PAPER Special Section on Dependable Computing

On-Chip Detection of Process Shift and Process Spread for Post-Silicon Diagnosis and Model-Hardware Correlation

A.K.M. Mahfuzul ISLAM[†a)], *Nonmember and* **Hidetoshi ONODERA**[†,††b)], *Fellow*

SUMMARY This paper proposes the use of on-chip monitor circuits to detect process shift and process spread for post-silicon diagnosis and model-hardware correlation. The amounts of shift and spread allow test engineers to decide the correct test strategy. Monitor structures suitable for detection of process shift and process spread are discussed. Test chips targeting a nominal process corner as well as 4 other corners of "slow-slow", "fast-fast", "slow-fast" and "fast-slow" are fabricated in a 65 nm process. The monitor structures correctly detects the location of each chip in the process space. The outputs of the monitor structures are further analyzed and decomposed into the process variations in threshold voltage and gate length for model-hardware correlation. Path delay predictions match closely with the silicon values using the extracted parameter shifts. On-chip monitors capable of detecting process shift and process spread are helpful for performance prediction of digital and analog circuits, adaptive delay testing and post-silicon statistical analysis.

key words: *process shift, process spread, monitor structure, post-silicon analysis, delay test, adaptive test*

1. Introduction

Testing is must for product chips to ensure correct operation. With the emerge of System-on-chip (SoC), testing has become more difficult. An SoC contains several blocks of digital, analog and RF circuits. Testing each part is extremely costly. Often the testing cost surpasses the manufacturing cost. On-chip solutions with the help of sensors can reduce test cost [1].

Delay testing is performed to make sure that the product chip operates at the desired target operating frequency. However, due to process variation and other manufacturing faults, some parts may not achieve the desired frequency, and therefore delays of all the paths must be tested. When some parts of the chip do not achieve the desired frequency, the biggest challenge is to debug the causes of the defect. In case of delay defects, there are mainly two reasons. One is the parametric variation which is often called as process variation and the other is random defect [2].

Normally, process variation is used to express both process shift and process spread effects. Process shift refers to the shifts of several process parameters from the target values. Process spread is the amount of variation within a chip. Process spread can be systematic or random. As process shifts occur at the chip level, they affect the gate delays globally and normally are accounted by considering their effects in the transistor models. Usually, several extreme cases are modeled which are called the corner models. For the designers, there are two strategies while selecting the operating frequency. One is to over design the circuits so that it operates even at the worst-case condition. Under the ever increasing process variation in the scaled technology, worst-case design is way too inefficient to accept in today's highly competitive market. Another strategy is to design the circuit for a typical process variation case, and categorize the products based on the performance [3]. This approach is often called speed binning. Monitor circuits capable of detecting process variation help to predict performances and diagnosis parametric faults.

In order to deal with process spread, several statistical design methods have emerged in recent years. STA (static timing analysis) with OCV (on-chip variation) and SSTA (statistical STA) are two methods which use statistical nature of process variation and try to reduce the extra margins caused by simple STA [4], [5]. STA with OCV is now being used in commercial tools for guard banding the effect of process spread on gate delays. In both cases, on-chip monitoring of process variation has become more important.

Generally, PCM (Process Control Module) and monitoring circuits are placed at the scribe lines to monitor process variation. However, they are not useful to debug the causes of timing failures in product chips. On-chip ring oscillators (RO) consisting of standard inverter or NAND gates can be used for monitoring variation [6]. On-chip monitors are also shown to be effective for screening delay defects [7]. The frequencies of these monitor structures give us useful information on the process variation to some extent. However, when there is mismatch between pMOSFET and nMOSFET performances, the conventional monitors fail to detect the mismatch. When pMOSFET and nMOSFET move to opposite directions, maximum operating frequency of digital circuit may not correlate to the ring oscillator frequency. This paper discusses the use of monitor structures for on-chip detection of process shift and process spread. Monitor structures designed to be sensitive to either pMOSFET or nMOSFET are used in the paper. Implementing the monitor structures with the conventional standard inverter based monitor, process shift and process spread are decomposed into the parameters of threshold voltages and gate length. These parameters are helpful for post-silicon

Manuscript received January 22, 2013.
Manuscript revised April 28, 2013.

[†]The authors are with the Department of Communications and Computer Engineering, Kyoto University, Kyoto-shi, 606–8501 Japan.
[††]The author is with the JST, CREST, Japan.
a) E-mail: mahfuz@vlsi.kuee.kyoto-u.ac.jp
b) E-mail: onodera@vlsi.kuee.kyoto-u.ac.jp

DOI: 10.1587/transinf.E96.D.1971

diagnosis.

The main contributions of this paper are as follows.

1. Detect process shift and process spread directly from on-chip monitor circuits.
2. An efficient model-hardware correlation methodology is presented.
3. Validation from silicon data.

The on-chip monitor circuits can be used to generate test pattern for exercising potential critical paths in the design because potential critical paths vary from corner to corner. Especially, when the process shifts, the monitor circuits provide direct information on the amount of that shift. This paper is an extension over [8]. Detailed discussion on the P/N-variation detection capability and example of several applications are key enhancements of this paper. Path delay predictions are done and compared with silicon values using the extracted parameter shifts. The predictions match closely with silicon values with maximum error of 1.3%.

The remainder of the paper is organized as follows. In Sect. 2, monitor structures capable of detecting process shift and process spread are discussed. A test chip has been fabricated to prove the use of monitor circuits for process detection. Section 3 shows the structure of the test chip and our measurement results. In Sect. 4, a methodology is presented to estimate process parameters for model-hardware correlation. Several application examples are shown in Sect. 5. Finally, Sect. 6 concludes the paper.

