Real Circuit Delay Measurement Method by Variable Frequency Operation with On-Chip Fine Resolution Oscillator

KOTARO SHIMAMURA^{1,a)} NAOHIRO IKEDA¹

Received: June 4, 2019, Revised: August 27, 2019,
Accepted: October 27, 2019

Abstract: With the progress of semiconductor process miniaturization, delay degradation by aging increases and threatens the reliability of fabricated chips. The amount of delay degradation is known to be circuit and workload dependent, but previous evaluations are based on simulations, and delay degradation measurement of real circuit under realistic workload has not been reported yet. This paper proposes real circuit delay measurement method, which achieves enough accuracy to measure circuit and workload dependent delay degradation. In the proposed method, on-chip oscillator supplies fine resolution variable frequency clock to internal circuit. Internal circuit execute test pattern to activate critical paths at various frequency and determine the maximum frequency at which correct results can be obtained. The maximum frequency corresponds to the delay of the critical paths activated by the test pattern. Clock multiplication improves delay resolution, and repetitive measurement reduces measurement error caused by time dependent random delay variation. The proposed method has been implemented on a 65 nm low power process test chip. Variable frequency oscillator utilizes only standard cells and is designed with automatic layout flow without any timing tuning. The area overhead of the proposed method is 0.09% of the total random logic. The evaluation result show that 0.18% average measurement accuracy has been achieved.

Keywords: delay measurement, variable frequency, high resolution, delay degradation, workload dependent

1. Introduction

With the progress of semiconductor process miniaturization, delay degradation by aging such as hot carrier injection (HCI), negative bias temperature instability (NBTI), and positive bias temperature instability (PBTI) increases. To maintain the reliability of fabricated chips, the amount of delay degradation should be estimated and enough margin to avoid malfunction by delay degradation should be maintained. Delay degradation by aging is known to be circuit and workload dependent [1], [2], [3], but previous evaluations are based on simulations, which use transistor degradation model derived from simple circuit measurements. To the best of our knowledge, delay degradation measurement of real circuit under realistic workload has not been reported yet.

This paper aims to realize real circuit delay measurement with enough accuracy for circuit and workload dependent delay degradation estimation. The amount of delay degradation by aging is reported to be typically several percent range [1], [2], [3]. So less than 1.0% delay measurement error is necessary to measure delay degradation.

There are many related works about critical path delay measurement. Critical path replica [4], [5] is the circuit that shows similar temperature and voltage dependency as the critical paths, and is used to estimate the delay of critical paths. But the

on-probability and the switching frequency of each transistor, which affect the amount of delay degradation, are different from critical path, so the information about critical path delay degradation cannot be obtained from critical path replica.

Timing error prediction sensors [6], [7], [8] monitor the input of the flip-flops (FFs) at the endpoint of critical paths, but they only show whether delay has exceeded pre-defined threshold or not, and fine resolution delay measurement is not possible by these sensors.

Time-to-digital converters (TDC) has the potential to measure circuit delay with high accuracy [14], and is used to measure the delay of critical path replica in Ref. [5]. But applying TDC to real circuit delay measurement with thousands of critical paths leads to large area overhead [15].

DART [9] uses built-in self-test (BIST) circuit for manufacturing test to measure critical path delay. The interval of input injection and output capture timing are shortened gradually and the minimum interval at which output is correct is determined. The minimum interval corresponds to the critical path delay. It realizes real circuit delay measurement with small area overhead, but it can only be applied to random logic. There are many cases where memory access paths are the most critical path, so it is preferable that circuit delay including memory can be measured. DART also has the problem of compatibility to off-the-shelf design for test (DFT) tools.

In order to measure real circuit delay including memory with

¹ Hitachi, Ltd., Chiyoda, Tokyo 100–8280, Japan

^{a)} kotaro.shimamura.tx@hitachi.com

small area overhead, on-chip oscillator supplies variable frequency to internal circuit in the proposed method. There are many related works about variable frequency operation [10], [11], but they are targeted to scale performance in accordance with power budget, so frequency resolution is insufficient for delay degradation measurement. Also, frequency inaccuracy caused by process, voltage and temperature (PVT) variation should be considered. To resolve these issues, the proposed method utilizes fine resolution variable frequency oscillator (VFO) combined with method to improve accuracy such as clock multiplication, VFO frequency calibration and repetitive measurement. With the proposed method, sufficient accuracy to measure delay degradation can be achieved without special circuit level tuning such as utilizing customized cells and manual placement optimization.

Section 2 describes the concept of delay measurement method. Section 3 describes the method to improve accuracy. Section 4 describes the test chip implementation. Section 5 describes test chip measurement results. Section 6 compares proposed method with TDC. Section 7 is the concluding remarks.

This paper is an extended version of Ref. [12]. Method to improve accuracy by repetitive measurement and its evaluation results is the main addition to Ref. [12].

2. Delay Measurement Method

2.1 Variable Frequency Operation

Figure 1 shows the concept of real circuit variable frequency operation. VFO contains register to store frequency setting and VFO output frequency is controlled by the setting. It should have resolution under 1% and cover the frequency range wide enough to measure the circuit delay of the target critical path. PVT variation is calibrated outside VFO, so VFO need not have calibration function. An example of VFO implementation is described later.

The multiplexer (MUX) select input clock or VFO output and supply to the clock generator. Input clock is fixed frequency and low enough to ensure that clock generator and internal circuit can operate correctly even in the PVT worst conditions. The accuracy of input clock frequency directly affects the delay measurement accuracy, so high accuracy clock source such as crystal oscillator is the preferable source of input clock.

