Post-silicon skew tuning with variability of programmable delay element

Daijiro MUROOKA†, Qing DONG†, Yasuhiro TAKASHIMA†, and Shigetoshi NAKATAKE†
† Graduate School of Environmental Engineering, The University of Kitakyushu
E-mail: †daijiro.murooka@is.env.kitakyu-u.ac.jp, ††{dongqing,takasima,nakatake}@kitakyu-u.ac.jp

Abstract  The post-silicon tuning introducing programmable delay elements (PDEs) to mitigate the manufacturing variability on the delay is promising. This work presents a novel PDE based on the channel-length decomposition, and reveals that it contributes to the low-power and low-variability comparing with a conventional inverter-chain-type. In addition, in a model of a clock tree along with the PDEs, we propose a mechanism for post-silicon tuning of a skew between a pair of flip-flops by a multilevel DLL employing our PDEs of multiple delay steps. In experiments, our proposed mechanism provides a high tunability even under the variability of PDE itself. Furthermore, we demonstrate our mechanism can be extended to the clock skew distribution to reduce the peak current.

Key words  manufacturing variability, post-silicon tuning, programmable delay element, clock skew, delay-locked-loop circuit

1. Introduction

A timing error caused by the manufacturing variability is becoming critical for digital circuits. Hence, it is used to resolve the timing error by the post-silicon tuning technology[1][7][8]. A programmable delay element (PDE) is one of fundamental technologies used in the post-silicon tuning, and it can generate various delays by controlling the number of active delay elements from the outside of the chip. If we place PDEs on a path from a clock source driver to each flip-flop (FF), we can change the arrival time at the FF after manufacturing. It is well known that the linearity and the relative accuracy are significant characteristics for PDEs.

For the post-silicon delay tuning, however, a PDE should be paid attention to its variability and should be quite a low-power when it is used in a clock network. In addition, because it is an important issue how to capture skews to be tuned after manufacturing, we need to address a mechanism for tuning the skews by PDEs including how to set the delay to PDE.

In this paper, we present a low-power and low-variability...
programmable delay elements as well as a mechanism for clock skew tuning consistent with our PDEs. The contributions of this paper are summarized as follows;

(A) - We propose a novel PDE employing channel length decomposed transistors, like a clocked inverter, which is composed of a series of pMOS subtransistors and that of nMOS. Compared to a conventional inverter-chain-type PDE which iterates charging and discharging, our PDE always only one charging or discharging when generating any delay, so that it contributes to the low-power. Besides, because a series of channel-length decomposed subtransistors is regarded as a large transistor, the size-dependent variability is suppressed to be low [2]. We verify the variability by the Monte Carlo simulation. Plus, we fabricate the PDEs with various sizes and demonstrate the measurement results related to the linearity and the relative accuracy.

(B) - We provide a post-silicon skew tuning mechanism introducing a multilevel delay locked loop (DLL) to synchronize the timing at a clock source and a specified FF. PDEs are placed from the source to the FF of a clock tree. The key point is that our multilevel DLL configures a loop by employing the series of PDEs placed on the clock tree.

(C) - We illustrate the tunability by our multilevel DLL even when the PDE has a large variability. To our best knowledge, this is the first work to consider the PDE itself variability. In experiments, the Monte Carlo simulations on 0.6um/3.3V and 0.18um/3.3V processes show that our tuning mechanism can reduce the clock skews even though the PDE in the clock tree is affected by the process variability. Furthermore, we demonstrate the applicability of our proposed mechanism to reduce peak current by the skew distribution.

The rest of this paper is organized as follows; Section 2. describes the variability and power consumption of the proposed PDE and the simulation and the measurement results. Section 3. is devoted into describing a mechanism to employ the multilevel DLL for reducing the clock skew and demonstrating the simulation results under the PDE variability. Section 4. concludes this work.

2. Programmable delay element

A timing design is essential for designs of digital circuits. A programmable delay element (PDE) is a technology to change the intrinsic delay after manufacturing, and it is used for the post-silicon tuning to resolve a timing issue due to the manufacturing variability.

Several types of PDEs have been proposed so far. The most popular type is an inverter-chain-type as shown in Fig. 1(a). It changes the number of buffers (B1-B4) from the input to the output by selecting one of switches S1 to S4.

It has a good linearity of the delay and a useful model for designs, but suffers from the power saving and the suppression of the variability. Another PDE is a capacitor-type as shown in Fig. 1(b). It changes the amount of the load capacitance by controlling the signal D1 to D4. This type can save the power consumption, but needs the additional manufacturing process for capacitor implementation. A MOS transistor can be substituted for a capacitor, but it causes a large variability.

This paper proposes a simple but new PDE with better characteristics with respect to the power and the variability.

