Clocked storage elements

Vladimir Stojanovic

Material in this presentation is adapted from
“Digital System Clocking: High-Performance and Low Power Aspects”,

Outline

- Latch and Flip-Flop
- Timing and Power Metrics
- High-Performance Issues
- Low-Energy Issues
- State-of-the-art circuits
- Microprocessor Examples
Recent Interest in Flip-Flops

- Trends in high-performance systems
  - Higher clock frequency
  - More transistors on chip
- Consequences
  - Increased flip-flop overhead relative to cycle time
    - Cycle time 10 - 20 FO4 delays, flop overhead 2 - 4 FO4
  - Difficult to control both edges of the clock
  - Higher impact of clock skew
  - Higher crosstalk and substrate coupling
  - Higher power consumption
    - expensive packages and cooling systems
    - limit in performance
  - Clock burns up to 40%, flops up to 20% of total power

Why are clocked storage elements important?

- Cycle time ~ 12-20 FO4 delays
- Flip-flop overhead 2-3FO4 (20% of cycle time!!!)

![Diagram of flip-flop and clock](image)

\[ T_{D-Q} = T_{Clk-Q} + T_{Setup} \]
Latch and Flip-Flop

Latch – “soft” edge clocking
Latch is “transparent” (clock-level sensitive)

Flip-Flop – “hard” edge clocking
After the transition of the clock, data change does not affect the output (clock-edge sensitive)

Flip-Flop and Master-Slave Latch

Operational behavior appears the same…

How can one recognize the difference without knowing what is inside the “black-box”? 
Flip-Flop and M-S Latch: Structural Difference

Flip-Flop

Clock

Input

Pulse Generator

No Clock

Slave Latch

Q

Q

M-S Latch

Clock: \( \Phi_1 \)

Master (L1) Latch

Q1

\( \bar{Q}_1 \)

Clock: \( \Phi_2 \)

Slave (L2) Latch

Q2

\( \bar{Q}_2 \)

Pulse Capturing Latch

T-G Master-Slave Latch

- PowerPC 603 (Gerosa, JSSC 12/94)
Flip-Flop Example 1: SAFF
(Sense-Amplifier-Based Flip-Flop)

D=1
D=0
pulse
D=1

Pulse generator
Capturing Latch

SAFF DEC Alpha 21264 (Madden & Bowhill, 1990, Matsui 1994)

Flip-Flop Example 2: HLFF
(Hybrid Latch Flip-Flop)

D=1
D=0
signal at node X

Pulse Generator
Second Stage Latch
Logic Diagram of HLFF

Performance Metrics:
Timing and Power
Timing Parameters in Latches

Setup and Hold Times are defined relative to closing the clock signal

(a) Early data arrival
\[ D_1, Q_1 \]
\[ t_{cQ} \]

(b) Late data arrival
\[ D_2, Q_2 \]
\[ t_{dQ} \]

Timing Parameters in Flip-Flops

Setup and Hold Times are defined relative to rising edge of the clock

\[ D, Q \]
\[ t_{cW} \]
\[ t_{cQ} \]
Data-to-Output Delay

- Sum of setup time and Clk-Q delay is the only true measure of performance w.r.t. system speed
- \[ T = T_{\text{Clk-Q}} + T_{\text{Logic}} + T_{\text{setup}} + T_{\text{skew}} \]

Clk-Q Delay is a function of D-Clk

\[ T_{D-Q} = T_{\text{Clk-Q}} + T_{\text{Setup}} \]
Data to Clock Delay

- Constant Clk-Q Region
- Variable Clk-Q Region
- Failure Region

Data arrives early → Data arrives late

Opt Setup

Setup Time vs. Data-to-Output (D-Q) Delay

Power Consumption

- Power related to a CSE can be divided into:
  - Input power
    - Data power ($P_D$)
    - Clock power ($P_{CLK}$)
  - Internal power ($P_{INT}$)
    - Depends on data activity and glitching activity
  - Load power ($P_{LOAD}$)
    - Can be merged into $P_{INT}$

$$P_{tot} = P_{internal\&load} + \sum_{inputs(D,CLK)} P_{driver}$$
High-Performance Issues

