Soft Error Rate Reduction Using Redundancy Addition and Removal

Kai-Chiang Wu and Diana Marculescu

ECE Department

Carnegie Mellon University

Electrical & Computer ENGINEERING

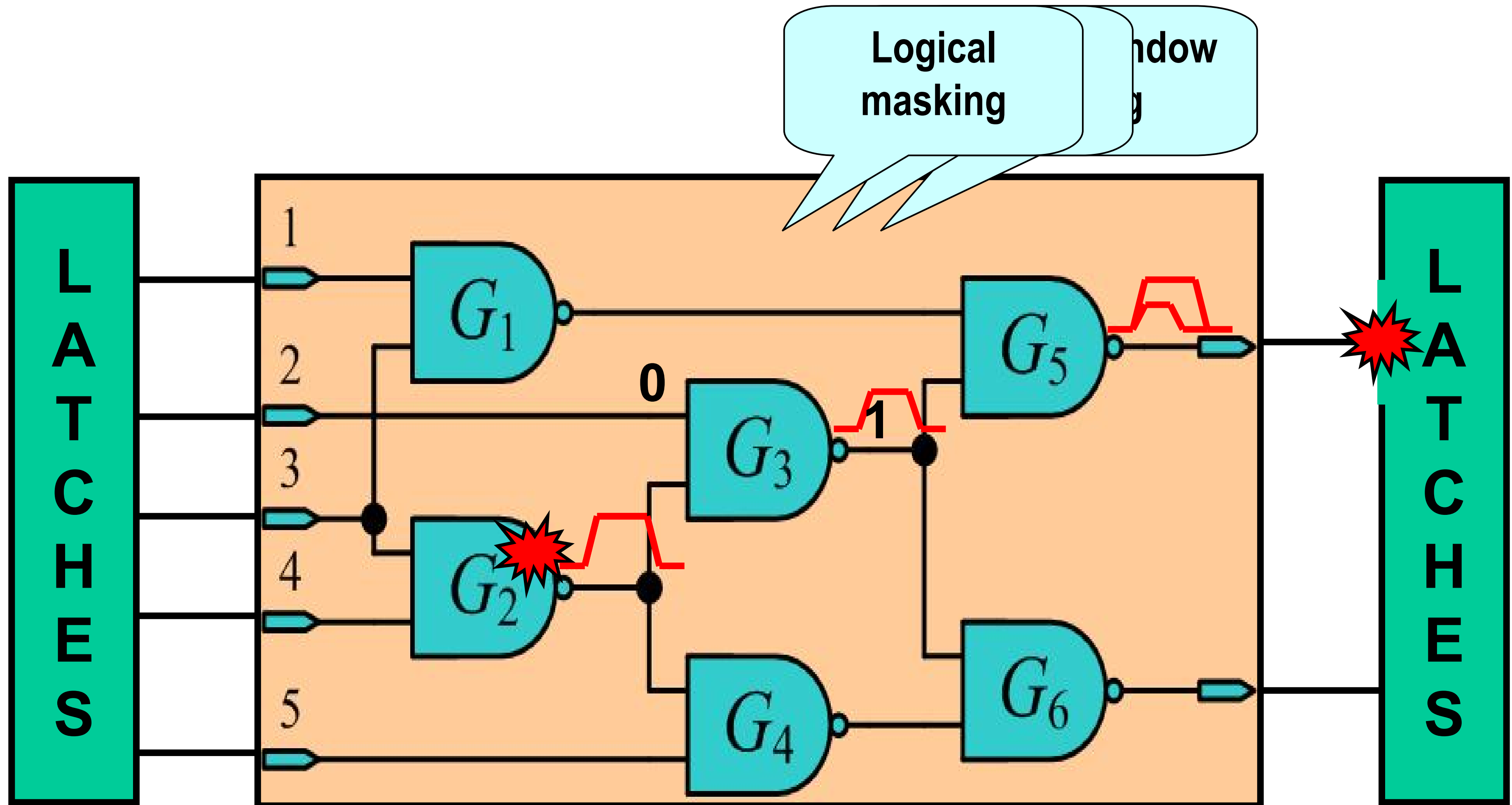
Soft Errors

- A soft error occurs when
  - a radiation-induced transient event causes a charge disturbance that flips the state of a storage element.
  - Such a bit-flip is called a *single-event transient* (**SET**) or a glitch.

- Definitions
  - A soft error is often referred to as a *single-event upset* (**SEU**).
  - The rate at which soft errors occur is called *soft error rate* (**SER**).

- Current technology scaling trends – shrinking feature sizes, etc.
  - Circuits become more susceptible to radiation-induced particle hits.
  - Particles with **less** energy could flip the states of storage elements.

Soft Errors in Combinational Logic



Scaling Trends of Masking Factors

- Logical masking is decreased due to
  - decreasing logic depth
- Electrical masking is decreased due to
  - faster logic gates
  - lower supply voltages
  - smaller node capacitances
- Latching-window masking is decreased due to
  - increasing clock frequencies

Outline

- Background and motivation
- **Related work**
- Proposed framework
  - Redundancy addition and removal (RAR)
  - Metrics for gate characterization
  - Constraints on RAR
- Results and conclusion

Related Work

- Triple Modular Redundancy (**TMR**)
  - consists of three identical copies and a majority voter.
  - incurs more than 200% overhead in terms of area and power.
- Partial duplication [Mohanram *et al.*, 2003] / gate sizing [Zhou *et al.*, 2004]
  - targets gates with high error impact.
  - incurs potentially large area overhead.
- Flip-flop selection [Joshi *et al.*, 2006]
  - increases the length of latching windows.
  - focuses only on latching-window masking.

Proposed Framework

- Using *Redundancy Addition and Removal* (**RAR**)
  - Iteratively add and remove redundant wires to minimize a circuit in terms of **literal count**.

- RAR for SER reduction
  - Estimate the effects of redundancy manipulations.
  - Accept only those with **positive impact** on SER.

