A Test Circuit for Measurement of Clocked Storage Element Characteristics

Nikola Nedovic, *Member, IEEE*, William W. Walker, *Member, IEEE*, and Vojin G. Oklobdzija, *Fellow, IEEE*

***Abstract*—We present a method, on-chip test circuitry, and an error analysis, for accurate measurement of timing characteristics and power consumption of clocked storage elements. The test circuit was fabricated in 0.11 μm CMOS technology and the measurements performed automatically using a serial scan interface. The precision and accuracy of the presented method are demonstrated by the ability to measure entire clock-to-output characteristics of flip-flops. Estimated data-to-output delay systematic measurement error is 6 ps (7%), and random error is 10 ps (11%). The method and the test circuit are applicable for delay measurements of other circuit blocks as well.**

***Index Terms*—Clocked storage elements, delay, measurement circuit, measurement error, on-chip measurement, power consumption.**

I. Introduction

FLIP-FLOPS and latches, which we collectively refer to as clocked storage elements (CSEs) in this paper, are among the most important components of high-performance VLSI systems. Due to increasing clock frequencies and the use of pipelines to achieve high throughput, CSEs occupy an increasingly large portion of the cycle time [1]–[3]. The complexity of VLSI systems requires a large number of CSEs on die, which consume a large share of system power, heavily load the clock, and generate large current spikes during switching [3]. Thus, CSEs affect various system characteristics such as power consumption, implementation of the clocking subsystem, signal integrity, chip area, global placement and routing, packaging, and heat removal.

In order to estimate the impact of a CSE on the system, its physical characteristics must be determined. The most valuable information is the timing, as it affects cycle time [4]–[6]. Timing parameters of logic gates and CSEs can be estimated using transistor models constructed from the measurements, and circuit simulators such as HSPICE. However, because of their importance in determining cycle time and races, precision hardware measurement of CSEs to validate the simulation models is highly desirable. A measurement can also help discriminate between alternative CSE implementations that may not show clear differences in simulation, particularly for high-performance applications [5], [6]. Measured CSE timing published in the literature frequently show only the highest achieved toggle frequency when the CSE is connected in a ring oscillator configuration, e.g., the T (toggle) flip-flop [8], [9]. Although these ring oscillator measurements are inherently precise, they can only be used to measure the maximum throughput of a CSE, whereas timing analysis for high-performance design needs a detailed specification of the setup time, hold time, and clock and data to output delays. A direct off-chip measurement of the delay between waveforms of CSE ports [10] can be used to validate the simulation models. However, an off-chip measurement approach has serious limitations, since the on-chip delays of CSEs in deep-submicron technologies are typically much smaller than that of the circuitry connecting the ports to the instrumentation. The measurement errors incurred by this circuitry can be comparable to the measured quantity.

In this paper, we present a circuit for the on-chip measurement of the timing characteristics and power consumption of CSEs. Performing measurements on-chip eliminates sensitivity to off-chip delay, noise, and bandwidth limitations, thus achieving lower measurement error than previously reported techniques. We also provide a detailed statistical error estimation of the presented measurement method [11] using our on-chip circuit. In Section II, we define the characteristics of a clocked storage element that we are trying to measure. The proposed test circuits are shown in Section III, and the analysis of the measurement error is given in Section IV. In Section V, we present measured results based on devices fabricated in a 0.11-μm CMOS technology. Section VI concludes the paper.

II. Clocked Storage Elements

This section describes timing characteristics of CSEs and their significance to the performance of a circuit. The particular CSE that we focus on is an edge-triggered flip-flop with a possible transparency window around the clock edge.

A. CSE Characteristics

The propagation time (clock-to-output delay, t_{Clk-Q}) of a CSE is the delay between the capturing clock edge and the output transition to a new value. Fig. 1 is a qualitative plot of the clock-to-output delay of a CSE as a function of data-to-clock delay (t_{D-Clk}). When t_{D-Clk} is large, t_{Clk-Q} is constant, which corresponds to normal operation of the CSE when its setup time is met. As data arrival approaches the capturing clock edge, t_{Clk-Q} increases. If the data arrival is pushed even further, the t_{Clk-Q} increase becomes severe enough to cause the capture to fail [5], i.e., the setup time is violated.

If subsequent data arrival is set long enough after the clock arrival, the initial data is safely captured in the CSE (e.g., if the data transition following the clock edge is "0" to "1", a "0" will be captured). As the subsequent data arrival approaches the

Manuscript received April 29, 2003; revised March 18, 2004.

N. Nedovic and W. W. Walker are with Fujitsu Laboratories of America, Sunnyvale, CA 94085 USA (e-mail: nikola@fla.fujitsu.com).

V. G. Oklobdzija is with the Department of Electrical and Computer Engineering, University of California, Davis, CA 95616 USA.

Digital Object Identifier 10.1109/JSSC.2004.831498

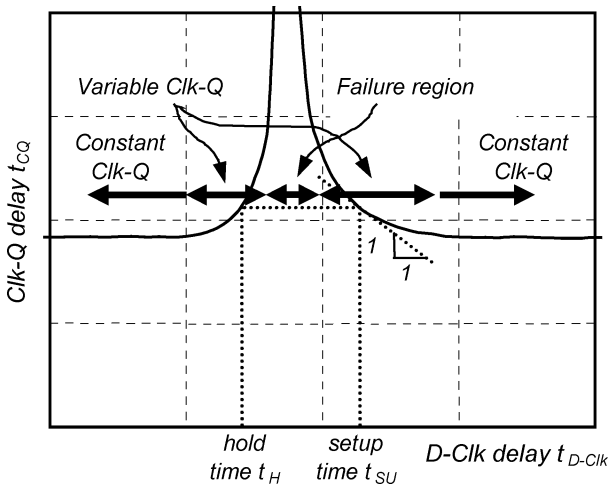


Fig. 1. Typical clock-to-output characteristics.

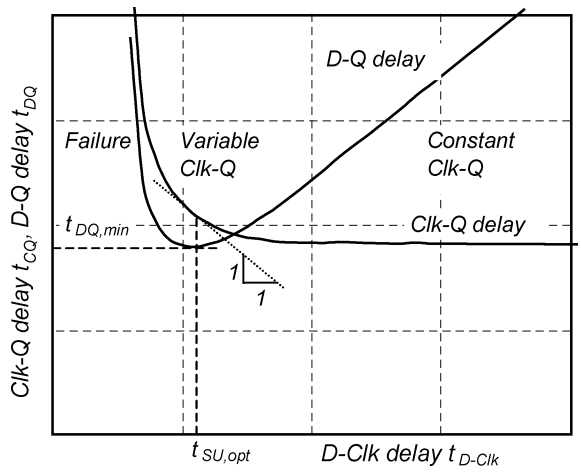


Fig. 2. Data-to-output characteristic.

clock edge, the clock-to-output time starts to degrade similar to the setup side. If the subsequent data arrives too early, the transition fails; i.e., a hold time violation occurs.

B. Optimal Setup Time

The conventional definition of setup time is the latest allowed data arrival with respect to the capturing clock edge in order to correctly capture data. An alternative view [5], which we adopt here, defines the setup time as the data arrival with respect to clock that minimizes the time that the CSE takes from the clock cycle. A CSE incurs clock-to-output delay, t_{Clk-Q}, at the beginning of a cycle and data-to-clock delay, t_{D-Clk}, at its end. Data-to-output delay is the sum $t_{DQ} = t_{D-Clk} + t_{Clk-Q}$, which we also refer to as *timing overhead* or *insertion penalty*, since it has to be accounted for in order to provide the synchronization and cannot be used for logic computation. Fig. 2 plots qualitatively t_{DQ} as function of t_{D-Clk}. The minimum value of data-to-output delay $(t_{DQ,min})$ optimizes performance, and the data-to-clock delay that corresponds to this point is called the optimal setup time $(t_{SU,opt})$ [5], The optimal setup time corresponds to the zero-slope point on the t_{DQ} versus t_{D-Clk} plot:

$$t_{SU,opt}, t_{DQ,\min} \Rightarrow \frac{\partial t_{DQ}}{\partial t_{D-Clk}} = 0 \Rightarrow \frac{\partial t_{Clk-Q}}{\partial t_{D-Clk}} = -1. \quad (1)$$

From (1), the optimal setup time occurs at the point where the reduction of t_{D-Clk} equals the increase of t_{Clk-Q}, i.e., where clock-to-output characteristic has the slope of -45°.

