Department of Electrical and Computer Engineering Faculty Publications

Department of Electrical and Computer Engineering

January 2008

A generic and reconfigurable test paradigm using low-cost integrated poly-Si TFTs

Jing Li

Swaroop Ghosh

Kaushik Roy

A Generic and Reconfigurable Test Paradigm using Low-Cost Integrated Poly-Si TFTs

Jing Li, Swaroop Ghosh and Kaushik Roy

Electrical and Computer Engineering, Purdue University, West Lafayette, IN47906

Abstract

In this work, we propose a novel low power, process tolerant, generic and reconfigurable test structure to reduce the test cost, improve diagnosability and verifiability of complex VLSI systems. The test structure contains a variety of configurable design-for-test units designed with low cost Low Temperature Polycrystalline Silicon Thin Film Transistors (LTPS TFTs) that are fabricated on a separate substrate (e.g., polymer, glass etc). The proposed test circuits do not consume any silicon area because they can be integrated on the chip using 3-D technology. This reconfigurable test paradigm eliminates the need to re-design the BIST components that may vary from one processor generation to another.

1. Introduction

In nanometer technologies, lithographic limitations may cause large parametric fluctuations leading to large spread in the overall system performance, even leading to functional failures. Shrinking transistor geometry allows the designers to put more functionality on chip but testing such complex systems becomes a challenge. Increasing die size not only introduces new fault sites but also new fault mechanisms. To maintain high coverage, a large number of test patterns are required, resulting in increased test-power, test-time, and test-cost. Apart from off-line test, on-line test also plays an important role in maintaining the reliability of nano-scaled systems. The need for on-line test has been growing with scaling of technology due to time dependent "hard failures" such as gate oxide breakdown, electro-migration and subtle manufacturing defects [1].

An on-chip tester may solve the above test problems but at the cost of design overhead and reliability of the tester itself. Another solution is addition of on chip design-for-test (DFT) [2] circuits for on-line/off-line test and calibration. For example, an on-chip leakage and delay monitoring scheme [3] has been used for calibration of body-bias to reduce parametric failures in memory. Similarly, delay sensing [4] has been used for delay test, diagnosis and speed binning. Another interesting technique called CrossCheck was proposed in [5] to completely address the testability problem. The authors suggest an array based test structure similar to bed-of-nails for drastic improvement in testability and diagnosis of the internal nodes. The design is laid-out in such a way that it can accommodate a grid of control and sense wires. A tiny NMOS transistor is used at the output of every logic cell which can be turned-on based on the control signal and the cell output can be observed on the sense line. This provides massive observability of the system under a given test sequence. The technique has been verified against a number of manufacturing defects. Although this technique improves test and diagnosis capability of the system, however, it comes at the price of die-area and test power associated with control/sense wires and NMOS transistors. Further, it does not address the issue of on-line test and diagnosis. Hence, there is a need to develop techniques which improve the on-line/off-line testability while reducing the area overhead and test cost.

Low Temperature Polycrystalline Silicon Thin Film Transistor (LTPS TFT) has been investigated as a promising candidate for low-power, low-cost applications with medium performance (e.g., 10-100MHz) [6]. Conventionally, LTPS TFTs have been widely used in liquid crystal display (LCD) as pixel-switching-elements due to its excellent optical properties, low fabrication cost, and manufacturability on flexible (e.g., polymers, glasses etc) substrates [7]. However, the performance of TFT devices is low and highly defective grain boundaries (GBs) regions are present in the channel. Proper device optimization strategy has been proposed to keep the number of GBs under control to improve the performance and reduce power, opening up plethora of new and interesting applications [6]. In this work, we develop a low cost and robust generic test architecture designed with such optimized TFTs that can be fabricated on a separate substrate and integrated on top of the Silicon die to address the testability issues of complex VLSI systems. Fig. 1 illustrates the conceptual diagram of the proposed test structure. The bottom surface shows a typical SoC containing analog, digital, DSP and mixed signal components. Each unit cell of the test structure is configurable and may function as (but not limited to) a set of process/reliability sensors, some discrete test components, storage elements etc. The TFT test structures can be used to partially or fully replace a DFT circuit.

