Comparative Analysis of Master–Slave Latches and Flip-Flops for High-Performance and Low-Power Systems

Vladimir Stojanovic and Vojin G. Oklobdzija, *Fellow, IEEE*

***Abstract*—In this paper, we propose a set of rules for consistent estimation of the real performance and power features of the flip-flop and master–slave latch structures. A new simulation and optimization approach is presented, targeting both high-performance and power budget issues. The analysis approach reveals the sources of performance and power-consumption bottlenecks in different design styles. Certain misleading parameters have been properly modified and weighted to reflect the real properties of the compared structures. Furthermore, the results of the comparison of representative master–slave latches and flip-flops illustrate the advantages of our approach and the suitability of different design styles for high-performance and low-power applications.**

I. Introduction

INTERPRETATION of published results comparing various latches and flip-flops has been very difficult because of the use of different simulation methods for their generation and presentation. In this paper, we establish a set of rules with the goal of making the comparisons fair and realistic. Simulation of the latches and flip-flops obtained by using different design styles makes this analysis more difficult in trying to achieve consistent and comparable results. There are two major results produced in the course of this work:

1) definition of the relevant set of parameters to be measured and rules for weighting their importance;
2) a set of relevant simulation conditions, which emphasize the parameters of interest.

The simulation and optimization procedures have high performance as the primary goal, but we have also paid attention to the possible reductions in power consumption, given that the limitation in performance is usually imposed by the available power budget.

II. Analysis

A. Power Considerations

Power consumption of a circuit depends strongly on its structure and the statistics of the applied data. Thus, power measurements should be conducted for the range of different data patterns comprising the worst and the best cases. Our goal is not only to find the range of power consumption for all the possible cases but also to define special patterns that are to reveal the major sources of power dissipation for the given family of master–slave latches and flip-flops (for specific applications where the statistics of data is known in advance, like in data processing, it is desirable to optimize the structure for the given data distribution).

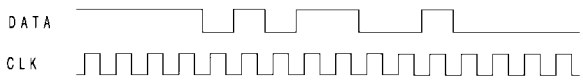

Fig. 1. The pseudorandom sequence.

In our simulation, the data activity rate α presents the average number of output transitions per clock cycle. We have applied four different data sequences where ...010 101 010..., $\alpha = 1$ reflects the maximum internal dynamic power consumption. However, depending on the structure, the sequence ...111 111... can in some cases dissipate more power. In general, a pseudorandom sequence with equal probability of all transitions (data activity rate $\alpha = 0.5$) is considered to reflect the average internal power consumption given the uniform data distribution; see Fig. 1. The sequence ...111 111 111..., $\alpha = 0$ reflects the power dissipation of precharged nodes. The sequence ...000 000 000..., $\alpha = 0$ reflects minimum power consumption.

Depending on the size of the precharged and static parts, internal dynamic power consumption can be estimated based on the analysis of Fig. 2.

The starting point is a well-known formula for estimation of dynamic power consumption

$$P_d = fC_{\text{eff}}Vdd^2, \quad \text{where } C_{\text{eff}} = \sum_{i=1}^{N} \alpha_i k_i C_i$$

where

- α_i switching probability of node i (in regard to the clock cycle);
- k_i swing range coefficient of node i ($k_i = 1$ for rail-to-rail swing);
- C_i total capacitance of node i;
- f clock frequency;
- Vdd rail-to-rail voltage range (supply voltage).

This strategy has been generalized in Fig. 2 to describe the differences in switching activity, and therefore power consumption, for different design styles.

Capacitances C_{total}, $C_{\text{precharge}}$, and C_{out} are calculated from node to node in the given topology taking into account the C_i and k_i coefficient of each node in the circuit. This

Manuscript received April 10, 1998; revised October 15, 1998.

V. Stojanovic is with the Faculty of Electrical Engineering, University of Belgrade, Serbia, Yugoslavia.

V. G. Oklobdzija is with Integration Corp., Berkeley, CA 94708 USA (e-mail: vojin@nuc.berkeley.edu).

Publisher Item Identifier S 0018-9200(99)02434-8.

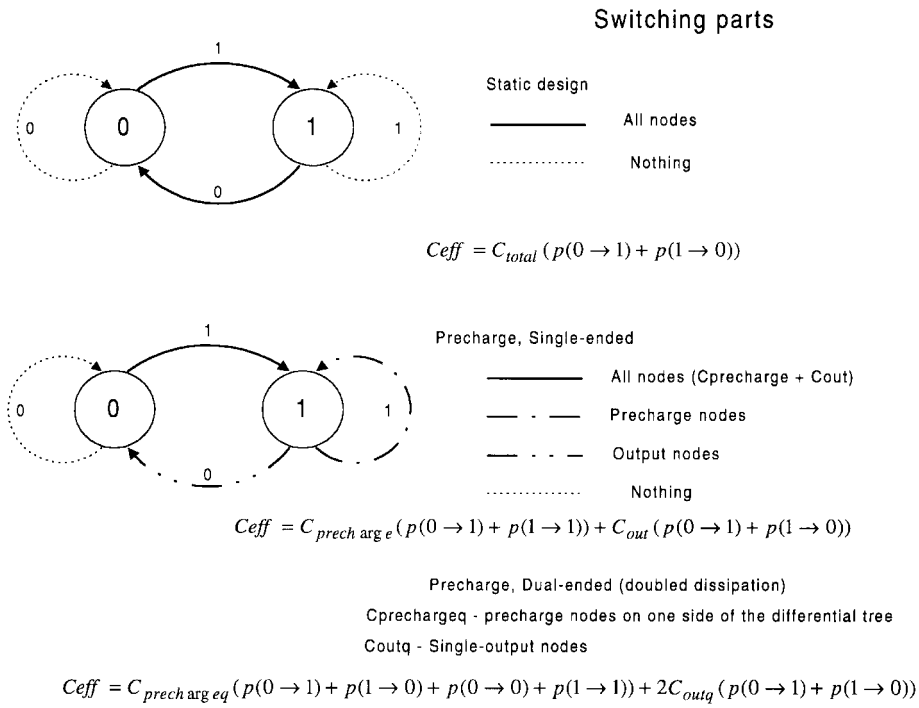

Fig. 2. Sources of internal, dynamic power consumption.

approach is essential for determining the strategy for the power budget needed to meet the performance requirements.

As all nodes in static structures switch uniformly and only for zero-to-one and one-to-zero transitions, C_{total} represents the total effective capacitance of the circuit charged and/or discharged in each cycle.

Semidynamic structures are generally composed of a dynamic (precharged) front end and static output part. This is why we designated two major effective capacitances: $C_{\text{precharge}}$ and C_{out}, each representing the corresponding part of the circuit. It is shown in Fig. 2 that these two capacitances have different charging and discharging activities.

In Fig. 2, semidynamic structures were further differentiated into single-ended and differential structures. This was done in order to emphasize the switching activity's independence from data statistics (and therefore higher average power consumption) in the case of differential structures. The total effective precharge capacitance is comprised of two effective capacitances of the same size, $C_{\text{prechargeQ}}$ and $C_{\text{prechargeQb}}$, which represent the two complementary halves of the precharged differential tree.

The second problem that we encountered was the measurement of the power dissipated by the latch. We used the .MEASURE average power statement in HSPICE to measure the power dissipation of interest. The results were compared with the earlier power-measurement method presented in [2] and showed the same level of accuracy.

We identified three main sources of power dissipation in the latch:

- *internal power dissipation of the latch,* excluding the power dissipated for switching the output loads;
- *local clock power dissipation,* which presents the portion of the power dissipated in the local clock buffer driving the clock input of the latch;
- *local data power dissipation,* which presents the portion of the power dissipated in the logic stage driving the data input of the latch.

This provides us with a good illustration of the total power dissipated in the latch and its surroundings. If we do not take into account all those sources of power consumption, the results will be misleading because of the possible tradeoffs among the three. Parameter *total power* refers to the sum of all three measured kinds of power. We excluded the power spent on switching of the output loads because its addition can make the results misleading in a way that for the load that we applied, it presented a large portion of the latch's intrinsic power consumption.

Another important detail is that we measured the power dissipated by the circuit driving the inputs of the latch to determine the *local clock and data power dissipation* parameters. Only the portion of that power (the one dissipated on driving the input capacitances) was calculated as relevant. But there is still a question of the overall power dissipated by the clock. Structures with large clock load require larger inverters in the clock tree and increase the power consumed by the clock. The problem is partially solved by the introduction of local clock regenerators like those used in the PowerPC 603 processor [12], which are usually used to generate the second phase of the clock and/or to incorporate the scan signal. Given the variety of clocking techniques, one cannot include these details in the overall comparison of different latches. However, one has to be aware of that fact and judge the results of the comparison accordingly.