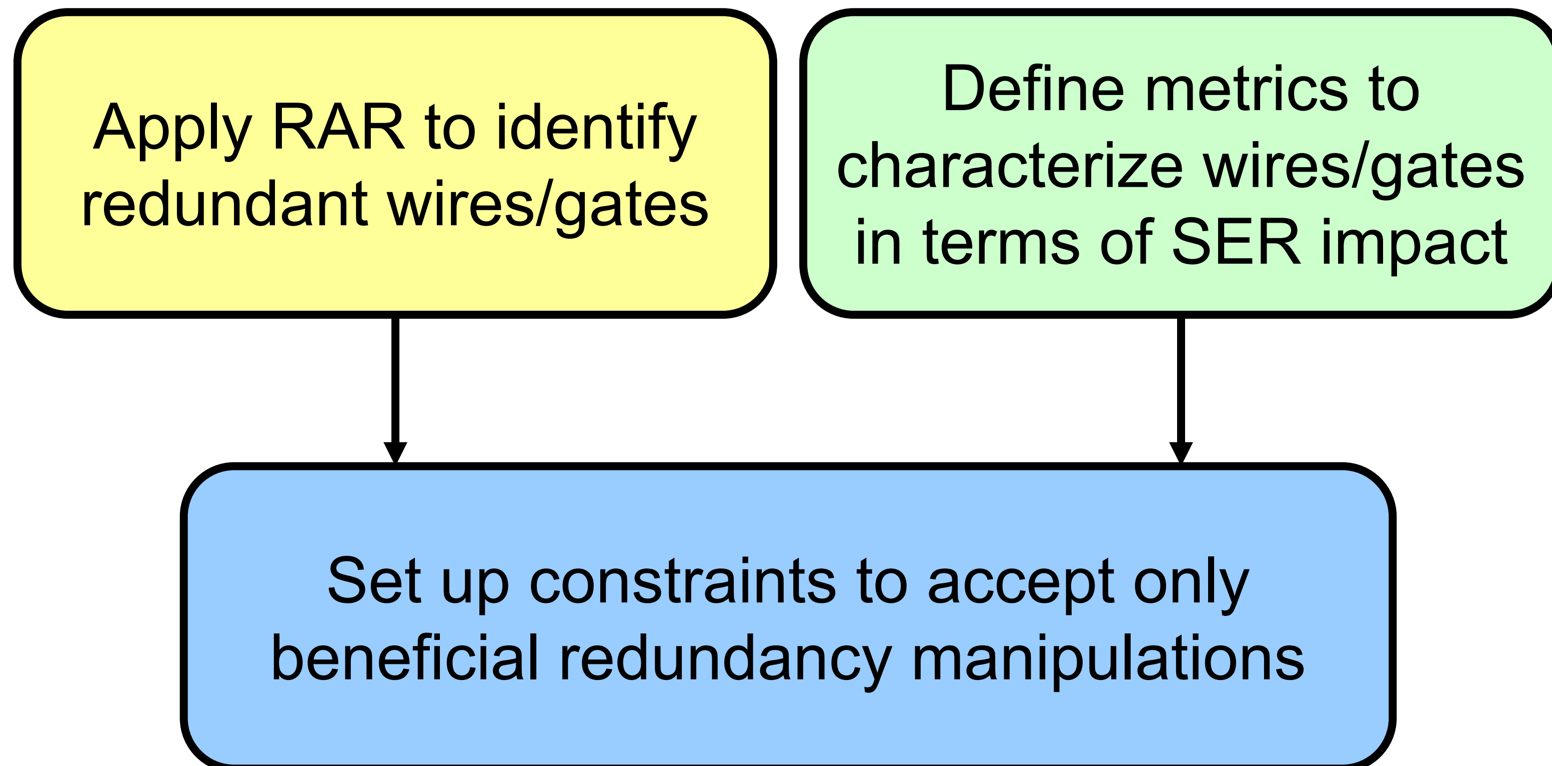
- Advantages over other techniques
  - Very little area overhead
  - **Unified treatment** of three masking factors via decision diagrams
  - **Precise estimation** of SER impact of added and removed wires

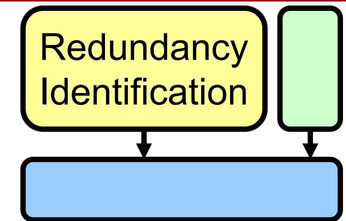
RAR: Combinational and Sequential Logic Optimization by Redundancy Addition and Removal [Entrena and Cheng, 1995]

Outline

- Background and motivation
- Related work
- **Proposed framework**
  - **Redundancy addition and removal (RAR)**
  - **Metrics for gate characterization**
  - **Constraints on RAR**
- Results and conclusion

Framework Overview

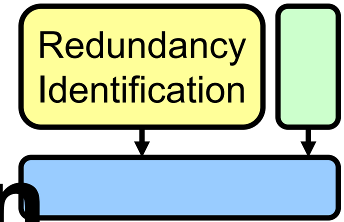




RAR – An Introduction

[Entrena and Cheng, 1995]

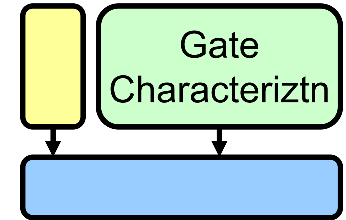
- Add a redundant wire found by mandatory assignments during *automatic test pattern generation* (**ATPG**).
  - The newly added wire could cause one or more originally irredundant wires/gates to become redundant (removable).
- Remove those redundant wires due to the added wire.
  - Delete gates with only one fanin and gates without any fanout.
- Repeat until no further improvement can be done.
  - The circuit will become **smaller** if the removed redundancies are more than the added redundancies.



RAR – For SER Reduction

- Apply RAR for SER reduction with little area penalty.
- **Unsystematic** RAR may increase SER by reducing the number of gates or the depth of circuits.
- Solution
  - Use **MARS-C**, based on BDDs and ADDs, to quantify the *error impact* and the *masking impact* of each gate.
  - Keep wires/gates with **higher** masking impact.
  - Remove wires/gates with **higher** error impact.

MARS-C: Modeling and Reduction of Soft Errors in Combinational Circuits [Miskov and Marculescu, 2006]



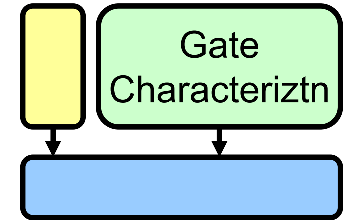
Mean Error Impact

Mean error impact (**MEI**) of each internal gate G_i:

$$\mathrm{MEI}(G_i^{d,a}) = \frac{\sum_{k=1}^{n_f} \sum_{j=1}^{n_F} \mathrm{P}(F_j \ fails | G_i \ fails \cap init_glitch = (d,a))}{n_F \cdot n_f}$$

- MEI quantifies the probability that at least one primary output is affected by a glitch originating at the gate.
- The **larger** MEI a gate has, the **higher** the probability that a glitch occurring at this gate will be latched.

Mean Masking Impact



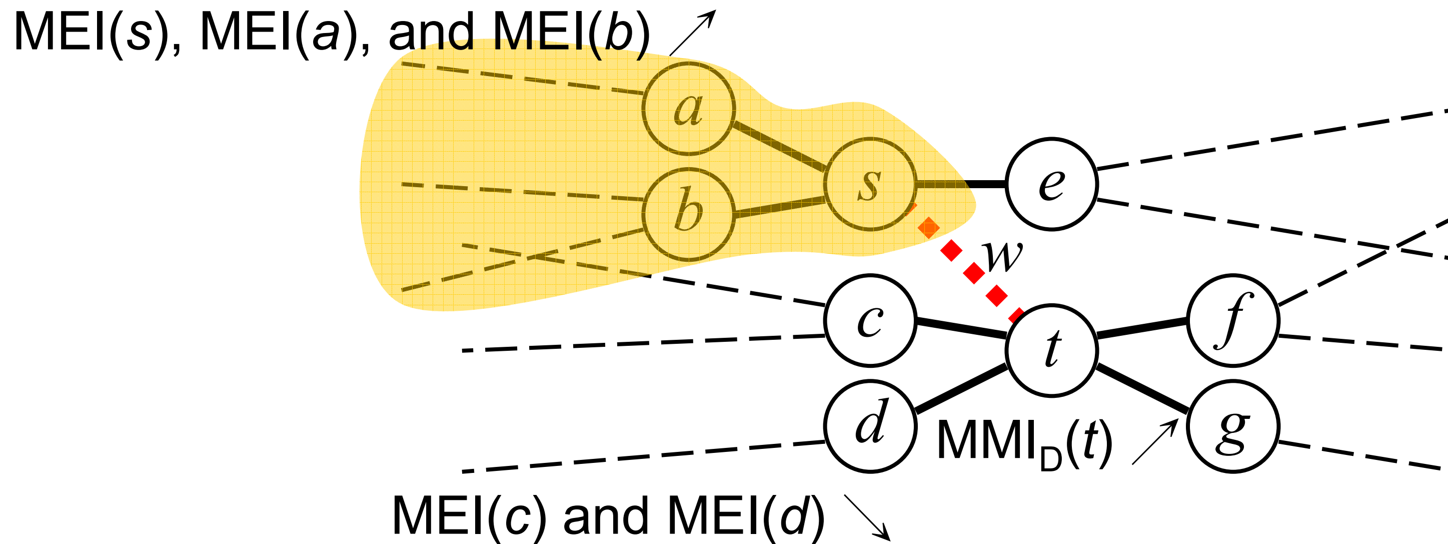
Mean masking impact on duration (**MMI$_D$**) of each internal gate G_i:

$$\mathrm{MMI}_{\mathrm{D}}(G_i^{d,a}) = \frac{\sum_{k=1}^{n_f}\sum_{j=1}^{n_G}\mathrm{MI}_{\mathrm{D}}(G_j^{d,a} \rightarrow G_i)}{n_G \cdot n_f \cdot d}$$

- MMI$_D$ denotes the normalized expected attenuation on the duration of all glitches passing through it.
- The **larger** MMI$_D$ a gate has, the **more** capable of masking glitches this gate is.

