A Fine-Grained, Uniform, Energy-Efficient Delay Element for FD-SOI Technologies

Ajay Singhvi[*], Matheus T. Moreira[‡†||], Ramy N. Tadros[†], Ney L. V. Calazans[‡||] , Peter A. Beerel[†§]

[*]Birla Institute of Technology and Science Pilani, Pilani Campus - Pilani, India.
[†]University of Southern California (USC) - Los Angeles, United States
[‡]Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS) - Porto Alegre, Brazil
ajaysinghvi93@gmail.com {matheus.moreira, ney.calazans}@pucrs.br, {rtadros, pabeerel}@usc.edu

Abstract—Contemporary digitally controlled delay elements trade off power overheads and delay quantization error. This paper proposes a new delay element that provides a balanced design that yields low power with low delay quantization error. The proposed element has a quasi linear delay characteristic, with uniform delay differences between adjacent codewords. The element employs and leverages the advantages offered by a 28nm FD-SOI technology, using its back body biasing feature to add an extra dimension to its programmability. To do so, a novel generic delay shift block is proposed, which enables incorporating both fine and coarse delays in a single delay element that can be easily integrated into digital systems, an advantage over hybrid delay elements that rely on analog design.

I. Introduction

Delay Elements (DEs) are used in a variety of applications in VLSI systems and are typically employed to provide precise timing control and/or satisfy timing constraints. In synchronous systems, DEs support clock distribution and synchronization across different blocks, dealing with clock skew and jitter problems [1], [2]. Other uses include phase locked loops, digitally controlled oscillators [3], time-to-digital converters [4] and poly-phase clock generators [5]. DEs are also widely used in bundled data (BD) asynchronous systems, to control the timing of request and acknowledge signals between different blocks [6]. For some of these applications, like control circuits of 2-phase BD asynchronous designs, DEs require balanced rise and fall delays [6], [7]. Moreover, a typical concern in the design of DEs in modern technologies are the effects of process, voltage and temperature (PVT) variations. To account for those, DEs must be conservatively designed to have extra timing margins that can compromise performance. The alternative is to use programmable DEs.

Programmable DEs alleviate the detrimental effects of PVT variations in deep sub-micron technologies by providing a range of attainable delays to which the DE can be tuned post-silicon. The delay granularity provided by programmable DEs is an important concern. For instance, systems that require precise timing control, such as phase shift compensators [8], timing generators [9] and timing verniers [10], used for delay fault testing in automatic testing equipment, employ fine-grained DEs to ensure correct operation. In essence, the precision to which these DEs can be tuned affects the amount of timing margin they can effectively avoid. DEs can be controlled by either analog voltages (or currents), or digitally. Traditionally, analog-controlled DEs provide fine delay tuning, while digitally-controlled DEs provide coarse-grained delays, with their combination forming hybrid DEs. However, since this work deals primarily with low-power applications and high-performance digital VLSI circuits, the use of hybrid DEs is not considered, to avoid the high power consumption of the required analog circuitry, the switching noise at high

frequencies, and the challenges in the distribution of global analog signals in predominantly digital systems [6].

In contrast, this paper proposes a new circuit architecture and the use of FD-SOI technology to design fine-grained programmable DEs with balanced rise and fall delays. Section II discusses the state of the art in digitally-controlled DEs. Next, section III explains the design of the proposed DE to provide a quasi linear and monotonic delay characteristic, reducing the delay quantization error (DQE) to 12.57% from 269.92% presented by a state of the art DE. It also proposes the architecture of the delay shift inverter (DSI), which utilizes the FD-SOI back-body biasing feature [11], to provide fine-grained delays in a single DE structure that can be easily incorporated into digital systems without any of the problems posed by hybrid DEs. Section IV then discusses the methodology adopted for optimizing power consumption of the proposed design, resulting in significantly lower energy consumption when compared to existing DEs [12], [13]. Section V presents and discusses our experimental results, while Section VI draws a set of conclusions.

II. Digitally-Controlled DEs

Different digitally-controlled DE architectures exist in the literature, exploring trade-offs in terms of delay range, power consumption, and area. Some of the existing DE topologies are the thyristor-based [14], the transmission gate-based [15], the current starved [12] and the cascaded inverter-based [16] designs. Among these, thyristor-based designs provide delays in ranges from μs to ms. This is beyond the scope of this paper, which focuses on DEs that provide shorter delay ranges (from ps to a few ns). Moreover, it is difficult to control both rise and fall transitions in thyristor-based designs. As for the transmission gate-based DE, it suffers from poor signal integrity and modifications that alleviate the problem [15] add significant costs in terms of area and power.