B. Timing

Many authors [5], [14] refer to *Clk-Q* delay, setup, and hold times as the timing parameters of flip-flops and master–slave latches. For the purpose of clarification, we will repeat the

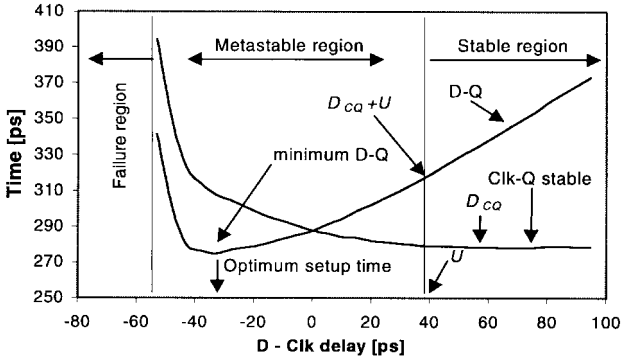

Fig. 3. StrongArm110 flip-flop, stable, metastable, and failure regions.

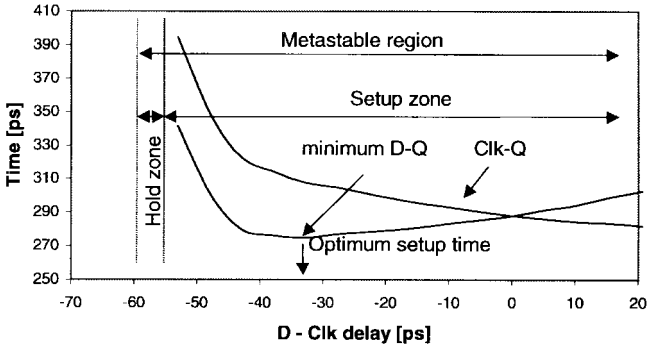

Fig. 4. StrongArm110 flip-flop, critical timing zones.

definitions of these parameters as stated by Unger and Tan in [1], where terminal C corresponds to Clk:

- D_{CQ} propagation delay from the C terminal to the Q terminal, assuming that the D signal has been set early enough relative to the leading edge of the C pulse;
- U *setup time*, the minimum time between a D change and the triggering (latching) edge of the C pulse such that, even under the worst conditions, the output Q will be guaranteed to change so as to become equal to the new D value, assuming that the C pulse is sufficiently wide;
- H *hold time*, the minimum time that the D signal must be held constant *after* the triggering (latching) edge of the C signal so that, even under worst case conditions, and assuming that the most recent D change occurred no later than U prior to the triggering (latching) edge of C, the Q output will remain stable after the end of the clock pulse (it is not unusual for the value of this parameter to be negative).

Unger and Tan made this concept more precise by requiring that the change in D occur sufficiently early so that making it appear any earlier would have no effect on when Q changes.

Given that the designer's interest is to use every possible fraction of the clock cycle, data evaluation is usually done as close to the raising edge of the clock as the setup time permits.

Regarding the nature of the latching mechanisms in flip-flops and master–slave latches, we tried to discover the relationship between the *Clk-Q* delay and the proximity of the last change in data that could cause the change at the outputs.

Before we proceed, we should define the stable, metastable and failure regions, illustrated in Fig. 3. The stable region is the region of the Data-Clk (the time difference between the last transition of data and the latching clock edge) axis in which *Clk-Q* delay do not depend on Data-Clk time. As Data-Clk decreases, at a certain point *Clk-Q* delay starts to rise monotonously and ends in failure. This region of the Data-Clk axis is the metastable region. The metastable region is defined as the region of unstable *Clk-Q* delay, where the *Clk-Q* delay rises exponentially as indicated by Shoji in [13]. Changes in data that happen in the failure region of D-Clk are unable to be transferred to the outputs of the circuit.

The question arises: how much we can let the *Clk-Q* delay be degraded in the metastable region and still benefit from the increase in performance (due to the decreased D-Q) while maintaining reliability of operation?

D_{CQ}, as defined previously, is the value of *Clk-Q* delay (Fig. 3) in the stable region, and U parameter is the minimum point on the D-Clk axis that is still a part of the stable region.

Parameters D-Q and *Clk-Q*, of the flip-flop used in the StrongArm110 processor [8] are presented in Fig. 3 as a function of D-Clk delay. In setup-time region, the *Clk-Q* curve rises monotonously from the value of *Clk-Q* delay measured in the stable region. On the other hand, the D-Q curve has its minimum as we move the last transition of data toward the latching edge of the clock. It is clear that beyond that *minimum* D-Q point, it is no longer applicable to evaluate the data closer to the rising edge of the clock. We refer to D-Clk delay at that point as the optimum setup time, which presents the limit beyond which the performance of the latch is degraded and the reliability is endangered.

Our interest is to minimize the D-Q delay (or $D_{\mathrm{CQ}}+U$, as defined by Unger and Tan [1]), which presents the portion of time that the flip-flop or master–slave structure takes out of the clock cycle. It will be shown later that the comparisons in terms of *Clk-Q* delay as a relevant performance parameter are misleading because they do not take into account the setup time and therefore the effective time taken out of the clock cycle. Since $D_{\mathrm{CQ}}+U >$ *minimum* D-Q (as defined in Fig. 3), it is obvious that the cycle time will be reduced if the change in data is allowed to arrive no later than the *optimum setup time* before the trailing edge of the clock.

In light of the reasons presented above, we accepted the *minimum* D-Q delay as the *delay* parameter of a flip-flop or master–slave latch.

As shown in Fig. 4, the metastable region consists of setup and hold zones. The last data transition can be moved all the way to the optimum setup time. The first or a late data transition is allowed to come after the hold zone. All the extractions of critical timing parameters should be done for the worst case corners and external conditions in order to ensure reliability.

Since the issue of timing is in fact the issue of reliability, we have to consider one more detail. Conventional master–slave structures do not have a positive hold-time requirement because of the positive setup time. This can be discussed in

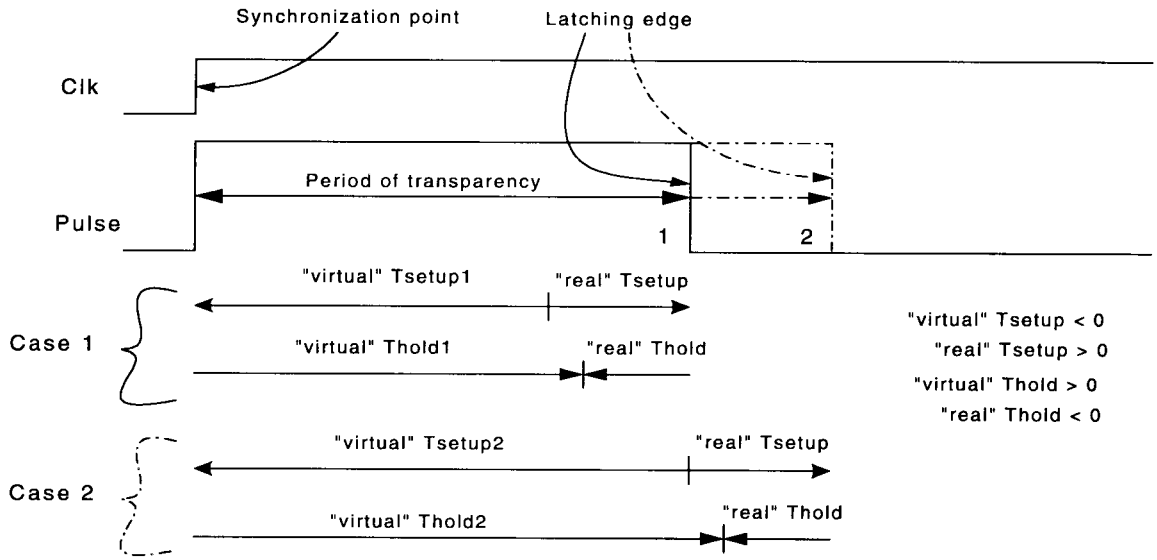

Fig. 5. Hybrid design, timing analysis.

the worst case for the hold time, i.e., a master–slave latch in a shift register. If the previous master–slave latch had the last change in data in the stable region (and thus the minimal *Clk-Q* delay after the rising edge of the clock), the following latch would have been critical if its hold time had been larger than the *Clk-Q* delay of the previous latch. As conventional master–slave structures have positive setup time, and thus in most cases negative hold time (or around zero), it is impossible for the hold time to become greater than the *Clk-Q* delay of the previous stage.