The proposed test structures have the following features: (a) reduction in test cost (area as well as power) associated with DFT circuits since they can be moved to the TFT layer with small performance overhead; (b) minimal overhead on silicon chips since the test structure can be integrated after the die has been fabricated; (c) ability to perform on-line test and diagnosis; (d) elimination of the need to re-design the BIST components that may vary from generation-to-generation for any processor or DSP core;
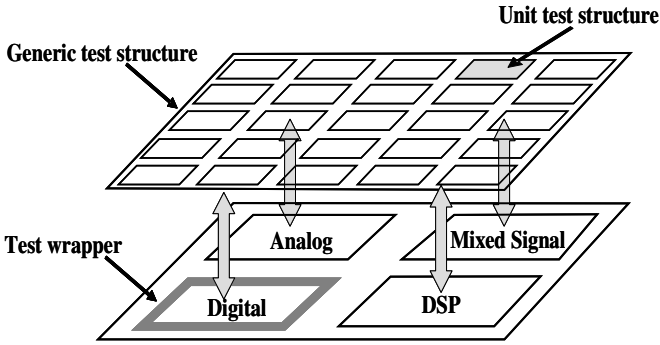
INTERNATIONAL TEST CONFERENCE

Fig. 1 A conceptual diagram of the generic test structure
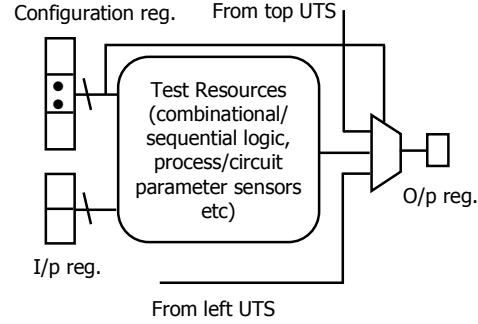
Fig. 2 Block diagram of a unit test structure (UTS)
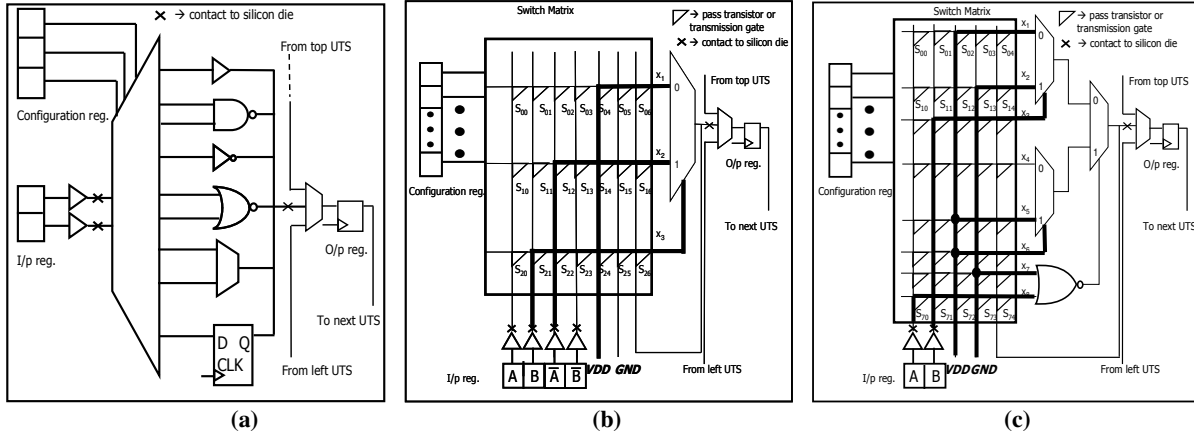

