Comparative Delay and Energy of Single Edge-Triggered & Dual Edge-Triggered Pulsed Flip-Flops for High-Performance Microprocessors

James Tschanz, Siva Narendra, Zhanping Chen, Shekhar Borkar, Manoj Sachdev*,
and Vivek De

Microprocessor Research Labs, Intel Corporation
5350 N.E. Elam Young Parkway, Hillsboro, OR 97124, USA
*Department of ECE, University of Waterloo, Canada

james.w.tschanz@intel.com

ABSTRACT
Flip-flops and latches are crucial elements of a design from both a delay and energy standpoint. We compare several styles of single edge-triggered flip-flops, including semidynamic and static with both implicit and explicit pulse generation. We present an implicit-pulsed, semidynamic flip-flop (ip-DCO) which has the fastest delay of any flip-flop considered, along with a large amount of negative setup time. However, an explicit-pulsed static flip-flop (ep-SFF) is the most energy-efficient and is ideal for the majority of critical paths in the design. In order to further reduce the power consumption, dual edge-triggered flip-flops are evaluated. It is shown that classic dual edge-triggered designs suffer from a large area penalty and reduced performance, prohibiting their use in critical paths. A new explicit-pulsed dual edge-triggered flip-flop is presented which provides the same performance as the single edge-triggered version with significantly less energy consumption in the flip-flop as well as in the clock distribution network.

Keywords
Flip-flops, latches, clocking, dual edge-triggered, low power.

1. INTRODUCTION
The number of logic gate delays in a clock period is reducing by 25% per generation in high-performance IA-32 microprocessors, and is approaching a value of 10 or smaller beyond 0.13μm technology generation [1]. As a result, latency of flip-flops or latches is becoming a larger portion of the cycle time. In addition, the energy consumed by low-skew clock distribution networks is steadily increasing and becoming a larger fraction of the chip power. In order to achieve a design that is both high-performance while also being power-efficient, careful attention must be paid to the design of the flip-flops and latches. In this paper, we compare latency and energy efficiency of different pulsed hybrid flip-flops

and edge-triggered flops for a 3Ghz microprocessor design in a 0.13μm, 1.3V, dual-Vt bulk CMOS process technology. Pulsed hybrid flops allow time borrowing and alleviate clock skew penalty [2-4], much like level-sensitive latches. At the same time, hold time requirements are easier to meet and the number of latches in logic cones can be reduced significantly. We consider both *semidynamic* and *static* pulsed flops with *implicit* and *explicit* pulse generation. We also present a *dual edge-triggered,* explicit-pulsed *static* flop that improves energy efficiency and preserves time-borrowing capability. This flip-flop allows the data throughput to remain constant while the clock frequency is reduced by 2X, resulting in significant total power savings.

The remainder of the paper is organized as follows. Section 2 describes the method used for flip-flop optimization and defines the delays and energies that are measured. Section 3 presents a comparison of several types of single edge-triggered flip-flops, describing the key differences in terms of both performance and power. Section 4 gives an overview of dual edge-triggered flip-flops and compares several dual edge-triggered designs against each other and against their single edge-triggered counterparts. Finally, section 5 concludes the paper.

2. FLIP-FLOP DESIGN OPTIMIZATION METHODOLOGY
A global optimizer, which uses a robust, steepest-descent algorithm, is used to determine transistor sizes in the various flip-flop topologies and minimize total energy per cycle (E) for different target values of data-to-Q (D-Q) delay. This process results in a plot of energy versus delay for each flip-flop, which simplifies comparisons between flops. Setup times and clock-to-Q delays for "low" and "high" values of input data are measured by sweeping the arrival time of data with respect to the rising clock edge and determining the point at which the data-to-Q delay is minimized [5]. Output storage nodes of all flops are protected from direct noise coupling by a single inverter. Therefore, some flip-flops are inverting while others are non-inverting. A constant output load of 20fF is used for all flops. Limiting the input capacitance value to 5fF sets maximum sizes of the inverters driving the data and clock inputs to the flops.