Clock generator generates clock to internal circuit. Modern microprocessor often uses different frequency for CPU and peripheral circuit [5], [10], and clock generator indicate the circuit to generate such clocks with different frequency. As mentioned later, higher clock multiplication factor at clock generator leads to higher delay measurement accuracy.

Figure 2 shows the concept of real circuit delay measurement. As VFO output clock frequency increases, clock period decreases. When VFO output is selected by multiplexer in Fig. 1,

VFO: Variable frequency Oscillator
MUX: multiplexer

Fig. 1 Variable frequency operation of real circuit.

internal circuit operates correctly at lower frequency and incorrectly at higher frequency. The maximum frequency at which internal circuit operates correctly correspond to the circuit delay of the most critical path activated by the operation. In order to measure the circuit delay of desired path, the path has to be activated by the operation. In addition to this, the path has to be the most critical path during the operation. This can be the limitation of the proposed method, and how to cope with this will be discussed later.

2.2 Test Pattern Execution

To measure circuit delay, the internal circuit has to operate at various VFO frequency. Example of the flow to measure circuit delay is shown in **Fig. 3**. First, VFO frequency setting S1 is set to VFO register and VFO output is selected for clock generator input. Under this condition, internal circuit executes test pattern and save the results to some memory. Test pattern contains several test cases, each of which activates one of the critical paths to be measured. Next, input clock is selected for clock generator input and internal circuit check the result of each test case stored in the memory. Input clock is low enough to ensure that internal circuit can operate correctly, so the validity of result check is assured. These two steps are repeated through VFO frequency setting S2, ..., Sn. VFO frequency settings (S1, S2, ..., Sn) are selected to cover the frequency range between a frequency at which internal circuit executes every test case correctly and a frequency at which internal circuit does not execute any test case correctly. After that, the maximum frequency at which each test case has been executed correctly is determined.

An example of the test pattern is shown in **Table 1**. The contents of the test pattern depend on the function of the internal

Fig. 2 Conceptual diagram of real circuit delay measurement.

S1, S2, ..., Sn: VFO Frequency Setting

Fig. 3 Clock selection method.

Table 1 Example of test pattern.

Category	Test Case
Integer	Multiplication
	Shift
Floating Point	Addition
	Multiplication
Memory	SRAM mat variation
	Address variation (inside mat)

circuit, and Table 1 is an example in which the internal circuit is a microprocessor. Microprocessors usual have several functional units such as integer multiplier, integer shifter, floating-point adder, and floating-point multiplier. In many cases, these functional units are mutually independent. For example, an instruction which activate integer multiplier does not activate integer shifter, floating-point adder and floating-point multiplier. So the circuit delay of the critical path in integer multiplier can be measured by a test case which includes instructions activating only integer multiplier. There are usually several critical paths in an integer multiplier and each path is activated by some input data. Ideally, test case for each path from several critical paths can be made with instructions activating integer multiplier with different input data. There are many cases in which memory access paths are the most critical paths in a microprocessor. If there are several memory mats in the microprocessor, circuit delay including access to each mat can be measured by a test case accessing the mat. Access time of a single memory mat may vary depending on the address inside the mat. This access time variation can be measured by test cases accessing different addresses in the mat.

2.3 Mitigation of Hang-up

As described above, the most critical path activated by each test case determines the maximum frequency at which the test case can be executed correctly. The delay of the paths other than the most critical path cannot be measured. In addition to this, if the internal circuit is a microprocessor, hang-up may occur during the execution of each test case if there is a most critical path in the circuit which carry out basic instruction execution function such as instruction fetch and instruction decode. Even though hang-up does not occur, functional units may not be activated as expected by the test case. Similar case may occur when the internal circuit is not a microprocessor. If there is a most critical path in the circuit which carry out basic control function, hang-up may occur, or functional units other than basic control function may not be activated as expected by the test case.

There are several methods to mitigate this problem. In the following description, the basic control circuit means circuit which carry out basic instruction execution function in microprocessor case, and basic control function in non-microprocessor case.

(a) Watchdog function to restore operation from hang-up

If hang-up occurs, the flow in Fig. 3 cannot be completed. To cope with this problem, there is a method to add a measure which detects hang-up and restores state of the internal circuit to normal state. For example, watchdog timer which reset the internal circuit when timer has not been cleared during predetermined period, can be used for this purpose. The multiplexer in Fig. 1

Fig. 4 Delay degradation measurement under realistic workload.

should select input clock after hang-up to assure the correct operation of the internal circuit. This ensures that the flow in Fig. 3 can be completed. Although the circuit delay of the paths faster than the critical paths in the basic control circuit cannot be measured, test cases which activate paths which are slower than the basic control circuit can be executed. So the circuit delay of the circuit slower than the basic control circuit can be measured.

(b) Additional timing margin for the basic control circuit

The circuit delay of the critical paths in the basic control circuit decides the measurable delay range. If the basic control circuit is designed to have additional timing margin than other circuit, it leads to wider measurable delay range. This may cause area overhead in the basic control circuit, so the trade-off between area overhead and width of measurable delay range should be considered.

(c) Combination with timing error prediction sensor

Timing error prediction sensors [6], [7], [8] can predict timing error at the lower frequency than the maximum frequency at which the targeted circuit can operate correctly. If these sensors are used for the circuit other than the basic control circuit, the malfunction of sensor-equipped circuit can be predicted at lower frequency than the frequency at which real malfunction occurs. This leads to wider measurable delay range. This causes area overhead for the sensors, and accuracy degradation because the sensors add delay to the critical paths. But this method has the merit to ensure that the paths to be measured can be certainly measured by adding the sensors to the paths.