Therefore, the focus here is on current starved and cascaded inverter-based designs. Attention is put on the directly-controlled current starved DE (DCCS-DE), analyzed in [12] and shown in Fig. 1. This DE falls under the category of current starved inverter (CSI)-based DEs, where current source transistors determine the current through an inverter. These transistors reduce the current through the inverter, or starve it, thereby increasing the delay of a signal propagating through it. This DE, as analyzed in [12] displays a few drawbacks, including a non-monotonic relationship between delay and the associated codewords which makes it difficult to predict the delay for a given codeword. Another problem is the non-uniform delay difference between successive codewords. Such non-linearity translates into large delay quantization error (DQE), which is problematic when one requires a delay not provided by any codeword.

In [6] the authors propose a modified DCCS-DE design to allow balanced rise and fall delays, which is beneficial
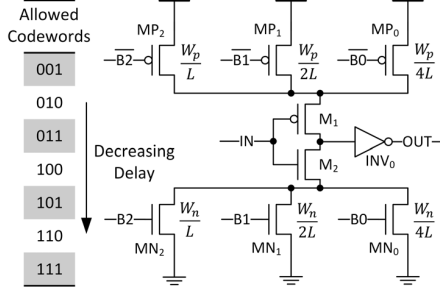
Fig. 1.   The directly controlled current starved DE (DCCS-DE) [12].

for 2-phase BD asynchronous circuits. The modified design comprises two replicated DCCS-DEs in series, with signal conditioning inverters added to their inputs, to provide an acceptable slew rate, and inverters at their outputs, to provide the same load to each of the replicated DCCS-DEs. Unfortunately, the modified DCCS-DE still exhibits the other disadvantages of the original DE. Maymandi-Nejad and Sachdev proposed in [13] an alternative to cope with this problem: a programmable current mirror-based current starved DE (CMCS-DE) which has a linear delay characteristic. However, the CMCS-DE suffers from very large static power consumption, as discussed in [6], and cannot be employed in low power applications. In fact, a significant advantage of the DCCS-DE is its better energy efficiency when compared to other DE architectures, as [6] discusses, with its static power consumption being three orders of magnitude lower than that of the CMCS-DE.

The multiplexer-based DE (MUX-DE), depicted in Fig. 2 is often used in designs. Its popularity arises from a relatively simple design that can be implemented using standard cells. It also presents a linear delay characteristic. The codeword provided to the MUXes fixes the number of inverters in the signal path and hence its delay. This paper uses the MUX-DE as a comparison baseline design, to analyze the impact of the proposed modifications on the energy efficiency of the DCCS-DE over the MUX-DE.

III.   PROPOSED DESIGN OF FINE-GRAINED DEs

A. Reducing Delay Quantization Error

Since the expected delay of a DE is uniformly distributed across the codewords, the delay quantization error (DQE), cited in Section II, is defined as the percentage deviation from the expected delay difference between any two adjacent codewords. In particular, the DQE is the maximum DQE across all adjacent codewords. The DQE is a handy metric for programmable DEs, since it encompasses the features of monotonicity, uniform delay distribution across codewords and the ability to predict the amount of delay provided by a particular codeword. Formally we have:

$$DD_{expected} = \frac{Delay\ Range}{N - 1}, \qquad (1)$$

$$DQE = \frac{max(|DD_{measured,i} - DD_{expected}|)}{DD_{expected}} * 100\%, \qquad (2)$$

where Delay Range is the delay difference between the minimum and maximum delay settings and DD refers to delay difference. $DD_{measured,i}$ is the delay difference between the ith and $(i + 1)^{th}$ adjacent codewords as observed in simulations, and $DD_{expected}$ is the ideal delay difference computed by (1). N is an integer representing the number of codewords that can be employed with a particular DE. A minimum DQE is required to enable the DE to be used efficiently across all codewords and possible delay values.
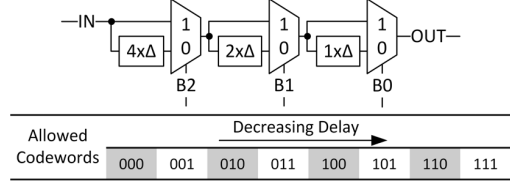

Fig. 2.   The multiplexer based DE.

As previously mentioned, both the DCCS design of [12] and the modified one in [6] have unpredictable and non-monotonic delay behavior, which results in a large DQE. To minimize DQE while taking into consideration low power and high density results in state-of-the-art DEs, a new version of the DCCS-DE is proposed herein, based on the design presented on [6]. The new architecture, illustrated in Fig. 3, uses a one-hot code scheme instead of the binary codes used in [12] and [6]. This imposes the constraint that for a particular codeword, only one of the current source transistors is ON. In moving from a binary to one-hot code, the lengths of the current source transistors were altered to increase linearly (1L, 2L, 3L, 4L, ..., nL), instead of exponentially (1L, 2L, 4L, 8L, ..., $2^{(n-1)}$L), where n is the number of current source transistors, chosen on the basis of the amount of delay needed. The above changes ensure a constant delay difference between any two adjacent codewords, thereby minimizing DQE. This can be demonstrated mathematically as follows:

$$t_{pd} = C_L \frac{V_{ds}}{I_{ds}} \text{ and } I_{ds}\ \alpha\ \frac{1}{R_{ds}}, \qquad (3)$$

with

$$R_{ds}\ \alpha\ \text{L} \implies t_{pd}\ \alpha\ \text{L}. \qquad (4)$$

Thus, as L increases linearly for different codewords, the delay also increases linearly. This is different from the binary scheme used in previous works, where multiple parallel current source transistors could simultaneously be active. This implies summing currents together, which produces a non-linear relation between the total current and the codewords, and hence results in a non-linear delay behavior.