The typical pulse width is set to 90ps for all pulsed flops so that the worst-case min-delay requirement in the logic cone feeding the flop is less than half the clock period for 3Ghz operation. Because the hold time of a pulsed flip-flop is roughly equal to the pulse width, this restriction provides a reasonable compromise between the pulsed flop's time-borrowing capability and logic design efforts needed to meet worst-case hold time requirements. In addition, designs which employ an explicit clocking pulse must ensure that the pulse width is large enough that data will correctly be captured across all process, voltage, and temperature corners. Maximum voltage droop criteria at intermediate and output storage nodes are used to size the keeper transistors for adequate robustness, and to determine hold times. Transition activity of input data is assumed to be one-tenth of clock signal activity and all simulations are conducted in a 0.13μm technology using low-Vt transistors at 1.11V supply and 110°C.

3. SINGLE EDGE-TRIGGERED FLIP-FLOPS

The simplest flip-flop designs are single edge-triggered, sampling data on only one clock edge (in this case, the rising clock edge). There are many different types of single edge-triggered flip-flops in use, each of which is particularly suited for a certain application. Here we compare the advantages and disadvantages of implicit-pulsed semidynamic flops, implicit-pulsed static flops, and explicit-pulsed flops.

3.1 Implicit-pulsed semidynamic flip-flops

For very high-performance applications, such as the most critical paths of a design, achieving a small flip-flop delay is crucial while power consumption is a secondary concern. *Semidynamic* flip-flops, which are composed of a dynamic stage coupled to a pseudo-static stage, are therefore appropriate for these types of applications. An *implicit-pulsed*, data-close-to-output, semi-dynamic hybrid flip-flop (**ip-DCO**, schematic in Figure 1) is compared with two other previously reported implicit-pulsed, semidynamic hybrid flops - **HFF** [2] and **SDFF** [3-4]. The energy vs. delay characteristics of these three semidynamic flops is shown in Figure 2a, while Figure 2b plots the energy*delay product (E*D) as a function of delay. Figure 2c summarizes the comparison in terms of D-Q delay, minimum E*D product point, total device width, and total energy. For an equal energy per cycle of 40fJ, **ip-DCO** offers 8% - 10% better D-Q delay than **HFF** and **SDFF** and better time-borrowing capability (more negative setup time). The primary reason behind this performance improvement is that while the transistor being driven by data in the 3-transistor stack of the input stage is located in the middle for **HFF** and **SDFF**, it is located close to the output node in **ip-DCO**. This improves the speed when sampling a '1' because the intermediate slack nodes are discharged when the data signal arrives. In addition, this arrangement allows a more negative setup '0' time because the stack node is initially precharged when the rising clock edge arrives, and this inhibits the (false) evaluation until data changes to '0'. The worse-case hold time of **ip-DCO** is significantly larger due to this different ordering of transistors in the input stage, but is still below the limit dictated by excessive design efforts needed to meet hold time requirements
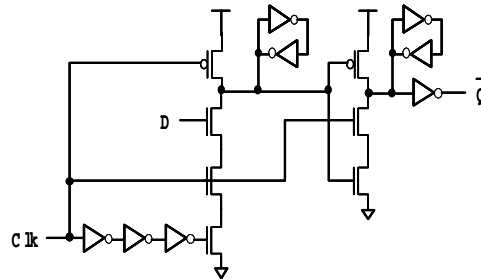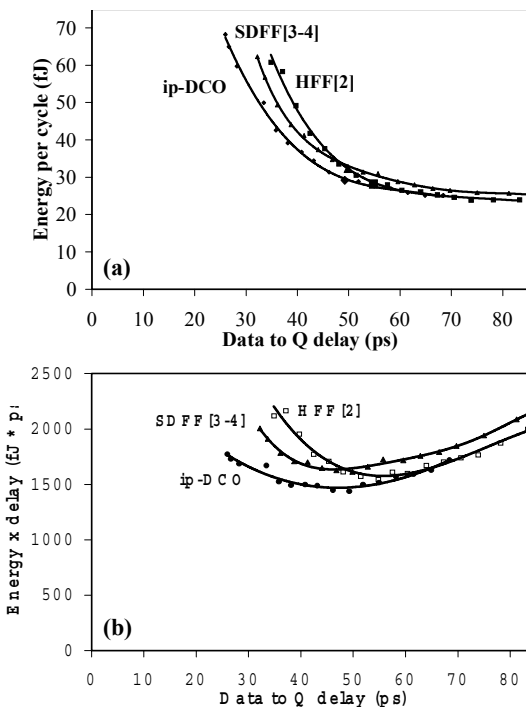