2. Monitor Circuits for Process Shift and Process Spread Detection

In this section, the concept and topology of on-chip monitor circuits are discussed. Monitor structures suitable for the concept are explored, and a suitable structure is presented.

2.1 Concept and Topology

Normally, process monitors and PCM monitors are placed on the scribe lines to track process variation. However, these monitors are not useful for chips that are located far away from these modules. ROs consisting of standard inverter cells are used as process monitors. These ROs give us useful information on the process. However, one disadvantage of using such ROs is that when the process moves to "SF" (slow nMOSFET and fast pMOSFET) or "FS" (fast nMOSFET and slow pMOSFET), the difference cannot be distinguished. These mismatches between nMOSFET and pMOSFET may cause unexpected timing failures. Besides, SRAM yield largely depends on the global and local mismatches between nMOSFET and pMOSFET. For post-silicon diagnosis, the location of the chip in the process space need to be known. We, therefore, propose to embed process monitors into the chip and use the monitor outputs for testing and post-silicon diagnosis. The purposes of on-chip monitor circuits are as follows.

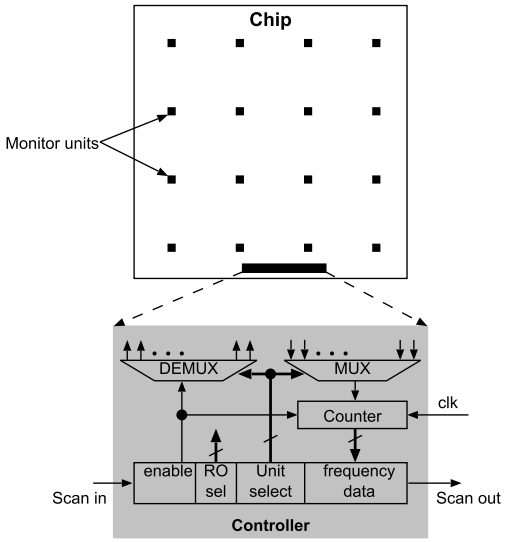


Fig. 1 One example of on-chip implementation of monitor circuits. Conventional scan-chain based interface is used in this example.

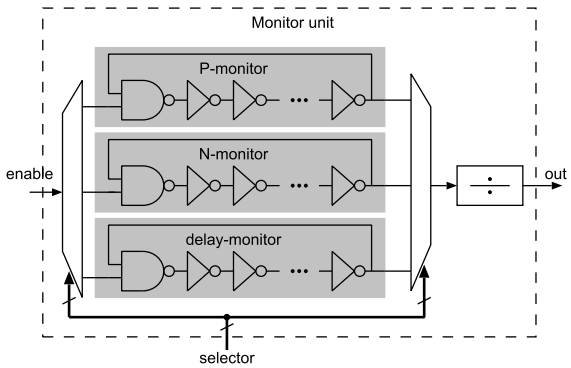
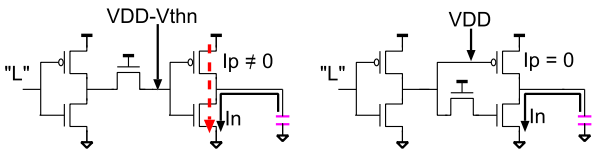


Fig. 2 One example of monitor unit. The monitor unit consists of three process monitors here to detect process shift and process spread.

1. Locate the chip in the process space.
2. Extract the amount of process spread.
3. Estimate parameter deviations.

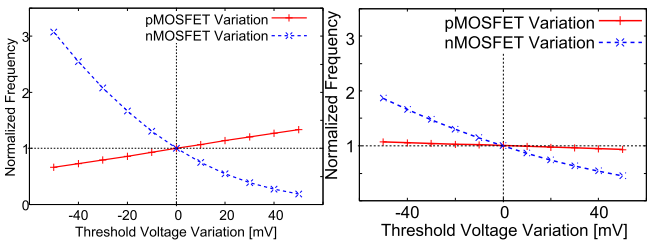
For 1, monitor circuits capable of detecting P/N-variation are needed. For 2, multiple instances of the same monitor circuit need to be distributed over the chip. The number of stages for the monitor should be small to capture process spread. For 3, a set of suitable monitor circuits is required.

Figure 1 shows an example of the concept of on-chip monitor circuits. Instead of placing single process monitor structure which is often an RO consisting of standard inverter cells, we propose to distribute the monitor structures across the chip. Figure 2 shows a block diagram where three ROs are used as monitor circuits. The P- and N-monitor circuits are sensitized to P- and N-variations. The delay-monitor can be used to capture delay deviation. In post-silicon, the ROs are measured and analyzed. The P- and N-monitors give us instant insight on the process condition. The delay monitor is used with the P- and N-monitors for model parameter extractions. The use of three ROs as monitor circuits for extracting three parameters of threshold volt-



(a) Pass transistor at output (b) Pass transistor at input

Fig. 3 Inverter with nMOSFET pass transistor at the input is sensitive to nMOSFET variation only.



(a) Pass transistor at output (b) Pass transistor at input

Fig. 4 Sensitivity to MOSFET threshold voltage variation. Structure of (b) is sensitive to particular MOSFET variation thus suitable for parameter extraction.

ages and gate length is shown in Ref. [9]. For practical purpose, we assume the extraction of these three parameters is enough as they are the dominant factors for performance deviations. Decoders and selectors can be used to select one of the ROs. Dividers can be used to reduce the frequency because the output of each monitor unit is routed globally to a controller. The monitors can also be integrated with on-chip test circuitry such as BIST for on-chip testing as the output is digital.