However, a new hybrid design technique introduced by Partovi in [10], featuring a negative setup time and short transparency period, has to take into account the hold time as a reliability requirement. This is the case not only with this new technique but also with some flip-flop structures featuring negative setup time, like the sense amplifier (SA)/F-F [7] or its modification used in the StrongArm110 processor [8].

The hybrid design technique shifts the reference point of hold- and setup-time parameters from the rising edge of the clock to the falling edge of the buffered clock signal, which presents the end of the transparency period. In this way, the setup and hold times measured in reference to the rising edge of the clock (as conventionally defined for flip-flops) are the functions of the width of the transparency period because their real reference point is the end of that period (just like in custom transparent latches). Since the hybrid design style is one of the best choices for high-performance systems, we will try to clarify this concept.

This advanced concept merges the good features of both transparent-latch and flip-flop (master–slave latch) design styles.

High-performance designs often use transparent latches, which provide the so-called cycle-stealing feature and reduce the length of the clock cycle [1]. These designs provide “soft clock edge”—i.e., absorption of the clock skew—but suffer from the racethrough and positive hold-time requirements, and thus require careful timing analysis.

Techniques that use flip-flops or master–slave latches suffer from positive setup time (which reduces the performance) and sensitivity to clock skew but eliminate hold-time violation and racethrough events.

The main idea of the hybrid design technique is to shorten the transparency period of a latch to a small time interval. In this way, the risk of racethrough is nearly eliminated, and the negative setup time and the absorption of the clock skew are retained.

The operation of the hybrid structures can be easily understood if they are regarded as the classical transparent latches with the shortened period of transparency. The pulse of the clock is not shortened, but the raising edge of the clock generates the pulse that enables the transparency of the structure.

We will consider the rising edge of the clock as the synchronization point (like in flip-flop or master–slave latch designs). The real latching edge is the falling edge of the transparency pulse. The timing analysis is presented in Fig. 5.

The “real” parameters are obtained with respect to the latching edge of the pulse and do not depend on the width of the transparency period.

The “virtual” parameters are obtained with respect to the synchronization point, i.e., the rising clock edge, and thus are functions of both the “real” parameters and the width of the transparency pulse.

Case 1 and *Case 2* illustrate the methods of calculation of all the mentioned parameters and their mutual dependence. Since the width of the transparency window affects the race analysis, hold-time requirements, and clock-skew tolerance, it is very important to have local control over the transparency period.

Setup and hold parameters mentioned in the following sections are addressing the “virtual” parameters, since the goal is to view the structures from the functional point of flip-flops and master–slave latches, which both have the synchronization point.

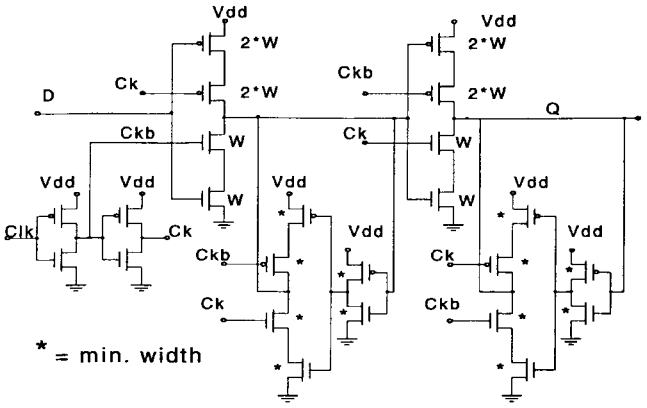

Fig. 6. Modified C^{2}MOS master–slave latch, power-delay tradeoff.

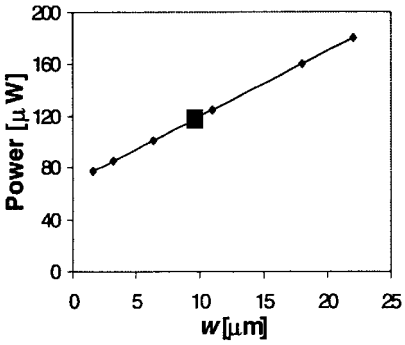

Fig. 7. Power versus generalized width w.

C. Power-Delay Product

A tradeoff between speed and power is always possible. In high-performance and low-power applications, both features are equally important. The point of minimum power-delay product is the point of optimal energy utilization at a given clock frequency.

To illustrate this, we show the power-delay tradeoffs in the case of a modified C^{2}MOS master–slave latch in Fig. 6.

Since this structure is relatively simple and symmetrical, because it consists of gated inverters, it was easy to express all critical transistor widths as functions of one variable. Given the variety of design styles, it is not always simple to express the power-delay tradeoff as a function of one common variable. This makes the optimization more difficult, along with the fact that *delay* and *power* parameters as defined in the previous analysis cannot be obtained at the same time, thus causing the optimization procedure to be iterative.

The power versus w diagram in Fig. 7 indicates the proportional relationship between power and generalized transistor width w, while Fig. 8 shows the nearly inversely proportional dependence of *delay* on generalized width parameter w. The rates of change of *delay* and power versus w are not the same, and lead to the minimum of the $\mathrm{PDP_{tot}}$ function, i.e., the point of optimal energy efficiency, marked as ■.

The symbol "■" in Figs. 7–11 marks the point of the best power-delay tradeoff. The approximate results could be derived from the first-order power and delay analysis, which will not be repeated here. On the basis of these assumptions, all the presented structures were optimized to reach the point

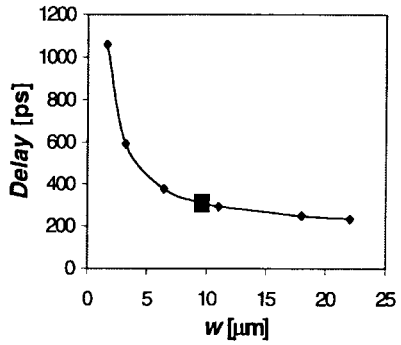

Fig. 8. *Delay* versus generalized width w.

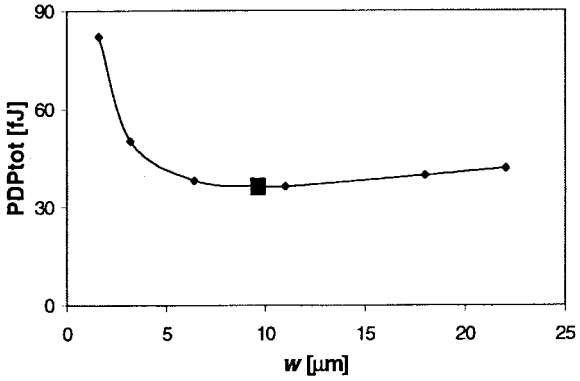

Fig. 9. Power-delay tradeoff, PDP versus generalized width w.

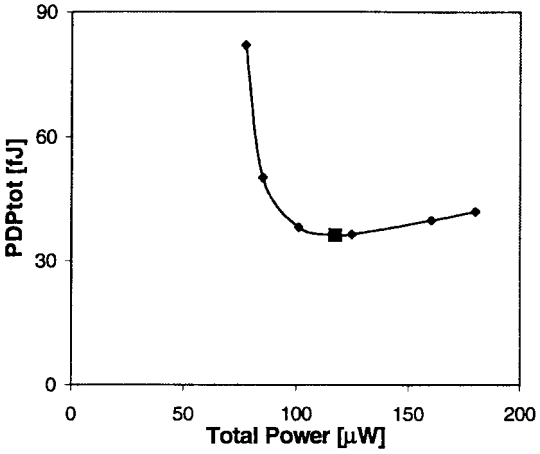

Fig. 10. Power-delay tradeoff, PDP versus power.

of minimal $\mathrm{PDP_{tot}}$. The simplified procedure presented above is not as accurate as the real optimization procedure, but was used here primarily to clarify the principal.

The $\mathrm{PDP_{tot}}$ parameter is the product of the *delay* and *total power* parameters. We have chosen $\mathrm{PDP_{tot}}$ as the overall performance parameter for comparison in terms of speed and power.

III. Simulation

A. Test Bench

To extract the parameters of interest, we defined simulation conditions, which include setting up a test bench and the selection of the device model used.

TABLE I
OUTPUT POWER DISSIPATION

$P_{Out} = \alpha \frac{fC_{Out}Vdd^2}{2}$, f=100MHz, Vdd=2V	Single-ended structures C_{Out}=C_l=200fF	Differential structures C_{Out}=2C_l=400fF, (except K6 which switches always and dissipates 80μW on output)
α=1	40μW	80μW
α=0.5	20μW	40μW

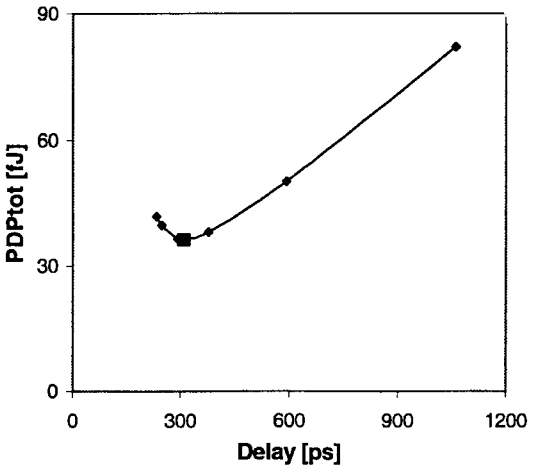

Fig. 11. Power-delay tradeoff, PDP versus *delay*.