2.4 Application to Delay Degradation Measurement

Figure 4 shows an example method to measure delay degradation under realistic workload. As shown in Fig. 4(a), frequency measurement as shown in Fig. 3 is executed periodically, and realistic workload is executed in the remaining time. This ensures that the internal circuit degrades just like when the circuit degrades in a real application. The maximum frequency at which each test case can be executed correctly is converted to the circuit delay of the most critical path activated by the test case. As shown

in Fig. 4(b), the relationship between hours of operation and circuit delay shows the delay degradation of the critical path. By calculating delay degradation for each test case, path dependent delay degradation can be measured.

Reference [3] reports supply voltage dependence of the amount of delay degradation. Although the amount of threshold voltage increase by aging is larger at higher supply voltage, circuit delay is more sensitive to threshold voltage at lower voltage. This results in larger delay degradation at lower voltage at the condition of Ref. [3] analysis. Similar analysis can be done by applying the method shown in Fig. 4 to several groups of chips supplied different voltages. For this purpose, supply voltages need not be changed dynamically. If the system under measurement support dynamic voltage change, valuable data may be obtained by executing workload and frequency measurement at different voltages. For example, executing workload at higher voltage and executing frequency measurement at lower voltage may lead to larger delay degradation than executing both workload and frequency measurement at the same voltage. Such measurement may provide information about appropriate delay margin at lower voltage. For those purposes, all components in a chip including VFO, clock generator and internal circuit operate at the same voltage. The proposed method does not pose any requirements on power circuit design.

In the measurement of Fig. 4(a), execution time of frequency measurement should be much less than the execution time of workload so that frequency measurement has small impact on total delay degradation. **Figure 5** shows an example method to estimate the impact of frequency measurement to total delay degradation. Frequency measurement is executed three times: once before workload and twice after workload. Delay degradation by NBTI and PBTI is known to recover during transistor is off: i.e., delay may decrease during transistor is off. This recovery effect should be considered when executing measurement in Fig. 5. In the following explanation, critical path delays of test case i at measurement a , b , and c are denoted as $d(i, a)$, $d(i, b)$, and $d(i, c)$ respectively. The difference between $d(i, a)$ and $d(i, b)$ is the sum of the result of delay degradation or recovery by executing workload and frequency measurement. If $d(i, b)$ is larger, it is the result of delay degradation. If $d(i, b)$ is smaller, it is the result of delay recovery. The difference between $d(i, b)$ and $d(i, c)$ is the result of delay degradation or recovery by executing frequency measurement. If Eq. (1) is satisfied for all test case i , frequency measurement interval is long enough to assure that the impact of frequency measurement on total delay degradation is small enough.

$$|d(i, a) - d(i, b)| \gg |d(i, b) - d(i, c)| \quad (1)$$

If Eq. (1) is not satisfied for some test cases, frequency measurement interval may have to be lengthened. But if the difference

Fig. 5 Measurement to decide frequency measurement interval.

between $d(i, a)$ and $d(i, b)$ does not increase even though frequency measurement interval is lengthened, delay degradation of the critical path activated by test case i will probably small and measurement result of test case i can be ignored.

3. Method to Improve Accuracy

3.1 Clock Multiplication

Modern high-performance microprocessors can operate at more than 1 GHz clock frequency. Creating fine resolution variable frequency clock over 1 GHz directly is not easy. For example, creating 1% resolution variable frequency clock around 1 GHz requires 10 ps delay resolution, which will not be possible without timing tuning at layout phase. On the other hand, creating 1% resolution variable frequency clock around 100 MHz requires 100 ps delay resolution, which may be possible without timing tuning at layout phase. If 100 MHz clock with 1% resolution is multiplied by 10, 1 GHz clock with 1% resolution can be indirectly created. So clock multiplication in the clock generator in Fig. 1 leads to higher resolution than directly generating clock at the same frequency.

3.2 VFO Frequency Calibration

CMOS circuit is vulnerable to PVT variation, and designing accurate oscillator is not easy. Instead of designing accurate oscillator, proposed method utilizes the method to calibrate VFO frequency. As shown in **Fig. 6**, two counters are used for VFO frequency calibration. Counter 1 operate with internal clock (CKS) generated by clock generator and make Enable signal to Counter 2. Enable signal becomes high during CS cycles at CKS. Counter 2 operate with VFO output clock (CKV) and count up during Enable signal is high. If the value of Counter 2 becomes CV after Enable signal becomes low, the ratio of CV vs CS is same as the ratio of CKV frequency and CKS frequency.

Table 2 shows the usage of VFO frequency calibration counter. When the multiplexer in Fig. 6(a) selects input clock, VFO output frequency (FV) can be calculated from CV, CS and CKS frequency (FS). In this mode, FS can be calculated by multiplying

(a) Block Diagram

(b) Count Timing

Fig. 6 VFO frequency calibration counter.

Table 2 Calibration counter usage.

Selected Clock	Usage	Note
Input Clock	VFO Frequency Calibration	$FV = \frac{CV \times FS}{CS}$
VFO Output	Clock Generator Validity Check	$\frac{CS}{CV} = \text{const}$

FS: frequency of CKS FV: frequency of CKV

Fig. 7 Quantization error of VFO frequency counter.

input clock frequency by the multiplication factor of clock generator. Using high accuracy input clock leads to high accuracy VFO frequency calibration. On the other hand, when the multiplexer in Fig. 6 (a) selects VFO output, the ratio of CS and CV normally becomes constant, which is the multiplication factor of the clock generator. But when VFO output frequency becomes too high for clock generator to operate correctly, the ratio of CS and CV becomes different values than the multiplication factor. For example, PLL (Phase-Locked Loop) often used for clock generator has the maximum output frequency, which may lead to frequency saturation above some frequency value of VFO output. Checking the ratio of CS and CV enables to check the validity of clock generator output. Please note that when input clock is selected, correct operation of clock generator is assured, so the malfunction of clock generator does not occur during VFO frequency calibration.

