

# A 40-Gb/s Decision Circuit in 90-nm CMOS

T. Chalvatzis\*, K. H. K. Yau\*, P. Schvan†, M. T. Yang‡ and S. P. Voinigescu\*

\*The Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Rd, Toronto, ON M5S3G4, Canada (Email: theo@eecg.toronto.edu)

†Nortel Networks, Ottawa, ON K1Y4H7, Canada

‡TSMC, 1 Creation Rd, Science-Based Industrial Park, Hsin-Chu, Taiwan 300-77, R.O.C.

**Abstract**—A low-power 40-Gb/s decision circuit for fiber-optic and mm-wave analog-to-digital converter applications was implemented in two 90-nm processes from two different foundries. The circuit uses a MOS-CML Master-Slave latch topology with only two vertically stacked transistors. It combines low and high- $V_T$  MOSFETs to allow for operation from a 1.2-V supply, without compromising speed. Full-rate retiming with jitter reduction and 7 ps rise/fall times is demonstrated at 37 Gb/s and 40 Gb/s from 1.2 V and 1.5 V, respectively. The entire decision circuit dissipates 130 mW from 1.2 V, with a record low power consumption of 10.8 mW per latch.

**Index Terms**—Decision circuit, retiming D flip-flop, MOS-CML.

## I. INTRODUCTION

Flip-flops and decision circuits are the most critical digital blocks used in high-speed wireline and fiber-optic transceivers [1], equalizers and mm-wave-sampling ADCs [2]. Full-rate retiming has been successfully demonstrated at speeds above 40 Gb/s in III-V [3], [4] and SiGe BiCMOS technologies [1], [5]. However, these circuits operate from 1.5 V or higher supplies and consume significant power. A MOS-CML implementation would permit 40 Gb/s serializer-deserializer (SERDES) chips to reach the same levels of digital integration, including FEC, as state-of-the-art 10 Gb/s chips, and to operate from a single 1.2-V power supply.

The main advantage of MOSFET scaling to nanometer gate lengths is the ability to reach device speeds exceeding 120 GHz with low supply voltages. However, to truly benefit from the lower power potential of 90-nm MOSFETs, one must simplify the latch topology to reduce the number of vertically-stacked transistors to allow for 1.2-V operation. In the past, this has been realized either by removing the current source [6] or using transformers [7] to couple the signal between the clock- and data-path differential pairs. The former has been demonstrated in 90-nm CMOS at speeds below 20 GHz. The latter has been used in a 60-Gb/s 2:1 MUX clocked at 30 GHz, but limits the bandwidth of operation to that of the transformer. This paper presents the first 40-Gb/s full-rate retiming D-type flip-flop (DFF) in CMOS. A combination of low and high- $V_T$  devices and optimal biasing of MOS-CML [8] are used in the topology without current source to overcome the speed limitations of earlier designs, while operating from 1.2 V. The design relies on biasing all MOSFETs at constant current density to capitalize on the invariance of the peak- $f_T$  current density with respect to threshold voltage [8]. This is immediately apparent from the measured  $f_T - V_{GS}$  and  $f_T - I_{DS}/W$  characteristics shown in Fig. 1.

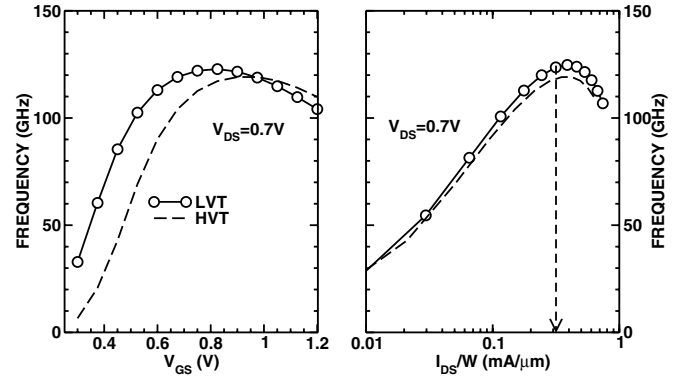


Fig. 1. Measured  $f_T$  vs a)  $V_{GS}$  and b)  $I_{DS}/W$  of 90-nm n-MOSFETs with low and high- $V_T$  showing that the peak- $f_T$  current density and peak- $f_T$  value do not depend on  $V_T$ .

