A Mathematical Solution to Power Optimal Pipeline Design by Utilizing Soft Edge Flip Flops

M. Ghasemazar, B. Amelifard, M. Pedram

University of Southern California

Department of Electrical Engineering

August 11, 2008

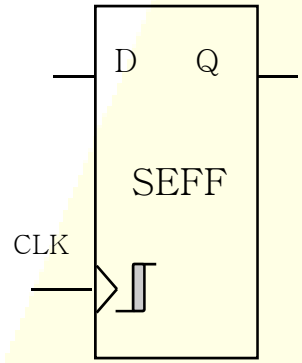
ISLPED 2008

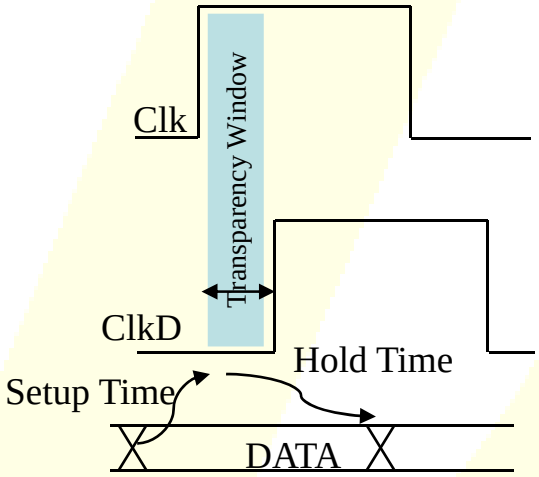
Outline

- Soft-Edge Flip Flops
- Power Optimal Pipeline Design
- Problem Formulation
- SEFF Modeling
- Experimental Results
- Conclusion

Soft Edge Flip Flop

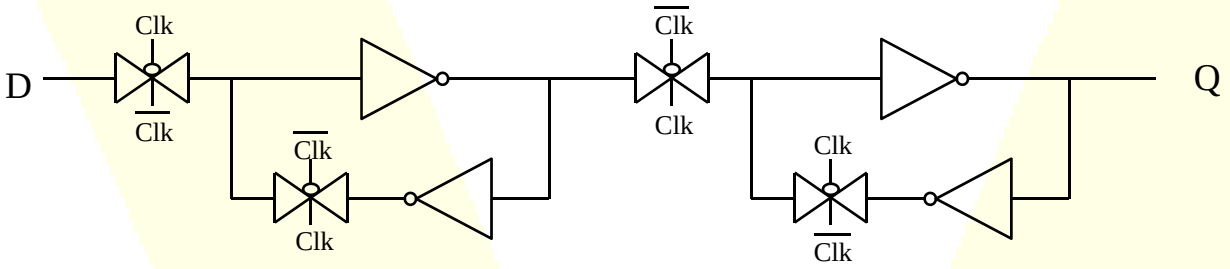
- Key idea: Allow the data to pass through a flip flop during a transparency window, instead of on a triggering clock edge
- Key advantage: Enable slack passing between adjacent pipeline stages which are separated by (master-slave) flip-flops
- Circuit implementation: Delay the clock of the master latch to create a window during which both the master and slave latches are ON