Clock Skew and Jitter

- Ref. Clock
- Received Clock
- $t_{DRVCLK}$
- $t_{skew}$
- $t_{skew}$
- $t_{jit}$
- $t_{jit}$
- $T$
The Idea of Clock Uncertainty Absorption

Change in D-Q delay is much smaller than the clock uncertainty (CSE absorbs a part of the uncertainty)

Clock Uncertainty Absorption

Nominal $D_{d-ck}$

Early $D_{d-ck}$

Late $D_{d-ck}$

Clock uncertainty $t_{CU}$

$T_{nominal}=0$

$D_{qnm}$

$D_{qqm}$

$Worst-case D_{d-q}$

Nominal $D_{d-ck}$

$D_{qnm}$

$D_{qqm}$
Example: HLFF

[Partovi et al, ISSCC’96]

State-of-the-Art CSEs in CMOS Technology
Requirements in the Flip-Flop Design

- Small Clk-Output delay, Narrow sampling window
- Low power
- Small clock load
- High driving capability (increased levels of parallelism)
  - Typical flip-flop load in a 0.18μm CMOS ranges from 50fF to over 200fF, with typical values of 100-150fF in critical paths
- Integration of logic into the flop
- Multiplexed or clock scan
- Crosstalk insensitivity
  - dynamic/high impedance nodes are affected

Low-Energy Designs: Master-Slave Latch Examples

\[ C^2MOS\ Latch \quad \text{Master-Slave Latch (MSL)} \]

\[ \text{[Suzuki et al, JSSC 1973]} \quad \text{[G.Gerosa et al, JSSC 1994]} \]

Feedback for pseudo-static operation
State node \( S_3 \) protection in PPC (decoupled Q)
Master-Slave Latches

- Positive setup times
- Two clock phases:
  - distributed globally
  - generated locally
- Small penalty in delay for incorporating MUX
- Some circuit tricks needed to reduce the overall delay

High-Performance Designs: Flip-Flop Examples 1/2

**HLFF (Hybrid Latch-Flip-Flop)**

**SDFF (Semi-Dynamic Flip-Flop)**

[Partovi et al, JSSC 1996]  [F.Klass et al, JSSC 1998]

Pulse-generating first stage (precharge-evaluate)
Keepers for pseudo-static operation
Output load decoupled from internal nodes
**SAFF (Sense-Amplifier-Based Flip-Flop)**

- Fully-differential circuit
- First stage sense-amp can take reduced-swing inputs
- 2nd stage is capturing latch delay to Q and !Q not equal

[SAFF with Improved S-R Latch](#)

- The first stage is unchanged sense amplifier
- Second stage is sized to provide maximum switching speed
- Driver transistors are large
- Keeper transistors are small and disengaged during transitions

[Nikolic & Stojanovic ISSCC '99](#)
Flip-Flops

- First stage is a pulse generator
  » generates a pulse (glitch) on a rising edge of the clock
- Second stage is a latch
  » captures the pulse generated in the first stage
- Pulse generation potentially results in a negative setup time and soft-edge property
- Must check for hold time violations

Note: power is always consumed in the clocked pulse generator

Delay Comparison: M-S Latches and Flip-Flops

![Min D-Q Delay Comparison](image)

- Flip-Flops are faster

0.18um, high load (14 min inverters)
Energy Comparison: M-S Latches and Flip-Flops

Energy breakdown (50% activity)

- **Latches are lower energy**

0.18um, high load (14 min inverters)