Figure 1. Implicit pulsed semidynamic flip-flop (ip-DCO).

(c)	% D-Q	% E*D	% total W	% total E
HFF[2]	*ref*	*ref*	*ref*	*ref*
SDFF[3-4]	2.5% better	5% worse	7.6% better	8.7% worse
ip-DCO	10.1% better	7.3% better	20% better	2.3% better

Figure 2. Comparisons of implicit-pulsed semidynamic flops. (a) Energy vs. delay. (b) Energy*delay product. (c) Comparison of D-Q delay @ E/cycle of 40fJ, min E*D, and total W, total E @ D-Q of 60ps.

in a 3Ghz clock cycle. It is also evident from Figure 2b that **ip-DCO** offers the best minimum energy*delay product - 7% better than **HFF** and 12% better than **SDFF**. For an equal target D-Q delay of 60ps, **ip-DCO** consumes less energy per cycle than either **HFF** or **SDFF**, and total transistor width is 12% - 20% smaller. As the target delay is reduced, the energy advantage of **ip-DCO** over **HFF** and **SDFF** increases.

3.2 Implicit-pulsed static flip-flops

The fast data-to-Q delay of the pulsed semidynamic flip-flops, however, comes at the expense of significant power consumption. The main reason for this high power consumption is the dynamic nature of the flip-flop: power may be consumed in the dynamic stage due to the precharge and evaluate cycle even when the input is held constant. Paths that are not critical in the design can achieve lower power consumption by employing static, rather than dynamic, flip-flops. Among static flip-flop designs, the most commonly used are the conventional static master-slave (**SMS**) and the time-borrowing master-slave (**tb-SMS**, schematic in Figure 3). Figures 4a and 4b show the energy-delay comparisons of these static flip-flops with the best of implicit-pulsed, hybrid semidynamic flip-flops (**ip-DCO**). It is apparent that **ip-DCO** provides significantly better D-Q delay (25% faster) than either **SMS** or **tb-SMS** and also offers more time-borrowing capability. However, the classic **SMS** flop is the most energy-efficient among these three – it provides 18% to 28% better minimum $E*D$ value than **tb-SMS** and **ip-DCO**, and consumes 34% smaller energy than **ip-DCO** at a target $D-Q$ delay of 60ps. **tb-SMS** adds time borrowing capability to **SMS** at a cost of 25% higher energy consumption, and thus offers an attractive trade-off between energy-efficiency and tolerance to clock skew.
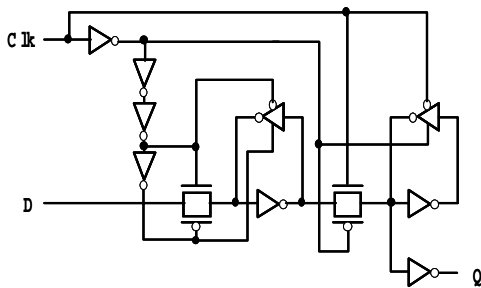

Figure 3. Time-borrowing static master-slave (tb-SMS).