Regarding the MUX-DE, the design proposed in [6] uses a sum-of-product MUX implementation. This has an intrinsically linear delay behavior, because changes in the number of cascaded inverters from one codeword to the next are constant, ensuring a low DQE. Note that the MUX-DE still utilizes a binary codeword, as opposed to the one-hot scheme employed in the proposed DCCS-DE design.

B. Allowing Fine-Grain Tuning

The UTBB (Ultra Thin Body and Buried oxide) FD-SOI technology provides devices with better performance, lower leakage, and several power management design techniques. The Si film (the FDSOI ultra thin body) is very thin so that the depletion region continues to its end. This fully depleted (FD) body results in devices with low sub-threshold slope and low drain-induced barrier lowering (DIBL) figures. Transistors are normally controlled by the high-κ metal gate, which is called the *front gate*. Also, due to the very small width of the ultra thin body and buried oxide (or box), applying a potential from the back-body (or the back gate) has a large influence on the transistor's threshold voltage [11]. This is what is called back-body biasing, or just body biasing. There are two ways to employ body biasing: (i) forward body biasing (FBB), which decreases the threshold voltage for a faster mode of operation; and (ii) reverse body biasing (RBB), which increases the threshold voltage and, hence decreases the leakage current for power management purposes.

28

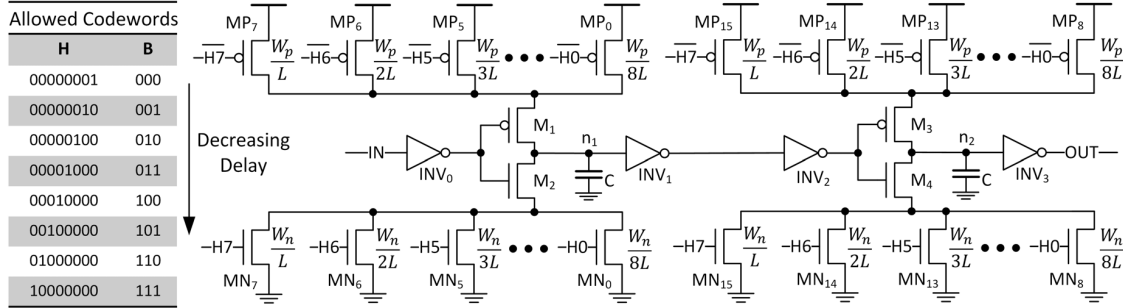| Allowed Codewords	
H	B
00000001	000
00000010	001
00000100	010
00001000	011
00010000	100
00100000	101
01000000	110
10000000	111

Fig. 3.   The proposed DCCS-DE.

While body biasing is conventionally used either to reduce power consumption or to provide a performance boost, this paper employs it to provide fine-grained delay control. Therefore, we focus on RBB, as increasing the threshold voltage enables not only increasing the delay of transistors, but also reducing their leakage power, a side benefit to our techniques. But RBB is not applied to all transistors of the DE because each of them would get affected differently depending on its size. Moreover, this adds to the complexity of the design and, hence, the delay characteristic can change significantly. Also, it would increase the load that the bias voltage generating circuitry has to drive, resulting in more power consumption. Therefore, instead of employing RBB in each separate transistor, we propose the use of a Delay Shift Inverter (DSI) as shown in Fig. 4. The DSI is a conventional CMOS inverter with a programmable back body voltage that adjusts the threshold voltage of the inverter transistors, altering the current flowing through the inverter, changing its delay. Under normal operating conditions, the back-gate of the inverter pMOS transistor is connected to the core supply, while the back-gate of the nMOS is connected to ground. As illustrated in Fig. 4, depending on availability, additional body biasing voltages can be applied to (a) both pMOS and nMOS (b) only pMOS (c) only nMOS transistors. The delay shift provided by the DSI depends on two factors: (i) the change in the back body voltage, and (ii) transistors size. The number of delay shifts can be increased by additional body biasing voltages or by using differently sized DSIs. Section V explores this further.