2.2 Monitor Structures

In order to realize P- and N-monitor circuits, monitor circuits need to be sensitized to P/N variation. ROs consisting of modified inverter structures to enhance sensitivity for capturing various process information are proposed in recent years [9]–[13]. The use of pass-gate based inverter structures are reported to be suitable for sensitivity enhancement [9], [10]. Figure 3 shows two types of inverter structures with pass-gates. In Fig. 3 (a), pass-gate is added at the output of the inverter cell. In Fig. 3 (b), pass-gate is before the input gate of MOSFET. nMOSFET pass-gate based structure increases the inverter delay sensitivity to nMOSFET parameter variations because pass-gate's delay depends on the nMOSFET performance directly. The voltage drop across the pass-gate reduces the gate overdrive for the nMOSFET of the inverter. Thus, any change in threshold voltage is amplified in the RO frequency. Figure 4 shows the change of RO frequency to MOSFET threshold variations for inverter structures of Fig. 3 (a) and Fig. 3 (b), which is obtained by circuit simulation assuming a commercial 65-nm process. Here, nMOSFET pass-gate based inverter structure is used. In Fig. 4, both types of structures show high sensitivity to nMOSFET variations. However, for the structure in Fig. 3 (a), the frequency also changes largely

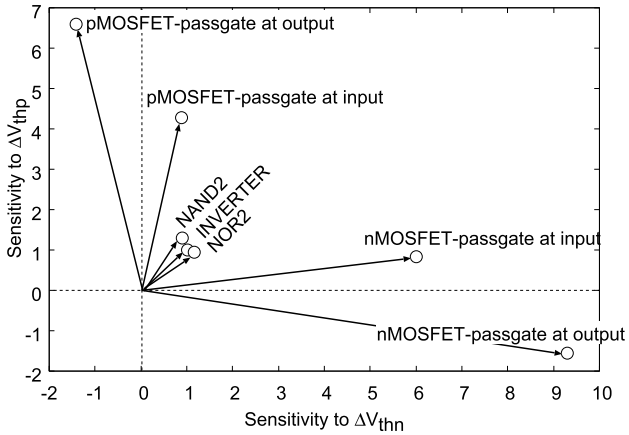


Fig. 5 Sensitivity vectors of pass-gate based process monitors and RO with standard inverter, NAND2 and NOR2 cells.

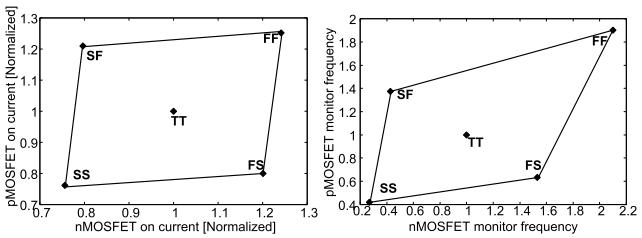
Table 1 Condition Numbers of sensitivity matrices for two structures of pass-gate based monitors along with standard inverter cell RO.

Monitor Structure	Condition Number
Monitor structures in Fig. 3 (a)	55
Monitor structures in Fig. 3 (b)	34

when pMOSFET threshold voltage varies. When pMOSFET threshold voltage lowers meaning pMOSFET becomes faster, the RO frequency decreases instead of increasing. This behaviour is the opposite to the behaviour found in conventional CMOS digital circuits. This is because when pMOSFET becomes faster, the through current increases resulting in large delay during the pull-down. So, RO consisting of this structure does not give us direct in-sight into the process variation. In Fig. 4 (b), the RO frequency is not sensitive to pMOSFET variation.

Figure 5 shows the sensitivity vectors of the pass-gate based process monitors. Sensitivity vectors for ROs consisting of inverter, NAND2 and NOR2 gates are also shown. X-axis and Y-axis refer to $\Delta V_{\rm thn}$ and $\Delta V_{\rm vthp}$ sensitivities respectively. The vector values are normalized with the sensitivity of the standard RO to nMOSFET threshold variation. Standard inverter based RO has similar sensitivities to nMOSFET and pMOSFET variations. NAND2 and NOR2 based ROs do not have enough sensitivity to separate pMOSFET and nMOSFET variation. In the figure, pass-gate based monitors have high sensitivity to nMOSFET or pMOSFET variations.

In order to choose the suitable monitor structures, analysis based on condition numbers can be used [9]. Condition number is an indicator of how robust the matrix is for estimation of unknown parameters. Smaller the number is, more robust the matrix is. In order to detect the process information of three parameters of pMOSFET threshold voltage, nMOSFET threshold voltage and gate length, three monitor circuits are needed. We add the standard inverter RO to the pass-gate based RO sets and evaluate their condition numbers. The result is shown in Table 1. The monitor structure in Fig. 3 (b) set has smaller condition number of 34



(a) MOSFET ON-current (b) p/n-monitor frequency

Fig. 6 MOSFET ON-currents (a) and monitor frequencies (b) simulated at the process corners. Monitor frequencies are distributed in the process space.

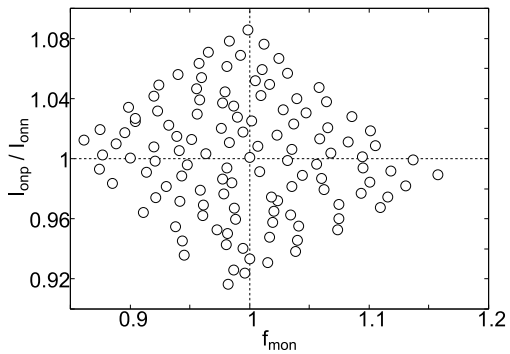


Fig. 7 Correlation between ON-current ratio of MOSFETs, I_{onp}/I_{onn} and standard inverter based process monitor frequency.

than the structure in Fig. 3 (a).