3.3 Explicit-pulsed flip-flops

While the semidynamic flip-flops and the **tb-SMS** static flip-flop achieve a transparency window through an implicitly-generated pulse (through the use of transistor stacks or transmission gates), it is also possible to control the flop with an explicitly-generated clocking pulse. An *explicit-pulsed*, hybrid *semidynamic* flop (**ep-DCO**, schematic in Figure 5a) does not offer any performance advantage over **ip-DCO**, and consumes larger energy due to the explicit pulse generator (Figure 6). However, the pulse generator power consumption can be significantly reduced by sharing a single pulse generator among a group of flip-flops. Thus both **ip-DCO** and **ep-DCO** with shared pulse generator are the best among all semidynamic flip-flops considered here for use in a minority of speed-critical paths. For reduced power consumption, an *explicit-pulsed*, hybrid *static* flip-flop (**ep-SFF**) is shown in Figure 5b. This flop has 29% better D-Q delay than **tb-SMS** while consuming 8% less energy than **ip-DCO** (Figure 6). In

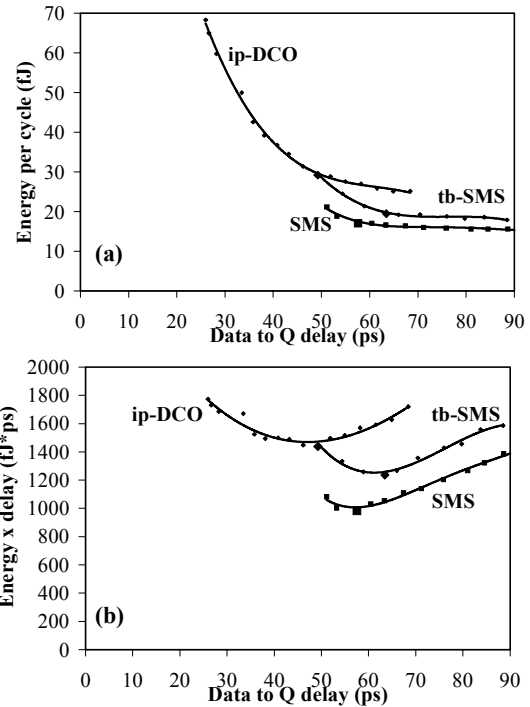| (c)	% D-Q	% E*D	% total W	% total E
SMS	*ref*	*ref*	*ref*	*ref*
tb-SMS	3% better	23.2% worse	35.6% worse	25.1% worse
ip-DCO	25.4% better	39.2% worse	3% worse	51.5% worse

Figure 4. Comparisons of implicit-pulsed semidynamic and static flops. (a) Energy vs. delay. (b) Energy*delay product. (c) Comparison of D-Q delay @ E/cycle of 40fJ, min E*D, and total W, total E @ D-Q of 60ps.
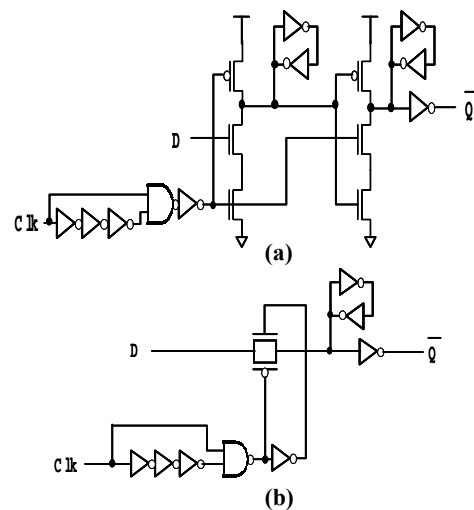

Figure 5. Explicit-pulsed flip-flops. (a) ep-DCO. (b) ep-SFF

(c)	% D-Q	% E*D	% total W	% total E
ip-DCO	*ref*	*ref*	*ref*	*ref*
ep-DCO	7.9% worse	16.5% worse	13.9% worse	81.5% worse
tb-SMS	29.9% worse	11.5% better	7.8% worse	17.4% better
ep-SFF	7.3% better	14.8% better	41.6% better	7.7% better

Figure 6. Comparisons of implicit and explicit-pulsed flops. (a) Energy vs. delay. (b) Energy*delay product. (c) Comparison of D-Q delay @ E/cycle of 40fJ, min E*D, and total W, total E @ D-Q of 60ps.