Redundancy Manipulation Checking

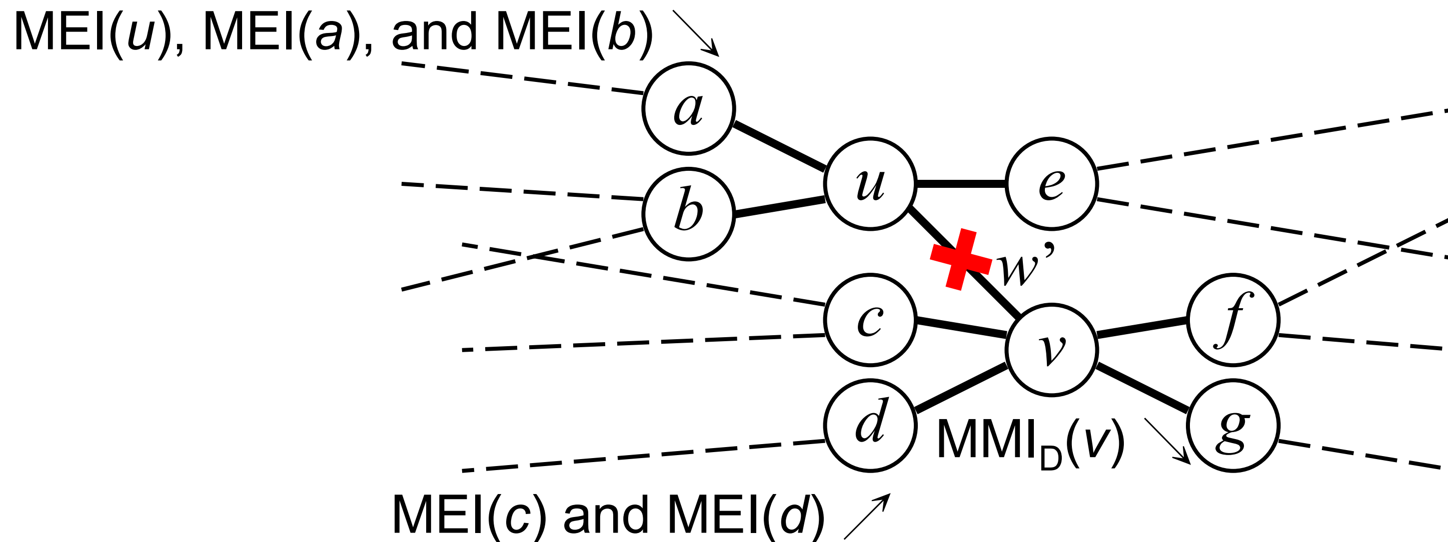
Wire Addition Constraint



- Δ MEI(s) = MEI(t) $\times$ [1-MMI$_D$(t)]
- Wire w ($s \rightarrow t$) can be added only if
  - MEI(t) < T_1
  - MMI$_D$(t) > T_2

Redundancy Manipulation Checking

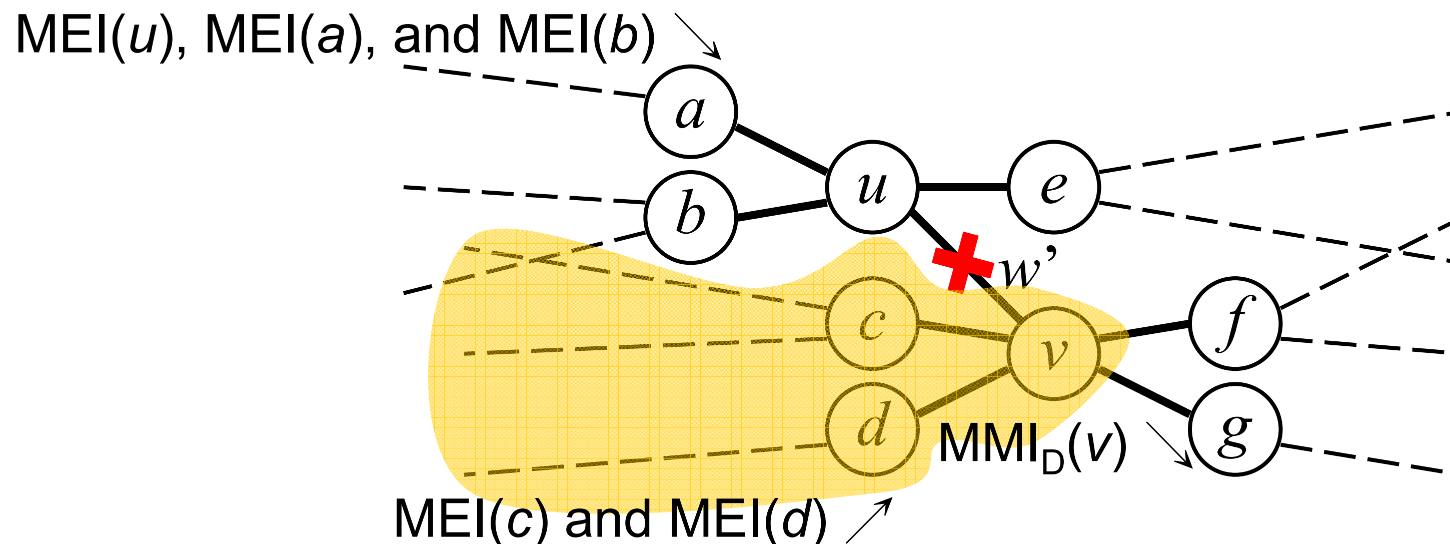
Wire Removal Constraint 1



- Δ MEI(u) = MEI(v) $\times$ [1-MMI$_D$(v)]
- Wire w' ($u \rightarrow v$) can be removed only if
  - MEI(v) > $T_3 \geqq T_1$
  - MMI$_D$(v) < $T_4 \leqq T_2$

Redundancy Manipulation Checking

Wire Removal Constraint 2



- Wire *w*' (*u* → *v*) can be removed only if
  - Wire *w*' is crucial in logical masking at gate *v*.
  - The probability that gate *u* goes to the controlling value of gate *v* is sufficiently low.