DSIs can be easily incorporated into any existing DE architecture, as Fig. 5 illustrates. The intrinsic rise and fall delay characteristic of the original DE can also be maintained by cascading two DSIs in series as shown in Fig. 5, with buffers used to provide identical loads to both DSIs. The novelty of using the DSI is thus threefold: (i) it does not alter the original delay characteristics; (ii) it leads to less overhead in terms of area, as compared to replicating the DE architectures to increase the delay range; and (iii) it can be applied to any DE architecture. Hence, it serves as a good candidate to cope with the problems of using hybrid DE architectures to achieve precise and fine-grained delays. Moreover, applying body biasing to specific inverters is well-suited to the proposed DCCS-DE, because the biasing can be directly applied to the existing signal conditioning inverters (INV0 and INV2) of the DCCS-DE design (Fig. 3), instead of using additional area- and power-expensive DSI blocks.

Despite its advantages, a complication for the DSI is the generation and control of voltages from a domain other than the core supply and ground. The two most practical solutions are level shifters and voltage charge pumps. A level shifter is a simple circuit that shifts an input signal from its voltage domain to the provided reference domain and is commonly used to interface off-chip and on-chip voltage domains. On
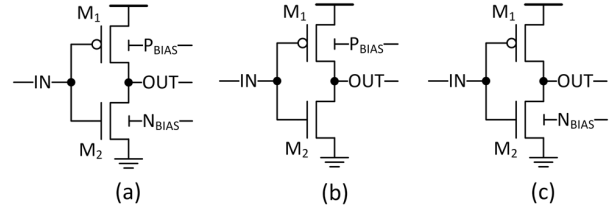

Fig. 4.   Delay shift inverters (DSIs) with RBB applied to: (a) both pMOS and nMOS (b) only pMOS (c) only nMOS.

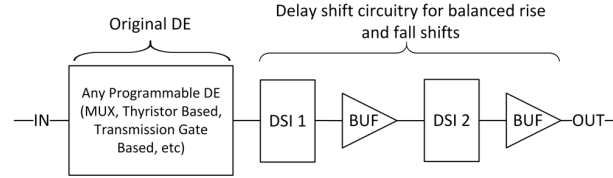Fig. 5.   The proposed architecture of fine-grained DEs.

the other hand, charge pumps use a complex arrangement of switching capacitors to pump charges up and generate higher voltage levels using only the input supply. Level shifters are small, simple, fast, and do not use any passive components. Charge pumps, on the other hand, consume significant amount of area due to the need of capacitors, power due to the need of switching clocks, and delay due to the time required to pump charges through several stages of capacitors. Also, its voltage output suffers from ripples [17]. The advantage of charge pumps is that they need no reference voltages.

Our target DSI, employs RBB only in the pMOS transistors (Fig. 4(b)), because the I/O voltage for the 28nm FD-SOI technology is 1.8V and it is easy to use this voltage internally to drive the pMOS RBB circuitry. In particular, we propose to use level shifters to actively switch the pMOS back-body from the normal supply VDD=1V to the I/O voltage Vhigh=1.8V.

Fig. 6 shows our level shifter. Tran et al. [18] proposed the contention mitigated level shifter (CMLS), which was used in [19] for body biasing the LVT (flip-well) devices in FD-SOI. Since the devices used in the design of this paper's DEs are RVT (normal well), the low voltage connected to the CMOS inverter is VDD and not ground as in [19]. The circuit works as follows: when the input (IN) is low, M4 and M5 are on, while M3 and M6 are off. Then, the gate of M2 is discharged to ground, resulting in the charging of the input of INV1 to Vhigh and hence the output is only VDD. When the input (IN) is high, symmetrically the input of INV1 goes low, and hence the output is Vhigh. A conventional level shifter does not have transistors M3 and M4, but this entails a serious contention between the cross-coupled pMOS devices and the input coupling nMOS. Adding M3 and M4 reduces this contention and results in lower switching energy
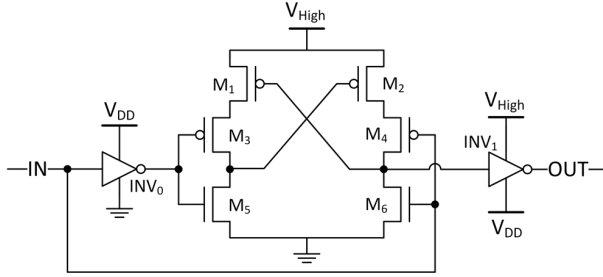
29

Fig. 6. The circuit diagram of the employed level shifter. Terminals shown across the CMOS inverters represent the connected high and low supply values.

Fig. 7. Leakage vs delay trade-off in gate length biasing.

and in faster switching. This is why this architecture is called CMLS [18]. It is worth mentioning that Vhigh=1.8V is the highest voltage value that can be used by such an architecture, because it results in 1.8V across the gates of the MOS devices, which is the maximum difference of potential allowed to avoid gate oxide breakdown. Regarding the overheads of the level shifter addition, leakage and area are relatively small, as is the switching delay, due to the use of contention mitigation.