VFO frequency calibration method shown in Fig. 6 and Table 2 has several sources of error such as CKI frequency inaccuracy, quantization error, systematic delay variation and random delay variation.

CKI frequency inaccuracy directly affects the length of time Enable signal in Fig. 6 (a) is high. If CKI frequency inaccuracy is $e(CKI)\%$, it leads to FVO frequency error of $e(CKI)\%$.

Figure 7 explains the quantization error of VFO frequency counter. In the following explanation, VFO frequency calculated by the equation in Table 2 is denoted as FV_{est} , and VFO frequency without quantization error is denoted as FV_{act-q} . In Fig. 7 (a), there are CKV rising edges just after Enable signal becomes high and just before Enable signal becomes low. In this case, FV_{act-q} is calculated by Eq. (2).

$$FV_{act-q} = \frac{(CV - 1) \times FS}{CS} = FV_{est} \times \left(1 - \frac{1}{CV}\right) \quad (2)$$

In Fig. 7 (b), there are CKV rising edges just before Enable signal becomes high and just after Enable signal becomes low. In this case, FV_{act-q} is calculated by Eq. (3).

$$FV_{act-q} = \frac{(CV + 1) \times FS}{CS} = FV_{est} \times \left(1 + \frac{1}{CV}\right) \quad (3)$$

FV_{act-q} is bounded by the minimum value of Eq. (2) and the maximum value of Eq. (3). So VFO frequency quantization error $e(qtz)$ is bounded as described in Eq. (4).

$$e(qtz) = 100 \times \frac{|FV_{est} - FV_{act-q}|}{FV_{est}} \leq 100 \times \frac{1}{CV} (\%) \quad (4)$$

Systematic delay variation is caused by the delay difference when Enable signal goes high and low. Both delays in Counter 1 and Counter 2 contributes to systematic delay variation. In the following explanation, the delay difference is denoted as $dd(sys)$, VFO frequency error caused by the delay difference is denoted as $e(sys)$, and VFO frequency without $e(sys)$ is denoted as FV_{act-s} . FV_{act-s} can be calculated by adding $dd(sys)$ to the length of time Enable signal is high as Eq. (5). Using (5), $e(sys)$ is calculated as Eq. (6).

$$FV_{act-s} = \frac{CV}{\frac{CS}{FS} + dd(sys)} = FV_{est} \times \frac{1}{1 + dd(sys) \times \frac{FS}{CS}} \quad (5)$$

$$e(sys) = 100 \times \frac{|FV_{est} - FV_{act-s}|}{FV_{est}} = 100 \times \frac{|dd(sys)| \times \frac{FS}{CS}}{1 + dd(sys) \times \frac{FS}{CS}} (\%) \quad (6)$$

Random delay variation can be mitigated by repetitive measurement method described in Section 3.3.

3.3 Repetitive Measurement

It is known that time dependent delay variation called random telegraph noise (RTN) [13] affect circuit delay in deep sub-micron CMOS process. This means that circuit delay under measurement as well as VFO output frequency varies from time to time. These variations lead to circuit delay measurement inaccuracy. VFO frequency calibration counter described in Section 3.2 only measures the mean value of VFO frequency and time dependent frequency variation cannot be calibrated. Also, when clock generator in Fig. 1 employs PLL to multiply frequency, PLL induced clock period jitter also leads to circuit delay measurement inaccuracy.

To cope with this problem, repetitive measurement method is employed. **Figure 8** shows the expected relationships between VFO frequency mean value measured by VFO frequency calibration counter at each VFO setting, and the probability of incorrect operation at that VFO setting when each test case is executed multiple times. Although the shape of the curve may different from Fig. 8, it is certain that as VFO frequency (mean value) increases, the probability of incorrect operation increases. In Fig. 8,

Fig. 8 Accuracy improvement by repetitive measurement.

$f_{\min}$ means the minimum VFO frequency (mean value) at which probability of incorrect operation is not zero. f_{mean} means the frequency at which probability of incorrect operation is 0.5. $f_{\max}$ means the maximum VFO frequency (mean value) at which probability of incorrect operation is less than one. At $f_{\min}$, the test case delivers correct result at most of the time and only in small number of cases the result is incorrect. This means that $f_{\min}$ is determined by a small number of events and is very vulnerable to statistic error. In other words, when measurement such as Fig. 8 is executed several times, the small number of cases in which the result is incorrect may vanishes by statistic error in some measurement, and measured $f_{\min}$ becomes adjacent VFO setting. Similarly, $f_{\max}$ is very vulnerable to statistic error. At f_{mean} both the number of cases with correct result and the number of cases with incorrect result are much larger, and the ratio of statistic error and the total number of cases is minimum. Actual measured data is discrete and will not probably contain the point whose probability of incorrect operation is equal to 0.5. Approximating the curve in Fig. 8 by, for example, straight line using measured data, and calculating the frequency at which the approximated curve crosses the 0.5 value of incorrect operation probability is the realistic method to determine f_{mean} .

4. Test Chip Implementation

4.1 Test Chip Overview

The proposed method has been implemented on a 65 nm low power process test chip. Table 3 shows the overview of the test chip. The chip contains 5.0M gates of random logic and 19.0M bits of memory. The overhead of the proposed method includes VFO and multiplexer in Fig. 1, VFO calibration counter in Fig. 6(a), and watchdog timer described in Section 2.3. Other methods described in Section 2.3 are not implemented, but the measurement results show that the most critical paths are memory access paths, and basic control circuit has substantial delay margin. Although memory access for instruction fetch and storing test case results can limit the measurable frequency range, selecting the fastest memory for this purpose broadens the range. Compared to the total random logic, the overhead of the proposed method is very small (0.09%). Clock generator generates three clocks and the multiplication factor of these clocks are 6, 3 and 1 respectively.