From the above discussions, we choose the pass-gate structure shown in Fig. 3 (b) for P- and N-monitor structures and standard inverter RO for the third monitor circuit for parameter estimation. Next section discusses the capability of the P- and N-monitors for process detection.

2.3 Process Detection Capability

Next, we show the capability of process corner detection for the monitor structures. Figure 6 (a) shows the simulated MOSFET ON-currents for the process corners. The ON-currents at the corner models form a rectangular shape. ON-currents in the silicon are expected to be within the corner boundary. Figure 6 (b) shows the monitor frequencies simulated for corner models. In Fig. 6 (b), the frequencies form a boundary within which the actual silicon values are expected to be. The main point here is that "SF" and "FS" corners are clearly distinguishable from the frequencies of the P/N-monitors.

As the P/N-monitors are sensitized to pMOSFET and nMOSFET variations, they are used to detect P/N-variation. A conventional process monitor which consists of standard inverter gates gives only one output that is the oscillation frequency. Figure 7 plots simulated MOSFET ON-current ratio against the simulated output frequency of standard inverter based process monitor f_{mon} for different process conditions. Different process conditions are realized by varying pMOSFET and nMOSFET threshold voltages in the range of ±50 mV. From Fig. 7, no correlation is found between

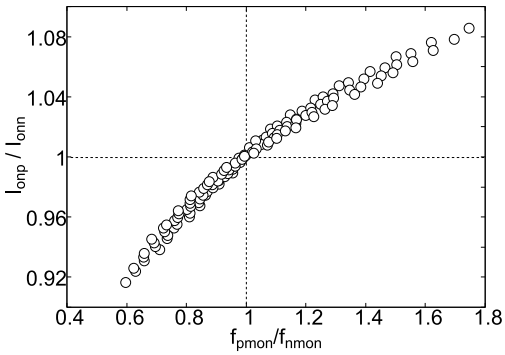


Fig. 8 Correlation between MOSFET ON-current ratio, I_{onp}/I_{onn} and monitor frequency ratio (f_{pmon}/f_{nmon}). High correlation coefficient of 0.98 is found. The p- and n-monitors can detect P/N-ratio with high accuracy.

P/N-ratio and f_{mon}. Conventional process monitor is not useful for debugging circuit failures if the failure is caused from unbalanced P/N.

Now, consider the use of the P-monitor and N-monitor circuits as P/N detectors. Figure 8 plots simulated P/N-ratio and the ratio of P/N-monitor frequencies, f_{pmon}/f_{nmon}. The monitor structures show high correlation coefficient of 0.98 meaning P/N-variation can be predicted with high accuracy. Thus, the P/N-monitors are capable of detecting process shift of pMOSFET and nMOSFET performances.

3. Test Chip Design and Measurement Results

The ability of the on-chip monitors to detect process shift and process spread are demonstrated here. Test chips have been fabricated targeting "TT" condition, as well as four corners of "SS", "FF", "FS" and "SF". Five chips from each of the process corners are measured.

3.1 Test Chip Design

A test chip in a 65-nm process has been designed to verify the use of monitor circuits for process detection. Figure 9 shows the chip micro-graph and overall structure of the test chip. 294 monitor units are implemented with an array of 14×21. In a product chip, the monitors will be distributed over the chip at several locations. Here, the purpose is to make in-depth analysis and confirm the validity of the monitor circuits. Each unit consists of P/N-monitor circuits, standard inverter cell RO, NAND2 RO and NOR2 RO. NAND2 RO and NOR2 RO are used for post-silicon timing validation which is discussed in Sect. 5. An NAND2 gate is used to control the oscillation of the ROs with an enable signal. From Fig. 5, the sensitivities of P-monitor and N-monitor gate is more than six times larger than the NAND2 gate. Thus, the use of NAND2 gate at one stage does not affect the capabilities of P- and N-monitors. A decoder selects an RO and a selector selects the output of the oscillating RO. The output frequency is captured outside the chip. Dividers are used to reduce the frequency. Select signals are provided from outside the chip. All the ROs are 13-staged. The aver-

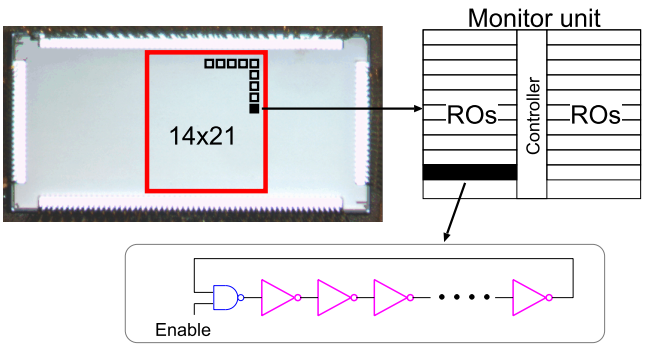


Fig. 9 Test chip structure. several types of ROs are implemented on-chip.