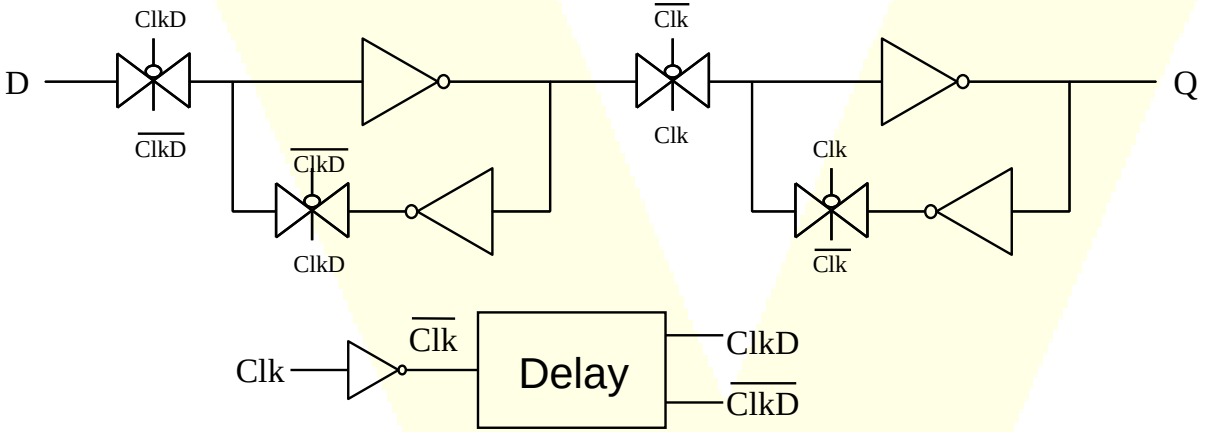


SEFF Implementation

Conventional (Hard Edge) Master-Slave FF



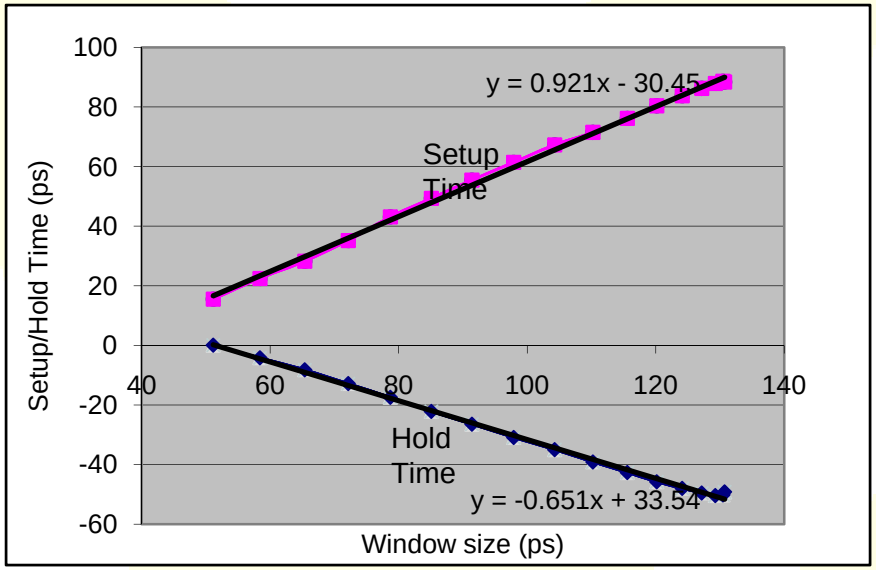
Soft Edge Master-Slave FF



SEFF Characteristics

- Setup and hold times, and clock-to-q delay of a soft-edge flip-flop are all functions of the transparency window width, *w*
- Simulations show a linear dependency on *w*

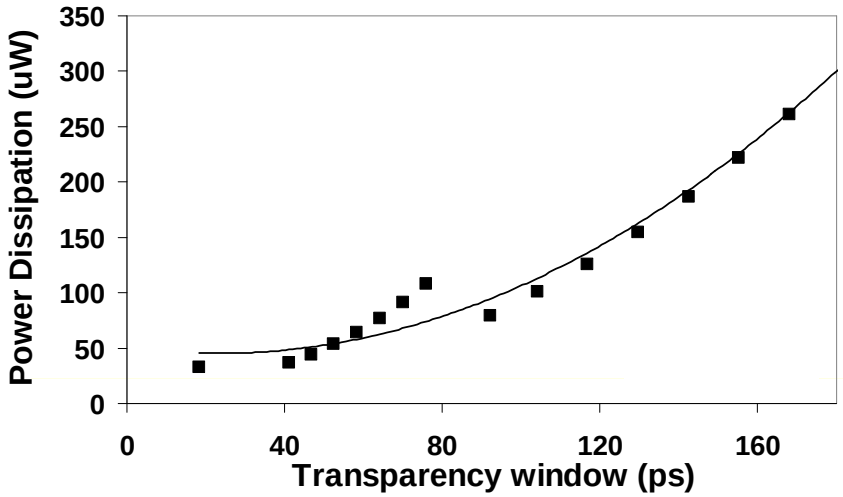
$$\begin{cases} t_{s,i}(w_i) = a_1 w_i + a_0 \\ t_{h,i}(w_i) = b_1 w_i + b_0 \\ t_{cq,i}(w_i) = c_1 w_i + c_0 \end{cases}$$