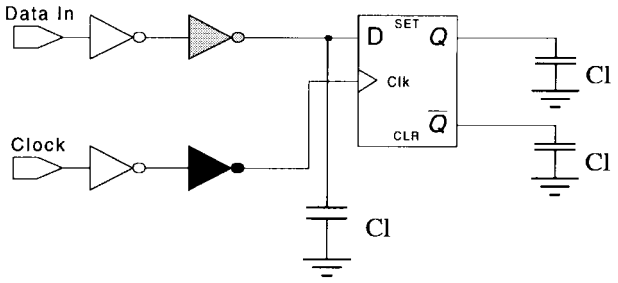

Fig. 12. The simulation test bench.

The role of the test bench shown in Fig. 12 is to provide the realistic data and clock signals, the fanout signal degradation from the previous and to the succeeding stage, and measurement of power dissipated on switching of the clock and data inputs.

Buffering inverters in Fig. 12 provide the realistic data and clock signals, which themselves are fed from ideal voltage sources. Furthermore, capacitive load at the data input simulates the fanout signal degradation from previous stages. Capacitive loads at the outputs simulate the fanout signal degradation caused by the succeeding stages.

As mentioned in the section on power considerations, there are three kinds of power dissipation that were measured in order to get the real insight in the amount of power consumed in and around the latch due to its presence.

- Local data power dissipation presents the portion of the gray inverter's power consumption (Fig. 12) dissipated on switching the data input capacitance.
- Local clock power dissipation presents the portion of the black inverter's power consumption (Fig. 12) dissipated on switching the clock input capacitance.
- Internal power dissipation includes the intrinsic power dissipated on switching the internal nodes of the circuit and excludes the power dissipated on switching the output load capacitances Cl.

Buffering inverters dissipate power even without any external load (due to their internal capacitances). Thus, we applied the following procedure. We interpolated the total measured power dissipation of the inverter over the wide range of loads. The *local clock power dissipation* parameter was then calculated as the difference in power dissipation of the black inverter when loaded with latch and when unloaded. The *local data power dissipation* parameter was calculated as the difference in power dissipations of the gray inverter when loaded with latch and Cl and when loaded only with Cl.

This approximation appeared to be fairly good for the wide range of load capacitances. Yuan and Svensson took a similar approach in [15] by setting the drain and source capacitance parameters to zero, thus minimizing the internal inverter power consumption. Since these are the major but not the only sources of internal power consumption, we applied the correction method presented above as the most accurate. The other reason was that we wanted the "real" inverters, and not the ones with degraded parameters, to drive the inputs of our latch.

The reason for exclusion of the power spent on output loads was that for the given load Cl = 200 fF, that portion of power reached the values presented in Table I. We decided to load the circuits with heavy load in order to estimate their driving capabilities and simulate the critical situations in pipelines. Thus, we decided to exclude the power spent on that load in order to get a fair picture of the circuit's power behavior.

All power measurements were conducted for 16-cycle-long data sequences. The pseudorandom data sequence used to illustrate the average power dissipation is presented in Fig. 1.

Due to the topological differences among the existing latches, some of them required a modified test bench, i.e., a dual input and/or a single output. However, these modifications did not alter the principal of the analysis approach based on the simulation conditions.

Parameters of the MOS model used in our simulations and simulation conditions are shown in Table II. For the given technology, the load capacitance Cl = 200 fF that was used as the input and the output load equals the load of the 22 minimal inverters (wp/wn = 3.2 u/1.6 u).

B. Transistor-Width Optimization

We optimized all structures in terms of both speed and power. All structures were optimized for the same fixed load in order to compare them under the same conditions. For

TABLE II
SIMULATION PARAMETERS

Technology:		Simulation:	
Channel length	.2 microns		
Min. gate width	1.6 microns	(1) Data/Clock slopes of ideal signals: 100ps	
Max. gate width	22 microns	(2) Clock duty-cycle: 50%	
| Vtp,n |	0.7V	(3) Delay calculation: between 50% points	
MOSFET Model:		(4) Data sequences: 16 clock cycles	
Level 28 modified BSIM Model		(5) Clock frequency: 100MHz	
MOS Gate Capacitance Model:			
Charge Conservation Model			
Conditions:			
Nominal	Vdd=2V T=25°C		

TABLE III
GENERAL CHARACTERISTICS

Nominal conditions	# of transistors	Total transistor width [u]	Internal power [uW]	Clock power [uW]	Data power [uW]	Total power [uW]	Delay [ps]	PDP_{tot} [fJ]
PowerPC 603	16	147	36	46	5	87	266	23
HLFF	20	162	106	18	3	127	199	25
SDFF	23	167	158	27	2	187	187	35
mC^2MOS	24	170	94	15	6	115	292	34
SA-F/F	19	214	97	18	3	118	272	32
StrongArm FF	20	215	101	18	3	122	275	34
K6 ETL	37	246	250	15	5	270	200	54
SSTC	16	147	94	22	4	120	592	71
DSTC	10	136	132	22	4	158	629	99

smaller loads, the differences in performance would become less apparent since the driving capability of the design would lose importance.

We used the Levenberg–Marquardt optimization algorithm embedded in HSPICE. The search direction of this algorithm is a combination of the steepest-descent and the Gauss–Newton method. It has a good feature of the optimization toward the goal stated in the .MEASURE statements.

The main point of the optimization is the minimization of the power-delay product, given the always-present tradeoff between power and speed. Instead of using PDP, which is the product of *Clk-Q* delay and internal power dissipation, we used PDP_{tot}.

The first step in the process is the optimization of both the *Clk-Q* delay and the *total power*, which essentially presents the optimization in terms of PDP with the addition of the *total power* parameter. The next step is the calculation of the *minimum D-Q* taken as the *delay* parameter. Now, when we obtain the second parameter *delay*, we have to make the correction in the next optimization iteration using the *delay* parameter instead of *Clk-Q* and optimizing the real PDP_{tot}. Since the measurements of the *delay* and *total power* cannot be made in one step, several iterations are needed to achieve satisfying results. We did not make the attempt in the direction of automating the process since that was not the main purpose of our research. However, the automated tools are needed especially because the existing ones consider the *Clk-Q* delay as a relevant parameter for the optimization. The following example illustrates the difference between the two approaches.

If we consider the master–slave latch and try to optimize it in terms of the classical PDP (*Clk-Q* * internal power) the result will be a minimal master latch optimized for low power and a slave latch optimized for both speed and power. The "optimized" structure will have an excessively large setup time, thus requiring the larger clock cycle to meet the timing requirements. The reason for such a result is that the optimizer does not "see" the real performance through *Clk-Q* delay. Our *delay* parameter would take that into account, and the master and slave latches would both be optimized in terms of PDP_{tot}.

The transistor widths of optimized structures are shown on each schematic. They are expressed relative to the minimum width in the given technology.

IV. RESULTS

We have chosen a set of representative latches and flip-flops. All of them have been designed for use in either high-performance or low-power processors.

The results of the simulations shown in Table III are measured for the pseudorandom data sequence (Fig. 1) with equal probability of all transitions. We assumed that the distribution of data transitions is uniform, and thus the above-mentioned sequence presents the illustration of average power consumption.

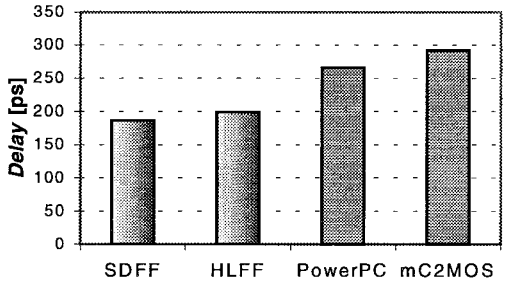

Fig. 13. Single-ended structures, overall *delay* comparison.

