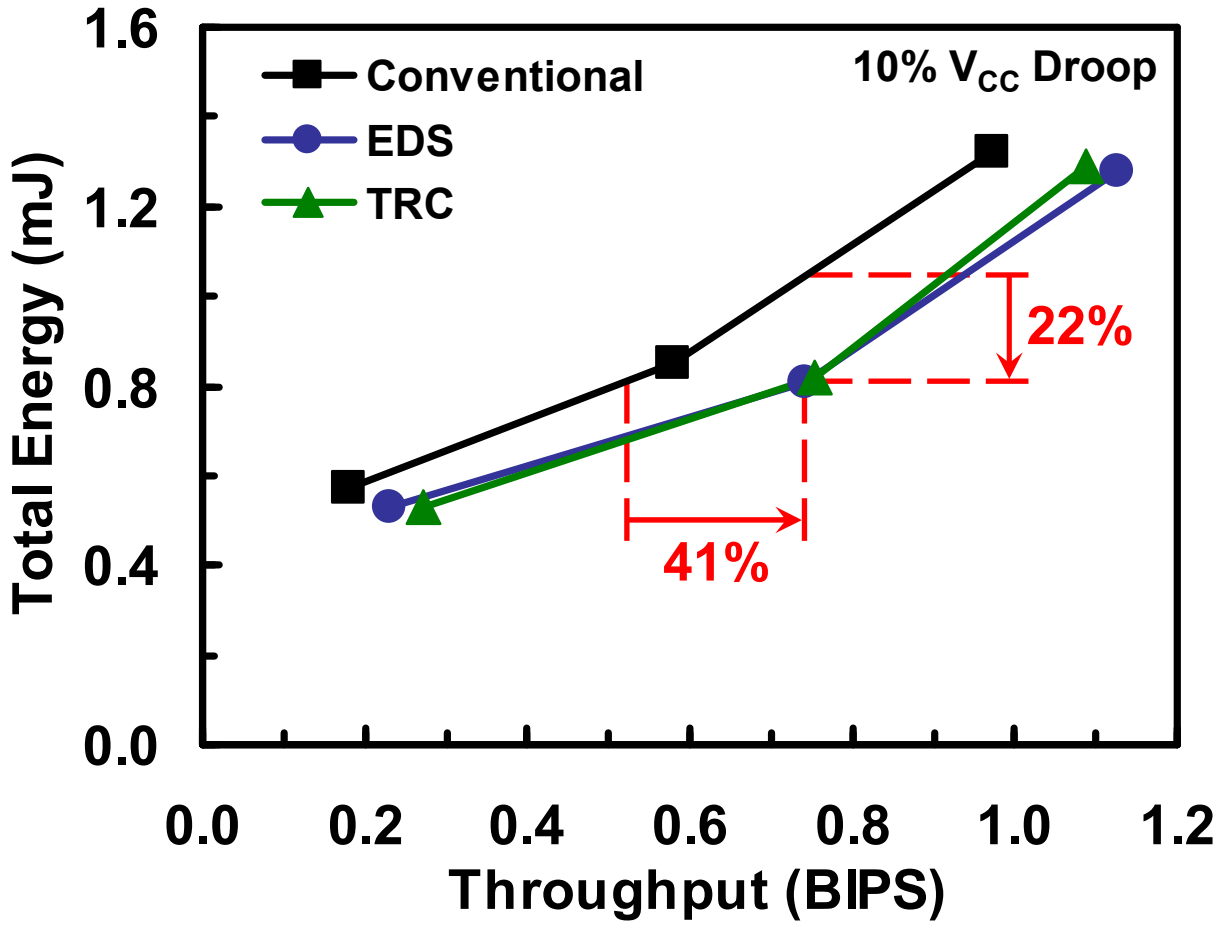
- Programs compiled from C code
- Caches and settings loaded via JTAG scan

Measured Throughput (TP) vs F_{CLK}



- 16% TP gain with EDS
- 12% TP gain with TRC

Measured Energy vs Throughput



- TRC & EDS resilient circuits enable:
  - 41% throughput gain at equal energy
  - 22% energy reduction at equal throughput