## II. CIRCUIT DESIGN

### A. Proposed Low-Power Latch

The proposed MOS-CML latch schematic can be found in Fig. 2, along with that of the entire decision circuit. The clock signal switches the transistors of the differential pair M1-M2 from 0 to  $2 \times I_{BIAS} = 0.3 \text{ mA}/\mu\text{m}$ , the peak- $f_T$  current density [8]. Equivalently, the current density through each device in balanced mode is  $0.15 \text{ mA}/\mu\text{m}$ . To fully switch the 90-nm MOS differential pair, a voltage swing  $> 300 \text{ mV}_{pp}$  per side is required [8]. For  $I_{BIAS} = 4.5 \text{ mA}$  and a  $30 \times 1 \mu\text{m} \times 0.09 \mu\text{m}$  device, a load resistance  $R_L = 40 \Omega$  produces a voltage swing at the output of each latch

$$\Delta V_{swing} = (I_{M1} + I_{M2})R_L = 9 \text{ mA} \times 40 \Omega = 360 \text{ mV}$$

which is enough to switch the next stage MOS differential pair and results in an inverter gain  $A_V = -1.2$ . The bandwidth of the latch is extended with shunt inductive peaking. For a fanout of  $k = 1$ , the total capacitance at the drain of M3 is

$$C_T = C_{db3} + C_{gd3} + C_{gd5} + C_{db5} + C_{gs6} + (1 - A_V)C_{gd6} + k(C_{gs} + (1 - A_V)C_{gd}) = 197 \text{ fF}$$

and the inductance  $L = 100 \text{ pF}$  extends the  $BW_{3dB}$  [9] to

$$BW_{3dB} = 1.6 \times \frac{1}{2\pi R_L C_T} = 32.3 \text{ GHz}$$

As shown in Fig. 2, devices with different  $V_T$  are employed in the data and clock paths of the latch. The data path differential pairs M3-M4, M5-M6 are designed with low- $V_T$ , while the clock differential pair has high- $V_T$  devices. This

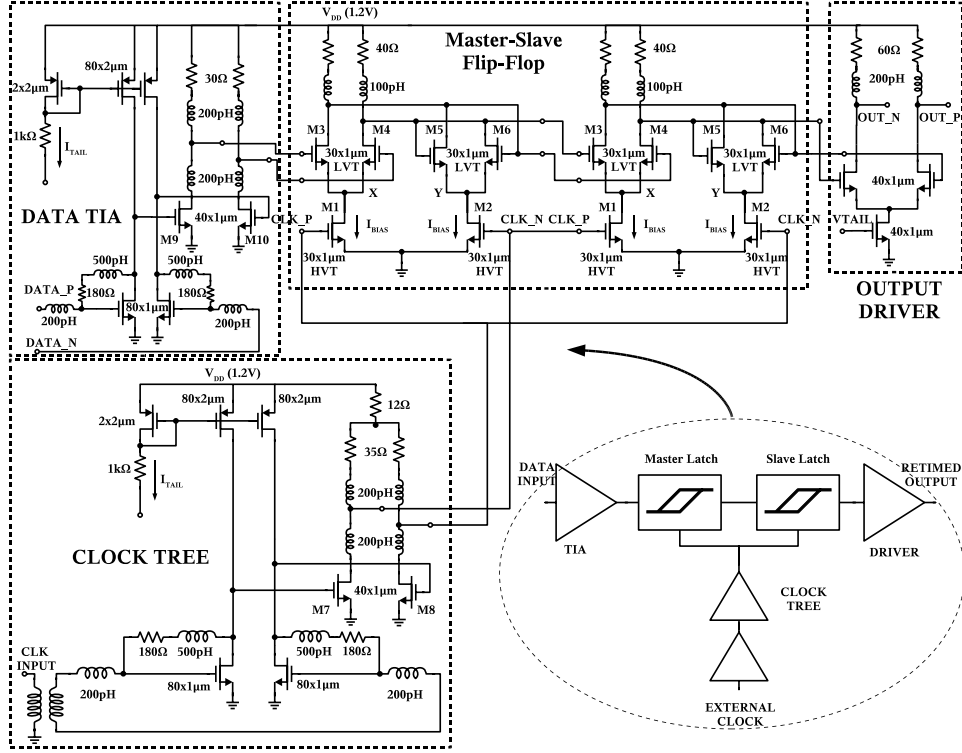


Fig. 2. Full transistor-level schematic of decision circuit.