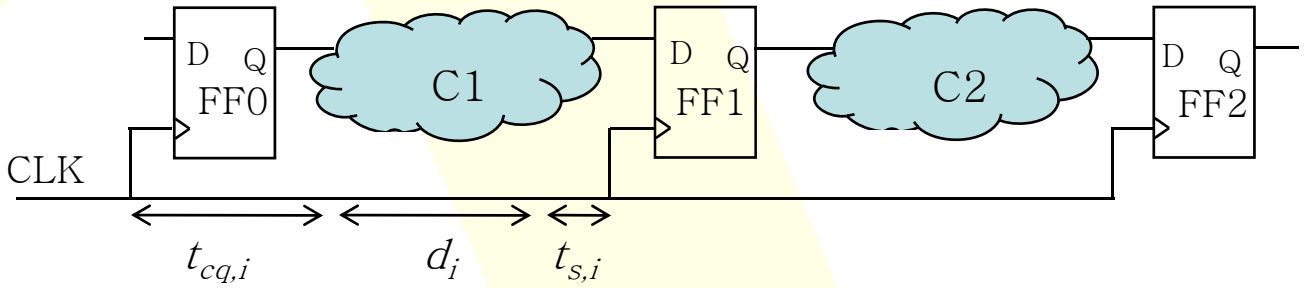
SEFF Characteristics – cont'd

- Power consumption of a SEFF is monotonically increasing with its window size (*w*). This is due to:
  - Higher switching activities in the internal nodes in the transparency window
  - Higher dynamic and leakage power consumption in the additional delay generation circuitry
- Experimental evaluation of <u>total</u> power consumption:

$$P_{FF,i} = d_2 w_i^2 + d_1 w_i + d_0$$



Pipeline Basics



- Timing constraints for a linear pipeline

$$d_i + t_{s,i} + t_{cq,i-1} \leq T_{clk} \quad 1 \leq i \leq N \qquad (1)$$

$$\delta_i + t_{cq,i-1} \geq t_{h,i} \quad 1 \leq i \leq N \qquad (2)$$

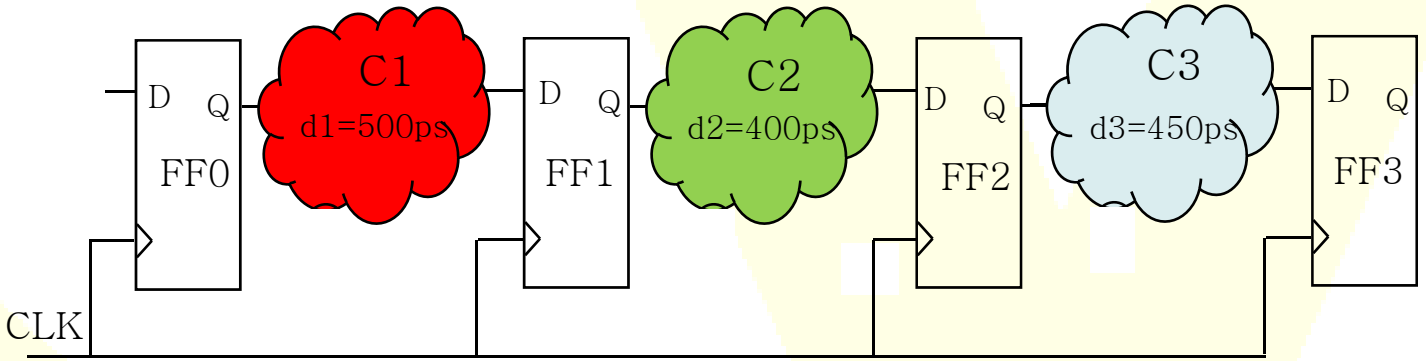
- Substitute FFs with SEFFs
  - First and Last FF's remain hard-edge ones
    - This is needed to avoid imposing constraints on the sender/receiver of data
  - Intermediate stage FF's may be substituted by SEFFs

$$d_i \leq T_{clk} - t_{s,i}(w_i) - t_{cq,i-1}(w_{i-1}) \quad 1 \leq i \leq N$$

$$\delta_i \geq t_{h,i}(w_i) - t_{cq,i-1}(w_{i-1}) \quad 1 \leq i \leq N$$

Power Optimal Pipeline

- Main Idea: Passing available slack of some stages to more timing critical stages to provide them with more freedom in power optimization through voltage scaling
- For example, let T_{clk}=$T_{clk,min}$=560ps and t_s=t_h=t_{cq}=30ps
  - If FF1 is replaced with a SEFF with a window size of 50ps
    - the first stage borrows 50ps from the second stage
    - the circuit can be powered with a lower supply voltage level
  - Ideally, 10% V_{dd} reduction ->19% power saving