Outline

- Background and motivation
- Related work
- Proposed framework
  - Redundancy addition and removal (RAR)
  - Metrics for gate characterization
  - Constraints on RAR
- **Results and conclusion**

Practical Considerations

Mean error susceptibility (**MES**) of each primary output F_j:

$$\mathrm{MES}(F_j^{d,a}) = \frac{\sum_{k=1}^{n_f} \sum_{i=1}^{n_G} \mathrm{P}(F_j\ fails \mid G_i\ fails \cap init_glitch = (d,a))}{n_G \cdot n_f}$$

Output failure probability of each primary output F_j:

$$\mathrm{P}(F_j) = \frac{\Delta d \cdot \Delta a}{(d_{\max} - d_{\min}) \cdot (a_{\max} - a_{\min})} \sum_n \sum_m \mathrm{MES}(F_j^{d_m, a_n})$$

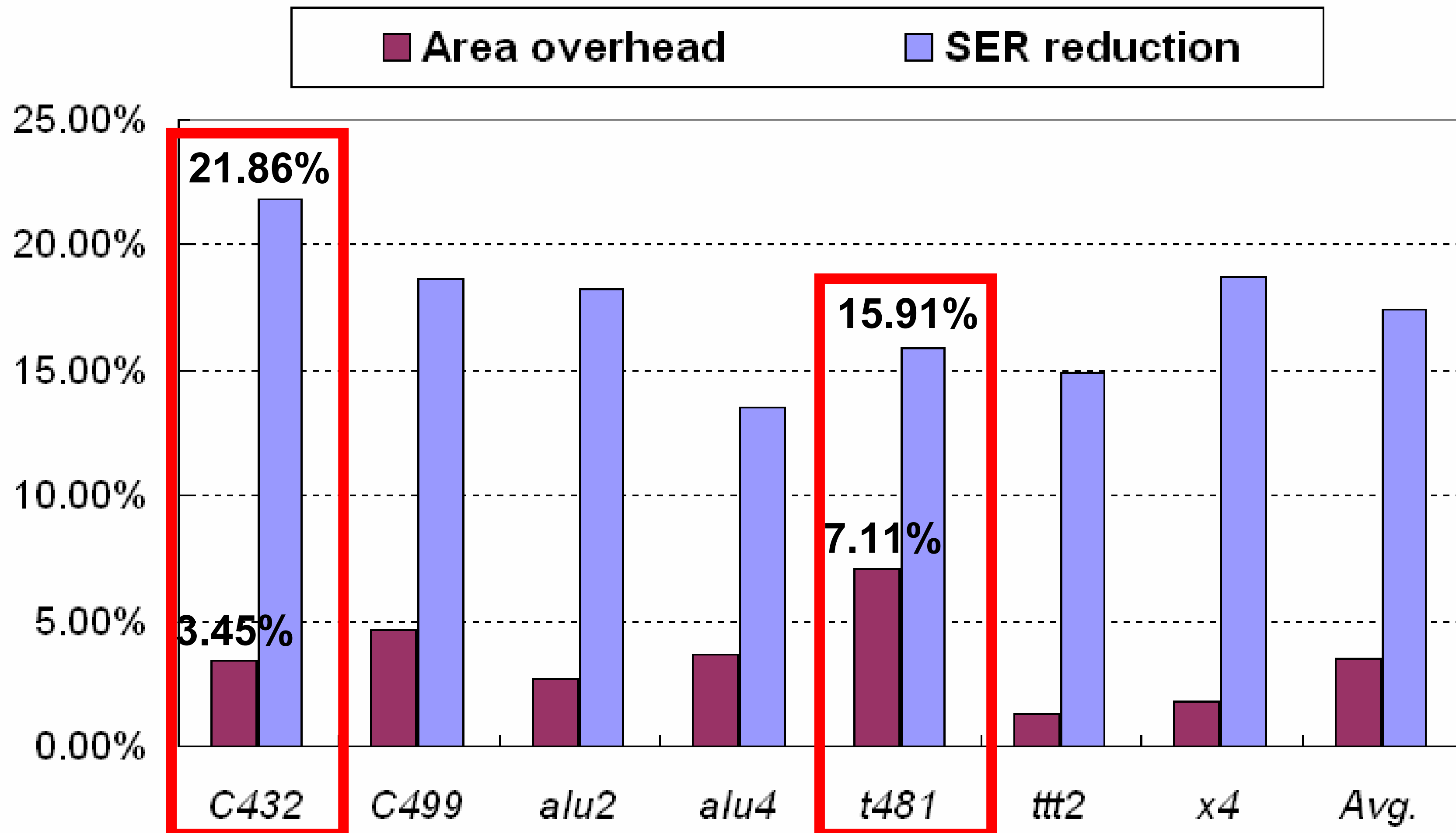
Soft error rate of each primary output F_j:

$$\mathrm{SER}(F_j) = \mathrm{P}(F_j) \cdot \mathrm{R}_{\mathrm{PH}} \cdot \mathrm{R}_{\mathrm{EFF}} \cdot \mathrm{A}_{\mathrm{CIRCUIT}}$$

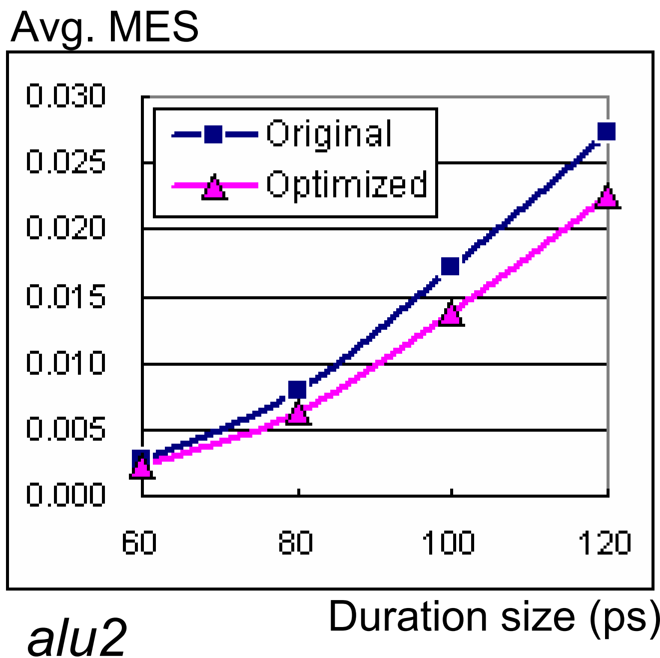
Experimental Setup

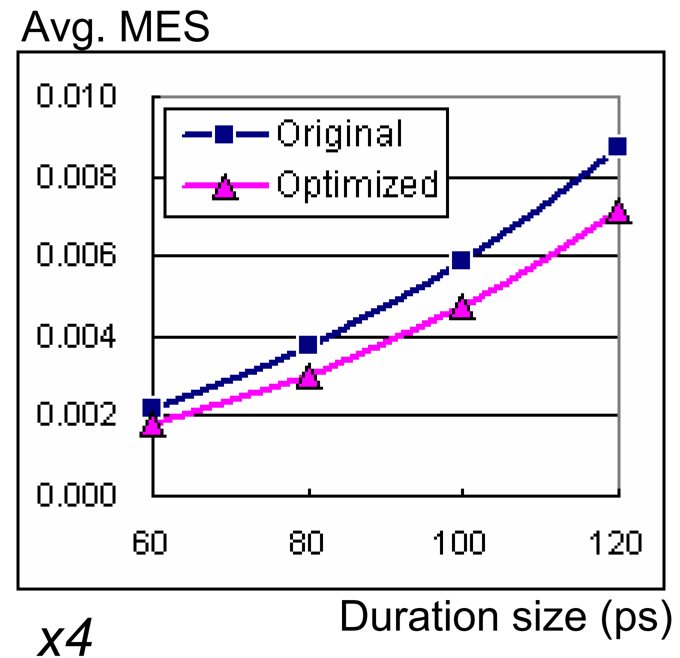
- Technology: 70nm, BPTM
- Clock period: 250ps
- Setup time/hold time: 10/10ps
- Supply voltage: 1.0V
- $(d_{min}, d_{max})/(a_{min}, a_{max})$: (60, 120)ps/(0.8, 1.0)V
- $\Delta d/\Delta a$: 20ps/0.1V
- R_{PH}: 56.5 $m^{-2}s^{-1}$
- R_{EFF}: 2.2×10^{-5}