approach is necessary to ensure that transistors on the clock path have  $V_{DS} > 0.3V$ , as needed for operation at 40 Gb/s. When M1 or M2 is turned off, its  $V_{GS}$ , which is set by the  $V_{DS}$  of M8-M7, must be equal or lower than  $V_T$ . A high- $V_T$  device (0.34V) on the clock path solves this problem, with a relatively small degradation of the  $f_T$  of M7/M8, which remains larger than 80 GHz at  $V_{DS} = 0.34$  V. On the other hand, when M1 or M2 is conducting,  $V_X = V_{DS,M1} = V_{DD} - \Delta V_{swing} - V_T$ . By choosing a low  $V_T$  (0.18V) device for M3-M6, the  $V_{DS}$  and speed of M1/M2 are maximized.

### B. Data and Clock Buffers

The buffer stages that drive the data and clock signals to the latches are implemented as low-noise transimpedance amplifiers (TIA) [8]. As illustrated in Fig. 2, to simplify testing, a 1:1 vertically stacked transformer is implemented on the die to convert the single-ended external clock to a differential signal applied to the TIA input. The transformer limits the bandwidth of the clock tree at low frequencies to about 20 GHz. The TIA consists of a NMOS inverter with shunt-shunt resistive and inductive feedback for bandwidth improvement and impedance/noise matching. The PMOS load is needed to increase the gain of the amplifier at low supply voltages, at the expense of higher capacitance at the output node. The latter effect is mitigated by the feedback inductor, which resonates out the parasitic capacitance of the NMOS and PMOS transistors. The 500-pH inductor in the TIA feedback loop is realized with vertically stacked windings in the two top metal layers of the process. It has

35  $\mu\text{m}$  diameter, 2  $\mu\text{m}$  conductor width and a self-resonance frequency exceeding 100 GHz. The NMOS and PMOS transistors are biased at the minimum noise current density of 0.15 mA/ $\mu\text{m}$  and 0.07 mA/ $\mu\text{m}$ , respectively [8]. The PMOS current mirrors control the bias currents of the MOSFETs in the TIA stages and in the following MOS-CML inverters, making them independent of temperature and power supply variations. MOS-CML inverter stages with inductive peaking are placed between the TIA stages and the latch to provide the proper DC and signal levels to the latches in the flip-flop. The common mode resistor at the clock tree output sets the appropriate DC voltage level at the gates of M1 and M2 such that they are biased at 0.15 mA/ $\mu\text{m}$  and switch from 0 to 0.3 mA/ $\mu\text{m}$ . It should be noted that a current source cannot be employed to bias the differential pair M7-M8 due to lack of voltage headroom. Since sensitivity is not a problem on the clock path, its input TIA stage could be replaced by a chain of inductively peaked MOS-CML inverters to further increase the 3-dB bandwidth of the clock tree.

### III. FABRICATION AND MEASUREMENT RESULTS

To verify the portability of the design, the retiming DFF was fabricated in two different 90-nm CMOS processes. Both dies (Fig. 3) occupy  $800 \times 600 \mu\text{m}^2$  including the pads. All transistor sizes are identical and passive components have been scaled to have the same value (R and L) in both technologies. Between the two designs, the circuit from foundry A employs low- $V_T$  devices only, while the circuit from foundry B has both low and high- $V_T$  MOSFETs, as in Fig. 2.

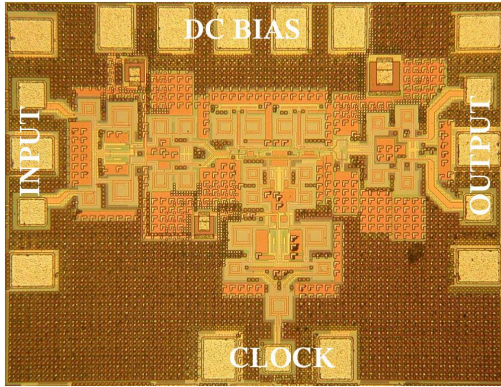


Fig. 3. Die photo of decision circuit.