PSLP Problem Statement

- Power-optimal Soft Linear Pipeline Design
  - Goal: Minimize the total power consumption of an N-stage linear pipeline circuit
  - Variables:
    - Optimal supply voltage level (1 variable)
    - Transparency windows size of the individual soft-edge FF-sets (N-1)
    - Delay elements to avoid hold time violations (N)
  - Constraints:
    - Setup/hold times
    - Window size limits
    - Single supply voltage

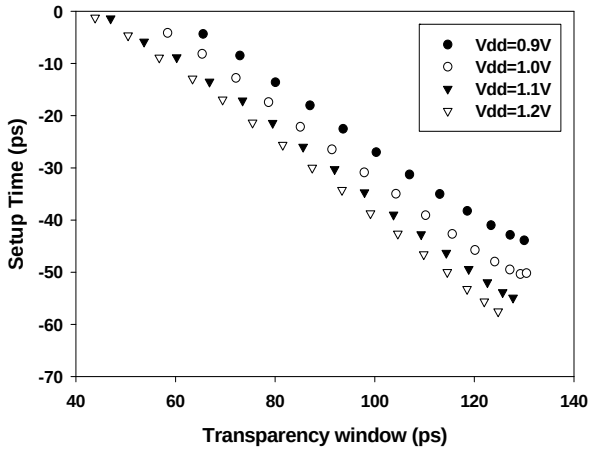
$$
\begin{cases}
Min.\ \ P_{total} = \sum_{i=1}^{N} P_{Comb,i}(v) + \sum_{i=1}^{N-1} P_{FF,i}(w_i, v) + \sum_{i=1}^{N} P_{DE,i}(z_i, v) \\
s.t.\ (I)\ \ d_i(v) \le T_{clk} - t_{s,i}(w_i, v) - t_{cq,i-1}(w_{i-1}, v);\ 1 \le i \le N \\
\quad (II)\ \ \delta_i(v) + z_i \ge t_{h,i}(w_i, v) - t_{cq,i-1}(w_{i-1}, v);\ 1 \le i \le N \\
\quad (III)\ w_{\min} \le w_i \le w_{\max};\ 1 \le i \le N-1 \\
\quad (IV)\ \ v \in \{V_0, V_1 ..., V_{m-1}\}
\end{cases}
$$

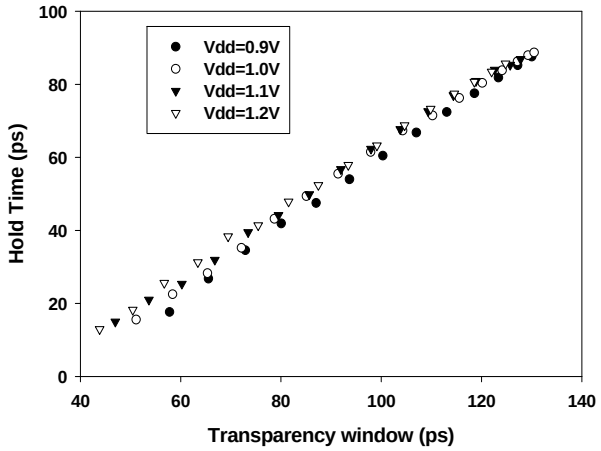
SEFF Modeling

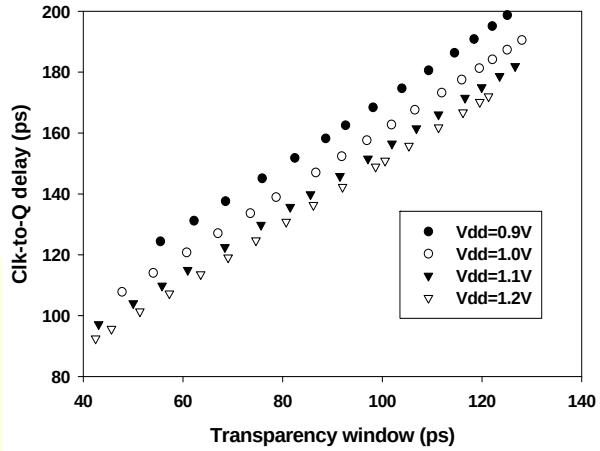
- Setup time, hold time, clock-to-q delay, and power dissipation are functions of both voltage and transparency window size
  - Voltage-dependent coefficients are determined from SPICE simulations