4.2 VFO Circuit

Figure 9 shows the circuit of VFO. VFO is a ring oscillator

Table 3 Test chip overview.

Random Logic (Overhead)	5.0M gates 4.5k gates (0.09%)
Memory	19.0M bits
Clock Generator Multiplication Factor	X6, X3, X1
Process	65nm Low Power

Fig. 9 VFO circuit.

designed with standard cells without any customized cell. The variable frequency is realized by two variable delay: one is coarse resolution variable delay and the other is fine resolution variable delay.

Coarse resolution variable delay uses inverter and NAND cells to realize delay variation. One of the control signals $S_0, S_1, S_2, \dots$, is set to high and the others are set to low. For example, when S_0 is high, input clock passes through two NAND cells. When S_1 is high, input clock passes through three NAND cells and one inverter cell. The delay difference between two cases is the delay of one NAND cell and one inverter cell. Clock signal passes through the coarse resolution variable delay twice in one clock cycle. The first time is rise transition and the second time is fall transition. So the clock period difference between two cases is the sum of rise and fall delay of one NAND cell and one inverter cell.

Fine resolution variable delay uses the delay difference between 2 input NAND cell and 3 input NAND cell. One of the control signal S_n and S_{n+1} ($n = 0, 1, \dots$) is high and the other is low. When S_n is high, input clock passes through 2 input NAND cell. When S_{n+1} is high, input clock passes through 3 input NAND cell. The delay difference between two cases is the delay difference of 2 input NAND cell and 3 input NAND cell. The clock period difference is the sum of rise and fall delay of 2 input NAND cell and 3 input NAND cell.

Fixed delay is composed of delay cells included in standard cell, which is usually used to fix hold violations.

Table 4 shows post layout delay calculation result at PVT worst condition (not the measured results described in Section 5).

Table 4 VFO specification.

Variable Delay (Coarse)	Delay Range	2.26ns - 205.14ns (512 stages)
	Resolution	0.40ns (average)
Variable Delay (Fine)	Delay Range	4.04ns - 4.77ns (8 stages)
	Resolution	0.10ns (average)
Fixed Delay	Delay	15.97ns

All delay value is the sum of rise and fall delay and shows the contribution to clock period. Average clock period resolution is 0.40 ns for coarse resolution variable delay and 0.10 ns for fine resolution variable delay. As no timing tuning has been done in layout design phase, the variation of the wire load of each cell in variable delay causes the variation of the delay of each cell. So the clock period resolution is not uniform.

As described in previous section, output of VFO is multiplied by at most 6 times. This means that clock period resolution is 0.067 ns for coarse and 0.017 ns for fine resolution. If 6-time multiplied clock is created directly by VFO, the finest resolution will be same as the designed VFO which is 0.10 ns. So multiplying clock by 6 times leads to 6 times finer frequency resolution.

5. Measurement Result

5.1 Accuracy Improvement by Repetitive Measurement

Figure 10 is an example of repetitive measurement result when one test case has been executed 100 times for each VFO setting. As previously described, the resolution of VFO frequency is not uniform. The VFO frequency (mean value) of seven plotted point in Fig. 10 is 65.13, 65.18, 65.29, 65.36, 65.45, 65.48 and 65.76 MHz. Minimum resolution is 0.05% between 65.45 and 65.48 MHz, and maximum resolution is 0.42% between 65.48 and 65.76 MHz. f_{mean} of the test case plotted in Fig. 10 resides in the region with rather fine resolution and repetitive measurement greatly improves accuracy. If f_{mean} is around 65.6 MHz, the incorrect operation ratio at 65.48 MHz may be 0.0 and the incorrect operation ratio at 65.76 MHz may be 1.0. The measurement error in such a case will be discussed in the next section.

As described in Section 3.3, f_{mean} of the test case plotted in Fig. 10 can be determined from the obtained data. The leftmost two points and rightmost three points seem to be out of the region approximated with straight line. Connecting the remaining two points with straight line and calculating the frequency at which the line crosses the 0.5 value of incorrect operation ratio leads to 65.31 MHz of f_{mean} value. As test case is executed 100 times for each VFO setting, static error (standard deviation) is 5 times. This mean 0.05 point in incorrect operation ratio which leads to 0.01 MHz (0.02%) frequency error.

Table 5 shows the estimated measurement error of non-repetitive measurement. In non-repetitive measurement, each test case is executed once for each setting. The frequency is determined by averaging the maximum frequency at which the result is correct (pass frequency in Table 5) and the minimum frequency at which the result is incorrect (fail frequency in Table 5). For example, in Fig. 10, the results are always correct at 65.13%, but the results are incorrect in 5 times out of 100 at 65.18%. This means that, at the probability of 0.05, the frequency is determined to

Fig. 10 Example of repetitive measurement result.

Table 5 Error of non-repetitive measurement.

Frequency (MHz)			Probability	Error (%)
Pass	Fail	Estimated		
65.13	65.18	65.15	0.05	-0.24
65.18	65.29	65.23	0.35	-0.12
65.29	65.36	65.33	0.34	0.02
65.36	65.45	65.41	0.19	0.15
65.45	65.48	65.47	0.05	0.24
65.48	65.76	65.62	0.02	0.47
Root Mean Square				0.14

be 65.15 MHz. If 65.31 MHz derived from repetitive measurement is assumed to be the correct value, the error in this case is -0.16 MHz (-0.24%). Table 5 lists the all possible cases, and the root mean square of measured frequency error is 0.14%. Repetitive measurement reduces error by 7 times (0.14% vs. 0.02 %).