IV. ENERGY AND LEAKAGE OPTIMIZATIONS

A. Energy Optimization

To minimize the DE energy consumption, an initial version was analyzed to determine the consumption in different parts of the circuit. Next, we redesigned the circuit in the most energy-efficient manner by optimizing each of these parts. As Equation (3) illustrates, for a given operating voltage the provided delay depends on two factors: the current through the CSI, and the output capacitive load, C_L. The former and the latter are controlled by the following parameters: *i)* the W and L of the current source transistors (MPx and MNx); *ii)* the W and L of the CSIs (Mx); *iii)* the external load capacitance, C; and *iv)* the input capacitance of the signal conditioning inverters. Parameters *i)-iv)* are to be tuned to get the required delay range, consuming the least energy per transition and having the lowest leakage.

To better understand the energy vs delay trade-offs for the above parameters, experiments were conducted in which each parameter was used to independently achieve a fixed delay range. The experiments show that increasing L of the current source transistors is the most energy-efficient manner to achieve the required delay range, as larger L results in lesser current and hence lesser energy. However, the maximum L that can be used is constrained by the layout rules of the technology and might not always be enough to get the desired delay range, especially for larger delays.

To increase the delay range, one may further decrease the current by increasing the L of the CSIs (parameter *ii*), increase the output capacitance, by adding an external shunt capacitor (item *iii*), or increase the size, and thus the input capacitance, of the signal conditioning inverters (parameter *iv*). However, increasing L of the CSIs must be done conservatively, ensuring that the CSIs do not dominate the current source transistors by constraining the maximum current that can flow. This leaves two options, both implying the increase of the output capacitance. An added advantage of using parameters *iii)* and *iv)* is that these help mitigate the charge sharing problem present in the DCCS-DE from [12].

From experiments we conclude that increasing the input capacitance of the signal conditioning inverters yields the largest energy overhead. This is because increasing the input capacitance for these results in a slow slew rate, which in turn generates more short circuit current through the inverters,
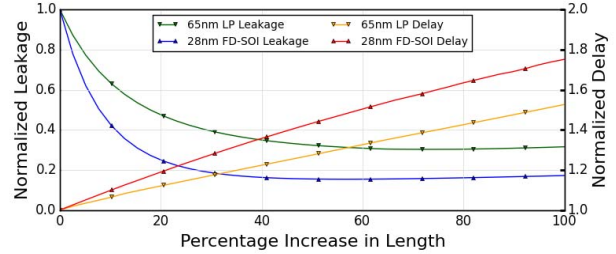
and hence leads to larger energy per transition (EPT). Thus, adding an external shunt capacitance at the output node is the preferred approach. In fact, an optimal combination of *i)* and *iii)* helps achieving the best energy-delay trade-off.

As for the MUX-DE, its delay range depends on two factors: *i)* the number of cascaded inverters; and *ii)* the W and L of these. The optimization of EPT for the MUX-DE is better done using approach *ii)*, i.e. increasing the lengths of the nMOS and pMOS transistors to meet the desired delay range, rather than adding more cascaded inverters. Approach *i)* is used only after reaching the maximum allowable L for a transistor, because a higher L would result in less current flowing through the inverters, increasing delay. Cascading smaller inverters results in additive current flowing through the DE, consuming more EPT.

B. Leakage Reduction

Gate length biasing [20] [21] is a promising technique for achieving substantial leakage reduction, and also requires no additional process steps. It involves increasing the length of the transistors to reduce leakage, at the cost of a delay increase. Gupta et al. [21] suggest a 10% upper bound on the increase in length to achieve the best trade-off for a bulk 130nm process. Experiments were run on an inverter in a bulk 65nm process as well as in the 28nm FD-SOI process to decide on a bound. The trade-off can be seen in Fig. 7, with greater reduction of leakage in the 28nm FD-SOI process as compared to the 65nm bulk CMOS technology, at the expense of an increase in delay.

This leakage vs delay trade-off is important when applying gate length biasing to transistors in the critical path, because of timing constraints that limit the increase in L to only 10%. However, the same limitation does not exist for DEs, because the delay range can always be tuned using other parameters like an external shunt capacitance or changing the number of DEs used to build the element. Thus, overlooking the 10% limitation advocated in [20], it can be observed from Fig. 7 that after roughly 40% increase in L, one obtains high leakage reduction, after which leakage reduction stagnates. Thus, while designing any of the DEs, the minimum L chosen is 40% greater than the technology's smallest L. Experiments on the DCCS-DE and the MUX-DE showed the same trend, with substantial leakage reduction.

V. EXPERIMENTS AND DISCUSSION

A 28nm FD-SOI CMOS technology with 1V supply was used for the DEs. All simulations employed the Cadence Spectre Simulator with a same environment across all designs, for fair comparisons. All simulations assumed an operating temperature of 27°C and typical process corners. DEs employed the techniques proposed in Section III as well as the power reduction optimizations of Section IV. Each DE was designed to have 8 different delay settings and provide an identical delay range.

Table I summarizes the trade-offs between DEs. The techniques from Section III-A significantly improve the delay

30

TABLE I. TRADE-OFFS BETWEEN DEs FOR A 400PS RANGE.