$$\begin{cases} t_{s,i}(w_i, v) = a_1(v)w_i + a_0(v) \\ t_{h,i}(w_i, v) = b_1(v)w_i + b_0(v) \\ t_{cq,i}(w_i, v) = c_1(v)w_i + c_0(v) \end{cases}$$

$$P_{FF,i} = d_2(v)w_i^2 + d_1(v)w_i + d_0(v)$$







Combinational Circuit Modeling

- Total power consumption at voltage level, *v:*

$$P_{comb,i}(v) = \left(\frac{v}{V_0}\right)^2 P_{dyn,i} + \left(\frac{v}{V_0}\right)^3 P_{leak,i}$$

- Max and Min combinational logic cell delays (calculated from the alpha power law):

$$d_i(v) = \left(\frac{V_0 - V_t}{v - V_t}\right)^\alpha d_i(V_0)$$

$$\delta_i(v) = \left(\frac{V_0 - V_t}{v - V_t}\right)^\alpha \delta_i(V_0)$$

- Power dissipation overhead of a delay element:

$$P_{DE}(z, v) = k(v) \cdot z$$

Solving the PSLP

- To solve PSLP
  - Enumerate all possible values for v
  - PSLP with fixed voltage (*PSLP-FV)*
    - $P_{comb,i}$ terms drop out of the cost function
    - Voltage constraint (IV) disappears
    - All other timing and power parameters become only dependent on w_i and z_i variables
  - For each fixed v, a quadratic program is set up and solved
    - We must minimize a quadratic cost function subject to linear inequality constraints
    - PSLP-FV can be solved optimally in polynomial time

USC

Experimental Setup

- Hspice simulations were used to extract parameters that are needed for the problem formulation
  - 65nm Predictive Technology Model (PTM)
  - Nominal supply voltage 1.2V
  - Die temperature 100^{o}C
- The SIS optimization package was used to synthesize a set of linear pipelines as test-bench circuits
- The MOSEK toolbox used to solve the mathematical optimization problem
- All results were collected on a 2.4GHz Pentium 4PC with 2GB memory

Benchmark Spec

Testbench (# of stages)	**(max, min) stage delays at nominal voltage (ps)**	**Clock freq. (GHz)**
TB1 (4)	(320,140), (332,150), (308,150), (320,170)	2.0
TB2 (5)	(320,140), (332,150), (308,150), (280,145), (320,170)	2.0
TB3 (3)	(325, 150), (310,155), (219,160)	2.0
TB4 (5)	(275,40), (235,40), (245,60), (275,50), (275,70)	2.5
TB5 (4)	(310,100), (245,40), (245,50), (245,60)	2.5

Experimental Results

Using slack passing to minimize power without degrading performance

TB	**Power Red. (%)**	**Optimum Vdd (V)**	**Optimum Window size (ps)**
TB1	32.1	1.0	40, 49, 22
TB2	33.8	1.0	40, 49, 46, 21
TB3	48.1	0.95	43,52
TB4	16.3	1.10	36, 35, 35, 20
TB5	25.4	1.05	60, 41, 36