5.2 Worst Case Error Estimation

As described in previous section, there may be cases in which incorrect operation ratio goes from 0.0 to 1.0 immediately due to the insufficient frequency resolution. In this case, the best way to determine the frequency is to average the frequency at which incorrect operation ratio is 0.0 (f_1) and the frequency at which incorrect operation ratio is 1.0 (f_2). Estimated value of f_{mean} ($f_{\text{mean-est}}$) is calculated by Eq. (7).

$$f_{\text{mean-est}} = \frac{f_1 + f_2}{2} \quad (7)$$

Actual value of f_{mean} ($f_{\text{mean-act}}$) can vary from f_1 to f_2 . So the f_{mean} estimation error ($e(f_{\text{mean}})$) is maximum when $f_{\text{mean-act}}$ is f_1 or f_2 . Upper bound of $e(f_{\text{mean}})$ is described as Eq. (8).

$$\begin{aligned}
 e(f_{\text{mean}}) &= 100 \times \frac{|f_{\text{mean-est}} - f_{\text{mean-act}}|}{f_{\text{mean-est}}} \\
 &\leq 100 \times \frac{f_2 - \frac{f_1 + f_2}{2}}{\frac{f_1 + f_2}{2}} \\
 &= 100 \times \frac{f_2 - f_1}{f_1 + f_2} \quad (\%) \quad (8)
 \end{aligned}$$

Figure 11 is the histogram of worst-case error over the VFO frequency range used for circuit delay measurement. VFO is designed with large margin in output frequency range and large portion of them is not used for real circuit measurement. The frequency range not used for real circuit measurement is excluded in calculating Fig. 11. The larger the frequency difference between adjacent VFO setting, the more the probability that measured

f_{mean} is reside between the frequency. So Fig. 11 is derived by calculating weighted average whose weight is the frequency difference between adjacent VFO setting. The mean value of worst-case error is 0.18% and maximum value is 0.46%.

Please note that repetitive measurement contributes to reduce the mean value of worst-case error. For example, f_{mean} of the test case plotted in Fig. 10 is between 65.29 MHz and 65.36 MHz. The worst-case error calculated by Eq. (1) is 0.05%. On the other hand, as shown in Table 5, error of non-repetitive measurement is 0.14%. Repetitive measurement reduces the error by about 3 times.

To confirm the worst-case error in Fig. 11, same chip has been measured 10 times. About 6,000 test cases are measured, but about 60% of them are not measurable because the path activated by the test cases is faster than memory access critical paths for instruction fetch or storing results. Total of 2,777 test cases can be measured. The histogram of peak to peak deviation and root mean square error is shown in Fig. 12. For more than half of the test cases peak to peak error is no more than 0.1%, which

Fig. 11 Distribution of worst-case error.

Fig. 12 Deviation of same chip measurement results.

Fig. 13 Inter-chip delay variation measurement.

is consistent with Fig. 11. On the other hand, there are several cases which deviation is more than worst-case error indicated by Fig. 11. The reason is not confirmed yet, but probably there are several critical paths which can be activated by the test case, and initial condition at the start of the test case affect which of the potential critical paths are actually affected during that execution. This means that different path is measured during the different measurement, which leads to different frequency. This is the limitation of the proposed method, but if we use the proposed method for delay degradation measurement of large number of paths, the result with large error can be discarded and there is no significant impact on the delay degradation measurement result.

5.3 Inter-Chip Delay Variation Measurement

For the demonstration of the effectiveness of the proposed method, the proposed method is applied to inter-chip delay variation measurement. Two sample chips are measured and the measured frequency is compared. Figure 13 shows the result. There are 2,661 test cases at which frequency can be measured on both chips. Delay deviation of minus value means the second chip is slower than first chip. On average, the second chip is about 2% slower than the first chip. On the other hand, delay of each test case varies from -5% to 1%. This means that local delay variation inside each chip is larger than the global delay variation between the two chips.

6. Comparison with TDC

The test chip evaluated in this paper gives priority to ease of implementation over accuracy. It has been designed without custom cell and manual placement optimization. To understand the effect of this design methodology, this chapter compares proposed method with TDCs, which is the most accurate on-chip delay measurement method. There are three causes of delay measurement result inaccuracy: resolution, calibration error and random delay variation (Table 6).

In the previous chapters, measurement error of the test chip is described as relative value to measured value, and the unit is

Table 6 Accuracy comparison of proposed method with TDC.

	Proposed	TDC				
		[5]	[16]	[17]	[19]	[14]
Process	65nm	32nm	700nm	130nm	180nm	22nm
Resolution	10.8ps (average) 27.6ps (maximum)	6ps	30ps	6ps (*1) 1ps (*2)	—	—
Calibration Error	0.2ps	—	—	—	4ps	0.58ps
Random Variation Considered?	yes	no	no	yes (*2)	yes	no
Customize Cell Used?	no	yes	yes	yes	yes	yes
Placement Optimized Manually?	no	yes	yes	yes	yes	yes

(*1) single measurement

(*2) repetitive measurement

%. To compare with TDCs, relative value should be converted to absolute value, whose unit is second. In the measurement described in chapter 5, VFO frequency from 60.0 MHz to 76.2 MHz is used. VFO output is multiplied by 6 times in clock generator, and the frequency of the clock supplied to critical paths is between 360 MHz and 457 MHz. This means that the delay of measured critical paths is between 2.19 ns and 2.78 ns. For simplicity, delay value of 3 ns is used to convert % to second. For example, relative error of 1% is converted 30 ps.