DE	Original Binary DCCS-DE [6]	Proposed One-hot DCCS-DE	MUX-DE
DQE (%)	269.92	18.94	3.19
Avg. EPT (fJ)	0.73	1.01	2.71
Avg. Idle Power (nW)	0.30	0.12	0.28
Active Area (μm^2)	1.18	1.96	0.35
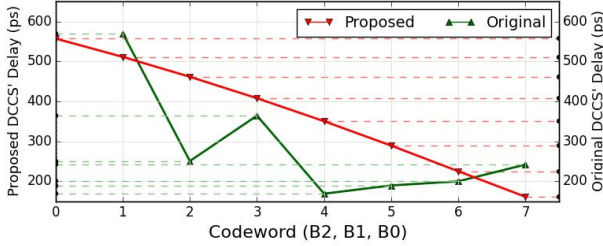
Fig. 8. Comparison of delay characteristics for the proposed DCCS-DE and the original DCCS-DE.

Fig. 9. Comparison of proposed DCCS-DE and MUX-DE: (a) EPT (b) $Energy/Delay$ (c) Leakage.

Fig. 10. Effect of L and body biasing voltages on delay shift.

characteristic of the DCCS-DE over the DCCS-DE proposed in [12] and [6]. Improvement is quantified using the definition of DQE in Equation 2. As Fig. 8 shows, the original DCCS-DE has a non-monotonic delay, which is problematic, as certain codewords might provide delays that are too close or too far from each other. This characteristic translates into a large DQE of 269.92%, making the original DCCS-DE unreliable for building a programmable DE. Note that the DQE for the original DCCS-DE was calculated after re-ordering the codewords, to provide a monotonically increasing delay characteristic; still, it presented high DQE. On the other hand, the proposed DCCS-DE has an almost linear delay characteristic, with nearly uniform delay difference between codewords, and does not require codeword reordering. This uniform delay difference enabled a much smaller DQE, 18.94%. Moreover, this DQE improvement comes without significant power overhead. The active area values in Table I are the sum of the W*L of all the transistors in the design. Furthermore, to make the comparison between the proposed and original DEs more pessimistic, any area and power overheads of the circuitry needed to re-order the original DE codewords are not considered when presenting results for the original binary DCCS-DE.

As Table I shows, the MUX-DE displays a better DQE of 3.19% as opposed to 18.94% of the proposed DCCS. This is due to the fact that the technique presented in Section III-A does not take non-idealities into account. For the 400ps delay range programmed using 8 codewords, this translates to an absolute error of 1.82ps, while the proposed DCCS-DE has a max deviation of 10.08ps from the ideal characteristic of having a uniform delay difference of $400/7 ps$ between adjacent codewords. However, with the aforementioned technique as a basis, the DQE achieved by the DCCS-DE can be improved by iteratively adjusting the L's of those current source transistors that contribute to the larger DQE. Moreover, as discussed later in this Section, the fine-graining technique proposed here further improves the DQE, and any issues arising due to minor deviations from the ideal characteristic can also be alleviated. On the other hand, the proposed DCCS-DE still consumes 2.68 times less energy than the MUX-DE.

The metric used for comparing energy efficiency is the average energy per transition for all codewords measured for a particular delay range. As Fig. 9(a) shows, the MUX-DE consumes nearly five times more energy than the proposed DCCS-DE for small delay ranges, due to more current being drawn by the cascaded inverters in the MUX-DE than the CSIs of the DCCS-DE. The disparity decreases as L of the cascaded
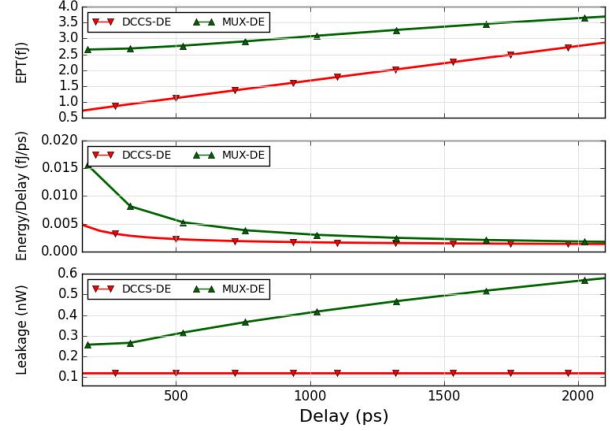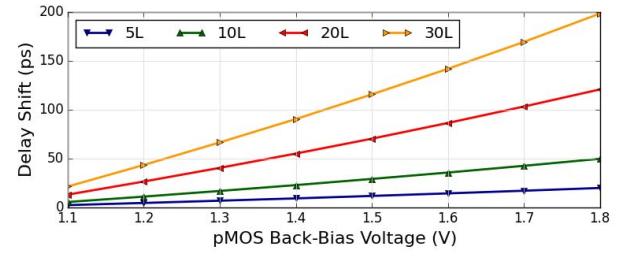
inverters increases to improve the delay range of the MUX-DE, with the energy advantage of the proposed DCCS-DE reducing by a factor of two for ranges larger than 2ns.