Experimental Results



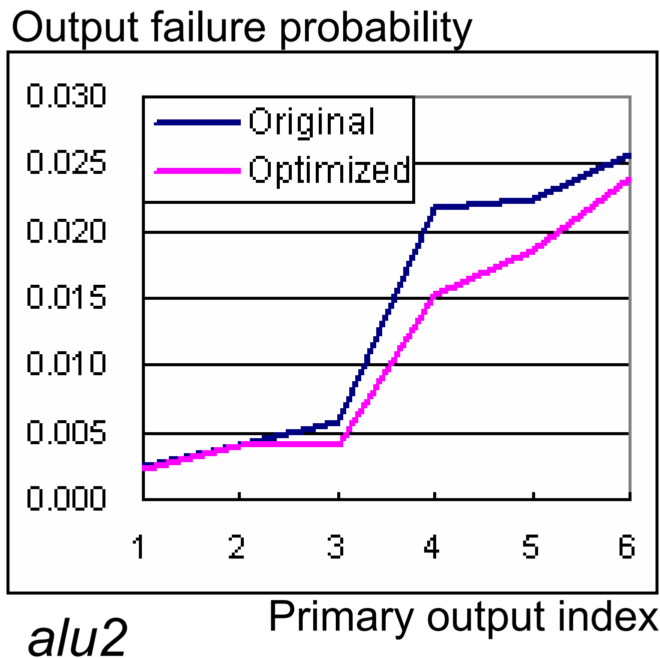
Average Mean Error Susceptibility



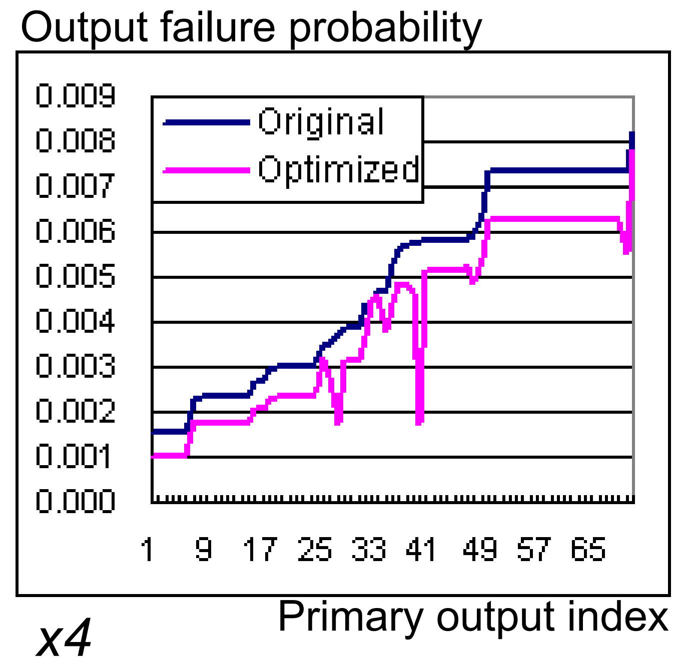


16-20% reduction in average mean error susceptibility

Output Failure Probability



alu2



x4

30-70% maximum reduction in output failure probability

Conclusion

- We propose a SER reduction framework
  - based on redundancy addition and removal (**RAR**)
  - using symbolic SER analysis (**MARS-C**)
  - for combinational logic
- Two metrics and three constraints are introduced to guide this framework towards SER reduction.
- Experiments on a subset of standard benchmarks reveal the effectiveness of our framework.

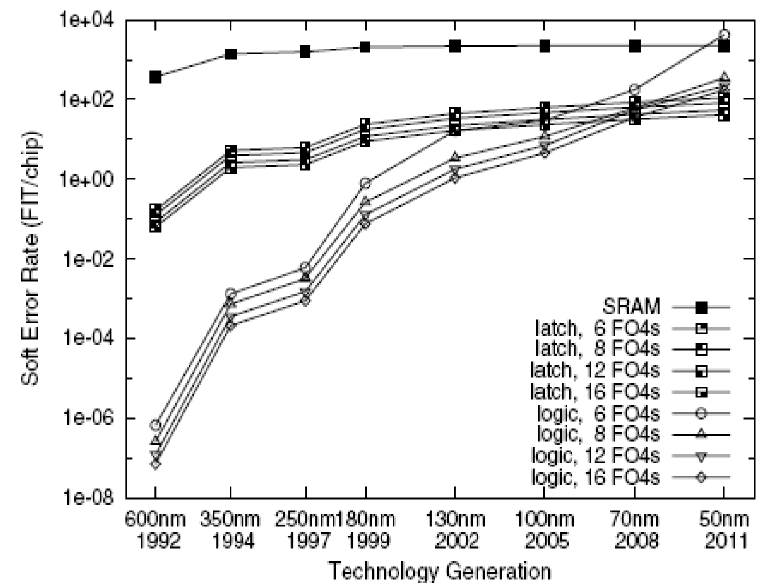
Thank you!

Backup Slides

Soft Errors – A New Great Concern in Logic Circuits

- Soft errors would significantly degrade the robustness of logic circuits, while the nominal SER of SRAMs tends to be nearly constant from the **130nm** to **65nm** technologies.
  - Source: Mitra *et al.*, “Robust System Design with Built-in Soft-Error Resilience,” *IEEE Computer Magazine*, Feb. 2005

- The SER of combinational logic is predicted to be comparable to that of memory elements by **2011**.
  - Source: Shivakumar *et al.*, “Modeling the Effect of Technology Trends on the Soft Error Rate of Combinational Logic,” *Proc. Int’l Conference on Dependable Systems and Networks*, Jun. 2002

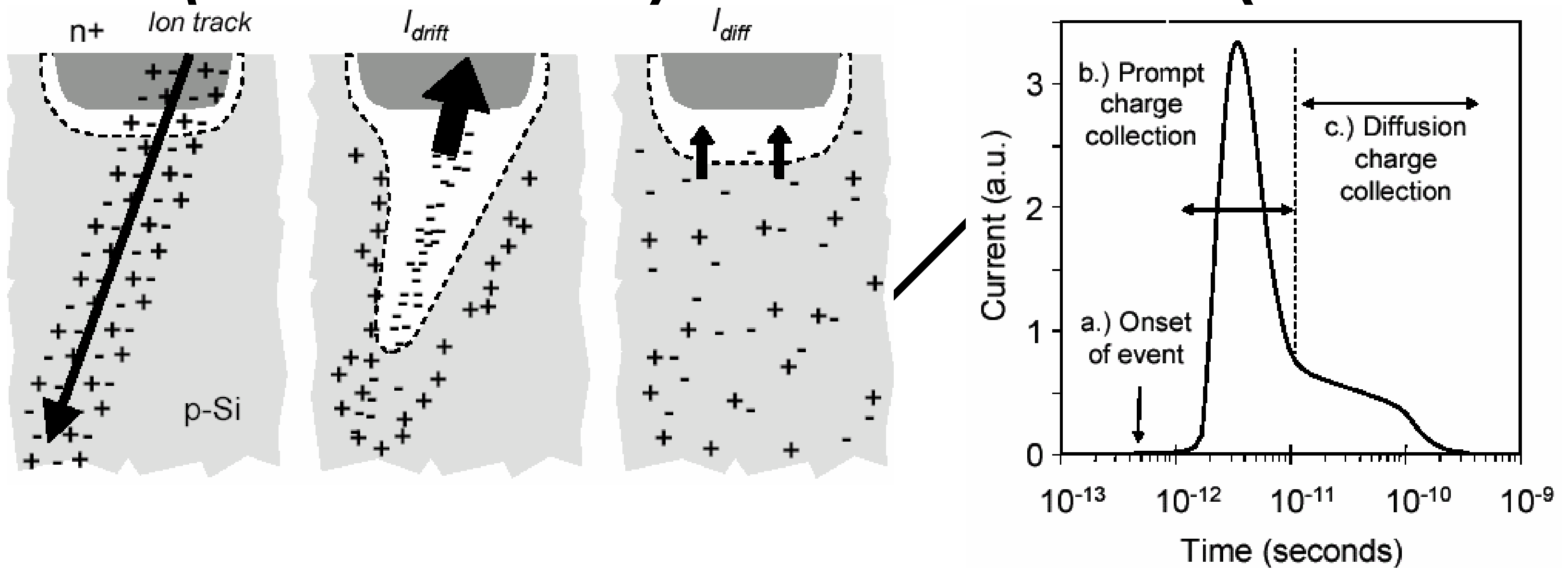


Soft Error Generation

Radioactive decay of ^{232}Th, ^{235}U, and ^{238}U in packaging materials

Cosmic rays react with atmosphere

Cascades of secondary particles (neutrons, protons, etc.)