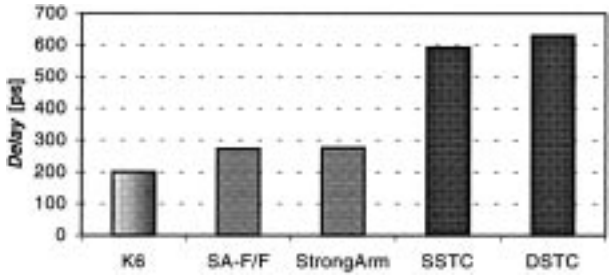

Fig. 14. Differential structures, overall *delay* comparison.

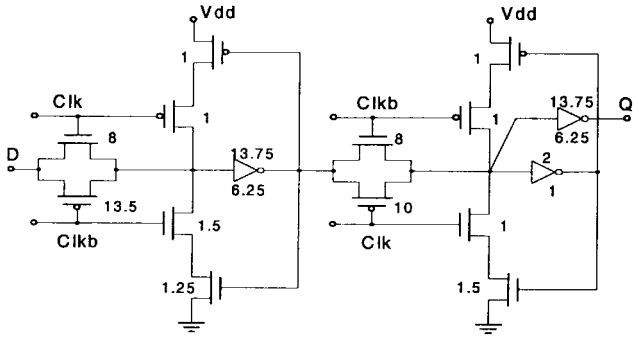

Fig. 15. PowerPC 603 master–slave latch.

Figs. 13 and 14 illustrate the speed characteristics of the presented structures. For the sake of a fair comparison, structures were differentiated in two groups, as single ended and differential. In Fig. 13, a group of hybrid structures features negative setup time and thus superior performance characteristics over the group of master–slave structures featuring positive setup time and therefore reduced performance. In Fig. 14, K6 ETL, SA-F/F, and StrongArm110 flip-flop outperform the static SSTC and DSTC master–slave latches due to the slightly negative setup-time property and shortened latency.

The PowerPC master–slave latch (Fig. 15), presented in [12], is one of the fastest classical structures. Its main advantages are a short direct path and low-power feedback. But one has to keep in mind another aspect of this structure—its large clock load, which greatly influences the total power consumption on-chip.

The modification of the standard dynamic C^2MOS master–slave latch (Fig. 16) has shown good low-power properties featuring small clock load, achieved by the local clock buffering, and low-power feedback, assuring fully static operation. The circuit is somewhat slower than the PowerPC

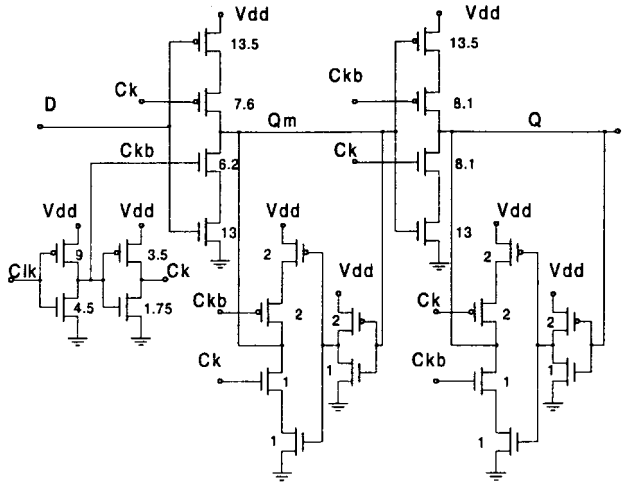

Fig. 16. Modified C^2MOS latch.

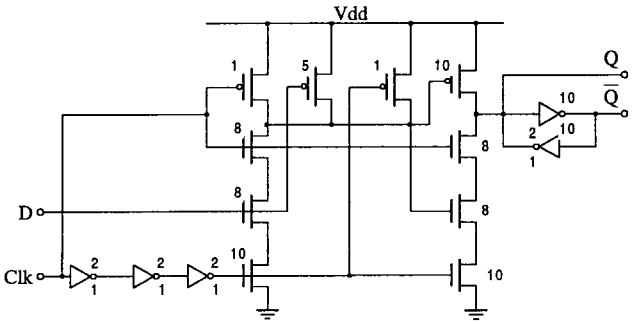

Fig. 17. HLFF.

603 master–slave latch. The faster pullup in the PowerPC 603 master–slave latch is achieved by the use of complementary pass-gates, which also increase the sensitivity to racethrough in the period of one gate delay in which the two phases overlap. Unlike the classical C^2MOS structure, modified C^2MOS is robust to clock-slope variation due to the local clock buffering.

Hybrid-latch flip-flop (HLFF) (Fig. 17), presented in [10], is one of the fastest structures presented. It also has a very small PDP_{tot}. The major advantage of this structure is its soft-edge property, i.e., its robustness to clock skew. One of the major drawbacks of the hybrid design in general is the positive hold time, discussed in Section II-B. Due to the single-output design, the power-consumption range of the HLFF is comparable to that of the static circuits. However, depending on the power distribution, precharged structures can dissipate more than static structures for data patterns with more "ones." Hybrid design appears to be very suitable for high-performance systems with little or no penalty in power when compared to classical static structures.

Another interesting approach to hybrid design is the semi-dynamic flip-flop (SDFF) structure (Fig. 18) presented in [16]. It is the fastest of all the presented structures. The significant advantage over HLFF is that there is very little performance penalty for embedded logic functions. The disadvantages are bigger clock load and larger effective precharge capacitance, which results in increased power consumption for data patterns with more "ones." This is still the most convenient structure

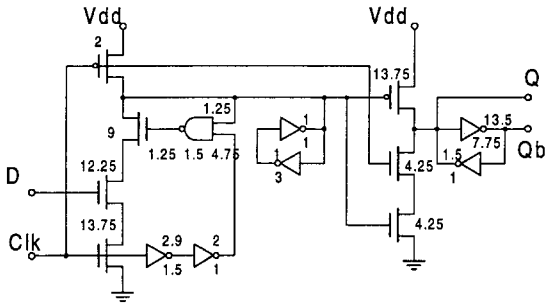

Fig. 18. Semidynamic flip-flop.

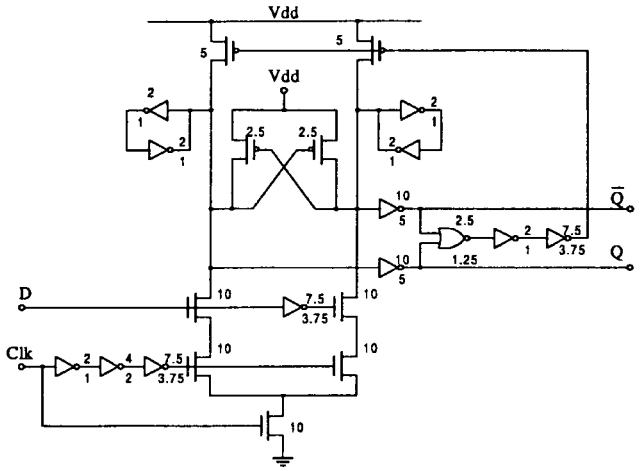

Fig. 19. K-6, dual-rail ETL.

for applications where speed is of primary importance, without a big penalty in power consumption.

The K6 edge-triggered latch (Fig. 19), presented in [11], is very fast but its differential structure, along with the precharge and self-reset property, causes very high power consumption independent of the data pattern.

We have to point out some details about the design of the SA-F/F [7] and the flip-flop used in StrongArm110 [14] (Figs. 20 and 21). The precharged sense-amplifier stage is very fast, but the set–reset latch almost doubles the delay because of asymmetric rise and fall times. This not only degrades speed but also causes glitches in succeeding logic stages, which increase total power consumption. Therefore, the flip-flop used in the StrongArm110 low-power processor lacks the delay budget, which can be traded for power. The results of its counterpart SA-F/F are little better when compared under the same conditions, but suffer from zero floating output nodes of the sense amplifier when data change during the high phase of the clock.

Both the precharged sense-amplifier in SA-F/F and StrongArm110 FF and the self-reset stage in K6 ETL have the very useful feature of monotonous transitions at the outputs, i.e., always from zero to one or from one to zero, which is essential for driving the succeeding fast domino stages. This property of the sense-amplifier stage was used in the high-performance WD21264 Alpha processor [9], where sense amplifiers were used as dynamic flip-flops. These structures also have very small clock load.

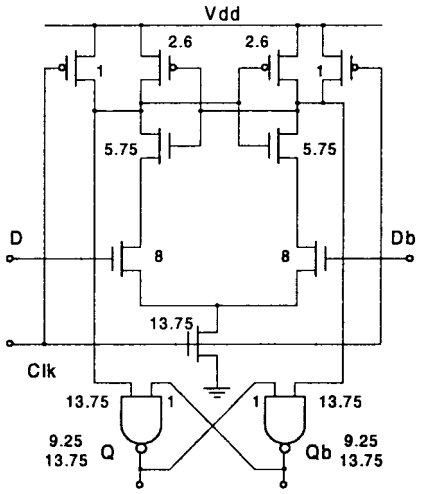

Fig. 20. Sense-amplifier flip-flop.