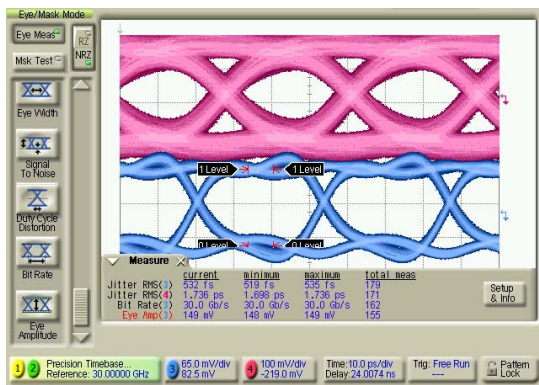


Fig. 4. Input (top, channel 4) and output (bottom, channel 3) eye diagrams at 30 Gb/s ( $V_{DD}=1.2V$ ).

The circuits were tested on wafer. In the absence of a full-fledged 40-Gb/s PRBS BERT, the 40-Gb/s PRBS data were generated by multiplexing 4 streams at 10 Gb/s each.

Figure 4 reproduces the input and output eye diagrams at 30 Gb/s, showing a significant reduction in jitter from 1.7 to 0.5 ps rms. The rise/fall times are improved to less than 7 ps (Fig. 5). Contributions from the test setup and scope have not been de-embedded from the measured jitter and rise/fall times. Full-rate retiming from 1.2 V was experimentally verified up to 37 Gb/s (Fig. 6). To reach 40 Gb/s (Fig. 7), the power supply was increased to 1.5 V. The resulting bathtub curve at 40 Gb/s can be found in Fig. 8. Error-free operation was confirmed for an input pattern of  $4 \times (2^7 - 1) = 508$  bits, by capturing the input and output bitstreams on the Agilent DCA-86100C sampling scope. Part of the captured bitstream at 40 Gb/s is shown in Fig. 9. Power dissipation at 1.2 V and 1.5 V is 130 mW and 240 mW, respectively. Measurements on the circuit with only low- $V_T$  devices show somewhat degraded performance. The latter worked up to 32 Gb/s due to the lack of high- $V_T$  transistors in the clock path. Table I compares this circuit to state-of-the-art latches in SiGe BiCMOS and InP technologies. The MOS-CML latch has the lowest power dissipation, while operating at 15 – 20% lower data rates than the SiGe circuits, tracking the 120-GHz to 150-GHz  $f_T$  ratio of these technologies.

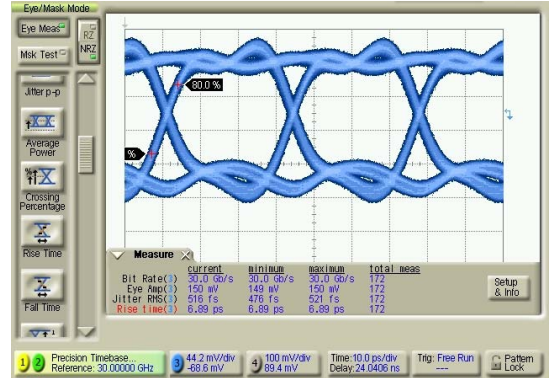


Fig. 5. Output eye diagram at 30 Gb/s ( $V_{DD}=1.2V$ ) showing rise time smaller than 7 ps.

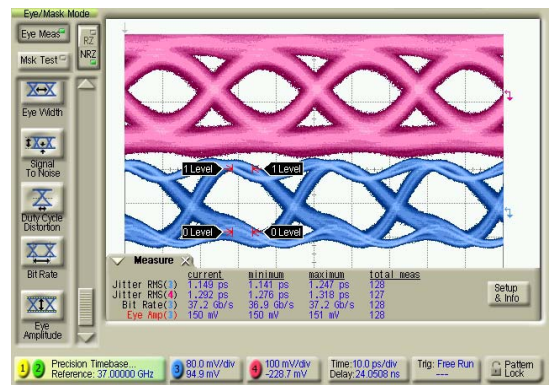


Fig. 6. Input (top, channel 4) and output (bottom, channel 3) eye diagrams at 37 Gb/s ( $V_{DD}=1.2V$ ).