Soft Error Modeling

- $\mathcal{A}$ (Amplitude condition):
  $A > V_{th}$ (if the correct output is "0") or
  $A < V_{th}$ (if the correct output is "1")
- $\mathcal{D}$ (Duration condition):
  $D > t_{setup} + t_{hold}$
- $\mathcal{T}$ (Timing condition):
  $t \in [T_{clk} + t_{hold} - T - D,\ T_{clk} - t_{setup} - T]$

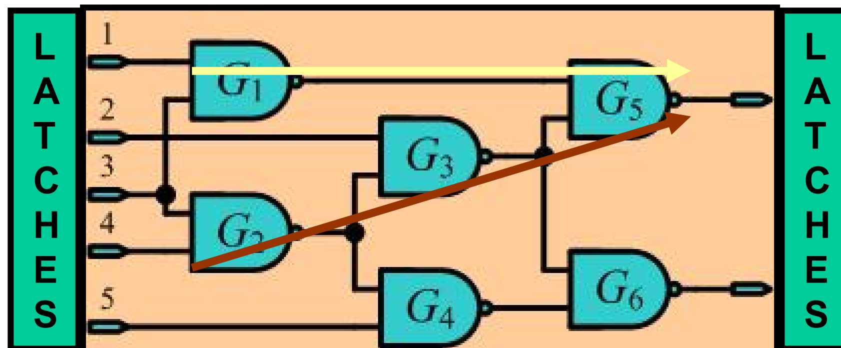
$d_{init} = t_2 - t_1$
$T = t_1' - t_1$
$D = t_2' - t_1'$
$t_1' \leqq T_{clk} - t_{setup}$
$t_2' \geqq T_{clk} + t_{hold}$

- $\mathrm{P}(\mathcal{A} \cap \mathcal{D} \cap \mathcal{T}) = \mathrm{P}(\mathcal{D} \cap \mathcal{T}) = \mathrm{P}(\mathcal{T} | \mathcal{D}) \cdot \mathrm{P}(\mathcal{D})$

$$= \sum_k \Big(\mathrm{P}(t \in [T_{clk} + t_{hold} - T - D, T_{clk} - t_{setup} - T] \,|\, D = D_k) \cdot \mathrm{P}(D = D_k) \Big)$$

$$= \sum_k \left(\frac{D_k - (t_{setup} + t_{hold})}{T_{clk} - d_{init}} \cdot \mathrm{P}(D = D_k) \right)$$

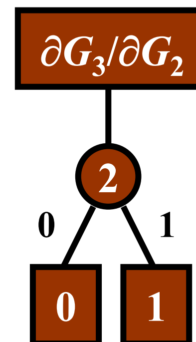
Sensitization BDDs

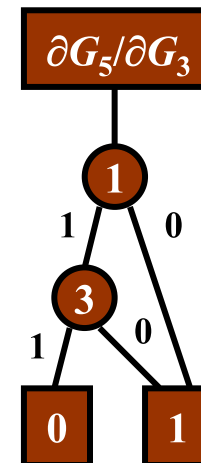


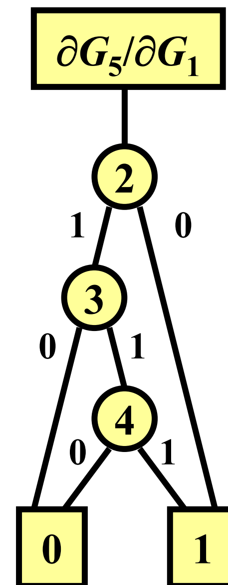
- Sensitization BDDs include information about logical masking.

- Sensitization BDD of $G_i \rightarrow G_j$ is Boolean difference of G_j w.r.t. G_i

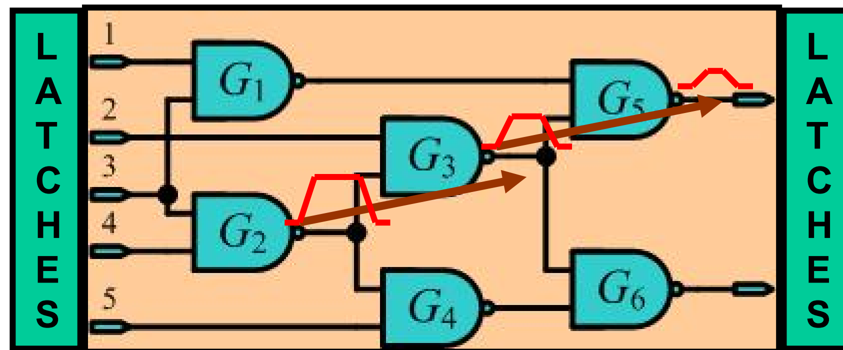
- $G_2 \rightarrow G_3$: Bool. diff. of G_3 w.r.t. G_2
- $G_3 \rightarrow G_5$: Bool. diff. of G_5 w.r.t. G_3
- $G_1 \rightarrow G_5$: Bool. diff. of G_5 w.r.t. G_1