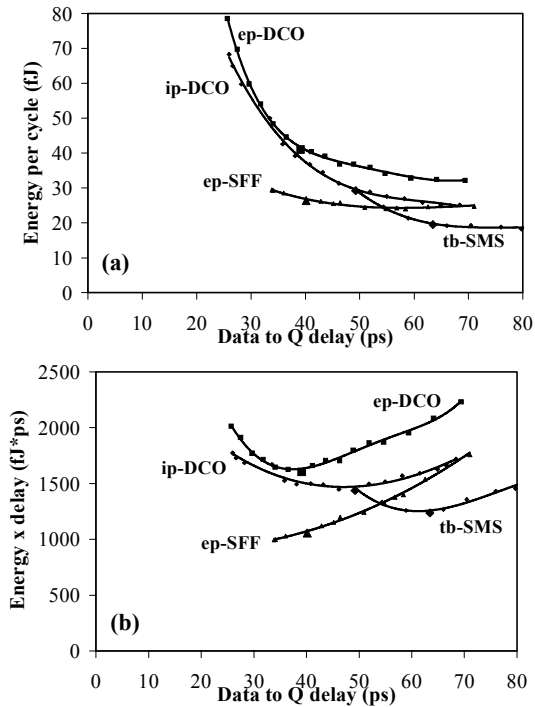addition, **ep-SFF** is the most energy-efficient of all the flops with time-borrowing capability: 15% better E*D value than **ip-DCO** and 4% better E*D value than **tb-SMS**. Thus while the minimum delay of **ep-SFF** is larger than the minimum delay of **ip-DCO**, **ep-SFF** is much more energy-efficient and is appropriate for the large number of paths on a chip which are speed-sensitive and can benefit from a fast delay and large amount of time-borrowing. Clearly, for speed-insensitive paths that will not benefit from time borrowing, classic **SMS** is the most energy-efficient choice.

In contrast to **ip-DCO** or **tb-SMS**, **ep-DCO** and **ep-SFF** can share a single pulse generator among multiple flops to improve energy efficiency. The degree of sharing possible is limited by additional pulse width variations due to transistor mismatches and noise coupling to the pulse distribution network. For example, with eight flops sharing a single pulse generator, the minimum $E*D$ value of **ep-SFF** improves by 39% and the energy consumption at a target D-Q delay of 60ps is 32% smaller (Figures 7a and 7b).
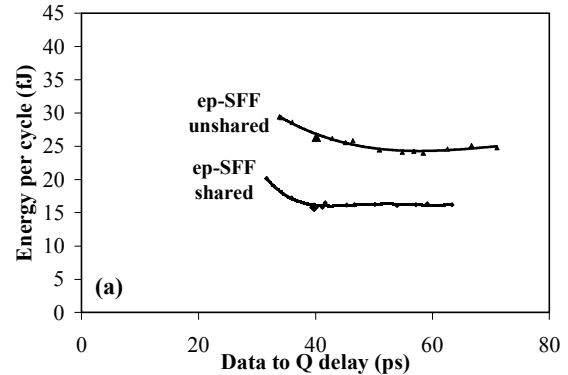
(b)	% D-Q	% E*D	% total W	% total E
ep-SFF unshared	*ref*	*ref*	*ref*	*ref*
ep-SFF shared	10.3% better	39.4% better	4.1% better	31.8% better

Figure 7. Comparisons of unshared and shared ep-SFF flop. (a) Energy-delay comparison. (b) Comparison of best achievable D-Q delay, min E*D, and total W, total E @ D-Q of 60ps.