(a)   **(b)**   **(c)**

Fig. 3 Three possible implementations of a UTS: (a) discrete component based (COMT), (b) one multiplexer based (MUX -1) and (b) three multiplexer based (MUX-3). The bold lines in MUX-1 and MUX-3 indicate input connections to realize a *nand* gate

and (e) nearly constant footprint because the test circuits do not consume any silicon area. In particular, we make following contributions in this paper:

- We proposed a novel, low cost, generic and reconfigurable TFT-based test paradigm that can assist off-line test by reducing test time and test overhead. The proposed TFT-based circuits can be implemented in a separate flexible substrate and integrated with the silicon die using 3-D hybrid integration. The test structure can be programmed to obtain MISR/LFSR of required length and characteristic polynomial. The test architecture also provides controllability and observability to the system by controlling and probing out the hard-to-control/observe nodes.

- The proposed test structure can be a part of high-volume manufacturing process where a family of processors or DSP cores can be tested efficiently just by re-configuration of the test structure. The test structures can be fabricated independently to reduce the time-to-market of the processors/DSP cores.

- We developed a device optimization methodology to make LTPS TFT operate at CMOS compatible supply voltage with sufficient current drive. We also developed Hspice model of TFT for circuit implementation. A statistical simulation technique to estimate and evaluate the variations in TFTs and their

impact on designing low power and robust test circuits is also developed.

- We proposed three possible implementations of the test structure that can be configured to realize different logic functionalities. We studied the basic logic gates (e.g., inverter, nand, nor, xor and latches obtained by configuring the test structures) and evaluated them in terms of power, performance and process variability.

- The proposed test architecture can be used for both off-line and on-line test, diagnosis and possibly self repair. It can not only monitor the basic functionality of any chip/system block but also detect/diagnose functional failures in individual blocks.

The rest of the paper is organized as follows. The overall test approach (using TFT test structures) is described in Section 2. Optimization of TFT for low power and high performance in presence of variability is discussed in Section 3. Various implementations of unit test structures are evaluated in terms of power, speed and variability in Section 4. Examples of potential applications using the proposed test structure are presented in Section 5. The practical challenges and issues are addressed in Section 6 and conclusions are drawn in Section 7.