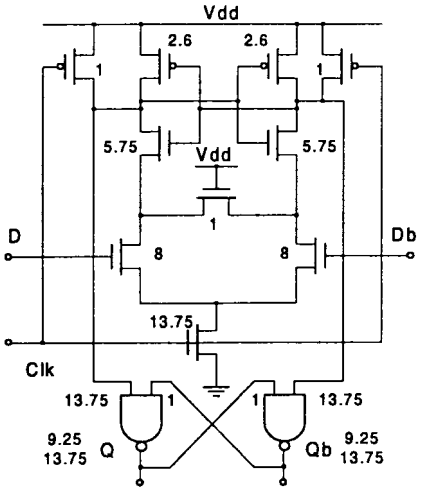

Fig. 21. StrongArm110 flip-flop.

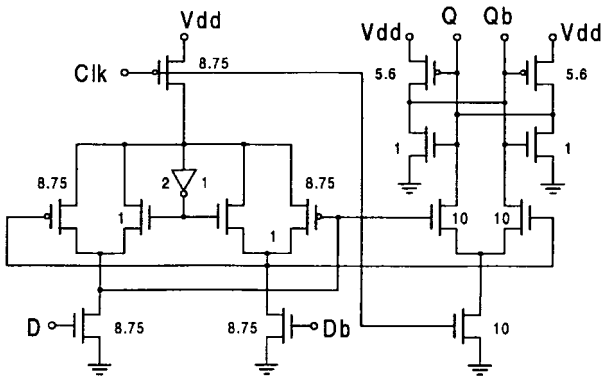

Fig. 22. SSTC master–slave latch.

The master latch in SSTC (Fig. 22) and DSTC (Fig. 23), presented in [5], suffers from substantial voltage drop at the outputs due to the capacitive coupling effect between the common node of the slave latch and the floating high output driving node of the master latch. This effect occurs on the rising edge of the clock, when the master latch becomes opaque and the

TABLE IV
TIMING PARAMETERS

Nominal conditions	Clk-Q(Qb)hl [ps]	Clk-Q(Qb)lh [ps]	Minimum D-Q(Qb)hl [ps]	Minimum D-Q(Qb)lh [ps]	Optimum Setup time [ps]
HLFF	195	191	199	155	-21
PowerPC	145	139	266	220	79
SDFF	176	176	187	143	-21
mC^{2}MOS	193	188	292	282	92
Strong Arm FF	262	162	275	171	-35
SA-F/F	262	162	272	168	-35
K6 ETL		168		200	-4
SSTC	97	301	374	592	267
DSTC	98	318	375	629	263

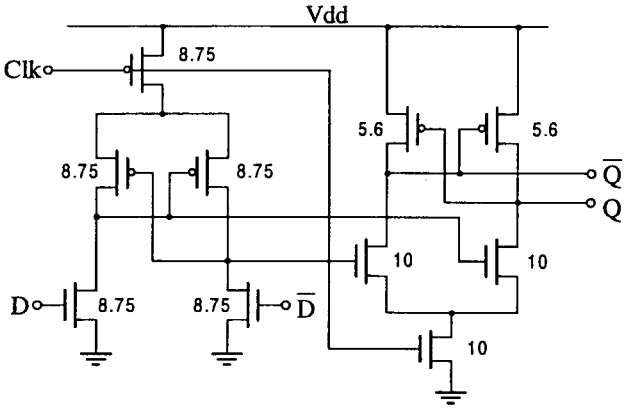

Fig. 23. DSTC master–slave latch.

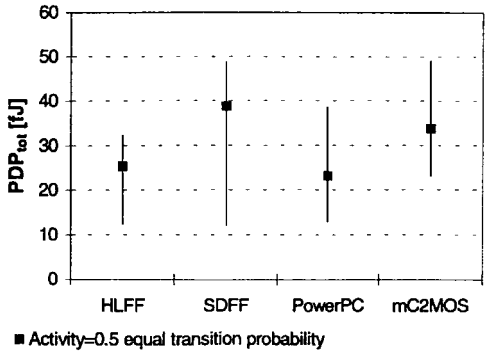

Fig. 24. Ranges of PDP_{tot}, single-ended structures.

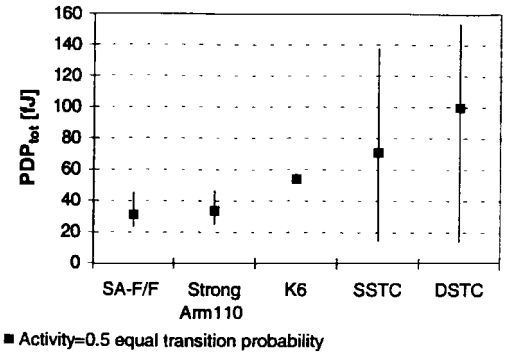

Fig. 25. Ranges of PDP_{tot}, differential structures.

slave latch transparent. One of the outputs of the master latch is left floating high and is capacitively coupled with the drain of the clocked transistor in the slave latch. The voltage drop that occurs on the common node of the slave latch when the clock goes high decreases the driving capability of the floating high output of the master latch. This causes an increase in delay and short-circuit power consumption in the slave latch, which tends to dominate dynamic power consumption.

The presented capacitive-coupling effect along with the problems associated with the glitches at the data inputs, noted by Blair in [6], result in relatively lower performance as compared with other presented latches. Due to the effects described above, SSTC and DSTC have somewhat weaker driving capabilities than the rest of the structures. In shift registers, where the circuits have a very light load, they perform very well, as shown in [5] and [15].

Detailed timing parameters of the presented structures are shown in Table IV. One can emphasize that the *Clk-Q* delay parameter does not illustrate the real performance of the circuit. A more relevant and severe constraint is *minimum D-Q(Qb)*. Timing results also show that some (mostly differential) design styles suffer from unequal low-to-high and high-to-low *delays*, which cause glitches and short-circuit power consumption in the succeeding logic stages, making the styles less desirable for low-power design.

Figs. 24 and 25 present the ranges and distribution of PDP_{tot} for different data patterns and for two major design styles, single ended and differential. The ranges of PDP_{tot} were obtained using different data sequences reflecting minimum, maximum, and average power consumption and corresponding PDP_{tot}. The symbol "■" designates the point of power dissipation PDP_{tot}) an for average-activity data pattern.

Among single-ended structures, the PowerPC 603 master–slave latch and HLFF show the best compromise between speed and power, followed by the classical modified C^{2}MOS master–slave latch and the hybrid SDFF. Hybrid structures are faster but consume more power than fully static master–slave designs due to the precharge nature of the front-end stage. The semidynamic nature of the hybrid circuits causes different power-dissipation dependence on data distribution than fully static master–slave structures.

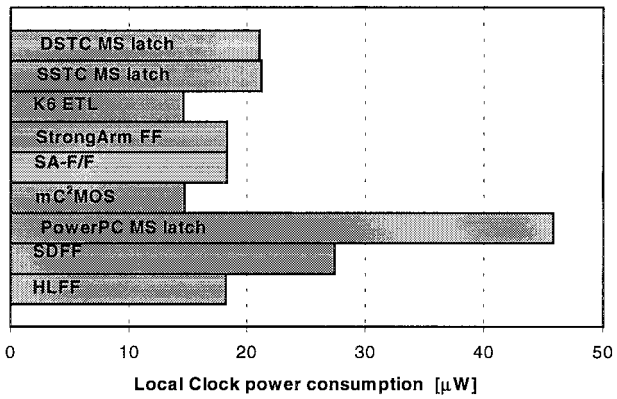

Fig. 26. Local clock power consumption.

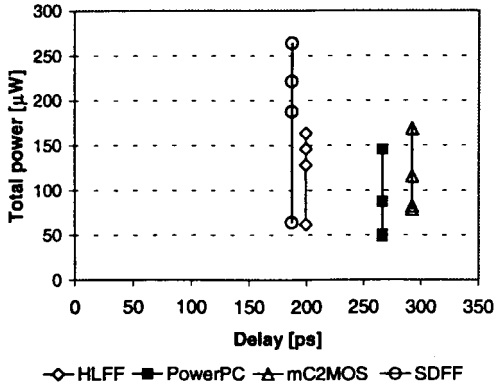

Fig. 27. Single-ended structures, *total power* range versus *delay*.

Differential latches based on differential cascode voltage switch logic style suffer from uneven rise and fall times, which can cause glitches and short-circuit power dissipation in succeeding logic stages.