For the proposed method, resolution corresponds to error caused by VFO frequency resolution. As described in Section 5.2, measurement error caused by VFO resolution is 0.18% average and 0.46% maximum. As measurement error is about half of the frequency resolution, frequency resolution is 0.36% average and 0.92% maximum, which corresponds to 10.8 ps and 27.6 ps respectively. Average resolution calculated in Section 4.2 is 17 ps, but it is the value at PVT worst condition and measured result is finer. Calibration error corresponds to the VFO frequency calibration error described in Section 3.2. In the measurement described in Section 5, calibration errors of three causes are as described below, and the sum is less than 0.0073%, which corresponds to about 0.2 ps calibration error.

- $e(\text{CKI}) \leq 0.0035\%$
- $\text{CV} > 30000 \rightarrow e(\text{qtz}) < 0.0034\%$
- $|d(\text{sys})| < 1 \text{ ns}, \text{FS} = 32 \text{ MHz}, \text{CS} > 10000$
 $\rightarrow e(\text{sys}) < 0.00032\%$

As described in Sections 5.1 and 5.2, repetitive measurement method mitigates random delay variation and keeps measurement error equal or less than error caused by VFO resolution.

Typical TDC connects start signal to buffer chain and monitor the output of each buffer by FFs clocked by stop signal. The delay between start signal and stop signal is determined by the position of last buffer start signal reached. The delay of one buffer in the buffer chain determines the resolution. TDC described in Ref. [5] utilizes 2-stage interpolating delay line as buffer chain to achieve finer resolution than simple buffer chain. TDC of Ref. [16] utilizes two buffer chains: one for start signal and the other for stop signal. Two buffer chains are implemented with voltage controlled buffers and controlled so that the buffer chain for the start signal is slightly slower than the buffer chain for the stop signal. The difference of the delay of the start signal buffer and the stop signal buffer determines the resolution, which is much less than the delay of the buffer itself. Although the resolution of Ref. [16] is much coarser than Ref. [5], it is because [16] utilizes much older process than Ref. [5]. Implementing [16] with the same process as Ref. [5] may result in finer resolution than Ref. [5]. TDC of Ref. [17] utilizes multi-path gated ring oscillator as the buffer chain. Multi-path ring oscillator contributes to achieve finer resolution for single measurement. Gated ring oscillator accumulates errors of multiple measurements. Measuring accumulation result effectively improves resolution and also contributes to mitigate random delay variation.

TDC calibration method is divided into indirect method and direct method. Indirect calibration method utilizes two independent clock sources [18]. Time between the rising edge of the 1st clock and the rising edge of the 2nd clock varies evenly between 0 and

the period of the 2nd clock. Indirect calibration makes use of this fact and decide the delay of each buffer in the buffer chain by measuring the time between two clock edges many times. This method can calibrate relative value of the buffer delay very accurately but absolute delay value calibration accuracy is not discussed in Ref. [18]. Direct calibration method utilizes variable frequency clock source. The accuracy of this method is limited by the resolution of the period of the clock source. Calibration method described in Ref. [19] adds random timing jitter to the clock source output and measures the clock period many times. Adding random timing jitter causes variation in measurement results, which enables to effectively improve accuracy just like the repetitive measurement method of this paper. Reference [19] achieves 4 ps calibration accuracy with 20 ps resolution clock source. Calibration method described in Ref. [14] adds a delay time controllable (DTC) inverter to TDC. By gradually increasing or decreasing the delay of DTC inverter and deciding the point where measured delay value increases or decreases, calibration accuracy improves. Reference [14] achieves 0.58 ps calibration accuracy with 5.2 ps resolution clock source.

As shown in Table 6, VFO frequency resolution limits the accuracy of the proposed method. Manual placement optimization of VFO will decrease both average wire length and wire length variation in VFO. This will lead to the improvement of both average and maximum resolution. Utilizing customized circuit such as 2-stage interpolating delay line in Ref. [5], multi-path ring oscillator in Ref. [17] or DTC inverter in Ref. [14] may also lead to resolution improvement.

7. Concluding Remarks

This paper proposes real circuit delay measurement method, which achieves enough accuracy to measure circuit and workload dependent delay degradation. In the proposed method, on-chip oscillator supplies fine resolution variable frequency clock to internal circuit. Internal circuit execute test pattern to activate critical paths at various frequency and determine the maximum frequency at which correct results can be obtained. The maximum frequency corresponds to the delay of the critical paths activated by the test pattern. Clock multiplication improves delay resolution. VFO frequency calibration counter calibrates the VFO output frequency and achieves high accuracy which is not influenced by PVT variation. Repetitive measurement reduces measurement error caused by time dependent random delay variation. The proposed method has been implemented on a 65 nm low power process test chip. VFO utilizes only standard cells and is designed with automatic layout flow without any timing tuning. VFO achieves average clock period resolution of 0.10 ns, and in combination with clock multiplication by 6 times, average internal clock period resolution of 0.017 ns is achieved. The area overhead of the proposed method is 0.09% of the total random logic. The repetitive measurement method is shown to achieve 7 times accuracy improvement (0.14% to 0.02%) for sample test case. Worst-case measurement error is analyzed to be 0.18% average and 0.46% maximum.

Acknowledgments The authors would like to thank Seiji Kajihara, Yasuo Sato, Yousuke Miyake, Michiko Inoue, Yukiya

Miura, Satoshi Ohtake, Kazumi Hatayama, and Tomokazu Yoneda for the discussion about real chip delay measurement; Tetsuya Shimomura, Takeshi Takehara, Shinnosuke Sato, and Hiroki Matsumoto for the support to the test chip implementation and evaluation.