C. Hold Time

The hold time of a CSE is defined as the earliest time after the triggering clock edge that the input D is allowed to change that still allows safe capture of its value at the clock edge. A definition of hold time commonly used in practice is the minimum clock-to-data delay that causes a voltage dip of less than a certain percentage of full swing at a specified internal CSE node. Such a definition is inadequate, since a comparison of different CSEs require applying the same criteria to all, and selecting appropriate internal nodes to measure for two different designs involves some degree of subjectivity. In order to provide a general definition of the hold time, we observe that it is safe to increase the clock-to-output delay as long as the timing of the subsequent stages is uncorrupted. Since the system operates correctly for any data arrival that does not violate the setup time, it still operates correctly when t_{Clk-Q} on the hold time side of the delay characteristic (determined by t_{D-Clk}) equals the maximum t_{Clk-Q} on the setup time side. Since this maximum t_{Clk-Q} occurs when t_{D-Clk} equals the setup time (Fig. 1), we define the hold time as the data arrival with respect to clock edge that produces the same clock-to-output delay of the previous data as for the setup time:

$$t_H = t^{-1}_{Clk-Q,H}\left(t_{Clk-Q,SU}(t_{SU})\right). \quad (2)$$

In (2), $t_{Clk-Q,SU}(t_{D-Clk})$ and $t_{Clk-Q,H}(t_{D-Clk})$ are parts of $t_{Clk-Q}(t_{D-Clk})$ before and after the failure, i.e., $t_{Clk-Q,SU}(t_{D-Clk})$ corresponds to setup $Clk-Q$ characteristic, and $t_{Clk-Q,H}(t_{D-Clk})$ is the $Clk-Q$ characteristic of previously captured data versus data-to-clock delay (Fig. 1). This definition of the hold time is independent of how the setup time is defined. If we use the setup time definition from [5], we refer to the corresponding hold time as the optimal hold time (Fig. 1).

III. Built-In Measurement Circuit

This section describes the design and operation of the on-chip test circuits used to measure the timing and power consumption of CSEs.

A. Delay Measurement Test Circuit

The test circuit used for the delay measurement is shown in Fig. 3. The primary blocks are a scan register to hold a setup vector and capture the test results, two delay lines that provide D and Clk inputs to the device under test (DUT, the CSE), a reference generator, a multiplexer, M_1, that selects the signal to be observed from clock Clk, data D or outputs Q *or* $\overline{Q}$), and the capturing storage element CSE_{OUT} that captures the multiplexer output. The delay of delay line Δ_D is variable, and the delay of Δ_{Clk} is constant, allowing a sweep of data-to-clock delay. The reference generator, shown in Fig. 4, provides variable delay between Ref_1 and Ref_2. For any port of the DUT, we can sweep the arrival time of the reference signal Ref_2 and

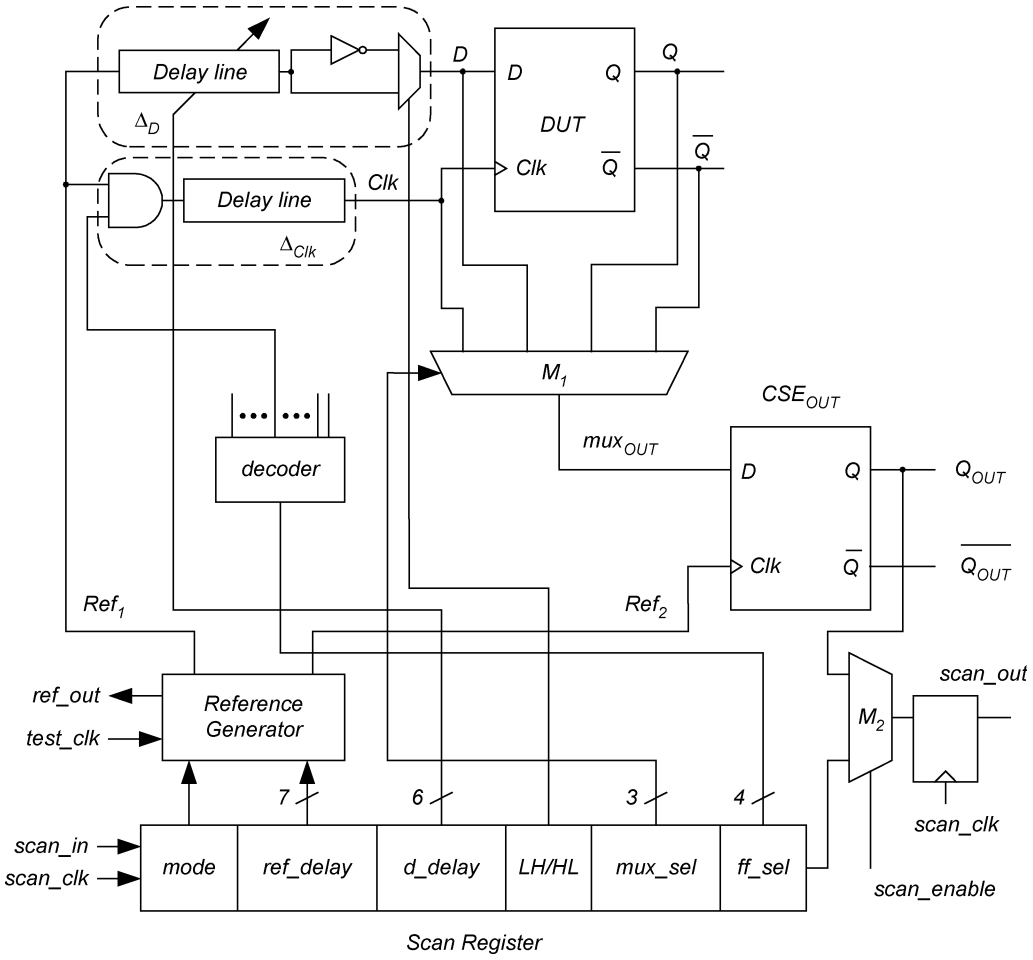


Fig. 3. Delay measurement test circuit.

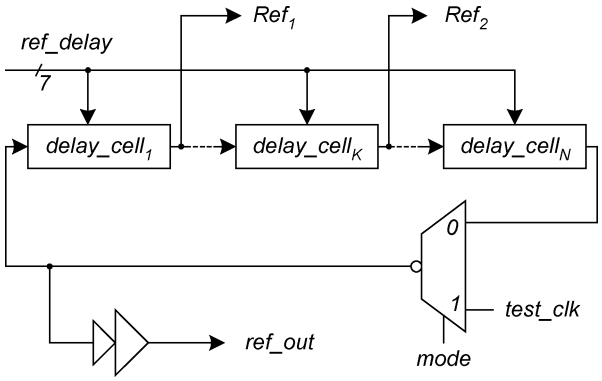


Fig. 4. Circuit for generating references for delay measurement.

observe where CSE_{OUT} just fails to capture the corresponding multiplexer output. The delay between any two signals is equal to the delay difference between Ref_1 and Ref_2 that correspond to the first capture failure of CSE_{OUT} for these two signals.

Use of Ref_1, Ref_2 and CSE_{OUT} as described above allows the entire measurement to be performed on chip. After a measurement is performed, the result, stored in CSE_{OUT}, is transferred to the scan-out flip-flop, and then scanned back into the tester. No chip-crossings are involved in the measurement. Thus, disregarding the error associated with the calibration of the reference generator, there is no error contributed to the measurement by chip I/O.

The circuit for generating references Ref_1 and Ref_2 is shown in Fig. 4. It consists of $N = 10$ identical programmable delay cells connected as a ring oscillator. The references Ref_1 and Ref_2 are obtained as input and output of one of the programmable delay cells. Each delay cell is loaded with a dummy buffer that matches the load of the references Ref_1 and Ref_2, ensuring equal delay of all the cells in the ring. The delay between Ref_1 and Ref_2 is characterized by setting $mode = 0$, which enables oscillation. The period of the oscillations is measured by connecting ref_out to an external frequency counter, which allows us to determine delay between Ref_1 and Ref_2. To perform the DUT measurement, a single pulse starting from $test\_clk$ is fed through the delay lines passing through Ref_1 and Ref_2. Since, as we will show below, any measured quantity is obtained as the difference of the two delays Ref_1 and Ref_2, all we need to know about the timing of these references is the difference in the delays, and not the actual delays.