Measurements were also performed across temperature for different supply voltages to verify the robustness of the biasing scheme. The circuit was tested for supply voltages between 1 V and 1.5 V and at temperatures up to 100°C. At 1-V supply and 100°C, the maximum rate with retiming and jitter reduction is 32 Gb/s. Figure 10 shows the 40 Gb/s eye diagram at 1.5 V and 100°C. Even though no errors were observed in this case, the output jitter is not improved over that at the input, indicating that the clock path does not have enough bandwidth to retime the data.

#### IV. CONCLUSION

A low-power decision circuit has been demonstrated at 40 Gb/s in 90-nm CMOS. The circuit achieves full-rate retiming at 37 Gb/s and 40 Gb/s from 1.2 V and 1.5 V, respectively. Measurements vs temperature prove the robustness of the proposed latch biasing scheme. At 32 Gb/s, retiming (with reduction in jitter) was verified for supply voltages as low as 1 V and 100°C.

#### ACKNOWLEDGMENT

This work was funded by Nortel Networks. The authors wish to thank ECTI, OIT, and CFI for equipment and CMC for CAD support. Chip fabrication was provided through Nortel Networks and by TSMC.

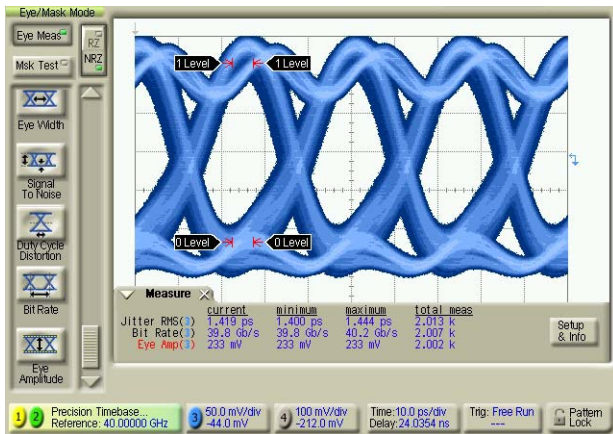


Fig. 7. Output eye diagram at 40 Gb/s and 25°C ( $V_{DD}=1.5V$ ) with  $2 \times 223mV_{pp}$  output swing.

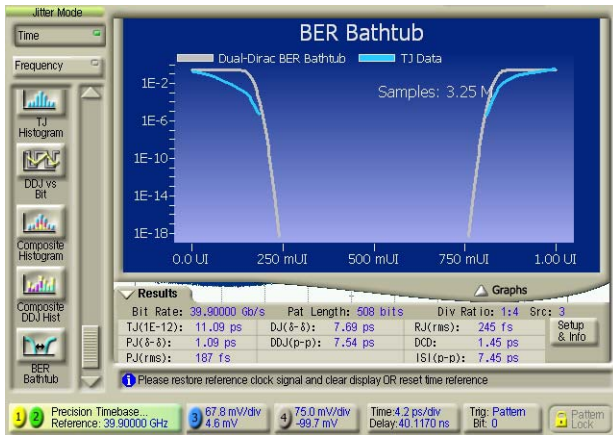


Fig. 8. Bathtub curve of output at 40 Gb/s and 25°C ( $V_{DD}=1.5$ ).

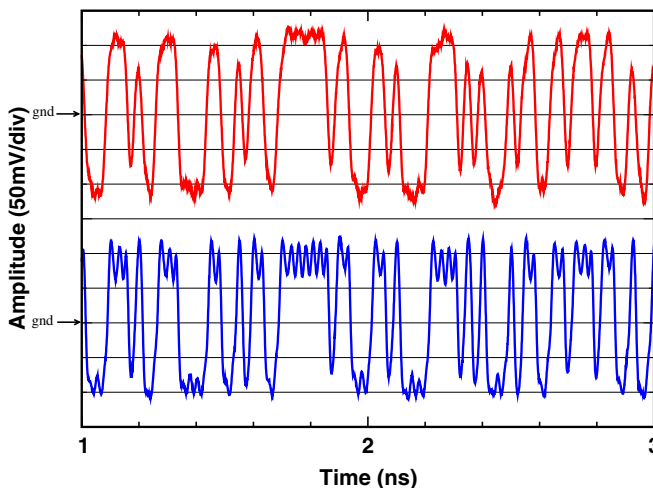


Fig. 9. Input (top) and output (bottom) signals for a 508-bit pattern at 40 Gb/s and 25°C ( $V_{DD}=1.5V$ ).

