28.3 Conditional Push-Pull Pulsed Latches With 726FJ-ps Energy-Delay Product in 65nm CMOS

Elio Consoli¹, Massimo Alioto², Gaetano Palumbo², Jan Rabaey³
¹University of Catania, Catania, Italy, ²University of Siena, Siena, Italy, ³University of Michigan, Ann Arbor, MI, ⁴University of California at Berkeley, Berkeley, CA

Flip-flops (FFs) are key building blocks in the design of high-speed energy-efficient microprocessors, as their data-to-output delay (D-Q) and power dissipation strongly affect the processor’s clock period and overall power [1]. From previous analyses [2], the Transmission-Gate Pulsed Latch (TGPL) [3] proved to be the most energy-efficient FF in a large portion of the design space, ranging from high speed (minimizing ED products with >1) to minimum ED product designs [2], while simple Master-Slave FFs (TGFF [1] and ACFF [4]) are the most energy efficient. TGPL also has the lowest D-Q delay along with STFF [5]. However, the latter has considerably worse energy efficiency [2], hence, the TGPL is the best reference for a comparison. In this work, two new FFs are introduced, the Conditional Push-Pull Pulsed Latch (CPPL), and a version with a Shareable (CSPPL) Pulse Generator (PG). The adoption of a fast push-pull second stage, which requires a conditional PG, enables 50-to-100% delay improvements compared to TGPL, and absolute D-Q up to 0.7F04. CPPL and CSPPL also exhibit superior energy efficiency to TGPL in terms of minimum ED and ED products. A test chip is fabricated in 65nm CMOS technology (VDD=1V) to measure delay and energy consumption of CPPL, CSPPL and TGPL in minimum ED and ED sizings. Different loadings are used in the minimum ED (16×) and the minimum ED (64×) cases.

Figure 28.3.1 shows schematics of the proposed FFs and TGPL [3]. In CPPL, the data-to-output path is broken in two parallel paths, capturing data rising and falling transitions. The output inverter is replaced by a push-pull stage, the first stage is split into two half-latches and the gated keeper is moved to the output. Since the push-pull stage is prone to current contention, only one of the signals S_Neg and R can be asserted at a time due to the conditional PG, depending on the previous Q value. The pulsed signals CP (rising) and CP (falling) are alternately enabled by employing pseudo-NOR/NAND that are gated by a delayed version of Q. If Q_1 (0), CP can change its state if D=0 (D=1), and does not change its state if D=1 (D=0). The pseudo-NOR (pseudo-NAND) does not change CP (CP) and also S_Neg (R) does not make any transition. Conversely, the pseudo-NAND (pseudo-NOR) is enabled and generates a pulse on CP (CP). Furthermore, if D remains equal to 1 (0), R (S_Neg) does not change. Q has to be sufficiently delayed in order to avoid the pseudo-NOR (pseudo-NAND) being enabled when the previous output was equal to 1 (0). Otherwise, an undesired transition on CP (CP) would occur, dissipating power. This would not affect correctness, as the input signal on D must remain stable until the end of the transparent window to avoid hold time violations. In CPPL, the PG cannot be shared among different latches, since pseudo-NOR/NAND are driven by Q, which is unique for each latch. CSPPL solves this by fully integrating the conditional logic in the latch so that the PG can be shared among several FFs. Two transmission gates and a few small keepers have to be added at the two pulsed nodes to achieve the same operation as before.

Figure 28.3.2 shows a block diagram of the delay and energy test circuits. The delay measurement setup is essentially the same as in [6]. Clock-to-data (CK-D), clock-to-output (CK-Q), and D-Q delays are measured as time differences by using a capturing FF, locked by a signal (CKMS) obtained via the propagation of a pulse through a programmable Delay Generator (DG), as is the case for the D and CK inputs provided to the FF under test (DUT) [6]. The average delay step of the DG is equal to 1.8ps (=F04/12). Since the delay measurement is carried out locally by way of the capturing FF coupled to each DUT, we employed a single DG to test an array of 384 FFs, i.e. 64 DUTs for each topology-sizing combination. The energy measurement approach permits extraction of transient energy (dynamic and short circuit) under different data activities [6].