Each programmable delay cell in the ring oscillator from Fig. 4 consists of coarse and fine-tuned segments (Fig. 5) which are controlled by $delay\_ref[6:0]$, providing a total of 128 oscillation periods. The coarse delay selects one of $0, 4, 8, \ldots, 28$ inverters in addition to the delay of the multiplexer. Fine delay tuning is achieved using three inverter stages whose load can

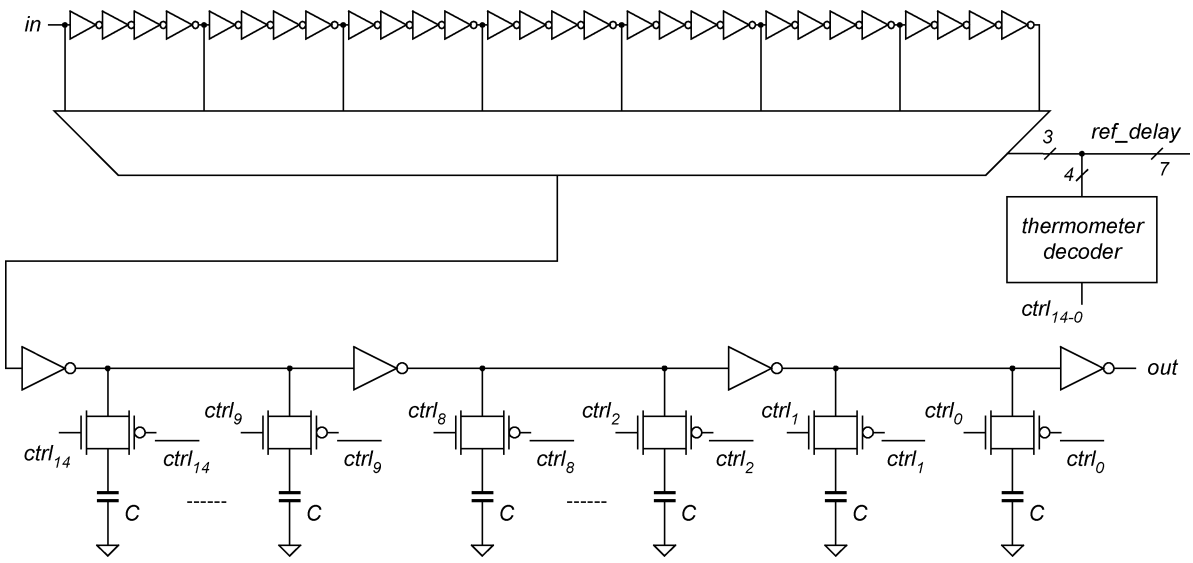


Fig. 5. Delay cell in ring oscillator.

be varied using four LSB control bits. In order to achieve as uniform dependence of the delay on the control code as possible, thermometer decoding is used so that a larger code results in a larger load on any inverter stage in the fine-tuning block. Thermometer encoding allows the delay between signals Ref_1 and Ref_2 to rise monotonically, varying between about 200 ps to about 850 ps with average step of 5 ps for typical process.

We briefly present the implementation details of the other blocks in the test circuit. The variable-delay block Δ_D from Fig. 3 is similar to the oscillator delay line (Fig. 5). Since Δ_D needs to provide a smaller delay span than the oscillator delay line, only six delay control bits are needed. No pre-characterization of the Δ_D block is required because its delay is measured each time a DUT measurement is taken. A multiplexer at the end of the Δ_D delay line is needed to selectively apply either rising or falling inputs to the DUT. If the DUT is differential, the Δ_D delay line must supply differential data inputs. The Δ_{Clk} delay line can be inverting or noninverting, depending on whether the DUT is rising-edge or falling-edge triggered. The capturing storage element CSE_{OUT} is differential, with symmetrical low-to-high and high-to-low setup and hold times, in order to minimize the error. The actual test circuit was expanded to accommodate delay measurement of up to 16 DUTs. The control input $ff\_sel[3:0]$ is used to select a particular DUT for the measurement.

A key feature of the proposed test circuit is use of the same components for measuring the arrival times of all ports of the DUT, rather than matching the delays of two different components. This approach eliminates the error due to process-variation-induced delay mismatch of identical circuits. In addition, the test circuit draws the same current from the power supply for all measurements. Thus, the voltage and temperature variation contribution to the measurement error is minimized. The delay mismatch of the programmable delay cell, which is the largest component of the error, is reduced by placing the oscillator cell used to generate references Ref_1 and Ref_2 in the middle of the layout of the oscillator (assuming an approximately linear variation of process across the die.)

B. Delay Measurement Test Procedure

A total of 2^{17} test vectors are needed to measure low-to-high and high-to-low setup and hold characteristics. The vectors were applied by software running on a PC equipped with an IEEE 1149.1 scan controller card.

The delay measurement for a single DUT consists of the following steps:

- characterization of the oscillator: for each value of the oscillator control word $ref\_delay[6:0]$, oscillating mode $(mode = 0)$ is chosen and the period of the oscillations T_{OSC} is measured. The measured delay of a single delay cell in the oscillator, which we designate t_{ref}, as a function of $ref\_delay[6:0]$ is calculated from measured period:

$$t_{ref}(ref\_delay) = \frac{T_{OSC}(ref\_delay)}{2N}. \tag{3}$$

- DUT measurement: the reference delay generator is set to nonoscillating mode $(mode = 1.)$ For each data-to-clock delay, determined by control bits $d\_delay[5:0]$, and for each delay between reference signals Ref_1 and Ref_2, the following is executed for each DUT port:
  - initialization: the initialization is performed by setting the data arrival early enough or late enough to assure that the initialization value will be captured, then applying a single pulse at the input $test\_clk$. The pulse propagates to the DUT and initializes it to the desired value.
  - measurement: low-to-high or high-to-low delay measurement (control bit LH/HL) and data-to-clock delay (control bits $d\_delay[5:0]$) are selected. The $mux\_sel[2:0]$ control signal is set to select the DUT port to be observed. A single pulse is applied to $test\_clk$. The value captured by CSE_{OUT} is written to the scan register and scanned into the tester. Note that measurements of setup and hold characteristics differ only in the initial value written to the DUT before the measurement.