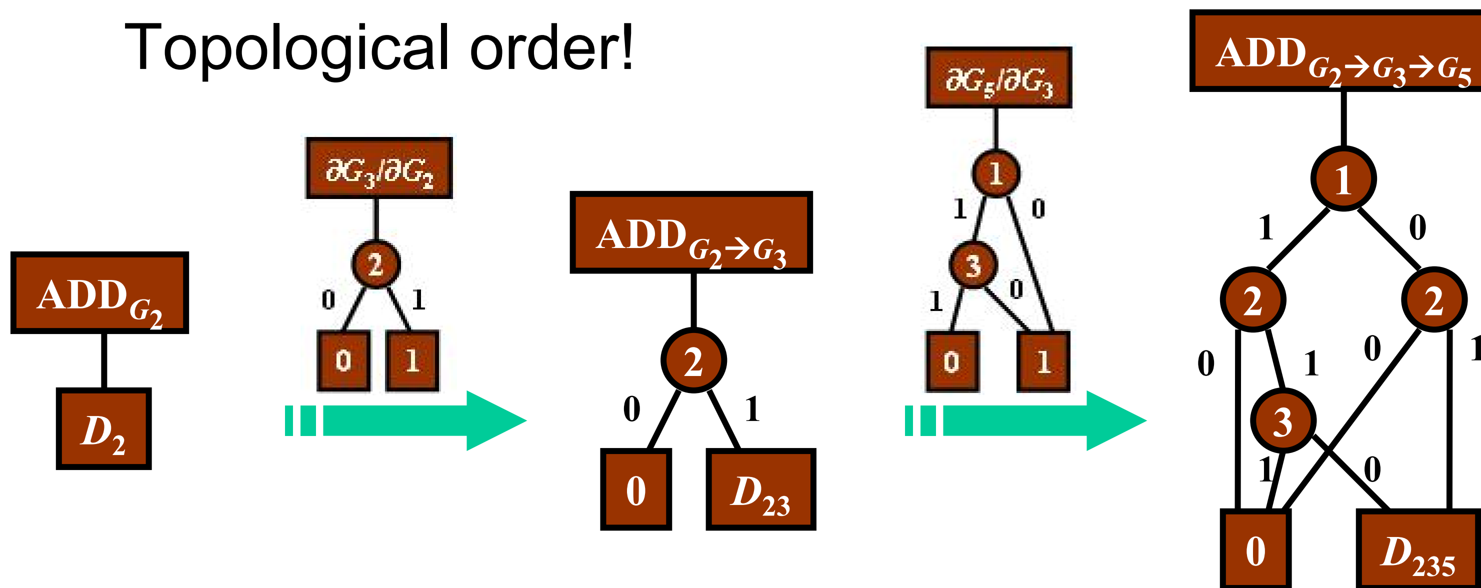


Duration ADDs

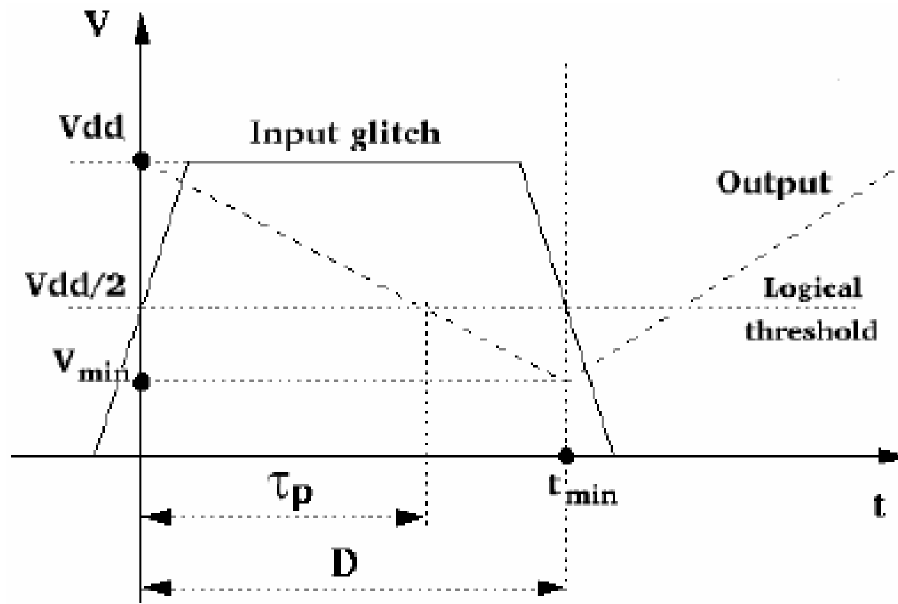


Topological order!

- Duration ADDs are created with respect to sensitization BDDs (logical masking) and attenuation model (electrical masking).



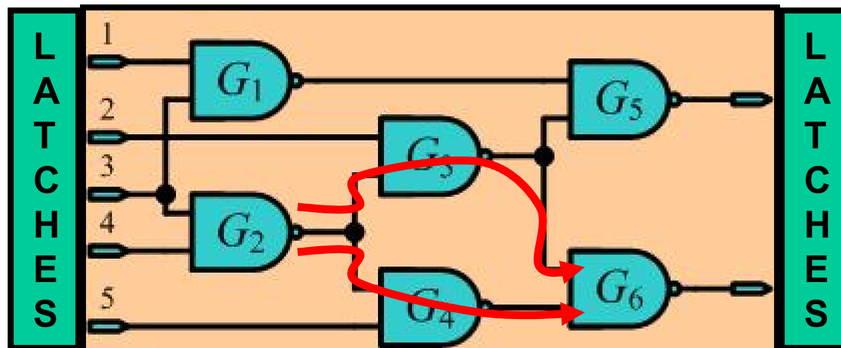
Attenuation Model

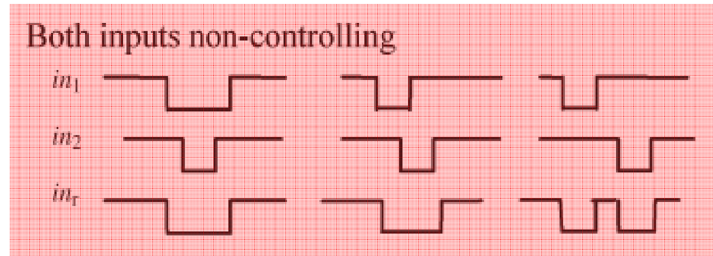


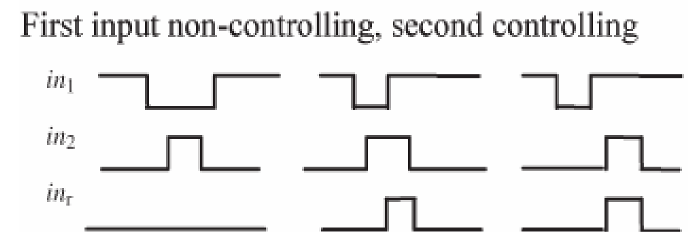
$$V_{min} = V_{dd} - \frac{V_{dd}}{2\tau_p} \cdot D \qquad \Rightarrow \begin{cases} \text{if } D \le \tau_p \text{, the glitch is masked.} \\ \text{if } \tau_p < D \le 2\tau_p \text{, the glitch is attenuated.} \\ \text{if } D > 2\tau_p \text{, the glitch remains the same.} \end{cases}$$

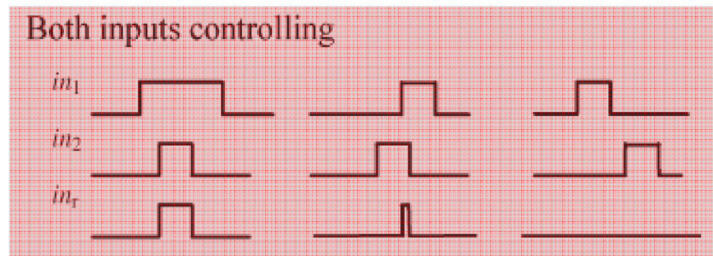
Reconvergent Glitches

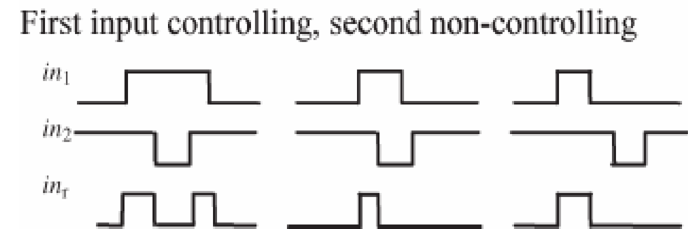
- Glitches on reconvergent paths arriving to inputs of a gate
  - can be merged into a new glitch.
  - can be masked by each other.