References

- [1] Wang, W. et al.: The Impact of NBTI Effect on Combinational Circuit: Modeling, Simulation, and Analysis, *IEEE Trans. Very Large Scale Integration (VLSI) Systems*, Vol.18, No.2, pp.173–183 (Feb. 2010).
- [2] Keukner, H. et al.: NBTI Aging on 32-Bit Adders in the Downscaling Planar FET Technology Nodes, *2014 17th Euromicro Conference on Digital System Design*, pp.98–107 (Aug. 2014).
- [3] Mintarno, E. et al.: Workload Dependent NBTI and PBTI Analysis for a sub-45nm Commercial Microprocessor, *IEEE International Reliability Physics Symposium (IRPS)*, pp.3A.1.1–3A.1.6 (Apr. 2013).
- [4] Tschanz, J. et al.: Tunable Replica Circuits and Adaptive Voltage-Frequency Techniques for Dynamic Voltage, Temperature, and Aging Variation Tolerance, *Symposium on VLSI Circuits*, pp.112–113 (June 2009).
- [5] Drake, A.J. et al.: Single-cycle, pulse-shaped critical path monitor in the POWER7+ microprocessor, *International Symposium on Low Power Electronics and Design (ISLPED)*, pp.193–198 (Sep. 2013).
- [6] Sato, T. and Kunitake, Y.: A Simple Flip-Flop Circuit for Typical-Case Designs for DFM, *8th International Symposium on Quality Electronic Design (ISQED)*, pp.539–544 (Mar. 2007).
- [7] Wirnshofer, M., Heis, L., Georgakos, G. and Schmitt-Landsiedel, D.: An Energy-Efficient Supply Voltage Scheme using In-Situ Pre-Error Detection for on-the-fly Voltage Adaptation to PVT Variations, *International Symposium on Integrated Circuits*, pp.94–97 (Dec. 2011).
- [8] Amouri, A. and Tahoori, M.: A Low-Cost Sensor for Aging and Late Transitions Detection in Modern FPGAs, *21st International Conference on Field Programmable Logic and Applications*, pp.329–335 (Sep. 2011).
- [9] Sato, Y. et al.: DART: Dependable VLSI Test Architecture and Its Implementation, *IEEE International Test Conference*, pp.1–10 (Nov. 2012).
- [10] Fischer, T. et al.: A 90-nm Variable Frequency Clock System for a Power-Managed Itanium Architecture Processor, *IEEE Journal of Solid-State Circuits*, Vol.41, No.1, pp.218–228 (Jan. 2006).
- [11] Bull, D. et al.: A Power-Efficient 32 bit ARM Processor Using Timing-Error Detection and Correction for Transient-Error Tolerance and Adaptation to PVT Variation, *IEEE Journal of Solid-State Circuits*, Vol.46, No.1, pp.18–31 (Jan. 2011).
- [12] Shimamura, K. and Ikeda, N.: Real Circuit Delay Measurement Method by Variable Frequency Operation with On-Chip Fine Resolution Oscillator, *Design Automation Symposium 2019*, pp.3–8 (Aug. 2019) (in Japanese).
- [13] Ito, K. et al.: Modeling of Random Telegraph Noise under Circuit Operation — Simulation and Measurement of RTN-induced delay fluctuation, *12th International Symposium on Quality Electronic Design*, pp.1–6 (Mar. 2011).
- [14] Cui, R. and Namba, K.: A Calibration Technique for DVMC with Delay Time Controllable Inverter, *IPJS Trans. System LSI Design Methodology*, Vol.9, pp.30–36 (Feb. 2016).
- [15] Katoh, K. et al.: A Low Area On-chip Delay Measurement System Using Embedded Delay Measurement Circuit, *19th Asian Test Symposium*, pp.343–348 (Dec. 2010).
- [16] Dudek, P. et al.: A High-Resolution CMOS Time-to-Digital Converter Utilizing a Vernier Delay Line, *IEEE Trans. Solid-State Circuits*, Vol.35, No.2, pp.240–247 (Feb. 2000).
- [17] Straayer, M.Z. and Perrott, M.H.: A Multi-Path Gated Ring Oscillator TDC With First-Order Noise Shaping, *IEEE Trans. Solid-State Circuits*, Vol.44, No.4, pp.1089–1098 (Apr. 2009).
- [18] Katoh, K. et al.: An Analysis of Stochastic Self-Calibration of TDC Using Two Ring Oscillators, *22nd Asian Test Symposium*, pp.140–146 (Nov. 2013).
- [19] Levine, P.M. and Roberts, G.W.: High-resolution flash time-to-digital conversion and calibration for system-on-chip testing, *Proc. IEE Computers and Digital Techniques*, Vol.152, No.3, pp.415–426 (May 2005).

Kotaro Shimamura received his M.S. degree in physics from Tokyo University, Tokyo, Japan, in 1990. In the same year, he joined Hitachi, Ltd. Since then, he has been involved in the research and development of high-performance RISC microprocessor for massive parallel supercomputers, fault-tolerant controller for power

converters, microcontroller soft core, and fail-safe microcontroller for railway systems. He is a member of the IPSJ, IEEE, ACM, IEICE, and IEEJ.

Naohiro Ikeda received his M.S. degree in engineering from Yokohama National University, Yokohama, Japan, in 1994. In the same year, he joined Hitachi, Ltd. Since then, he has been involved in the research and development of train propulsion control system and controller for railway signaling control system, such as

inverter controller, fail-safe CPU board for ATC (Automatic Train Control), transmitter and receiver for digital ATC, railway signaling equipment controller, fail-safe microcontroller, and Eurobalise.

(Recommended by Associate Editor: *Satoshi Komatsu*)