2. Basic Idea

In this section, first we present a *testable* TFT-based test architecture. Next, we elaborate the proposed test and
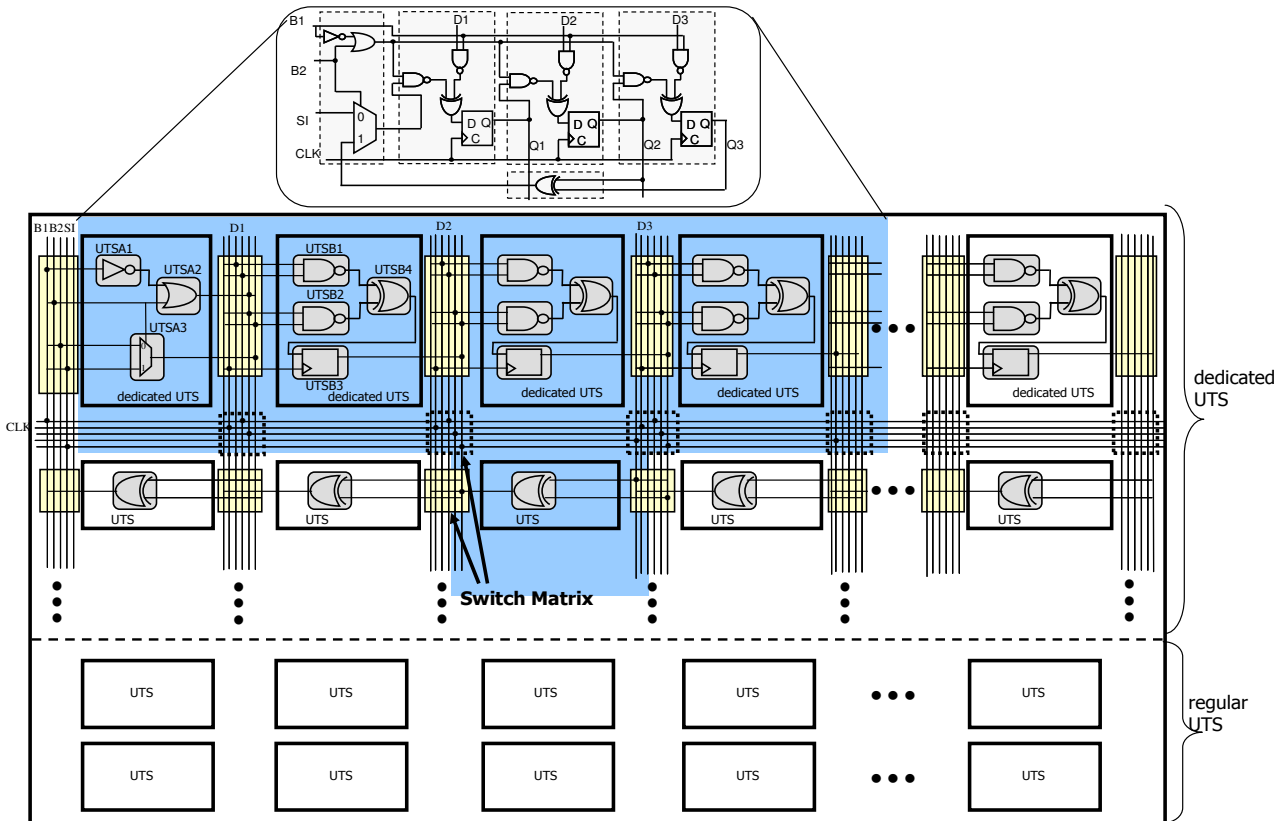
Fig. 4 Block diagram of the test architecture consisting of two blocks: dedicated UTS and regular UTS. The dedicated UTS may consist of a number of regular UTSs (e.g., one dedicated UTS in this figure is constituted of UTSA1-UTSA3) and it can be configured as any test module e.g., MISR, LFSR, scan chain etc. For example, the shaded part demonstrates the dedicated UTSs configured as a 3-bit BILBO circuit.

diagnosis scheme using this architecture. We also address the 3-D integration of the test structure with silicon die.

2.1 Test Architecture

Fig. 1 illustrates the basic concept of test using the TFT based reconfigurable test structure that is fabricated on a separate substrate and integrated with the silicon die using 3-D die-to-die vias. The test structure consists of an array of cells (i.e., Unit Test Structures (UTSs)) each of which can be configured to perform different combinational/sequential logic operations. The UTSs may also contain sensors to measure various circuit parameters (e.g., slew rate, delay, leakage etc). Therefore, each UTS may consist of various test resources, a configuration register and input/output registers as shown in Fig. 2. The UTS can be configured to perform a particular function by loading appropriate pattern in the configuration register. Inputs to the test resources of UTS can either come from the silicon die or from the input registers depending on the operating mode of the test structure. During the *test mode* of the test structure, the inputs of the UTS are loaded by the tester in the input registers, however, during *normal mode* of operation; the inputs come from the node of interest on silicon die. The output of the UTS can propagate along the two directions on the 2-D plane (in TFT layer), namely, top-to-down (in vertical fashion) or left-to-right (in horizontal fashion) as shown in Fig. 2. The direction of propagation of output can be controlled by

providing proper select inputs to the multiplexer from the configuration register.

Possible implementations of the UTS

The UTS can be implemented in several ways. Fig. 3 shows three possible implementations of the UTS as discussed below:

Component based (COMT): In this implementation, the UTS consists of a set of combinational and sequential logic as test resources. The de-multiplexer connects the inputs to one of the gates based on the value stored in configuration register (Fig. 3(a)). The outputs of the gates and the outputs of the adjacent UTSs (from horizontal and vertical directions) are fed to a multiplexer. The select line of the multiplexer can be controlled by the configuration register. This feature can be used for test data propagation (in horizontal or vertical direction). The output of the UTS is latched and can be used for test/diagnosis purposes. The advantage of this UTS implementation is the requirement of only a few configuration registers and its simplicity. However, the area may increase to realize complex functionalities.