4. DUAL EDGE-TRIGGERED FLIP-FLOPS

Dual edge-triggered (DET) flip-flops provide an effective technique for reducing the power consumption of a large design by reducing the power consumed in the clock distribution network. An ideal dual edge-triggered flip-flop allows the same data throughput as a single edge-triggered (SET) flip-flop while operating at half the clock frequency and sampling data on both edges of the clock. If the clock load of the DET flip-flop is not significantly larger than the single edge-triggered version, the power in the clock distribution network is reduced by a factor of two. Because the clock distribution power is a large fraction of the total power of a microprocessor, significant overall power savings are possible.

4.1 Conventional DET flip-flops

Conventional implementations of the dual edge-triggered **SMS** or **tb-SMS** (schematic in Figure 8a) flip-flops rely on latch duplication to achieve operation on both clock edges. This roughly doubles the area of the flip-flop and also increases the load on the data and clock inputs, which affects performance. Because the maximum size of the inverter driving the data input is fixed, the dual edge-triggered flip-flop cannot achieve the same delay as the single edge-triggered version. An alternate structure (**DET SMS**, schematic in Figure 8b) attempts to reduce the clock load by sharing the clocking transistors between the two latches [6], but still suffers from a large data load and area penalty. Figure 8c shows a comparison of these flip-flops against their respective SET versions. It is evident that while these DET flip-flops may be attractive for low-performance (large delay) applications, the energy consumption becomes much larger than SET as the delay is reduced. Therefore these flip-flops are not appropriate for use in critical paths.
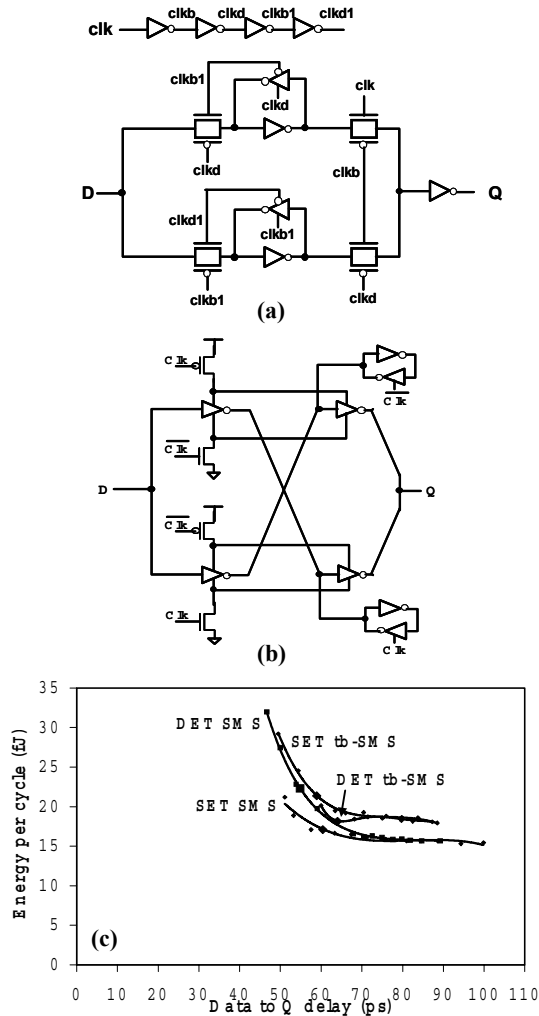
(a)

(b)

(c)

Figure 8. Comparisons of conventional DET flip-flops. (a) DET tb-SMS schematic. (b) DET SMS schematic. (c) Energy-delay comparison of SET and DET SMS and tb-SMS.