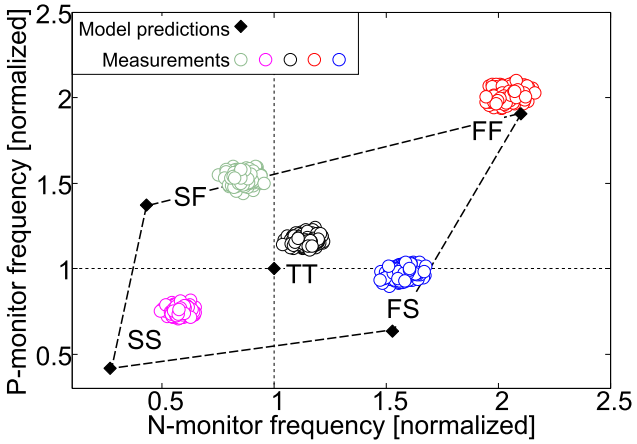


Fig. 10 Measured monitor frequencies from 5 chips. 5 chips represent 5 process corners. Each chip contains 294 instances of each monitor.

age value of frequencies of an RO type gives the information on process shift. Deviations in the frequencies give us the information on process spread.

3.2 Measurement Results

3.2.1 Process Shift

Figure 10 shows the measured frequencies of P-monitors and N-monitors from 5 chips (open circles). The chips represent five process corners of "TT", "SS", "FF", "FS" and "SF". The values are normalized by the values estimated with the "TT" corner model. Frequency values estimated using the corner models are also plotted in the figure (closed squares). In Fig. 10, process shifts from the "TT" model prediction are observed. In Fig. 10, deviations are observed between the predicted and measured corner frequencies. Clear deviations are observed for "TT", "SS", "SF" and FS" corners. The silicon values are higher than the model predictions. With comparison with the models, we can have quick understanding of process shift for each chip. This information allow us to take decisions for silicon debug and test pattern generation. By doing further analysis, model-hardware correlation can be performed which allow us to tune the designs. Model-hardware correlation results are presented in Sect. 4.

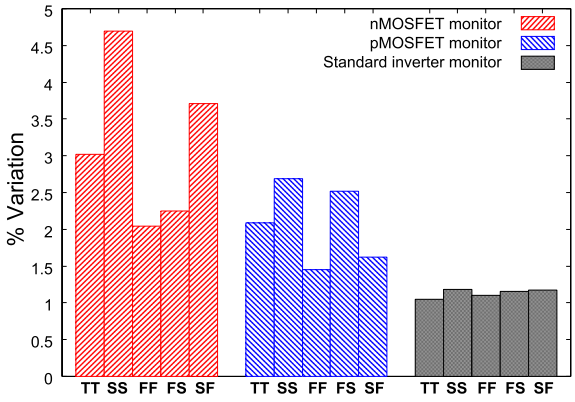


Fig. 11 WID variation observed in nMOSFET monitor, pMOSFET monitor and standard inverter ROs at the corner chips.

3.2.2 Process Spread

In Fig. 10, 294 measured values for P- and N-monitors are plotted for each chip. Variation among the frequencies in a chip is observed which represent the process spread for that chip. The amount of spread can be different from chip to chip depending on the location of the chip in the process space. In order to evaluate the amount of spread, we have calculated the standard deviations for the frequencies in a chip. The results are shown in Fig. 11. The frequency variation is the function of the sensitivity coefficients and the amount of process variation. As P- and N-monitor ROs have larger sensitivities than the standard inverter RO, they are showing larger variations in Fig. 11. The difference in the variations between the P-monitor and N-monitor ROs also reflects the difference in the pMOSFET and nMOFET variations of the chip. From standard inverter RO measurements, similar WID variation is observed for all the corner chips. However, for the P-monitor and N-monitor, significant differences in the amount of spread between the chips are observed. For "FS" corner chip, nMOSFET monitor's variability becomes smaller and pMOSFET monitor's variability becomes larger than the "TT" corner chip. This indicates that the intrinsic variability in the nMOSFET and pMOSFET performances are different in the unbalanced corners which may cause the yield to decrease drastically [14]. The extend of process spread needs to be accurately monitored and feedback into the design. Our monitors are suitable for distinguishing nMOSFET and pMOSFET variations. In order to model the effects of WID variations, the variation needs to be expressed by transistor model parameters so that designers can use them. In Sect. 4, obtained WID variations are decomposed and modeled into variations of three transistor model parameters of threshold voltages and gate length.

4. Parameter Estimation for Model-Hardware Correlation

In this section, we show that how the measured on-chip fre-

Table 2 Estimation parameter deviation using the monitor circuit measurements against the typical model parameter values.

Corner	delvt0 [mV]		ΔL [nm]
	nMOSFET	pMOSFET	
TT	−17 ∼ −12	−32 ∼ − 17	−1.8 ∼ −0.4
SS	43 ∼ 55	7 ∼ 34	−0.6 ∼ 3.0
FF	−96 ∼ −85	−130 ∼ −111	−1.4 ∼ 1.2
FS	−70 ∼ −57	27 ∼ 49	−3.5 ∼ −1.2
SF	10 ∼ 23	−123 ∼ −104	1.0 ∼ 3.1

quencies can be used to estimate several process parameter deviations so that correlation between model and silicon can be obtained during the manufacturing process of a product.

4.1 Process Shift

After detecting the process shift of the product chips, those shifts can be decomposed into the model parameters. Reference [13] shows a methodology to decompose RO frequencies into several key process parameters such as threshold voltages and gate length. An iterative estimation procedure is used in this methodology where the differences between measured frequencies and predicted frequencies are decomposed into model parameters using the sensitivities to each parameter. In this test chip, we have extracted $\Delta V_{\rm thn}$, $\Delta V_{\rm thp}$ and ΔL against the values in the "TT" corner model for all the chips using the above method.