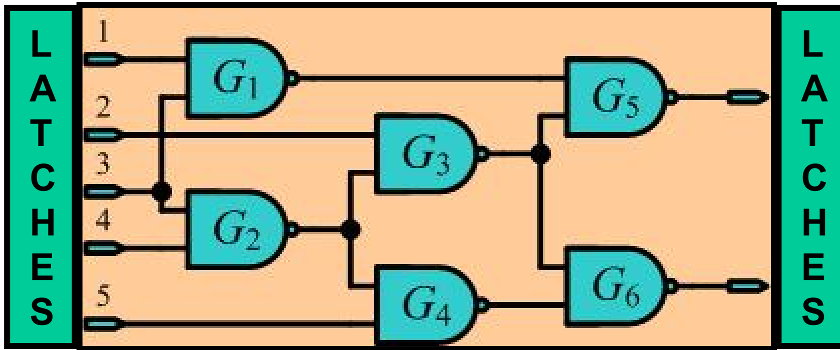


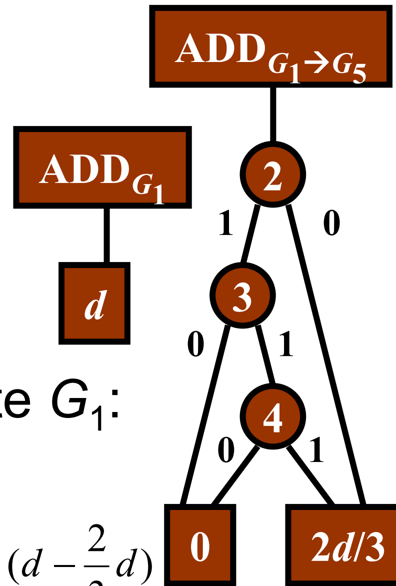


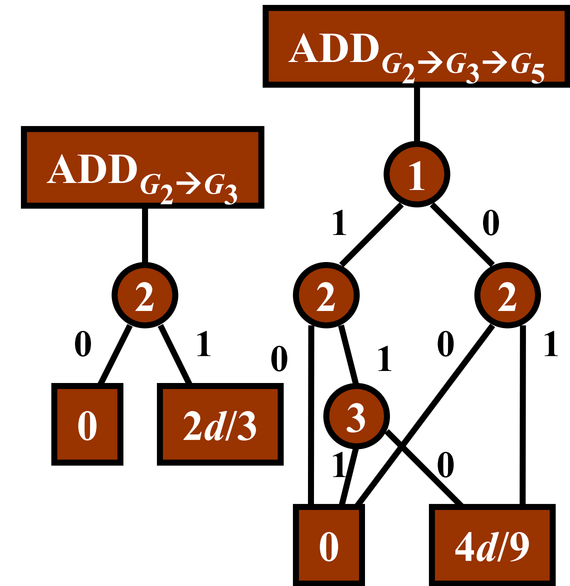




MMI Computation – An Example





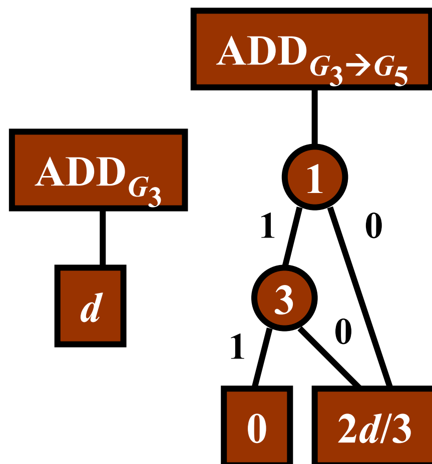


Masking Impact of gate G_5 w.r.t. gate G_1:

$$\mathrm{MI}_{\mathrm{D}}(G_1^{d,a} \to G_5)$$

$$= \mathrm{P}(ADD_{G_1 \to G_5} \to 0) \cdot (d - 0) + \mathrm{P}(ADD_{G_1 \to G_5} \to \frac{2}{3}d) \cdot (d - \frac{2}{3}d)$$

$$- \mathrm{P}(ADD_{G_1} \to d) \cdot (d - d) = \frac{3}{8}(d - 0) + \frac{5}{8}(d - \frac{2}{3}d) = \frac{7}{12}d$$



Masking Impact of gate G_5 w.r.t. gate G_2:

$$\mathrm{MI}_{\mathrm{D}}(G_2^{d,a} \to G_5)$$

$$= \mathrm{P}(ADD_{G_2 \to G_3 \to G_5} \to 0) \cdot (d - 0) + \mathrm{P}(ADD_{G_2 \to G_3 \to G_5} \to \frac{4}{9}d) \cdot (d - \frac{4}{9}d)$$

$$- \mathrm{P}(ADD_{G_2 \to G_3} \to 0) \cdot (d - 0) - \mathrm{P}(ADD_{G_2 \to G_3} \to \frac{2}{3}d) \cdot (d - \frac{2}{3}d)$$

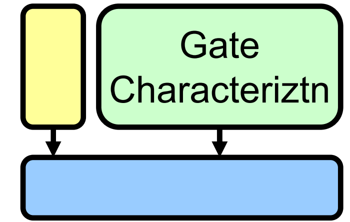
$$= \frac{5}{8}(d - 0) + \frac{3}{8}(d - \frac{4}{9}d) - \frac{1}{2}(d - 0) - \frac{1}{2}(d - \frac{2}{3}d) = \frac{5}{6}d - \frac{2}{3}d = \frac{1}{6}d$$

MMI Computation – An Example

Mean masking impact on duration (**MMI$_D$**) of gate G_5:

$$\begin{aligned}&\mathrm{MMI}_{\mathrm{D}}(G_5^{d,a})\\&=\frac{\mathrm{MI}_{\mathrm{D}}(G_1^{d,a}\rightarrow G_5)+\mathrm{MI}_{\mathrm{D}}(G_2^{d,a}\rightarrow G_5)+\mathrm{MI}_{\mathrm{D}}(G_3^{d,a}\rightarrow G_5)}{3d}\\&=\frac{7d/12+d/6+d/2}{3d}=\frac{5}{12}\end{aligned}$$

- The duration of a glitch is proportional to the probability of a soft error being latched, but the amplitude **is not**.
  - Use **only** mean masking impact on duration (MMI$_D$) as a guideline for SER reduction.



Why MEI & MMI?

- MEI
  - Glitch generation
  - Probability of generated glitches being registered
- MMI
  - Glitch propagation
  - Capability of filtering propagated glitches

Constraints on RAR

- $\Delta \mathrm{MEI}(s) = \mathrm{MEI}(t) \times [1-\mathrm{MMI}_D(t)]$
  - Worst-case estimation
    - $\mathrm{MMI}_D(t)$ increases after adding wire $s \rightarrow t$
      - More logical masking due to the new connection
      - More electrical masking due to larger gate delay
- $\Delta \mathrm{MEI}(u) = \mathrm{MEI}(v) \times [1-\mathrm{MMI}_D(v)]$
  - Average-case estimation

Experimental Results

Circuit	(# PIs, # POs, # Gates)	Dur. size (ps)	Ori. Avg. MES	Opt. Avg. MES	# Add. wires	# Rem. wires	Area over-head	SER reduc-tion
C432	(36, 7, 156)	60	3.57e-3	2.74e-3	37	29	3.45%	21.86%
		100	1.87e-2	1.35e-2	24	12		
		120	2.95e-2	2.43e-2	24	12		
C499	(41, 32, 458)	60	1.65e-3	1.40e-3	78	41	4.67%	18.64%
		100	7.12e-3	5.77e-3	47	21		
		120	10.9e-3	8.99e-3	52	27		
alu2	(10, 6, 339)	60	2.67e-3	2.22e-3	58	46	2.67%	18.27%
		100	1.71e-2	1.38e-2	36	28		
		120	2.74e-2	2.26e-2	28	23		
alu4	(14, 8, 660)	60	9.26e-4	8.10e-4	82	60	3.69%	13.54%
		100	8.70e-3	7.57e-3	77	42		
		120	1.46e-2	1.26e-2	72	42		
t481	(16, 1, 566)	60	10.5e-4	7.80e-4	162	76	7.11%	15.91%
		100	7.30e-2	6.21e-2	136	57		
		120	1.78e-1	1.59e-1	84	23		
ttt2	(24, 21, 166	60	4.11e-3	3.42e-3	28	20	1.30%	14.88%
		100	1.34e-2	1.13e-2	13	9		
		120	2.01e-2	1.71e-2	14	11		
x4	(94, 71, 288)	60	2.21e-3	1.80e-3	34	17	1.79%	18.75%
		100	5.89e-3	4.76e-3	18	10		
		120	8.72e-3	7.18e-3	19	7		
Avg.							**3.53%**	**17.41%**