Yet, differential structures have the unique property of differential signal amplification. Since dynamic power consumption depends on the square of voltage swing, great power savings can be made by reducing the voltage swing of the signals in the circuit. In that case, the logic in the pipeline operates with reduced voltage swing signals, and latches have the role of signal amplifiers, i.e., swing recovery circuits [7]. Thus, the logic in the pipeline, and not the latches themselves, is the party that saves power. The overall power dissipation of such pipeline structures is decreased, but the latches themselves are not ideal low-power structures when tested solely. This is why they appear to have a bad compromise between power and delay in comparison with single-ended structures.

In high-performance systems, clock power consumption is an important issue because of the portion that it takes from the total power budget. The *local clock power* parameter illustrates the clock load imposed by the latch. The amount of power consumed for driving the clock inputs of each structure is shown in Fig. 26.

Results presented in Figs. 27–30 clearly show that the assumptions made regarding *Clk-Q* as a relevant performance indicator can be fairly misleading, especially in the analysis of master–slave structures, which tend to have positive setup times.

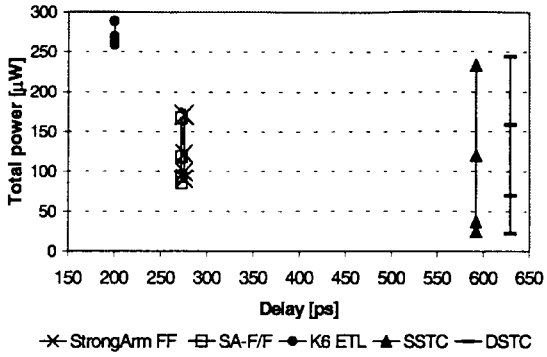

Fig. 28. Differential structures, *total power* range versus *delay*.

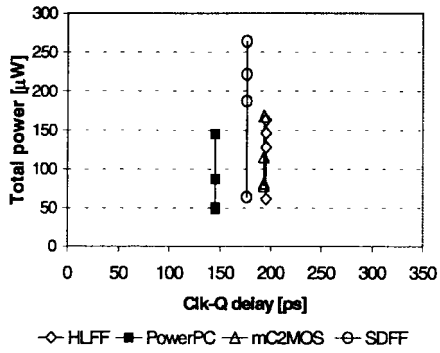

Fig. 29. Single-ended structures, *total power* range versus *Clk-Q delay*.

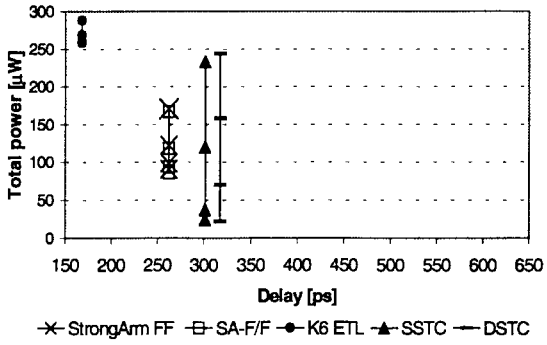

Fig. 30. Differential structures, *total power* range versus *Clk-Q delay*.

This is illustrated in Fig. 29, where the PowerPC 603 master–slave latch becomes the "fastest," the modified C^2MOS master–slave latch becomes as "fast" as HLFF and DSTC, and the SSTC master–slave latches become comparable to other structures in terms of "speed."

Figs. 31 and 32 illustrate the distribution of major parameters over the four chosen data patterns. The different design styles exhibit different power and PDP distributions depending on the specific features of the design. As mentioned in Section II, four different data patterns reveal the sources of power dissipation dominant in different design styles.

- ..010 101 01.. causes maximum switching activity in fully static circuits and in static parts of semidynamic structures.
- A random sequence utilizing equal transition probability gives the general estimate of average power dissipation, given the uniform data distribution.

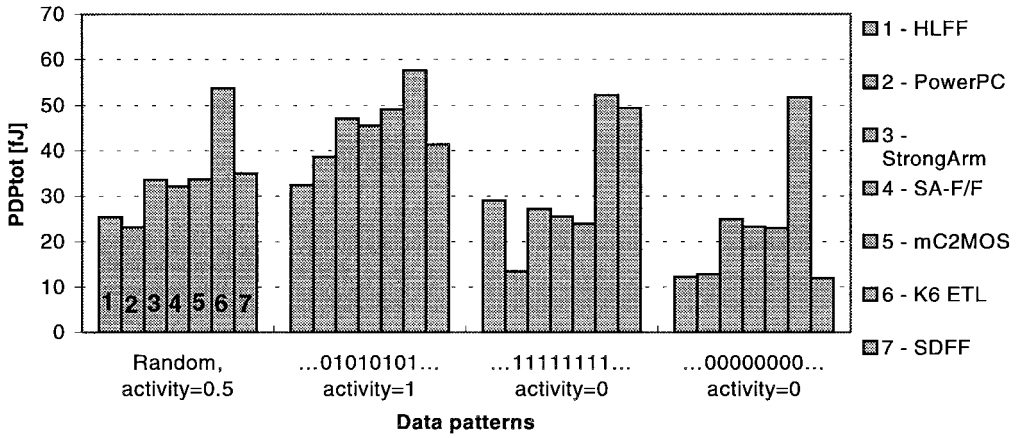

Fig. 31. PDP_{tot} dependence on data statistics.

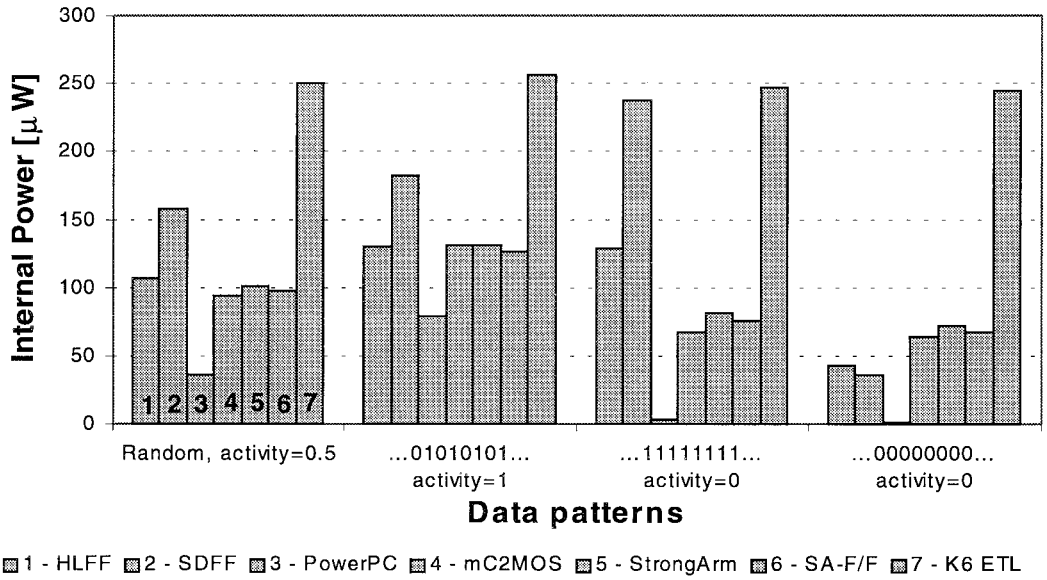

Fig. 32. Internal power dependence on data patterns.

- ..111 111 11.. causes maximum switching activity in precharge, dynamic parts of the semidynamic structures.
- ..000 000 00.. reveals the amount of leakage current power consumption and the amount dissipated in local clock processing, like in hybrid structures and modified C^2MOS master–slave latches.

The example of SDFF in Fig. 32 shows that, depending on the size of the precharge node capacitance, power consumption can be bigger for the ..111 111.. sequence (zero data activity) than for the ..010 101.. sequence (maximum data activity). This is because the precharge node is charged and discharged during the clock cycle only if data are high, which enables the discharge. If data are low, no discharge will occur, the high level of the precharged node will remain high, and it will not be charged again in the precharge phase of the next clock cycle.

Similarly, the differential dynamic nature of K-6 ETL causes data-independent power dissipation because for any data pattern, each of the sides in a differential tree is switching, as are the corresponding output parts, since the outputs are dynamic too.

Measurements of *internal power consumption* (especially in the case of PowerPC 603 master–slave latch) show that it can be far less than *total power consumption* due to the large amount of power dissipated in driving the clock and data inputs.

Modified C^2MOS master–slave latches and single-ended hybrid structures (HLFF and SDFF) dissipate a large amount of power in local clock processing (buffering) in order to decrease the clock load seen from the clock tree. This is shown in Fig. 32 for the ..000 000 00.. data pattern.

Similarly, precharge activity of the sense-amplifier stage in the SA-F/F and StrongArm110 flip-flop can be determined from internal power consumption for ..000 00.. and ..111 11.. data patterns.