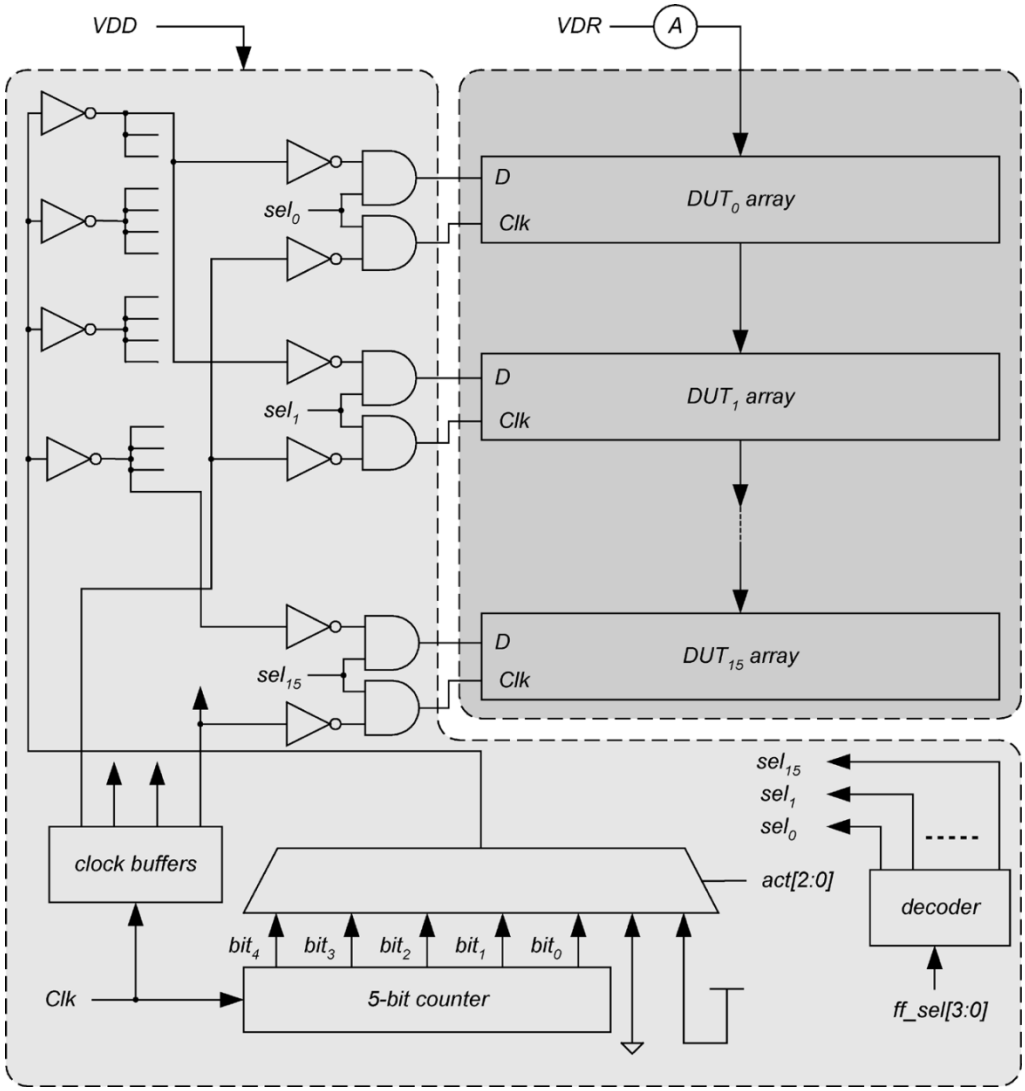


Fig. 6. Power consumption measurement setup.

- The maximum value of $ref\_delay[6:0]$ is recorded at which the capture of each port of the DUT fails, ($t_{ref,Clk}$, $t_{ref,D}$, $t_{ref,Q}$), and the data-to-clock and clock-to-output delays are calculated according to (4) and (5):

$$t_{D-Clk} = t_{ref,Clk} - t_{ref,D} \tag{4}$$

$$t_{Clk-Q} = t_{ref,Q} - t_{ref,Clk}. \tag{5}$$

Note that t_{ref} is slightly larger than the delay between reference signals Ref_1 and Ref_2 for the portion of the delay of the multiplexer in Fig. 4. This multiplexer delay need not be measured, as it cancels out in (4) and (5), so each of t_{D-Clk} and t_{Clk-Q} is exactly the difference between two $Ref_1 - Ref_2$ delays.

C. Power Consumption Test Circuit and Test Procedure

The power consumption is measured for an array of identical CSEs with outputs loaded and inputs buffered. The test circuit used in the measurement is shown in Fig. 6. Up to 16 different DUTs can be tested by the circuit. Within one DUT array, M identical CSEs are connected in parallel, and control circuitry is used to select the array to test and apply the clock and data inputs. The loading and buffering of each individual CSE within an array is shown in Fig. 7. The supply voltage of the storage element arrays is separated from the main supply to isolate their power. Only the array selected by $ff\_sel[3:0]$ is supplied with clock and data, while inputs to all other arrays are inactive. All arrays can be shut off by selecting control signals to keep the input clock and data low. Input $act[2:0]$ selects the data activity of the D input from 100%, 50%, 25%, 12.5%, 6.25%, and 0%. D can be selected as either 0 or 1 at 0% data activity. Data activity is defined as the percentage of the maximum data throughput. For example, 100% data activity means the data toggle rate is half the clock toggle rate. Inputs $ff\_sel[3:0]$ and $act[2:0]$ are supplied by the scan register.

All arrays share a common supply, requiring subtraction of their leakage power from the measurements. The estimate of leakage current can be obtained by measuring the power supply current when all arrays are disabled. Thus, power consumption of an individual storage element can be obtained by subtracting the measured supply power for two cases: 1) only the selected array which is currently under test receives toggling clock and

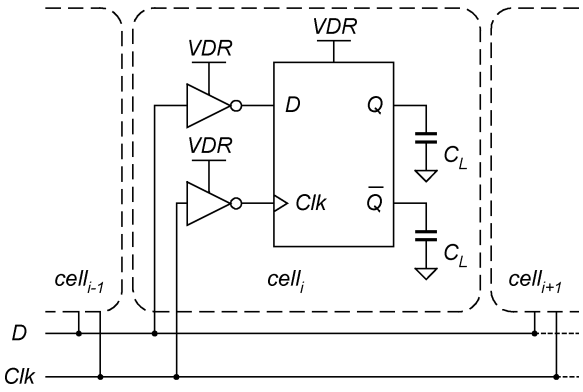


Fig. 7. Single storage element cell in array in power consumption measurement circuit.

data and all other arrays are shut off, and 2) all arrays are shut off.

$$P_{CSE} = \frac{VDR \cdot [I_{VDR}(CSE) - I_{VDR}(Clk = D = 0)]}{M} \tag{6}$$

where VDR is the supply voltage, $I_{VDR}(CSE)$ and $I_{VDR}(Clk = D = 0)$ are measured VDR supply currents when the CSE is selected using the $ff\_sel[3:0]$ control bits, and when all arrays are shut off, and $M = 120$ is the total number of CSEs in the array.

IV. Measurement Error Estimation

In any measurement, we can distinguish a DUT, a physical quantity to be measured, and the measurement system (MS) that produces the measurement. The MS introduces measurement errors due to noise, nonideal environmental parameters, and instrument accuracy and precision. Errors can be categorized as [11]:

Systematic: the difference between the physical quantity and the measured quantity that occurs consistently. Systematic error is associated with the instrument or particular measurement technique used. It is usually specified by its upper bound, since exact systematic error is never known (if known, it could be removed.) Associated with systematic error is the property of a system called accuracy, i.e., systems with small systematic error have high accuracy.

Random: the difference between a physical quantity and its measured value as a result of unpredictable fluctuations in the properties of the DUT and the MS. The source of this type of error is any random occurrence that affects the measurement: power supply noise, temperature variation, process variation, etc. Due to random errors, different measurements of the same quantity give different results. The random error is modeled by treating a measured quantity as a statistical variable that complies with a Gaussian probability distribution, characterized by its mean μ and standard deviation σ. The precision of a system is proportional to random error, i.e., small random error is equivalent to high precision.

We derive the random and systematic error in the delay measurements statistically from the parameter variations in the measurement circuit. These derivations are used in Section VI, where we estimate each component of the error and calculate the delay measurement error. The following notation is used in the derivations to describe the actual delays of each block (also refer to Fig. 3).

- t_{D-Clk} and t_{Clk-Q} are the data-to-clock and clock-to-output delays of the DUT.
- $t_{\Delta D}$ and $t_{\Delta Clk}$ are the propagation delays of the Δ_D and Δ_{Clk} delay lines.
- t_{mux} is the delay of the multiplexer.
- $t_{CSEout,SU}$ is the setup time of CSE_{OUT}.
- $t_{osccell,i}$ is the delay of the ith cell of the reference generator.
- Δt_{12step} is the delay step between Ref_1-to-Ref_2 delays for two consecutive ref_delay control words.
- t_{OUT} is the delay of the output buffer and off-chip cable ending with the instrument used for the frequency measurement.
- When a circuit is excited more than once, the variation of its actual delay is recorded in a superscript. For example, the multiplexer delays when measuring each of D, Clk, and Q delay from the reference are t^{D}_{mux}, t^{Clk}_{mux}, and t^{Q}_{mux}, respectively.