WID random variation affects the accuracy of the estimation. As random variation follows gaussian distribution, the error can be reduced by increasing the number of stages. The relationship between the number of stages and the estimation accuracy is discussed in Ref. [9]. In our test chip, 294 ROs each having 13 stages are used for global variation monitoring. The number of stages thus becomes $294 \times 13 = 3822$ which we consider to be large enough to cancel the random variation effect. In case of product chips, however, the number of stages is limited. The number need to be chosen based on the amount of random variation and the desired tolerable range of estimation [9].

In this extraction, $\Delta V_{\rm th}$ is expressed by the transistor model parameter "delvt0" which is a dedicated parameter for modeling threshold voltage shift in HSPICE [15]. The values are shown in Table 2. Here, negative values of threshold voltage deviation refers that the absolute value is lower than that in the "TT" model. One key characteristics derived from the real chip measurement is that the gate length does not vary much compare to the variations in the threshold voltages. Large threshold voltage shifts are observed for the "FF" corner chip.

Figure 12 shows the threshold voltage corner in the model and the estimated corner from the RO frequencies. Threshold voltage values from the PCM data of "TT" corner wafer as well as other corners are also plotted. In this lot, some deviations are observed between extracted values and those by the corner models. Some amounts of deviations are also observed between the estimated values and those in the PCM data. These deviations can be caused by a number of

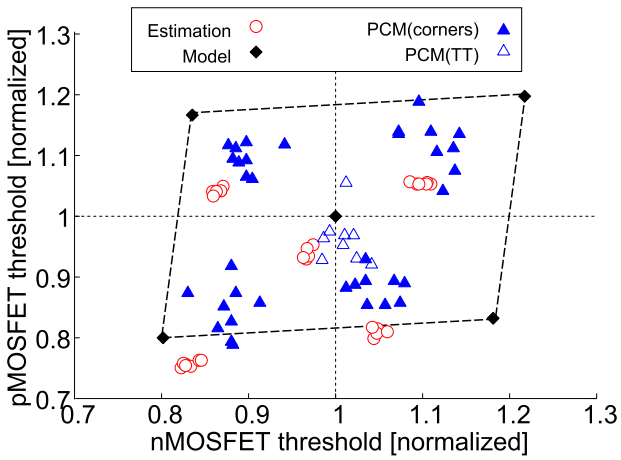


Fig. 12 Comparison between threshold voltages in corner model and in estimations. PCM data for "TT" corner wafer as well as other corners are also plotted.

Table 3 Extracted standard deviation of MOSFET threshold voltages and gate length from RO frequency measurements.

Corner	$\sigma_{V_{\rm thn}}$ [mV]	$\sigma_{V_{\rm thp}}$ [mV]	σ_L [nm]
TT	16.6	11.9	0.89
SS	18.3	14.5	0.53
FF	20.9	16.6	1.14
FS	18.2	13.3	0.99
SF	18.2	13.6	0.99

factors. One possible reason can be the difference between the layouts in the PCM circuits and our test circuits as differences in poly and diffusion densities can affect MOSFET performances.

4.2 Process Spread

If we put multiple instances of an nMOSFET monitor, a pMOSFET, and a standard inverter cell RO on the chip, then standard deviations can be derived for the RO frequencies. In the test chips, we assume that random variation is the most dominant component in the WID variation. Next, we extract the amount of variations in threshold voltages and gate length from the RO frequency variations using frequency sensitivities to each transistor in the RO. The sensitivity coefficients are calculated with circuit simulation. Sensitivity-based method to extract parameter variations from process-sensitive RO frequencies is discussed in Refs. [16], [17]. In Ref. [17], a system of linear equations is built using the sensitivity coefficients. Then, from the measured frequency deviations, the unknown amounts of parameter deviations are estimated with a maximum likelihood method. Using the sensitivity coefficients, the standard deviations of $V_{\rm thp}$, $V_{\rm thn}$ and L are derived from the measured standard deviations in Fig. 11. Derived values are shown in Table 3. nMOSFET variation is larger than that of pMOSFET. The extracted amounts of variations for nMOSFET and pMOSFET threshold voltage match closely with the values reported in Ref. [18] where a similar 65-nm process

Table 4 Delay mismatch between silicon and model prediction for NAND2 and NOR2 delay paths. Delays between silicon and prediction match close when process calibration is done with the estimated process parameter shifts.

	NAND2 delay mismatch [%]		NOR2 delay mismatch [%]	
Process Corner	w/o calibration	w/ calibration	w/o calibration	w/ calibration
TT	7.6	1.0	8.7	0.6
FS	9.7	0.0	5.8	0.8
SF	7.7	1.3	14	0.9
SS	−13	1.1	−12	0.5
FF	33	1.2	36	1.1

is used. These information are extremely useful for statistical design, yield analysis and post-silicon timing analysis. In product chips, limitations in the numbers of samples may cause some errors in the estimations. These limitations need to be considered during practical applications.

5. Application

In this section, several applications are discussed which can be benefited from on-chip monitor structures.

5.1 Performance Prediction

Predicting the performance of the designed circuit in real silicon is difficult under the presence of large process variation. Depending on the location of the chip in the process space, the critical paths change which affects the maximum operating frequency. The effect of process variation is more severe for analog and RF circuits. For analog and RF circuits, the mismatch between pMOSFET and nMOSFET may cause the circuits to fail. In this section, we show that how on-chip monitors that gives us information on the process shift can be useful for predicting performances for both the digital and analog circuits.