Figure 28.3.3 shows measured setup and hold curves of the proposed FFs and TGPL. CPPL and CSPPL achieve a minimum D-Q close to 15ps (0.7F04) and (ED) case. TGPL is known to exhibit better performance than other existing FFs thanks to its small logical effort, branching effort and low parasitic delay in its data-to-output path [2]. Observe that the D-Q delay of the proposed FFs is further reduced, since the S_Neg and R nodes have approximately half the load vs. TGPL. Hence, the half-latches are fast and the push-pull stage size can be increased without degrading energy efficiency. Also the branching due to fixed size gates is reduced and the load due to interconnect parasitics is smaller. Although the PGs of all FFs were designed to achieve a 4F04 pulse width, the flat D-Q region is more pronounced in TGPL than in CPPL and CSPPL. Note that though it may seem that the proposed FFs do not suffer from the setup/hold time tradeoff (when the previous Q value is equal to 1 (0), only a data falling (rising) transition can be captured in the upcoming clock cycle), the hold time requirement has to be honored even if the Q value is unchanged. That is, the usual setup/hold time tradeoff is present for CPPL and CSPPL (the critical hold time is the last CK-L leading to finite D-Q).

Figure 28.3.4 shows the measured energy dissipated per clock cycle vs. data activity and the energy/leakage-delay tradeoffs. CPPL and CSPPL show an energy 40-to-60% higher than TGPL, due to their higher complexity. Nevertheless, the adoption of the conditional technique reduces energy by avoiding unnecessary internal transitions when the data input does not change. The energy-efficiency of CPPL and CSPPL is indeed higher than TGPL. In particular, CPPL has 1.3× (2.3×) better ED (ED) product than TGPL at 25% data activity, and 2.7× (5.4×) better LEAK (LEAK) product, with LEAK being the average leakage current of the FF (estimated through simulations).

Figure 28.3.5 shows the variability of timing parameters from measurements on 256 FFs taken from four different dies, and leakage variation results from Monte Carlo simulations. Inspection of Fig. 28.3.5 reveals that CPPL and CSPPL bear similar variability to TGPL, except for the slightly worse (<1.4×) values in D-Q and CK-Q cases. Nevertheless, even including the impact of process variations, CPPL and CSPPL are largely faster than TGPL.

Figure 28.3.6 summarizes the figures of merit and also the simulation results for TGFF [1] and STFF [5], and measurements for ACFF [4]. Although they target different applications (TGFF/ACFF), or have much worse energy efficiency (STFF), these FFs represent valuable references for comparison because of their small ED product (TGFF/ACFF) and D-Q (STFF). Hence, data for these FFs is included for completeness. The improvements achieved with the proposed FFs are highlighted in the table, where it is shown that TGPL and STFF have >50% higher D-Q than CPPL and CSPPL. Regarding energy efficiency, CPPL and CSPPL have the best ED product (>120% reduction) and outperform TGPL, TGFF and ACFF, also in terms of ED product (>20% reduction).

Acknowledgements:

The authors thank STMicroelectronics for chip fabrication.

References:

Figure 28.3.1: Schematic of CP\(_3\)L, CSP\(_3\)L, their operation and TGPL [3].

Figure 28.3.2: Block diagram of delay and energy test circuits.

Figure 28.3.3: Setup (D-Q/CK-Q vs. CK-D) and Hold (CK-Q vs. CK-D) time characteristics.

Figure 28.3.4: Transient energy vs. data switching activity and energy/leakage-delay tradeoff.

Figure 28.3.5: Variability of timing parameters and leakage.

Figure 28.3.6: Performance summary and comparison with other FFs.