The data-to-clock delay measurement error can be found by examining the relationship between the actual data-to-clock delay t_{D-Clk} and the measured data-to-clock delay t_{D-Clk}. The actual data-to-clock delay is

$$t_{D-Clk} = t^{Q}_{\Delta Clk} - t^{Q}_{\Delta D}. \tag{7}$$

The measured data-to-clock delay is the difference between the largest Ref_1-to-Ref_2 delays of the reference generator for which CSE_{OUT} fails to capture the transition of Clk and D, as discussed in Section III. The Ref_1-to-Ref_2 delays are found by characterizing the reference generator prior to the measurement. Thus the difference between measured and actual data-to-clock delay consists of two components: the error associated with the characterization of the reference generator, and the error made by approximating actual data-to-clock delay by the delay between Ref_1 and Ref_2. The measured data to clock delay, which we denote as Δt^{char}_{D-Clk}, is

$$\Delta t^{char}_{D-Clk} = t_{ref,Clk} - t_{ref,D}. \tag{8}$$

In (8), $t_{ref,Clk}$ and $t_{ref,D}$ are the characterized delays between Ref_1 and Ref_2 that correspond to the largest control word $ref\_delay[6:0]$ for which CSE_{OUT} in Fig. 3 fails to capture Clk and D, respectively. From (3):

$$t_{ref,Clk} = \frac{1}{2N}\left[\left(t^{Clk}_{oscmux} + \sum_{i=1}^{N} t^{Clk}_{osccell,i}\right)_{LH} + \left(t^{Clk}_{oscmux} + \sum_{i=1}^{N} t^{Clk}_{osccell,i}\right)_{HL} \pm \Delta t_{OUT,Clk}\right] \tag{9}$$

$$t_{ref,D} = \frac{1}{2N}\left[\left(t^{D}_{oscmux} + \sum_{i=1}^{N} t^{D}_{osccell,i}\right)_{LH} + \left(t^{D}_{oscmux} + \sum_{i=1}^{N} t^{D}_{osccell,i}\right)_{HL} \pm \Delta t_{OUT,D}\right]. \tag{10}$$

In (9) and (10), $\Delta t_{OUT,Clk}$ and $\Delta t_{OUT,D}$ designate the variation of the delay of the output buffer and off-chip circuit during the characterization of the oscillator for the control words $ref\_delay[6:0]$ that corresponds to $t_{ref,Clk}$ and $t_{ref,D}$.

The data-to-clock delay measurement is performed by subtracting the longest Ref_1-to-Ref_2 delays smaller than or equal to the sum of the delay of Δ_D, the delay of the multiplexer and the setup time of CSE_{OUT}, from the longest Ref_1-to-Ref_2 delays smaller than or equal to the sum of the delay of Δ_{Clk}, the delay of the multiplexer and the setup time of CSE_{OUT} (Fig. 3). As the Ref_1-to-Ref_2 delays are discrete with a step size Δt_{12step}, the precision error introduced into each measurement is $\pm\Delta t_{12step}$. Assuming that the Ref_1-to-Ref_2 delay is determined exactly, the measured data-to-clock delay disregarding reference generator characterization error, which we denote by $\Delta t^{measxr}_{D-Clk}$, is given by

$$\Delta t^{measxr}_{D-Clk} = \left(t^{Clk}_{\Delta Clk} + t^{Clk}_{mux,Clk} + t^{Clk}_{CSEout,SU}\right) - \left(t^{D}_{\Delta D} + t^{D}_{mux,D} + t^{D}_{CSEout,SU}\right) \pm \Delta t_{12step}. \quad (11)$$

The overall measurement error is the difference between the measured and the actual data-to-clock delay:

$$\Delta t^{char}_{D-Clk} - t_{D-Clk} = \left(\Delta t^{char}_{D-Clk} - \Delta t^{measxr}_{D-Clk}\right) + \left(\Delta t^{measxr}_{D-Clk} - t_{D-Clk}\right). \quad (12)$$

The systematic and random error of data-to-clock delay can now be found by combining (7)–(12) and identifying constant and random component of the error expression. Systematic error is

$$\begin{aligned} sys\_err(t_{D-Clk}) = {} & \overline{(t_{mux,Clk} - t_{mux,D})} \\ & + \overline{\left(t^{Clk}_{CSEout,SU} - t^{D}_{CSEout,SU}\right)} \\ & + \frac{1}{2}\left[\overline{\left(t^{Clk}_{osccell} - t^{D}_{osccell}\right)_{HL}} \right. \\ & \left. - \overline{\left(t^{Clk}_{osccell} - t^{D}_{osccell}\right)_{LH}}\right] \\ & + \frac{1}{2N} sys\_err(instr). \end{aligned} \quad (13)$$

The first term in (13) is the systematic error of the multiplexer due to the separate paths and different low-to-high and high-to-low delays. The second is due to the different low-to-high and high-to-low delays of the capturing flip-flop. The third term is due to the mismatch between low-to-high and high-to-low delay of the delay cells in the ring oscillator. To minimize this source of systematic error, all the circuits were designed for equal rising and falling delays in the typical process. As the delay cells are noninverting, their low-to-high and high-to-low delays consist of individual delays of equal number of "up" and "down" transitions. Thus, low-to-high and high-to-low delays of the oscillator delay cells are similar, which keeps the third term in (13) small even in a skewed process corner. The last term in (13) is the accuracy of the frequency counter.

Random error of the delay measurement is caused by process, voltage and temperature variation and it is characterized by its standard deviation σ_{D-Clk}. From (7)–(12), the random error of the data-to-clock delay has two components: the random error associated with the characterization of the reference generator σ_{char}, and the random error associated with the measurement σ_{measxr}:

$$\sigma^2_{D-Clk} = \sigma^2_{char} + \sigma^2_{measxr}. \quad (14)$$

The random error due to the characterization of the reference generator in ring oscillator mode can be obtained by inspecting the mismatch between $\Delta t^{measxr}_{D-Clk}$ and Δ^{char}_{D-Clk}, (7)–(10), (12). Assuming that the reference signals Ref_1 and Ref_2 are taken from the kth cell in the ring oscillator, this mismatch can be expressed as follows:

$$\begin{aligned} & \Delta t^{measxr}_{D-Clk} - \Delta t^{char}_{D-Clk} \\ & \quad = t^{meas}_{osccell,k,Clk} - t^{meas}_{osccell,k,D} \\ & \qquad - \frac{1}{2N}\left(\sum_{i=1}^{N} t^{char}_{osccell,i,Clk} - \sum_{i=1}^{N} t^{char}_{osccell,i,D}\right)_{LH} \\ & \qquad - \frac{1}{2N}\left(\sum_{i=1}^{N} t^{char}_{osccell,i,Clk} - \sum_{i=1}^{N} t^{char}_{osccell,i,D}\right)_{HL} \\ & \qquad - \frac{1}{2N}\left(\left(t^{Clk}_{osc,mux} - t^{D}_{osc,mux}\right)_{LH} \right. \\ & \qquad\qquad \left. + \left(t^{Clk}_{osc,mux} - t^{D}_{osc,mux}\right)_{HL}\right) \\ & \qquad \pm \frac{1}{2N}\Delta t_{OUT,Clk} \pm \frac{1}{2N}\Delta t_{OUT,D}. \end{aligned} \quad (15)$$

Note that the error associated with the difference between $\Delta t^{measxr}_{D-Clk}$ and Δt^{char}_{D-Clk} does not depend on the mismatch between individual oscillator cells, but rather on the difference between these mismatches for different ref_delay control words. We assume that all oscillator cells have the same delay variation $\sigma_{osccell,D-Clk}$, and that all components in (15) are independent of each other. The delay variation of the multiplexer in the oscillator (Fig. 4) is denoted as $\sigma_{osc,mux}$, and that of the off-chip circuit is denoted as σ_{OUT}. Under these assumptions, and using elementary statistical relations, we obtain the contribution of the characterization of the reference generator to the data-to-clock delay error:

$$\sigma^2_{char} = \left(1 + \frac{1}{2N}\right)\cdot\sigma^2_{osccell,D-Clk} + \frac{1}{2N^2}\sigma^2_{oscmux} + \frac{1}{2N^2}\sigma^2_{OUT}. \quad (16)$$

The first term in (16) is the result of the delay difference mismatch between individual oscillator cells. The second term comes from the delay fluctuation of the multiplexer in the ring oscillator. The third term is the random error due to the variation of the output buffer, off-chip cable, and the measurement instrument during the characterization of the reference generator.