2.1 Channel-length decomposition PDE

A schematic of a PDE proposed in this paper is illustrated in Fig. 2. It is constituted, like a clocked inverter, such that pMOS transistors of the same size are connected in serial, as well as nMOS transistors. By the control signals (D1-D8), the number of active transistors can be changed, and it results in changing of the intrinsic delay. Focusing on the series of pMOS transistors, we can regard this series such that a large transistor is decomposed into a set of subtransistors along the channel-length. The nMOS transistors can be also regarded analogously. This is why we call our PDE channel-length decomposition type.

Besides, since each subtransistor has the same channel-length, the delay is proportional to the number of active subtransistors passed by the current when charging or discharging. Note that, in a DC model, a subtransistor is regarded as a resistor, and the delay is proportional to the RC value in the model when considering the first moment of the circuit, like the Elmore delay model.

2.2 Variability analysis

We analize the variability for our proposed PDE and an
inverter-chain-type as the existing PDE by the Monte Carlo simulation. In the simulation, both PDEs have the same size for each transistor, we observe the difference with respect to the deviation of the variability distribution. The transistor sizes are shown in Table 1. In addition, the variability to the control switch blocks for both PDEs are not given.

The simulation is set by $5\sigma$ normal distribution where $(\mu = 12\mu m, \sigma = 3\mu)$ of pMOS and $(\mu = 6\mu m, \sigma = 3\mu)$ of nMOS with respect to the channel width of each transistor. $\mu$ and $\sigma$ are the average and the standard deviation, respectively. The resultant distribution of the delay of each PDE is shown in Fig. 3. It can be observed that $\sigma$ of the distribution for our proposed PDE is smaller than that for the inverter-chain-type. This is because the transistor numbers of the the inverter-chain-type PDE and our proposed PDE corresponding to the same delay are 64 and 34, respectively.

![Diagram](image)

**Table 1** Channel-size of each transistor of PDEs

<table>
<thead>
<tr>
<th></th>
<th>pMOS</th>
<th>nMOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>our PDE</td>
<td>$L=0.6[\text{um}], W=12[\text{um}]$</td>
<td>$L=0.6[\text{um}], W=6[\text{um}]$</td>
</tr>
<tr>
<td>Inverter-chain-type</td>
<td>$L=0.6[\text{um}], W=12[\text{um}]$</td>
<td>$L=0.6[\text{um}], W=6[\text{um}]$</td>
</tr>
</tbody>
</table>

### 2.3 Power consumption analysis

The inverter-chain-type, current-stave-type, DCDE-type, thyristor-type, and capacitor-type are known as the conventional PDEs [3] [4] [5] [6]. Each circuit mechanism can achieve a good linearity as shown in Fig. 4, although it is hard for each to generate a specified absolute delay value.

The inverter-chain is the most frequent type, and a charging and a discharging are repeated at every time when the signal passes through each inverter of multi-stage. This means the power consumption increases according to the amount of delay to be generated. On the other hand, in our PDE, a charging and a discharging occur only once even when increasing the delay. As as result, our PDE achieves the further low power consumption as shown in Fig. 5.

![Diagram](image)

**Figure 3** Variability of PDE and Inverter-chain

**Figure 4** The linearity of the delay values generated by the inverter-chain-type and our PDE

### 2.4 Measurement result

We demonstrate the measurement results of various size PDEs in Fig. 6. Each PDE provides a good linearity along the code. These results imply the skews should be tune in terms of not absolute delay values but relative values.
3. Post-silicon skew tuning

In this section, we present a mechanism how to capture a clock skew and how to tune the skew after manufacturing by using our proposed PDE of channel-length decomposed type.

3.1 Multilevel delay locked loop

A key component of our mechanism for skew tuning is DLL (Delay Locked Loop). Also, we propose a DLL circuit incorporating channel-length decomposed PDEs, which is called multilevel DLL. A diagram of the proposed multilevel DLL is shown Fig. 7. The DLL detects the delay difference between two input signals by PD (Phase Detector), and control the number of PDEs and the number of active channels of each PDE according to the delay difference. The output the PDE is fed back as one of the input signals, and the DLL gradually decreases the delay difference. Because a total delay of PDEs should cover the maximum delay difference, we introduce a series of eight PDEs for delay tuning. We call it 8-level DLL shown Fig. 7.

3.2 Mechanism for skew tuning

In post-silicon tuning, we place the PDE on a clock tree, and align skews from a clock source to different FFs. Since it is hard to measure the arrival time from the clock source to the FF after chip fabrication, we need a circuit mechanism to capture the skew.

In a conventional tuning mechanism shown in [7] [8], a connection with a phase detector (PD) to synchronize the timing at FFs placed adjacently by using adjustable delay buffer (ADB) on the clock tree. An example of the conventional tuning mechanism is shown in Fig. 8(a). All FFs are connected by a chain of input connections to PDs. This idea contributes to reduce the number of ADBs and PDs as well as the wire-length of the connection for the PD. However, it could occur that a slight skew between adjacent FFs after tuning is transmitted and accumulated, and it results in a large skew between FFs apart from each other.