V. Conclusion

For systems where high performance is of primary interest within a certain power budget, hybrid, semidynamic designs present a very good choice, given their negative setup time and small internal delay.

Fully static MS latches are suitable for low-power applications, where speed is not of primary importance. Among different single-ended design styles, our work has stated two as the most suitable for low-power applications: the low-power pass-gate style used in PowerPC 603 and the low-power modified C^2MOS style.

Among differential structures, SA-F/F and StrongArm flip-flop offer the best compromise in terms of both power and speed despite the speed bottleneck in the output stage. These structures offer interface to the low swing logic families and probably present the mainstream of future design styles.

The problem of consistency in analysis of various latch and flip-flop designs was addressed. A set of consistent analysis approaches and simulation conditions has been introduced. We strongly feel that any research on latch and flip-flop design techniques for high-performance systems should take these parameters into account. The problems of the transistor width optimization methods have also been described. Some hidden weaknesses and potential dangers in terms of reliability of previous timing parameters and optimization methods were brought to light.

ACKNOWLEDGMENT

The authors gratefully acknowledge suggestions by and numerous discussions with B. Nikolic, which greatly improved this paper and helped them clarify some concepts. They are thankful for the technical support provided by Hitachi America Research Labs, Dr. Y. Hatano, and Dr. R. Bajwa for their insight in simulation and proper evaluation environment as well as numerous discussions on this subject. They thank S. Oklobdzija for improving the language and presentation of this paper.

REFERENCES

[1] S. H. Unger and C. Tan, "Clocking schemes for high-speed digital systems," *IEEE Trans. Comput.*, vol. C-35, pp. 880–895, Oct. 1986.
[2] G. J. Fisher, "An enhanced power meter for SPICE2 circuit simulation," *IEEE Trans. Computer-Aided Design*, vol. 7, May 1988.
[3] J. Yuan and C. Svensson, "CMOS circuit speed optimization based on switch level simulation," in *Proc. Int. Symp. Circuits and Systems (ISCAS'88)*, 1988, pp. 2109–2112.
[4] ——, "Principle of CMOS circuit power-delay optimization with transistor sizing," in *Proc. Int. Symp. Circuits and Systems (ISCAS'96)*, 1996, vol. 1, pp. 62–69.
[5] ——, "New single-clock CMOS latches and flipflops with improved speed and power savings," *IEEE J. Solid-State Circuits*, vol. 32, Jan. 1997.
[6] G. M. Blair, "Comments on new single-clock CMOS latches and flip-flops with improved speed and power savings," *IEEE J. Solid-State Circuits*, vol. 32, pp. 1610–1611, Oct. 1997.
[7] M. Matsui, H. Hara, Y. Uetani, K. Lee-Sup, T. Nagamatsu, Y. Watanabe, A. Chiba, K. Matsuda, and T. Sakurai, "A 200 MHz 13 mm^2 2-D DCT macrocell using sense-amplifier pipeline flip-flop scheme," *IEEE J. Solid-State Circuits*, vol. 29, pp. 1482–1491, Dec. 1994.
[8] J. Montanaro, R. T. Witek, K. Anne, A. J. Black, E. M. Cooper, D. W. Dobberpuhl, P. M. Donohue, J. Eno, W. Hoeppner, D. Kruckemyer, T. H. Lee, P. C. M. Lin, L. Madden, D. Murray, M. H. Pearce, S. Santhanam, K. J. Snyder, R. Stephany, and S. C. Thierauf, "A 160-MHz, 32-b, 0.5-W CMOS RISC microprocessor," *IEEE J. Solid-State Circuits*, vol. 31, pp. 1703–1714, Nov. 1996.
[9] B. A. Gieseke, R. L. Allmon, D. W. Bailey, B. J. Benschneider, S. M. Britton, J. D. Clouser, H. R. Fair, J. A. Farrell, M. K. Gowan, C. L. Houghton, J. B. Keller, T. H. Lee, D. L. Leibholz, S. C. Lowell, M. D. Matson, R. J. Matthew, V. Peng, M. D. Quinn, D. A. Priore, M. J. Smith, and K. E. Wilcox, "A 600 MHz superscalar RISC microprocessor with out-of-order execution," in *ISSCC Dig. Tech. Papers*, Feb. 1997, pp. 176–177, 451.
[10] H. Partovi, R. Burd, U. Salim, F. Weber, L. DiGregorio, and D. Draper, "Flow-through latch and edge-triggered flip-flop hybrid elements," in *ISSCC Dig. Tech. Papers*, Feb. 1996, pp. 138–139.
[11] D. Draper, M. Crowley, J. Holst, G. Favor, A. Schoy, J. Trull, A. Ben-Meir, R. Khanna, D. Wendell, R. Krishna, J. Nolan, D. Mallick, H. Partovi, M. Roberts, M. Johnson, and T. Lee, "Circuit techniques in a 266-MHz MMX-enabled processor," *IEEE J. Solid-State Circuits*, vol. 32, pp. 1650–1664, Nov. 1997.
[12] G. Gerosa, S. Gary, C. Dietz, P. Dac, K. Hoover, J. Alvarez, H. Sanchez, P. Ippolito, N. Tai, S. Litch, J. Eno, J. Golab, N. Vanderschaaf, and J. Kahle, "A 2.2 W, 80 MHz superscalar RISC microprocessor," *IEEE J. Solid-State Circuits*, vol. 29, pp. 1440–1452, Dec. 1994.
[13] M. Shoji, *Theory of CMOS Digital Circuits and Circuit Failures*Princeton, NJ: Princeton Univ. Press, 1992.
[14] U. Ko, A. Hill, and P. T. Balsara, "Design techniques for high-performance, energy-efficient control logic," in *ISLPED Dig. Tech. Papers*, Aug. 1996.
[15] C. Svensson and J. Yuan, "Latches and flip-flops for low power systems," in *Low Power CMOS Design*, A. Chandrakasan and R. Brodersen, Eds. Piscataway, NJ: IEEE Press, 1998, pp. 233–238.
[16] F. Klass, "Semi-dynamic and dynamic flip-flops with embedded logic," in *1998 Symp. VLSI Circuits Dig. Tech. Papers*, Honolulu, HI, June 11–13, 1998, pp. 108–109.

Vladimir Stojanovic was born in Kragujevac, Serbia, Yugoslavia, on August 2, 1974. He received the Dipl.Ing. degree in electrical engineering from the University of Belgrade, Yugoslavia, in 1998. He currently is pursuing the M.Sc. degree in the Department of Electrical Engineering, Stanford University, Stanford, CA.

In 1997–1998, he was a Visiting Scholar in the Advanced Computer Systems Engineering Laboratory, University of California, Davis, where he conducted research in the area of latches and flip-flops for low-power and high-performance VLSI systems. His current research interests include analog and digital IC's and architectures for low-power, portable systems.

Vojin G. Oklobdzija (S'78–M'82–SM'88–F'95) received the M.Sc. and Ph.D. degrees in computer science from the University of California, Los Angeles (UCLA), in 1982 and 1978, respectively. He received the Dipl.Ing. (M.Sc.E.E.) degree in electronics and telecommunication from the Electrical Engineering Department, University of Belgrade, Yugoslavia, in 1971.

He was on the Faculty at the University of Belgrade until 1976. Currently he is President and CEO of Integration Corp., Berkeley, CA. He spent eight years as a Research Staff Member with the IBM T. J. Watson Research Center, Yorktown Heights, NY, where he made contributions to the development of RISC and superscalar RISC architecture and received a patent on the IBMRS/6000-PowerPC. From 1988 to 1990 he was with the University of California, Berkeley, as Visiting Faculty from IBM. His industrial experience includes positions with the Microelectronics Center of Xerox Corp. and consulting positions with Sun Microsystems Laboratories, AT&T Bell Laboratories, Siemens, Hitachi Laboratories, Sony Corp., and various others. His interests are in computer design and architecture, VLSI, fast circuits, and efficient implementations of algorithms and computation. He has received four U.S. and four European patents in the area of circuits and computer design and has six patents pending. He is the author or coauthor of more than 100 papers in the areas of circuits and technology, computer arithmetic, and computer architecture and of a book on high-performance system design.

Dr. Oklobdzija is a member of the American Association of University Professors and the IEICE of Japan. He is a member of the editorial board of the IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATED (VLSI) SYSTEMS and the *Journal of VLSI Signal Processing*. He is a Program Committee Member of the International Solid-State Circuits Conference and was a General Cochairman of the 1997 Symposium on Computer Arithmetic. He was a Fulbright Scholar at UCLA in 1976 and a Fulbright Professor in Peru in 1991.