To better understand the relationship between delay range and EPT, Fig. 9(b) shows the $Energy/Delay$ relation between DEs. The results are consistent with the above discussion, as for delay ranges bigger than 2ns the energy spent per unit of delay becomes nearly equal for both. Next, the DEs idle power is compared. Leakage reduction is achieved for both DEs using the gate length biasing strategy from Section IV-B. As can be seen from Fig. 9(c), the DCCS-DE has a very low leakage power consumption of 0.12nW, which remains constant across delay ranges, due to the fact that the extended delay ranges were met using external shunt capacitors rather than more transistors. On the other hand, it was noticed that the MUX-DE has substantially higher leakage power consumption when compared to the DCCS-DE. This is attributable to the large transistor count of the MUX-DE, compared to the DCCS-DE.

The next set of experiments target enabling a fine-grained delay range of 400ps for the DCCS-DE and MUX-DE, In other words, the idea is to reduce the delay difference between two adjacent delay settings. As Section III-B mentioned, the amount of delay shift achieved by the DSI shown in Fig. 4 is controlled by the size as well as the magnitude of the additional body biasing voltage. Experiments were run to determine the optimal sizing and voltage. As Fig. 10 shows, the amount of delay shift increases as the body biasing voltage or the length of the transistor increases. Depending on the delay range and the application, the appropriate number and magnitude of body biasing voltages and transistor sizes can be chosen.

For the reasons elaborated in Section III-B, only an additional body biasing voltage of 1.8V is generated, using the contention mitigated level shifter shown in Fig. 6 to add an extra dimension of programmability. Moreover, for this application,
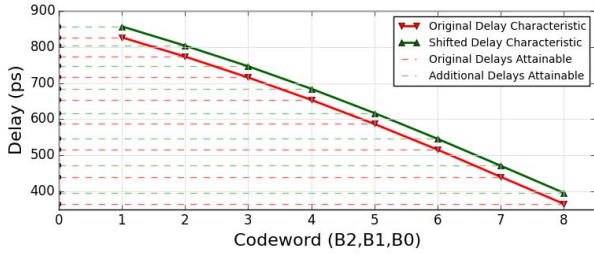
31

Fig. 11. Fine-grained delay characteristic.

TABLE II.   TRADE-OFFS BETWEEN FINE-GRAINED DEs FOR A 400PS RANGE.

DE	One-hot DCCS-DE with 16 current sources	One-hot DCCS-DE with 8 current sources + DSI	MUX-DE with 15 buffers
DQE (%)	26.81	12.57	4.55
Avg. EPT (fJ)	1.03	1.57	5.08
Avg. Idle Power (nW)	0.22	0.16	0.40
Active Area (μm^2)	3.72	3.42	0.28

only the length of the DSI transistors was increased, as it would be the more energy-efficient solution. For a 400ps delay range, across eight codewords, the normal delay difference between each codeword would be $400/7\,ps \approx 57\,ps$. Thus, the length of the transistor chosen is one which corresponds to a delay shift of $400/14\,ps \approx 29\,ps$ for each codeword. The final fine-grained delay characteristic for the DCCS-DE can be seen in Fig. 11. Similar results were also observed for the MUX-DE. As the figure shows, the addition of a single body biasing voltage level doubles the resolution of the discrete delays offered by the DE. In the above experiment a delay step of $\approx 29ps$ is achieved as one moves from one setting to the next. Additional body biasing voltages can be used to further reduce the delay step size and make the achieved delay granularity finer.

In order to study the effect of using the DSI on the DQE, experiments were conducted on the DCCS-DE and MUX-DE. To ensure a fair comparison, the fine-grained structure was implemented in two flavors: one with the DSI and another without it. The two finer grained DCCS-DE designs were implemented by: (a) using 16 current source transistors sized as (1L, 1.5L, 2L, ... , 7.5L, 8L) instead of the original eight sized as (1L, 2L, ... , 7L, 8L); and (b) body biasing the signal conditioning inverters INV0 and INV2 of Fig 3 while still using only eight current source transistors. The MUX-DE design was re-implemented to have 16 codewords instead of the original eight (Fig. 2),so as to have a fair comparison and reduce the delay difference between adjacent codewords. The comparison of these designs appears in Table II.

As Table II shows, using the DSI with the DCCS-DE enables a DQE of 12.57%, which is less than half of that achieved when using additional current source transistors. Compared to the DCCS-DE with sixteen current source transistors, the one with the DSI has lower area and also consumes 3.23 times less energy than the fine-grained MUX-DE, further improving the energy efficiency of the DCCS-DE over the MUX-DE that was observed in Table I. Thus, adding the DSI to the DCCS not only presents a better DQE but also does not result in excessive overheads.