In this paper, we introduce a mechanism employing the multilevel DLL with our PDE. In our mechanism, a series of PDEs is shared by the clock tree and the DLL. An example of our mechanism is shown in Fig. 8(b). The PDEs are placed on the clock tree, but they are low-power as described above. Besides, unlike [7] [8], we make a connection with a PD of a multilevel DLL between each FF and the junction of the tree as well as between each junction and the clock source. This means each FF is synchronized with the timing at the clock source, and it does not occur a large skew even for FFs apart from each other.
3.3 Simulation for skew tuning

In this section, we illustrate the tunability by our multilevel DLL under the variability of the PDE itself. In our simulation, we compare two cases, Case1 and Case2 as shown in Fig. 9 and Fig. 10, respectively.

Case1: This case assumes a clock tree has no tuning mechanism. In a path from the clock source to the FF (FF-A), buffers are placed Buf-A1 and Buf-A2. As well, Buf-B1 and Buf-B2 are placed on a path to FF-B.

Case2: This case is for the tunability by our mechanism. A signal from the clock source is input to the multilevel DLL, and the output of the DLL is connected to FF-A via the PDE-A. In addition, a signal from a point just after PDE-A is also input to the DLL for the synchronization. An analogous connection is configured for PDE-B and FF-B. The mechanism provides a real-time tunability for clock skews.

As a preliminary, we give the size of each buffer is shown in Table 2 and observe the waveforms at the input of FF-A and FF-B in both cases. As shown in Table 3, the skew between FF-A and FF-B in Case2 is much smaller than that in Case1.

3.4 Simulation under PDE variability

One of the most important motivations of the post-silicon tuning is to resolve the variability issue. However, most of studies related to the post-silicon tuning do not pay attention to the PDE itself variability. This is the first work to consider the PDE itself variability.

As for two cases shown in Fig. 9 and Fig. 10, we apply the Monte Carlo simulation under the assumption that Buf-A1, Buf-A2, Buf-B1 and Buf-B2 have the variation, and observe the skew distribution between FF-A and FF-B.

We use the channel-width as the variability parameters for the buffers(Case1) and PDEs(Case2) on the clock tree. The numbers of transistors given the variability is 112 and 128 for Case1 and Case2, respectively. The transistor size of the buffers (Case1) and PDEs (Case2) are the same as pMOS is W/L = 12u/1.5u, and nMOS is W/L = 6u/2u. Besides, for both cases, the fluctuation of channel-width is described as PMOS(W) = AGAUSS(12u 4u 5) and NMOS(W) = AGAUSS(6u 2u 5)) in the HSPICE simulation. Changing the parameters, 1000 times simulations are executed.

The resultant distributions of the skews by the simulation on 0.6um/3.3V and on 0.18um/3.3V processes are shown in Fig. 12 and Fig. 13, respectively. Our mechanism can obviously reduce the deviation of the skew for both processes, but the effect for the tunability (i.e. the difference between Case1 and Case2) on the finer process (0.18um) is much larger that on the 0.6um process. On the other hand, even for every case, we cannot achieve a complete zero skew. This means a conventional tuning mechanism in [7] [8] has a risk of increasing a skew between FFs apart from each other.

3.5 Skew distribution mechanism

It is known that a clock speed can be accelerated by con-
trolling of skews of FFs [9][10]. Furthermore, an appropriate skew distribution to FFs can contribute to the peak current reduction [11]. As described in this paper, however, the variability of PDE itself makes it difficult to distribute appropriate skews to FFs. Then, we introduce PDEs nearby the clock source to generate the specified skews. At the point $x_0$ in Fig. 14, four PDEs, $d_1$, $d_2$, $d_3$, and $d_4$ are added as shown in the figure. Note that a phase difference of $x_0$ and $x_1$ is a skew. It means that the delays of $d_1$, $d_2$, $d_3$, and $d_4$ correspond to the skews of FF groups, $G_1$, $G_2$, $G_3$, and $G_4$, respectively. These PDEs also have the variability, but they are placed closely each other as well as a large size to reduce the variability.

4. Conclusion

We propose a low-power and low-variability programmable delay elements and clarify the performance by simulation and measurement. Introducing the PDE, we present a post-silicon tuning mechanism along with the multilevel DLL. The DLL configures a loop by employing the series of PDEs placed on the clock tree, and attains the low variability even though the PDE itself in the clock tree is affected by the variability. In the comparison with a conventional clock skew tuning mechanism, the simulation result demonstrates the higher tunability of our mechanism.

In addition, we demonstrate our proposed mechanism can apply to various clock systems by easy extensions.

文 献


©2015 Information Processing Society of Japan