4.2 Explicit-pulsed DET flip-flop

A more efficient dual edge-triggered flip-flop may be realized by replacing the pulse generator in the single edge-triggered **ep-SFF** with an explicit dual edge-triggered pulse generator. This pulse generator may be local to each flop or shared among multiple flops. Because the entire latch is not duplicated, the area overhead for this technique is much less than for the conventional **DET SMS** and **DET tb-SMS**. In addition, implementing features such as scan, reset, or enable for this flip-flop may be easier than for the duplicated-latch designs since there only exists one path from data input to output. There are many possible implementations of flip-flops using dual edge-triggered pulse generators; an energy-efficient *dual edge-triggered, explicit-pulsed* static hybrid flop (**ep-DSFF**) is shown in Figure 9a. Because the path from data to output of the flip-flop is identical to the **ep-SFF**, latency and throughput of **ep-DSFF** are the same as **ep-SFF**, while the clock frequency is halved. As a result, the
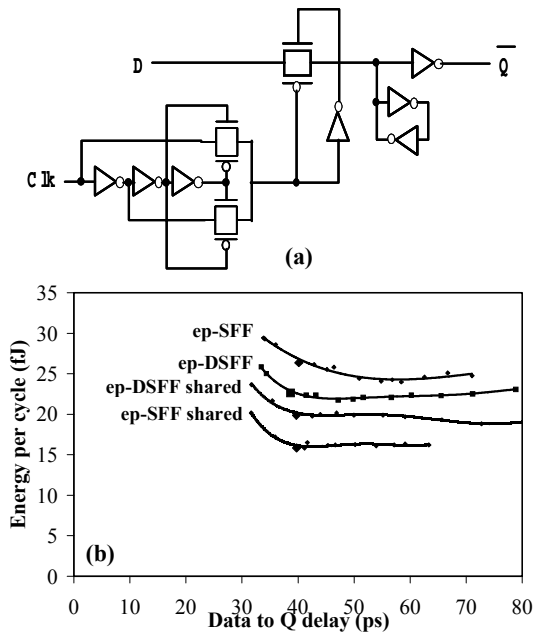

(a)

Figure 9. Comparisons of single and dual edge-triggered ep-SFF. (a) ep-DSFF schematic. (b) Energy-delay comparison of ep-DSFF for both unshared and shared pulse generator.

power dissipation of **ep-DSFF** with a local pulse generator is 21% less than **ep-SFF** at a target D-Q delay of 60ps (Figure 9b). Sharing the pulse generator is not as effective for the **ep-DSFF** as for the **ep-SFF** since the transistor sizes are larger; therefore if sharing is possible the single edge-triggered **ep-SFF** has the lowest energy consumption. These comparisons reflect only the energy of the flip-flop itself and do not include power in the clock distribution network.

4.3 Total DET power savings

In order to estimate the impact of dual edge-triggered flip-flops on the clocking power of an entire design, it is necessary to determine the power savings in the clock distribution network. For these calculations it is assumed that approximately half of the total clock power is consumed in the final flip-flop load while the other half is dissipated in the clock distribution network. Figures 10a and 10b compare the power consumption of SET and DET designs for two cases: low-power (low-performance) and high-speed. The height of each bar gives the total power of sequential elements in the design, including data power (power to drive the flip-flop output load), clock power (internal to the flip-flop), and clock distribution power. In the low-power case (Figure 10a), all flip-flops have a target D-Q delay of 70ps. If all **SMS** flip-flops in a design are replaced by **DET SMS** flops, the total power reduces by 20% due to the 2X reduction in clock distribution power. Similarly, a design employing **DET tb-SMS** flip-flops consumes 21% less energy than a **SET tb-SMS** design. Thus overall power savings are possible even if the DET flip-flop itself consumes more power than the SET version. The **ep-SFF** and **ep-DSFF** have larger energy consumption than the DET static flops

and are not attractive for a low-performance application unless the pulse generators are shared.