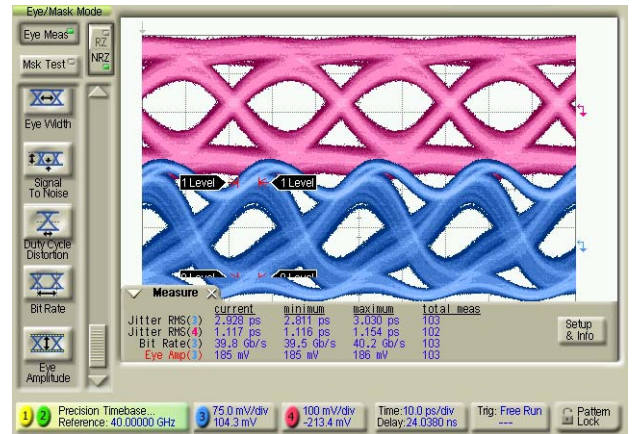


Fig. 10. Input (top, channel 4) and output (bottom, channel 3) eye diagrams at 40 Gb/s and 100°C ( $V_{DD}=1.5V$ ).

TABLE I  
COMPARISON OF HIGH-SPEED LATCHES IN DECISION CIRCUITS

Ref	Technology	Rate (Gb/s)	Supply (V)	$P_{latch}$ (mW)	$P_{dec,circuit}$ (mW)
[3]	245-GHz InP HEMT	80	5.7	N/A	1200
[4]	150-GHz InP HBT	50	1.5	20	125
[8]	150-GHz SiGe BiCMOS	48	2.5	29	288
[1]	120-GHz SiGe HBT	43	3.3	N/A	N/A
This work	120-GHz CMOS	37	1.2	10.8	130
		40	1.5	20	240

REFERENCES

- [1] M. Meghelli, "A 43-Gb/s full-rate clock transmitter in 0.18 $\mu$ m SiGe BiCMOS technology," *IEEE J. Solid-State Circuits*, vol. 40, pp. 2046–2050, Oct. 2005.
- [2] T. Chalvatzis and S. P. Voinigescu, "A low-noise 40-GS/s Continuous-Time Bandpass  $\Delta\Sigma$  ADC centered at 2GHz," in *Radio Frequency Integrated Circuits (RFIC) Symposium*, 2006, to appear.
- [3] T. Suzuki *et al.*, "A 80-Gbit/s D-type flip-flop circuit using InP HEMT technology," *IEEE J. Solid-State Circuits*, vol. 39, pp. 1706–1711, Oct. 2004.
- [4] Y. Amamiya *et al.*, "Low supply voltage operation of over-40-Gb/s digital ICs on Parallel-Current-Switching latch circuitry," *IEEE J. Solid-State Circuits*, vol. 40, pp. 2111–2117, Oct. 2005.
- [5] T. O. Dickson *et al.*, "A 2.5-V, 45-Gb/s decision circuit using SiGe BiCMOS logic," *IEEE J. Solid-State Circuits*, vol. 40, pp. 994–1003, Apr. 2005.
- [6] K. Kanda *et al.*, "A 40Gb/s 4:1 MUX/1:4 DEMUX in 90nm standard CMOS," in *ISSCC Dig. Tech. Papers*, 2005, pp. 152–290.
- [7] D. Kehrer and H.-D. Wohlmuth, "A 60-Gb/s 0.7-V 10-mW monolithic transformer-coupled 2:1 multiplexer in 90 nm CMOS," in *IEEE Compound Semiconductor IC Symp. Tech. Digest*, 2004, pp. 105–108.
- [8] T. O. Dickson *et al.*, "The invariance of characteristic current densities in nanoscale MOSFETs and its impact on algorithmic design methodologies and design porting of Si(Ge) (Bi)CMOS high-speed building blocks," *IEEE J. Solid-State Circuits*, Aug. 2006, to be published.
- [9] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*, 2nd ed. Cambridge, UK: Cambridge Univ. Press, 2004, ch. 9.