5.1.1 Digital Circuit

Post-silicon statistical processing can be used to identify faulty paths and predict the delays of paths [19]. Process calibration is needed for post-processing. Specific information of key process parameters are required for process calibration. For example, in SSTA, the delay, d of a path is modeled using the sensitivity coefficients as follows.

$$d = \mu_d + \sum k_{p_i} \Delta p_i + rnd. \quad (1)$$

Here, μ_d is the mean value, Δp_i is the $i-$th parameter that has variation and k_{p_i} is the sensitivity coefficient to parameter p_i. rnd is the random component which is the result of random variation. In Sect. 4, it is shown that process parameters can be extracted with on-chip monitor circuits. After the process parameters are known for a chip, the delay can be calculated using Eq. (1). The accuracy of the estimation of Δp_i in Eq. (1) holds the key for accurate timing analysis. Here, we show that the delays of different paths can be predicted with high precision using the estimated parameters.

Delays of two path types are tested on silicon. The first path consists of NAND2 cells and the second path consists of NOR2 cells. We have 294 instances of identical paths across the chip to evaluate mean delay value and the standard deviation. Table 4 shows the amount of mismatch between predicted mean delay values and real silicon mean delay values for 5 corner chips. Mismatches are shown for two cases. The first case is when process calibration is not considered. The second case is when process calibration is performed. When the process calibration is not considered, the prediction is done with the "TT" corner model provided by the foundry. As expected, large mismatches are found for the path delays when their is large variation which is represented by the corner chips here. Next, process calibration is performed by considering the parameter shifts from the "TT" model. In Table 4, the predictions with the estimated parameters and silicon values match closely with maximum error of 1.3%. This proves the validity of the monitor circuits and the estimation procedure. Thus, the monitor circuits are extremely helpful for post-silicon timing validation and performance prediction.

5.1.2 Analog Circuit

On-chip monitors can be useful to predict analog circuit performances. For analog circuits mismatch between transistors and P/N-ratio is important for deciding circuit parameters. If the process variation is not controlled very well, large amount of yield loss can happen. In [20], use of on-die monitor circuits are proposed to identify process variations. The information of process variations are then used to guide the test and to allow the estimation of selected performance figures. In their approach, replica of several parts of the DUT is used for monitor circuits. The use of replica of DUT parts gives us direct information on the DUT performance, but it is not useful for debugging the cause of failure. For example, when the DUT fails, the reason can be either unbalanced P/N-ratio or excessive random variation causing neighborhood transistor mismatch or manufacturing defect or so on. P/N-monitors give us information on P/N-ratio and help us debugging the causes.

5.2 Model Mismatch Detection

In analog circuit, often wider gate lengths are used to get better linearity. However, using wider gate lengths have risks of large mismatch between transistor model and silicon. Characterizing single transistor is expensive and data

processing takes time. Often we are interested on the key parameters such as threshold voltage and ON-current. The estimation method using the proposed set of monitor circuits can be used for the detection of model mismatch. Monitor circuits need to be designed with the same gate length as used in the DUT. In post-silicon, the monitor circuits will be measured and the measured values will then be used for the estimation of threshold voltages and gate lengths. The estimated parameters will give us the information on model mismatch. This method is particularly useful when the design is implemented on multiple fabs and processes.

5.3 Adaptive Testing

In the scaled technology, any path in a design has the potential to be critical depending on the process condition. Testing each path is impossible in today's SoCs where there are millions of paths. Delay test takes time and thus smaller the number of paths to be tested, smaller the test cost is. In order to reduce test cost as well as increase product quality adaptive testing is very attractive [21], [22]. In adaptive testing, instead of using a fixed test strategy and fixed test patterns, they are changed over time based on the information from the products. On-chip monitor circuits come into play a very important role here as they provide information of process parameters for each product chip.

There are various approaches of adaptive testing. One use of the on-chip monitor circuits which tell us the location of the chip in process space is for parametric fault testing. In [22], an adaptive test flow is proposed to detect parametric fault for SSTA based designs. The idea is to cluster the critical paths for several process conditions and generate test patterns for each process condition during the design phase. During the test phase, the monitor circuits are measured first. From the monitor circuit outputs, the process condition of the target chip is identified and the corresponding test patterns are used for path delay testing. This approach can save a large amount of test time which in turn saves test cost. The discussed on-chip monitors structures are ideal for this kind of application.

6. Conclusion

This paper proposes the use of on-chip monitor circuits for detection of process shift and process spread for post-silicon diagnosis, process characterization and model-hardware correlation. Special inverter structures are used to make the ring oscillator frequency sensitive to either nMOSFET or pMOSFET variation. The proposed monitors along with the standard inverter RO embedded in the product chips, enable quick detection of process shift and process spread in the transistor model parameters such as threshold voltage and gate length. Extraction techniques for model-hardware correlation are presented. Global and WID variations in key transistor model parameters are successfully derived from chip measurements for several process corners designed in a 65-nm process. The proposed monitor structures can be used for post-silicon diagnosis of parametric fault and performance mismatch of digital and analog circuits. The monitor structures are also suitable for adaptive testing based on process condition.

Acknowledgements

The VLSI chip in this study has been fabricated in the chip fabrication program of VDEC, the University of Tokyo in collaboration with STARC, e-Shuttle, Inc., and Fujitsu Ltd. This work was partly supported by JSPS KAKENHI Grant Numbers 22300016, 25·6432.