Summary

- CSE topology depends on target application
  - Master-Slave Latches for low-energy
  - Flip-Flops & Pulsed latches for high-performance
- Delay is critical in high-speed systems, although minimizing Clk energy is of increasing importance
- Methods for reducing Clk energy
  - Clock gating (more effective in high-performance than in low-energy designs)
  - Reduced-swing clocking
  - Dual-edge clocking
Microprocessor Examples

Sun UltraSPARC-III

Basic Flip-Flop (SDFF)

[Klass, 1998]
Sun UltraSPARC-III Dynamic Flip-Flops

Sun UltraSPARC-III Dynamic Flip-Flops

Sun UltraSPARC-III Flip-Flops with Logic Embedding

2-input XOR
Latches used in 20164 Alpha

[21064 modified TSPC Latches]

Logic Embedding in 21064/21164 Alpha

1 level of logic
(21064 Alpha)

2 levels of logic
(21164 Alpha)
Flip-Flop used in 21264 Alpha

Differential

IBM Processors: Level-Sensitive Scan Design (LSSD)

Hazard-free level-sensitive polarity-hold latch

[Eichelberger 1983]
The PowerPC 603 Master-Slave Latch

[Gerolla et al, 1994]

IBM Power4™ processor

Scannable Split Latch with LSSD Capability  [Warnock et al, 2002]
Low-Energy Issues:
Clock & Clocked elements burn more than 60% of the processor power

Low-Swing Clocking: Clock Driver Re-design

50% power reduction with half-swing clock
(minus some penalty in clock drivers)

[H. Kojima, JSSC, April 1995]
Low-Swing Clocking: CSE Re-design

PMOS does not fully turn off

[H. Kawaguchi and T. Sakurai, JSSC, May 1998]

Low-Swing Clocking: N-only CSEs

N-only clocked transistors, M-S Latch Example
(N₁ and N₂ improve pull-up on S_M)

[D. Markovic, J. Tschanz, V. De, 2001, patent pending]
Clock Gating: Global Clock Gating

(a) (b)

<table>
<thead>
<tr>
<th>In</th>
<th>0</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Load</td>
<td>Clk</td>
<td></td>
</tr>
<tr>
<td>REG</td>
<td>D Q</td>
<td></td>
</tr>
</tbody>
</table>

Time-mux (no gating!) 
Global Clk Gating

Used to save clocking energy when data activity is low

[M. Nogawa and Y. Ohtomo, JSSC, May 1998]

Clock Gating: Local Clock Gating

<table>
<thead>
<tr>
<th>D</th>
<th>CP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clk</td>
<td></td>
</tr>
</tbody>
</table>

Data-Transition Look-Ahead

Used to save clocking energy when data activity is low
Dual-Edge Triggering: Latch-Mux

[R.P. Llopis and M. Sachdev, ISPLED Aug. 1996]

Concept Circuit Example

Used to save clocking energy regardless of data activity!

Dual-Edge Triggering: Pulsed-Latch

Concept Circuit Example

Single-Edge Dual-Edge
Dual-Edge Triggered Flip-Flop

Concept

Circuit Example

Pulse-generating latches trigger capturing latch

Design goals

- **Apply**
  - Small clock load
  - Short direct path
  - Reduced node swing
  - Low-power feedback
  - Pulsed design
  - Optimization of both Master and Slave latch

- **Avoid**
  - Positive setup time
  - Sensitivity to clock slope and skew
  - Dynamic (floating) nodes
  - Dynamic Master latch

Conduct Energy - Delay optimizations
Take into account all sources of power dissipation
ALWAYS use Clk-Q + setup time for max delay

For more details on storage elements check prof. Oklobdzija’s ISSCC’02 workshop:
http://www.ece.ucdavis.edu/acsel under Presentations
What to Expect in the Future?

- Incorporating logic into the CSE
- Absorbing clock skew
- Pipeline boundaries will start to blur – pulsed latches
- Latch-less domino style clocking, signals used to clock
- Synchronous design only in a limited domain
- Asynchronous communication between synchronous domains