The circuit components that participate in the measurement of data-to-clock delay of the DUT and that contribute to the data-to-clock random error are delay lines Δ_D and Δ_{Clk}, multiplexer, and capturing flip-flop (Fig. 3). In addition, the precision (i.e., the delay step size) of the reference generator causes random error of standard deviation σ_{step}, given by $\sigma_{step} = (t_{\Delta 12step}/\sqrt{2})$:

$$\sigma^2_{measxr} = \sigma^2_{\Delta D} + \sigma^2_{\Delta Clk} + \sigma^2_{MUX} + \sigma^2_{CSEout,SU} + \sigma^2_{step}. \quad (17)$$

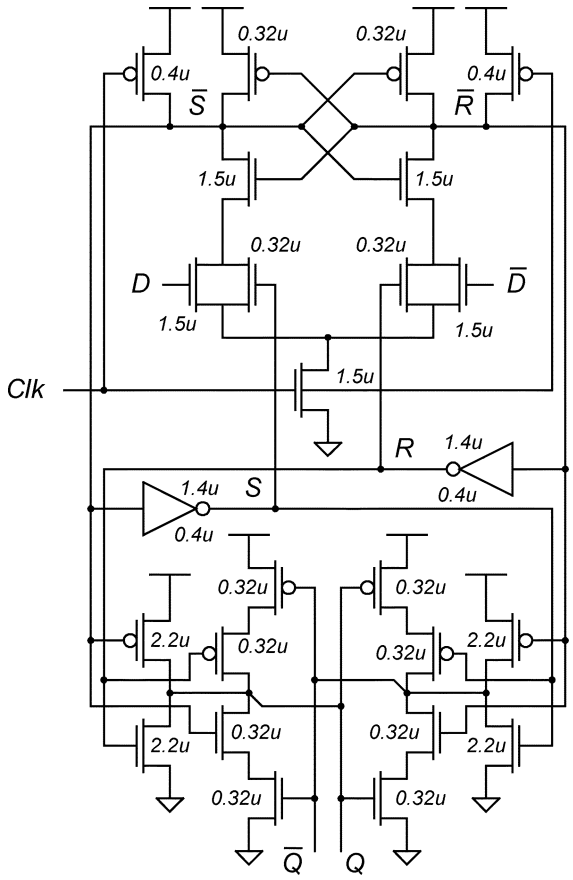


Fig. 8. Improved sense-amplifier flip-flop (im-SAFF) with transistor widths.

Combining (14), (16), and (17), we obtain the random error of data-to-clock delay:

$$\sigma^2_{D-Clk} = \sigma^2_{\Delta D} + \sigma^2_{\Delta Clk} + \sigma^2_{MUX} + \sigma^2_{CSEout,SU} + \sigma^2_{step} + \left(1 + \frac{1}{2N}\right) \cdot \sigma^2_{osccell,D-Clk} + \frac{1}{2N^2}\sigma^2_{oscmux} + \frac{1}{2N^2}\sigma^2_{OUT}. \quad (18)$$

The process of determining the error of clock-to-output delay is similar to that of the data-to-clock time, described above. Systematic error of clock-to-output delay is

$$sys\_err(t_{Clk-Q}) = \overline{(t_{mux,Q} - t_{mux,Clk})} + \overline{\left(t^{Q}_{CSEout,SU} - t^{Clk}_{CSEout,SU}\right)} + \frac{1}{2}\left[\overline{\left(t^{Q}_{osccell} - t^{Clk}_{osccell}\right)_{HL}} - \overline{\left(t^{Q}_{osccell} - t^{Clk}_{osccell}\right)_{LH}}\right] + \frac{1}{2N} sys\_err(instr). \quad (19)$$

Random error of clock-to-output delay is

$$\sigma^2_{Clk-Q} = \sigma^2_{\Delta D} + \sigma^2_{\Delta Clk} + \sigma^2_{MUX} + \sigma^2_{CSEout,SU} + \sigma^2_{step} + \left(1 + \frac{1}{2N}\right) \cdot \sigma^2_{osccell,Clk-Q} + \frac{1}{2N^2}\sigma^2_{oscmux} + \frac{1}{2N^2}\sigma^2_{OUT}. \quad (20)$$

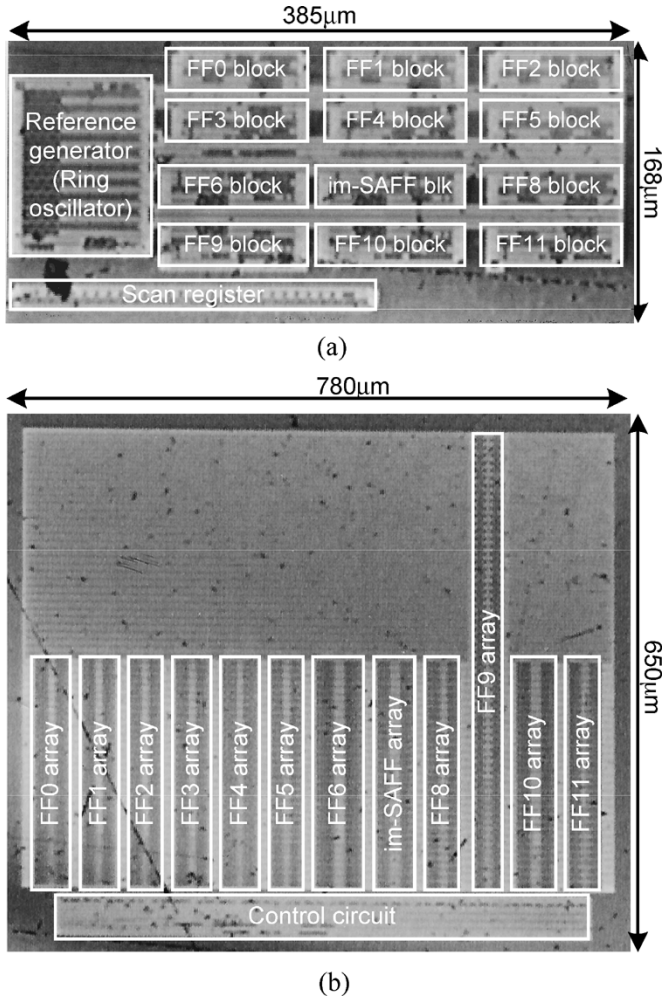


Fig. 9. Chip micrographs. (a) Micrograph of delay measurement test circuit. (b) Micrograph of power consumption measurement test circuit.

The only off-chip-related error is due to measuring the ring oscillator period. This error is reduced by a factor of $1/N$ due to the multiple stages in the ring oscillator. In the next section, we plug numbers into (13), (18), (19), and (20) to quantify the errors in our measurement circuit.

V. Results

We present the delay and power consumption measurements for the improved sense-amplifier flip-flop (im-SAFF, Fig. 8) [10], [12], which we fabricated in a 0.11-μm CMOS process [13] operating at 1.2 V. The chip micrograph of the test circuit used for measuring delay and power consumption of im-SAFF and 11 other CSEs is shown in Fig. 9. The flip-flop is loaded with 14 minimal inverters at both outputs Q and $\overline{Q}$, and its drivers are sized so that their input capacitance is four time smaller than the load they drive (fan-out of 4 load).

Fig. 10(a)–(d) shows the im-SAFF clock-to-output delay, setup, and hold characteristics for both high-to-low and low-to-high transitions from a single wafer measured at room temperature ($T = 25$ °C). Fig. 11 shows the measured power consumption at a clock frequency of 257 MHz. A summary of the measured characteristics of the im-SAFF is given in Table I.