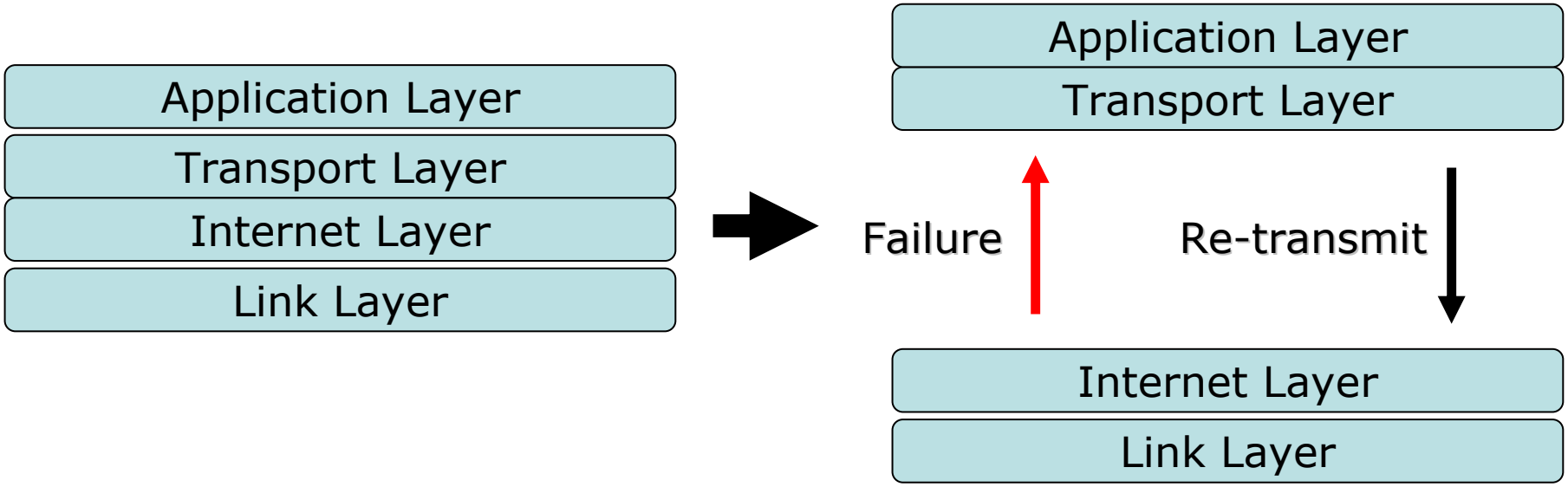
Summary of This Example

- Simple microprocessor core employs resiliency to mitigate dynamic variation guardbands
- Error-detection circuits:
  - Error-detection sequential (EDS)
  - Tunable replica circuit (TRC)
- Error-recovery circuits:
  - Instruction replay at ½F_{CLK}
  - Multiple issue instruction replay at F_{CLK}
- Silicon measurements indicate:
  - 41% throughput gain at iso-energy
  - 22% energy reduction at iso-throughput
- Resilient & adaptive circuits enable the microprocessor to adjust to operating variations for maximum efficiency

Networking Approach

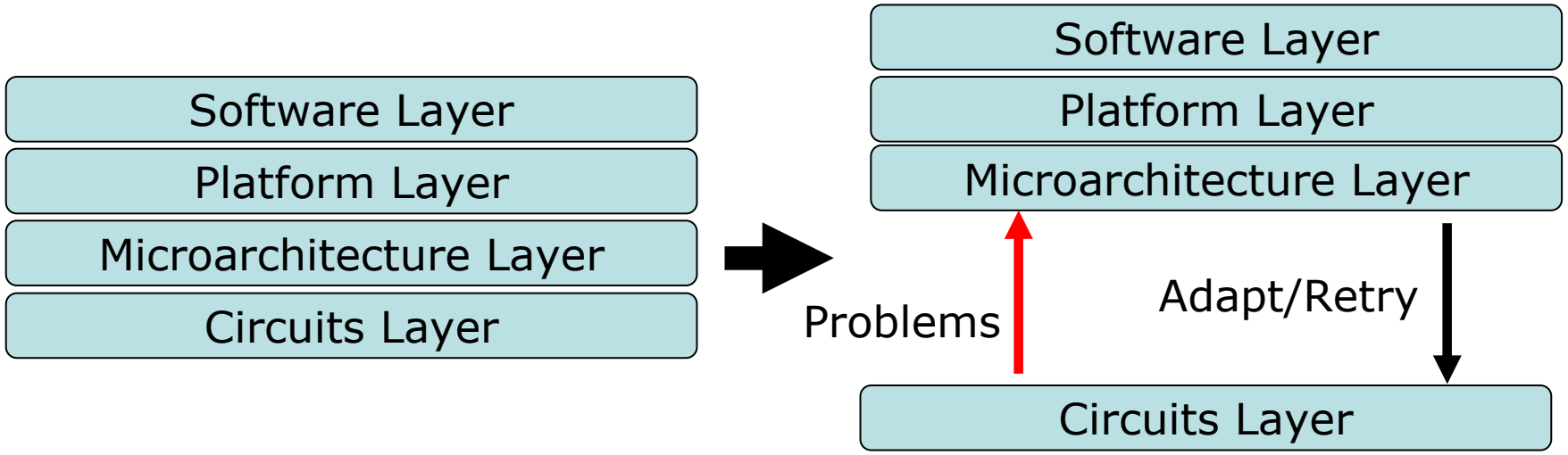
- In networking failures at each layer may be dealt with within the layer or passed to layer above.

Example: Internet Protocol



Unified Adaptive Design Framework

- Adaptive design proposes to handle failures in each layer by reporting failures to the next layer which delivers a response.



Global Optimization By Reconfiguring at the Appropriate Layer

Conclusion

- Resiliency as part of the optimization equation for performance/energy
- Memory is easier
- Logic is much harder
  - We addressed a solution for timing faults
  - Other transient faults?
  - Permanent fault?
  - Reconfigurable logic helpful?
- Cross-layer resiliency