References

[1] N. Aghaee, Z. Peng, and P. Eles, "Process-variation and temperature aware soc test scheduling using particle swarm optimization," IEEE 6th International Design and Test Workshop (IDT), pp.1–6, Dec. 2011.
[2] S. Mitra and P. Ryan, "Delay defect characteristics and testing strategies," IEEE Design & Test of Computers, vol.20, no.5, pp.8–16, Sept. 2003.
[3] A. Datta, S. Bhunia, S. Mukhopadhyay, and K. Roy, "Speed binning aware design methodology to improve profit under parameter variations," Asia and South Pacific Design Automation Conference, pp.712–717, 2006.
[4] C. Visweswariah, "First-order incremental block-based statistical timing analysis," IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems, vol.25, no.10, pp.2170–2180, 2006.
[5] A. Nardi, E. Tuncer, S. Naidu, A. Antonau, S. Gradinaru, T. Lin, and J. Song, "Use of statistical timing analysis on real designs," Design, Automation & Test in Europe Conference, pp.1–6, April 2007.
[6] K. Kuhn, "Reducing variation in advanced logic technologies: Approaches to process and design for manufacturability of nanoscale CMOS," IEEE Electron Devices Meeting, pp.471–474, 2007.
[7] S. Mitra, E. Volkerink, E. McCluskey, and S. Eichenberger, "Delay defect screening using process monitor structures," 22nd IEEE VLSI Test Symposium, pp.43–48, 2004.
[8] I.A. Mahfuzul and H. Onodera, "On-chip detection of process shift and process spread for silicon debugging and model-hardware correlation," 21st IEEE Asian Test Symposium, pp.350–354, Nov. 2012.
[9] A. Islam, A. Tsuchiya, K. Kobayashi, and H. Onodera, "Variation-sensitive monitor circuits for estimation of global process parameter variation," IEEE Trans. Semiconductor Manufacturing, vol.25, no.4, pp.571–580, 2012.
[10] M. Ketchen and M.B. Bhushan, "Product-representative "at speed" test structures for CMOS characterization," IBM Journal of Research and Development, vol.50, no.4, pp.451–468, 2006.
[11] A. Gattiker, M. Bhushan, and M.B. Ketchen, "Data analysis techniques for CMOS technology characterization and product impact assessment," IEEE International Test Conference, pp.1–10, 2006.
[12] V. Zolotov, D. Lackey, P. Ha, and C. Visweswariah, "Variation-aware performance verification using at-speed structural test and statistical timing," IEEE/ACM International Conference on Computer-Aided Design, pp.405–412, Nov. 2007.
[13] I. Mahfuzul, A. Tsuchiya, K. Kobayashi, and H. Onodera, "Variation-sensitive monitor circuits for estimation of Die-to-Die process variation," IEEE Intl. Conference on Microelectronic Test Structures, pp.153–157, 2011.
[14] M. Yamaoka and H. Onodera, "A Detailed Vth-Variation Analysis for Sub-100-nm Embedded SRAM Design," IEEE International SOC Conference, pp.315–318, Sept. 2006.
[15] I. Synopsys, HSPICE User's Manual: Simulation and Analysis, 2010.
[16] M. Bhushan, M.B. Ketchen, S. Polonsky, and A. Gattiker, "Ring

oscillator based technique for measuring variability statistics," IEEE International Conference on Microelectronic Test Structures, pp.87–92, March 2006.
[17] S. Fujimoto, I.A. Mahfuzul, T. Matsumoto, and H. Onodera, "Inhomogeneous ring oscillator for WID variability and RTN characterization," IEEE Intl. Conference on Microelectronic Test Structures, pp.25–30, 2012.
[18] T. Tsunomura, A. Nishida, F. Yano, A.T. Putra, K. Takeuchi, S. Inaba, S. Kamohara, K. Terada, T. Hiramoto, and T. Mogami, "Analyses of 5 σ Vth fluctuation in 65nm-MOSFETs using Takeuchi plot," Symposium on VLSI Technology, pp.156–157, 2008.
[19] Q. Liu and S.S. Sapatnekar, "Confidence scalable post-silicon statistical delay prediction under process variations," 44th ACM/IEEE Design Automation Conference, pp.497–502, June 2007.
[20] A. Zjajo, M. Barragan Asian, and J. de Gyvez, "BIST method for die-level process parameter variation monitoring in analog/mixed-signal integrated circuits," Design, Automation & Test in Europe, pp.1–6, 2007.
[21] P. Maxwell, "Adaptive testing: Dealing with process variability," IEEE Design & Test of Computers, vol.28, no.6, pp.41–49, 2011.
[22] M. Shintani, T. Uezono, T. Takahashi, H. Ueyama, T. Sato, K. Hatayama, T. Aikyo, and K. Masu, "An adaptive test for parametric faults based on statistical timing information," IEEE Asian Test Symposium, pp.151–156, Nov. 2009.

A.K.M. Mahfuzul Islam received B.E. in Electrical and Electronics Engineering in 2009, and M.E. in Communications and Computer Engineering in 2011 from Kyoto University, Kyoto, Japan. At present, he is working for the Ph.D. degree at the same university. His research interest includes process variation monitoring and variation-aware design techniques for low power LSI. He is a member of IEEE and IPSJ.

Hidetoshi Onodera received the B.E., and M.E., and Dr. Eng. degrees in Electronic Engineering from Kyoto University, Kyoto, Japan, in 1978, 1980, 1984, respectively. He joined the Department of Electronics, Kyoto University, in 1983, and currently a Professor in the Department of Communications and Computer Engineering, Graduate School of Informatics, Kyoto University. His research interests include design technologies for Digital, Analog, and RF LSIs, with particular emphasis on low-power design, design for manufacturability, and design for dependability. Dr. Onodera served as the Program Chair and General Chair of ICCAD and ASP-DAC. He was the Chairman of the IPSJ SIG-SLDM (System LSI Design Methodology), the IEICE Technical Group on VLSI Design Technologies, the IEEE SSCS Kansai Chapter, and the IEEE CASS Kansai Chapter. He served as the Editor-in-Chief of IEICE Transactions on Electronics and IPSJ Transactions on System LSI Design methodology.