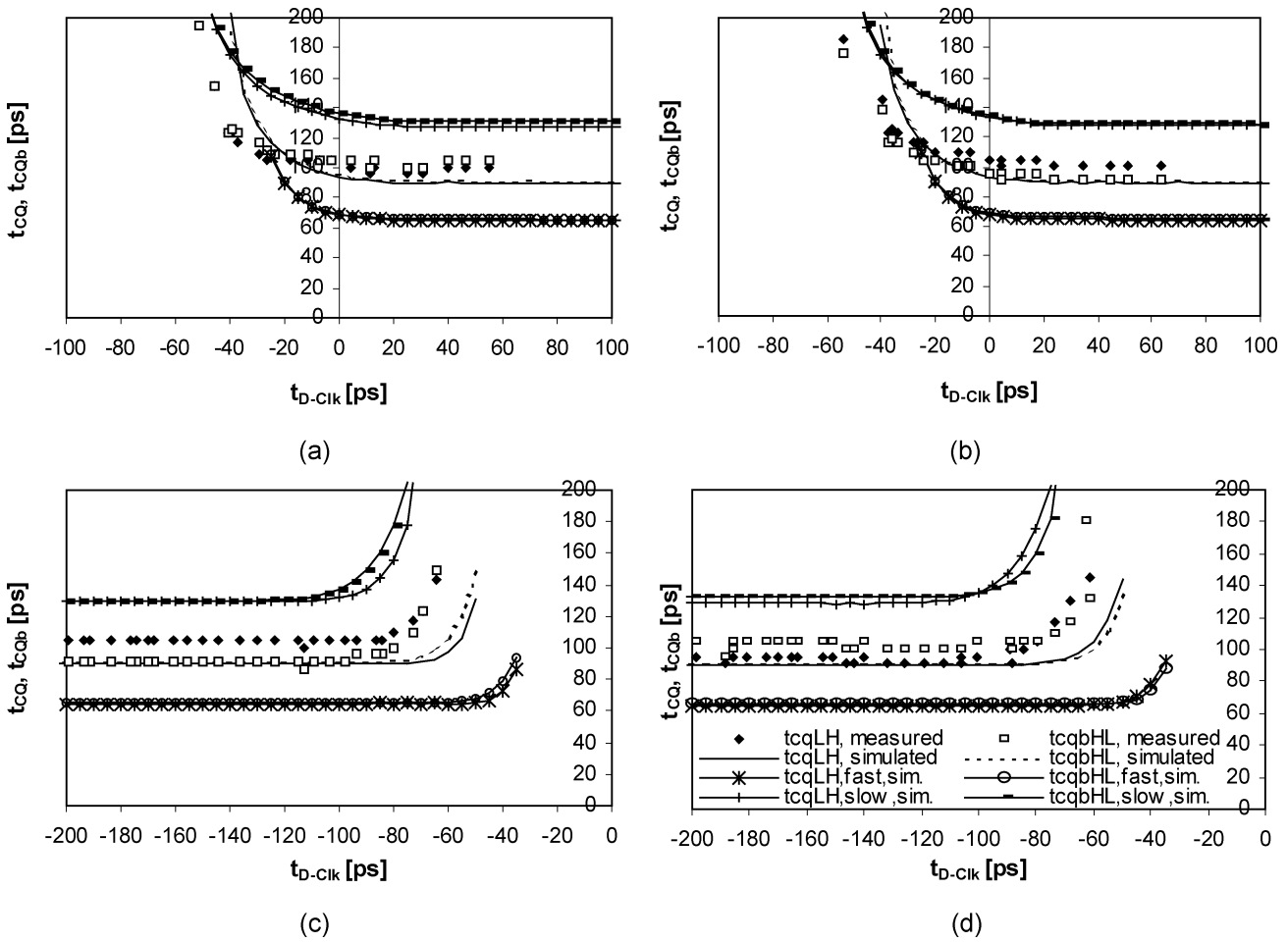


Fig. 10. Measured and simulated clock-to-output characteristic of im-SAFF (fast/slow corner designates fast/slow both nMOS and pMOS transistor models, $V_{DD} = 1.2$ V (typical), $T = 25°$C (typical): (a) setup, low-to-high input transition; (b) setup, high-to-low input transition; (c) hold, low-to-high input transition; (d) hold, high-to-low input transition.

In order to find the measurement error, we estimate all error components described in Section IV. A portion of the systematic error occurs due to unmatched multiplexer delay and CSE_{OUT} setup time for low-to-high and high-to-low D and Q transition (Fig. 3), (13), (19). The estimate of this error, obtained from SPICE simulation of the multiplexer and CSE_{OUT}, is 2 ps. The mismatch between the low-to-high and high-to-low delay of a single ring oscillator cell, estimated using simulation, is typically of the order of 1–2 ps, and always below 4 ps for all data-to-clock and clock-to-output delays of interest. For the characterization of the oscillator, we used an Agilent 54845B oscilloscope, which has a built-in frequency counter, with accuracy of 1.1 ps in the range of measured oscillation periods. From (13) and (19), the effect of this accuracy to both data-to-clock and clock-to-output systematic error is 0.055 ps.

An estimate of the process variation impact to the ring oscillator random error can be obtained from the distribution of the oscillation periods over a large number of dies on the same wafer. Measured standard deviation of this distribution over 20 dies from the same wafer is $\sigma_{D-Clk} = \sigma_{Clk-Q} = 8.9$ ps for oscillator periods typically used to measure t_{D-Clk} and t_{Clk-Q}. This estimation is pessimistic, since the variation of the transistor parameters is typically larger over different dies than for transistors on the same die.

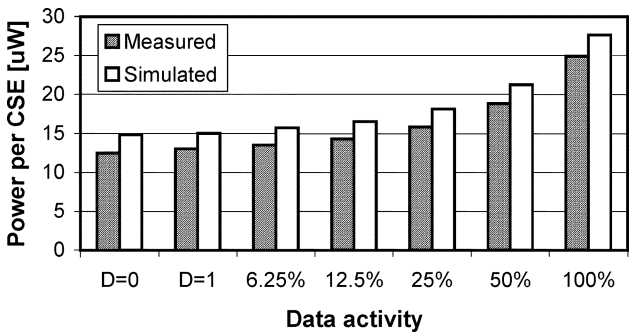


Fig. 11. Measured and simulated power consumption of im-SAFF.

TABLE I
SIMULATED AND MEASURED DELAY PARAMETERS OF IM-SAFF

Parameter	Measured
Constant Clk-Q delay, $t_{CQ,constant}$	105ps
Optimal setup time, $t_{SU,opt}$	-27ps
Minimum Data-to-Q delay, $t_{DQ,min}$	88ps
Hold time, t_{HOLD}	73ps

The random error due to the delay mismatch of the different inputs to the multiplexer M_1 (Fig. 3) cannot be measured in our circuit. In order to estimate this portion of the error, we

TABLE II
ESTIMATED ERRORS OF THE COMPONENTS OF THE TEST CIRCUIT

Error	Description	Systematic error	Standard deviation
$(t^{Clk}_{mux}-t^{D}_{mux})+(t^{Clk}_{CSEout,SU}-t^{D}_{CSEout,SU})$, $(t^{Q}_{mux}-t^{Clk}_{mux})+(t^{Q}_{CSEout,SU}-t^{Clk}_{CSEout,SU})$, σ_{mux}	Multiplexor error and setup time error of capturing storage element CSE_{OUT}	2ps	1.7ps
$\sigma_{CSEout,SU}$	Standard deviation of the setup time of CSE_{OUT}		0.6ps
$\sigma_{\Delta D}$, $\sigma_{\Delta Clk}$	Standard deviation of delay lines Δ_D and Δ_{Clk}		<0.5ps
σ_{prec}	Precision error due to minimal delay step between Ref_1 and Ref_2		3.5ps
$\sigma_{osccell,D-Clk}$, $\sigma_{osccell,Clk-Q}$	Standard deviation of oscillator cell delay difference		8.9ps
$\frac{1}{2}\left[\overline{(t^{Clk}_{osccell}-t^{D}_{osccell})_{HL}}-\overline{(t^{Clk}_{osccell}-t^{D}_{osccell})_{LH}}\right]$, $\frac{1}{2}\left[\overline{(t^{Q}_{osccell}-t^{Clk}_{osccell})_{HL}}-\overline{(t^{Q}_{osccell}-t^{Clk}_{osccell})_{LH}}\right]$	Difference of the mismatches between LH and HL delays of oscillator cell, when measuring *Clk* and *D*, or *Q* and *Clk*	4ps	
σ_{oscmux}	Standard deviation of delay of the multiplexor in the oscillator		<0.5ps
$sys\_err(instr)$, σ_{OUT}	Systematic error and standard deviation of delay of output buffers and off-chip circuitry	1.1ps	18ps
$sys\_err(t_{D-Clk})$, σ_{D-Clk}	**Data-to-clock error**	**6ps**	**10ps**
$sys\_err(t_{Clk-Q})$, σ_{Clk-Q}	**Clock-to-output error**	**6ps**	**10ps**

performed SPICE Monte Carlo simulation assuming Gaussian distributions of channel length, channel width, and threshold voltage of all transistors in the separate paths through the multiplexer. The standard deviations of the Gaussian distributions used in Monte Carlo simulation were estimated to be one third of the standard deviation of the process obtained from the foundry. The simulated standard deviation of the delay of multiplexer M_1 in Fig. 3 is 1.7 ps.

Overdesign of the power grid was used to minimize supply voltage drops. In order to estimate the effect of supply voltage and substrate voltage variations to the random error, we found by simulation that supply and ground voltage bounce due to IR drop is 0.5% V_{DD} and 0.3% V_{DD}, respectively. In addition, the estimated IR drop is very close to deterministic as it always occurs in the same scenario and no other circuit on the chip switches during the measurement. The external voltage supply tolerance is 0.1% for the range of current and supply voltage used. Each percent of supply voltage change causes 0.8% delay change, and each percent of ground voltage change causes 0.5% percent delay change. As a result, even the largest random error due to the supply voltage tolerance is less than 0.5 ps.