Figure 10b shows a comparison for the high-performance case (target D-Q delay of 40ps). **SMS** and **tb-SMS** are not included in this comparison since they cannot meet this aggressive target delay. If local pulse generators are used in each flip-flop, **ep-DSFF** provides 30% energy savings over **ep-SFF**. If pulse generators are shared among groups of flip-flops, it is evident that the energy savings are not as significant. However, sharing pulse generators introduces additional complexities into the design regarding pulse distribution and margining for pulse width variation. Figure 10c shows a summary of the SET and DET flip-flop designs in terms of minimum E*D point and total device width, as compared with a design using only **SET SMS** flip-flops. Both the **DET SMS** and **DET tb-SMS** designs employ latch duplication and therefore have large area penalties over the SET designs. **ep-DSFF** is the only dual edge-triggered design considered here with a better minimum energy*delay value than classic **SMS**, a smaller total area, and significantly faster achievable delay.

Actual designs consist of a combination of critical paths where high-performance flip-flops are required, and non-critical paths where low power is more important. This analysis shows that both types of paths can benefit from the use of dual edge-triggered flip-flops. As a result, employing dual edge-triggered flip-flops throughout the chip and distributing the clock signal at one-half the frequency has the potential to significantly lower the total power consumption of the chip.

5. CONCLUSIONS

Pulsed flip-flops offer an attractive method of meeting delay and energy requirements of a design while providing time-borrowing capability to mitigate clock skew effects. For high-speed operation, **ip-DCO** has the fastest delay of any flip-flop considered, along with a large amount of negative setup time. However, **ep-SFF** is the most energy-efficient due to its static design and low transistor count. Therefore this flip-flop is ideal for the majority of paths in a design. In order to further reduce the total power consumption, dual edge-triggered flip-flops may be used to reduce the clock frequency by 2X. The highest-performance dual edge-triggered flip-flop examined here is the **ep-DSFF**, which provides the same delay as **ep-SFF** with significantly less energy consumption in the flip-flop as well as in the clock distribution network.
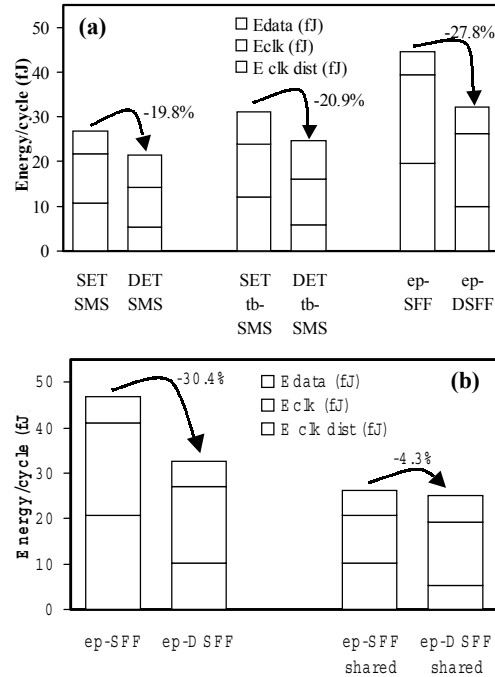
(c)	% D-Q	% E*D	% totalW	% totalE
SET SMS	ref	ref	ref	ref
DET tb-SMS	17.5% worse	17.4% worse	71.2% worse	9.2% better
ep-DSFF unshared	34.6% better	0.5% better	1.5% better	1.4% better
ep-DSFF shared	37.9% better	11.8% better	7.6% better	10.6% better

Figure 10. Comparisons of total clocking and flip-flop power for single and dual edge-triggered designs. (a) Low-power design (D-Q = 70ps). (b) High-performance design (D-Q = 40ps). (c) Comparison of minimum E*D point and total device width for target D-Q of 60ps.

6. REFERENCES

[1] V. De et al., 1999 ISLPED, pp. 163-168, 1999.

[2] H. Partovi et. al., 1996 ISSCC: Dig. Tech. Papers, pp. 138-139.

[3] F. Klass *et. al*, *IEEE JSSC*, pp. 712-716, May 1999.

[4] F. Klass, *1998 Symp. VLSI Circuits*, pp. 108-109.

[5] V. Stojanovic *et. al.*, *1998 ISLPED*, pp. 227-232.

[6] A. Gago *et. al.*, *IEEE JSSC*, pp. 400-402, March 1999