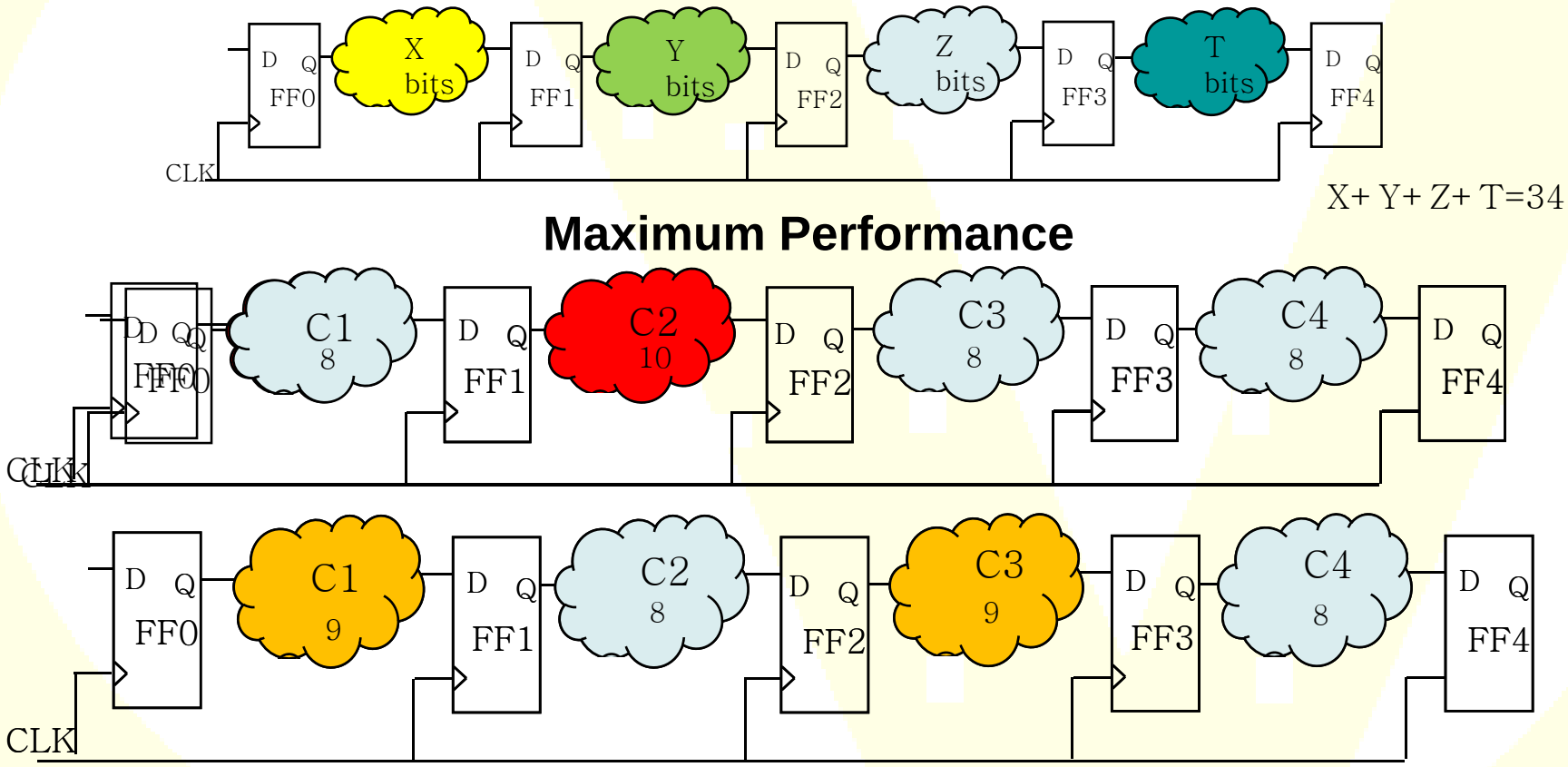
- Area overhead: Negligible compared to size of the rest of the pipeline circuit
- Runtime for all benchmarks: Less than one second

Utilizing slack passing to improve performance

Testbench	**Performance Improvement (%)**
TB1	14%
TB2	15%
TB3	20%
TB4	5%
TB5	10%

A Case Study: 34-bit Adder

- Problem: How to partition a 34-bit adder into 4 stages of pipeline to achieve maximum performance?



A Case Study: 34-bit Adder

Maximum Performance

Configuration	Vdd (V)	Min Clock Period (ps)	Power Consumption (mW)
10–8–8–8	1.2	450	6.42
8–10–8–8	1.2	472	6.50
8–8–10–8	1.2	472	6.51
8–8–8–10	1.2	486	6.55
9–9–8–8	1.2	455	6.42
9–8–9–8	1.2	433	6.51

A Case Study: 34-bit Adder

- Problem: How to partition a 34-bit adder into 4 stages of pipeline to achieve minimum power at target performance level?

Minimum Power @ 2.0GHz

Configuration	Vdd (V)	Power Consumption (MW)
10–8–8–8	1.05	4.9
8–10–8–8	1.15	5.1
9–9–8–8	1.05	4.9
9–8–8–9	1.10	4.9

Conclusion

- We presented a new technique to minimize the total power consumption of a linear pipeline circuit by utilizing soft-edge flip-flops and choosing the optimal supply voltage level for the pipeline
- We formulated the problem as a mathematical program and solved it efficiently
- Our experimental results demonstrate that this technique is quite effective in reducing the power consumption of a pipeline circuit under a performance constraint
- Future work will focus on problem of minimizing the energy cost of throughput in a linear pipeline circuit with dynamic error detection and correction capability

USC