The estimation of the setup time error of the capturing flip-flop CSE_{OUT} requires special attention due to high sensitivity of the flip-flop operating near the metastable region to power supply fluctuations. SPICE simulations show that the data-to-clock delay for which the flip-flop definitely captures the incoming data varies by ± 0.6 ps when the supply voltage changes by $\pm 1\%$ around the nominal voltage. We use this value for the random error in the setup time of the capturing flip-flop, $\sigma_{CSEout,SU}$, which provides a safety margin for both external supply voltage variation and on-chip IR drop.

The variation of the junction temperature due to the power consumption is negligible since the test circuit operates at a low switching frequency (less than 1 MHz).

The precision of the measurement (σ_{step}) is obtained using the average measured step size $\Delta t_{12step} = 5$ ps.

The component of the random error due to the delay fluctuation of the output buffer pads and off-chip circuits (σ_{OUT}) was estimated by observing the statistics of repeated measurement of the oscillation periods of the ring oscillator. For all measured oscillation periods, the standard deviation of this measurement is between 10.6 and 18 ps. The precision of oscilloscope is 2 ps. From (18), the effect of the output buffer pads and off-chip circuits is characterized by standard deviation of $\sigma_{OUT} = 1.3$ ps.

Table II summarizes the individual components of the measurement error. Using (13), (18), (19), and (20), we obtain systematic and random errors of data-to-clock and clock-to-output delay, given in Table II.

VI. CONCLUSION

We presented a test circuit and automated procedure for the measurement of delay and power consumption of clocked storage elements. The test circuit is useful to validate simulation models and select between alternative CSEs that show only small differences in simulation. The accuracy and precision for the measurement of relatively small delays is enhanced because all measurements are performed on-chip using common signal paths wherever possible. In addition, the error introduced by chip I/O and instrumentation is minimized. The estimated systematic and random errors, obtained from the measurements and SPICE simulation, are 6 and 10 ps, respectively. The accuracy and precision are demonstrated by the ability of the test

circuit to measure the variable portion of the clock-to-output delay characteristic. The power consumption is measured by isolating the power supply of an array of identical CSEs with variable input data activity. The presented circuit is extendable for inexpensive, yet precise and accurate, measurement of arbitrary on-chip delays.

REFERENCES

[1] P. Hofstee *et al.*, "A 1-GHz single-issue 64b PowerPC processor," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2000, pp. 92–93.

[2] A. Jain *et al.*, "A 1.2 GHz alpha microprocessor with 44.8 GB/s chip pin bandwidth," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2001, pp. 240–241.

[3] R. Heald *et al.*, "A third generation SPARC V9 microprocessor," *IEEE J. Solid-State Circuits*, vol. 35, pp. 1526–1538, Nov. 2000.

[4] S. H. Unger and C. J. Tan, "Clocking schemes for high-speed digital systems," *IEEE Trans. Computers*, vol. C-35, pp. 880–895, Oct. 1986.

[5] V. Stojanovic and V. G. Oklobdzija, "Comparative analysis of master-slave latches and flip-flops for high-performance and low-power systems," *IEEE J. Solid-State Circuits*, vol. 34, pp. 536–548, Apr. 1999.

[6] G. Oklobdzija *et al.*, *Digital System Clocking: High-Performance and Low-Power Aspects*. New York: Wiley, 2003.

[7] J. U. Horstmann, H. W. Eichel, and R. L. Coates, "Metastability behavior of CMOS ASIC flip-flops in theory and test," *IEEE J. Solid-State Circuits*, vol. 24, pp. 146–157, Feb. 1989.

[8] Q. Huang and R. Rogenmoser, "Speed optimization of edge-triggered CMOS circuits for gigahertz single-phase clocks," *IEEE J. Solid-State Circuits*, vol. 31, pp. 456–465, Mar. 1996.

[9] T. Maeda *et al.*, "An ultra-low-power-consumption high-speed GaAs quasidifferential switch flip-flop (QD-FF)," *IEEE J. Solid-State Circuits*, vol. 31, pp. 1361–1363, Sept. 1996.

[10] B. Nikolic *et al.*, "Improved sense-amplifier-based flip-flop: Design and measurements," *IEEE J. Solid-State Circuits*, vol. 35, pp. 876–884, June 2000.

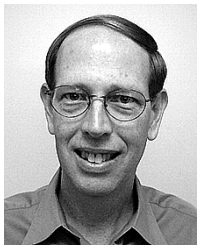
[11] H. D. Young, *Statistical Treatment of Experimental Data*. New York: McGraw-Hill, 1962.

[12] V. G. Oklobdzija and V. Stojanovic, "Flip-flop," U.S. Patent 6,232,810, May 15, 2001.

[13] Y. Takao *et al.*, "A 0.11 μm CMOS technology with copper and very-low-k interconnects for high-performance system-on-a-chip cores," in *IEDM Tech. Dig.*, Dec. 2000, pp. 559–562.

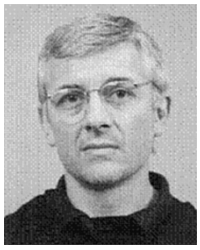
Nikola Nedovic (M'03) was born in Belgrade, Yugoslavia, in 1973. He received the Dipl.Ing. degree in electrical engineering from the University of Belgrade, Yugoslavia, in 1998, and the Ph.D. degree from the University of California at Davis in 2003.

In 2001, he joined Fujitsu Laboratories of America, Inc., Sunnyvale, CA, where he works in the area of high-performance and low-power VLSI circuits. He is a coauthor of ten papers and one book, and holds one U.S. patent with two other patents pending. His research interests include circuit design and clocking strategies for high-speed and low-power applications.



William W. Walker (M'79) received the A.B. degree in physics and applied math in 1976 and the M.S.E.E. degree in 1978, both from the University of California at Berkeley.

From 1978 to 1981, he was a Staff Engineer with IBM Corporation, East Fishkill, NY, and from 1981 to 1983, at IBM, Burlington, VT. At IBM, he was involved in the development of the LDD MOS transistor. From 1984 to 1991, he was a Senior Engineer with Integrated CMOS Systems, Inc., Sunnyvale CA. From 1991 to 2000, he was an Engineering Manager with Hal Computer Systems, Inc., Campbell, CA, where he developed circuits for the first 64-bit SPARC microprocessors. Since 2000, he has been with Fujitsu Laboratories of America, Sunnyvale, where he is currently Director of the LSI Technology Development Laboratory. His research interests include high-speed and low-power digital circuit design for microprocessors, and radio-frequency CMOS circuits for telecommunications.



Vojin G. Oklobdzija (M'82–SM'88–F'96) received the Dipl. Ing. degree from the Electrical Engineering Department of the University of Belgrade, Yugoslavia, in 1971, and the Ph.D. degree from the University of California at Los Angeles in 1982.

From 1982 to 1991, he was with the IBM Thomas J. Watson Research Center, where he made contributions to the development of RISC processors and supercomputer design. In the course of this work, he obtained several patents, the most notable one on register renaming, which enabled a new generation of computers. From 1988 to 1990, he was an IBM Visiting Faculty Member at the University of California at Berkeley. Since 1991, he has been a Professor at the University of California at Berkeley, and has served as a consultant to many companies, including Sun Microsystems, Bell Laboratories, Hitachi, Fujitsu, SONY, Intel, Samsung, and Siemens, where he was Principal Architect for the Infineon TriCore processor. He holds 12 U.S. and seven international patents, with five patents pending. He has published more than 140 papers, three books, and 12 book chapters in the areas of circuits and technology, computer arithmetic, and computer architecture. He has given over 150 invited talks and short courses in the U.S., Europe, Latin America, Australia, China, and Japan.

Prof. Oklobdzija serves as Associate Editor for the IEEE TRANSACTIONS ON COMPUTERS, IEEE TRANSACTIONS ON VLSI SYSTEMS, *Journal of VLSI Signal Processing* , and IEEE MICRO. He served on the IEEE ISSCC program committee from 1996 to 2003, among numerous other conference committees. He was a General Chair of the 13th Symposium on Computer Arithmetic. He is a Distinguished Lecturer of the IEEE Solid-State Circuits Society.