VI. CONCLUSION

This work presents and analyzes design modifications to the DCCS-DE. The proposed design has a linear monotonic delay behavior and low DQE, which is a considerable improvement over the previously discussed DCCS-DEs in [12] and

[6]. Additionally, the proposed DCCS-DE is significantly more energy-efficient than the current mirror based design proposed in [12] and [13]. And, it consumes less energy than the MUX-DE for delay ranges smaller than 2ns. The paper also proposes a generic DSI architecture, which utilizes the body biasing feature in 28nm UTBB FD-SOI technology to obtain fine-grained delays in a single DE structure, allowing the proposed architecture to be easily integrated into digital systems. This DSI architecture can be used to further improve the DQE of the delay element. Such advances enable leveraging the advantages of UTBB FD-SOI technologies for circuit design, and allow better design space exploration for applications that need low power DEs.

REFERENCES

[1] A. Chakraborty et al., "Dynamic Thermal Clock Skew Compensation using Tunable Delay Buffers," IEEE Trans. on VLSI Syst., vol. 16, no. 6, pp. 639–649, Jun. 2008.

[2] Y.-J. Jung et al., "A dual-loop delay-locked loop using multiple voltage-controlled delay lines," IEEE JSSC, vol. 36, no. 5, pp. 784–791, May 2001.

[3] B.-M. Moon et al., "Monotonic Wide-Range Digitally Controlled Oscillator Compensated for Supply Voltage Variation," IEEE Trans. on Circ. and Syst. II: Express Briefs, vol. 55, no. 10, pp. 1036–1040, Oct. 2008.

[4] G. Li et al., "A high resolution time-to-digital converter using two-level vernier delay line technique," in NSS/MIC, 2007, pp. 276–280.

[5] H. Lin et al., "New four-phase generation circuits for low-voltage charge pumps ," in ISCAS, 2001, pp. 504–507.

[6] G. Heck et al., "Analysis and Optimization of Programmable Delay Elements for 2-Phase Bundled-Data Circuits," in VLSID, 2015, pp. 321–326.

[7] P. Beerel et al., A Designer's Guide to Asynchronous VLSI.   Cambridge University Press, 2010.

[8] T. Dogsa et al., "Precision Delay Circuit for Analog Quadrature Signals in Sin/Cos Encoders," IEEE Trans. on Instr. and Meas., vol. 63, no. 12, pp. 2795–2803, May 2014.

[9] K. Ryu et al., "All-digital process-variation-calibrated timing generator for ATE with 1.95-ps resolution and a maximum 1.2-GHz test rate," in ESSCIRC, 2013, pp. 41–44.

[10] B. Arkin et al., "Realizing a production ATE custom processor and timing IC containing 400 independent low-power and high-linearity timing verniers," in ISSCC, 2004, pp. 348–349.

[11] B. Pelloux-Prayer et al., "Fine grain multi-VT co-integration methodology in UTBB FD-SOI technology ," in VLSI-SoC, 2013, pp. 168–173.

[12] M. Maymandi-Nejad and M. Sachdev, "A Digitally Programmable Delay Element: Design and Analysis," IEEE Trans. on VLSI Syst., vol. 11, no. 5, pp. 871–878, Oct. 2003.

[13] M. Maymandi-Nejad and M. Sachdev, "A Monotonic Digitally Controlled Delay Element," IEEE JSSC, vol. 40, no. 11, pp. 2212–2219, Nov. 2005.

[14] G. Kim et al., "A Low-voltage, Low-power CMOS Delay Element," IEEE JSSC, vol. 31, no. 7, pp. 966–971, Jul. 1996.

[15] N. Mahapatra et al., "An Empirical and Analytical Comparison of Delay Elements and a New Delay Element Design," in VLSI, 2000, pp. 81–86.

[16] N. Mahapatra et al., "Comparison and Analysis of Delay Elements," in MWSCAS, 2002, pp. 473–476.

[17] G. Palumbo and D. Pappalardo, "Charge Pump Circuits: An Overview on Design Strategies and Topologies," IEEE Circ. and Syst. Mag., vol. 10, no. 1, pp. 31–45, Mar. 2010.

[18] C. Tran et al., "Low-power High-speed Level Shifter Design for Block-level Dynamic Voltage Scaling Environment," in ICICDT, May 2005, pp. 229–232.

[19] J. Hamon and E. Beigne, "Automatic Leakage Control for Wide Range Performance QDI Asynchronous Circuits in FD-SOI Technology," in ASYNC, 2013, pp. 142–149.

[20] C. Lazzari et al., "An automated design methodology for layout generation targeting power leakage minimization," in ICECS, 2009, pp. 81–84.

[21] P. Gupta et al., "Gate-length biasing for runtime-leakage control," IEEE Trans. on CAD of Int. Circ. and Syst., vol. 25, no. 8, pp. 1475–1485, Aug. 2006.