Multiplexer based (MUX): This implementation contains a set of multiplexers that can produce a wide range of combinational and sequential logic functions. Fig. 3(b) shows single multiplexer based UTS (MUX-1) that can be used to provide as many as 10 different logic functions [8].
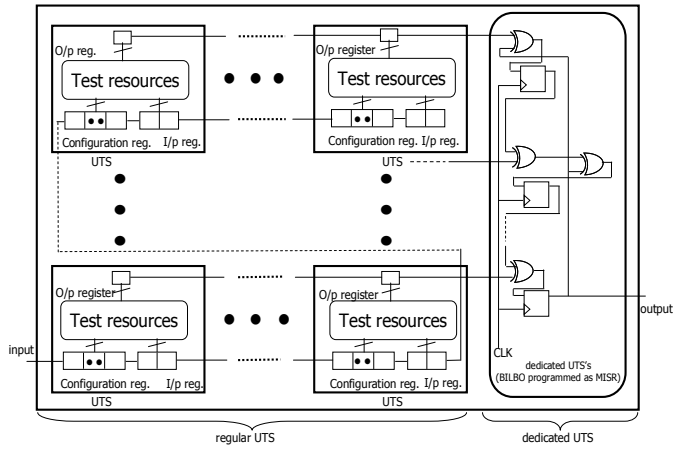
Fig. 5 Testing of the test structures. Test patterns are applied by external tester while the test outputs can either be compacted (using dedicated UTS) or fed back to the tester.

Fig 6 (a) 3-D integration of the TFT test structure with the silicon wafer, (b) model of die-to-die via

Inputs x_1-x_3 can be configured (using a switch matrix) to generate the required logic function by connecting them to the actual inputs A, B, $\overline{A}$, $\overline{B}$, VDD and GND. The input connections to realize a two-input *nand* gate are highlighted by bold lines in Fig. 3(b). In this figure, multiplexer inputs (x_1, x_2, x_3) are connected to (*VDD*, $\overline{A}$, *B*). Note that, the input connections to realize a desired logic function can be determined by first expanding the required logic function using Shannon's expansion theorem [9] with respect to one of the inputs, and then using the multiplexer to implement it. For example, the above mentioned two-input *nand* function can be expanded with respect to input *B* to get $\overline{A.B} = (1).\overline{B} + (\overline{A}).B$ which leads to the input connection as shown in Fig. 3(b). Programming of the input connections can be done by proper configuration of the switch matrix. As shown in Fig. 3(b), the switch matrix contains a set of pass transistors or transmission gates (S_{00}-S_{26}) to connect one of the inputs (*A, B, $\overline{A}$, $\overline{B}$, VDD or GND*) to the multiplexer inputs (x_1-x_3). To generate the *nand* gate mentioned above, switches S_{04}, S_{12} and S_{21} should be turned on. The control vector to turn-on and off a switch can be stored in the configuration register. Fig. 3(c) illustrates another structure of multiplexer-based UTS (MUX-3) which consists of three multiplexers and a two-input *nor* gate. The configuration process is similar as the MUX-1 implementation except that the size of the switch matrix and the configuration register are larger than MUX-1 implementation. MUX-1 requires only 21 transmission gates while MUX-3 implementation requires 40. However, note that we need only a subset of all possible logic functions that can be generated using MUX-3 implementation. Therefore, the number of switches are less than $n \times m$ (where *n* is the number of rows and *m* is the number of columns in the switch matrix). This reduces the size of the configuration register considerably. In Section 4, we will evaluate the implementations mentioned above in terms of various circuit metrics.
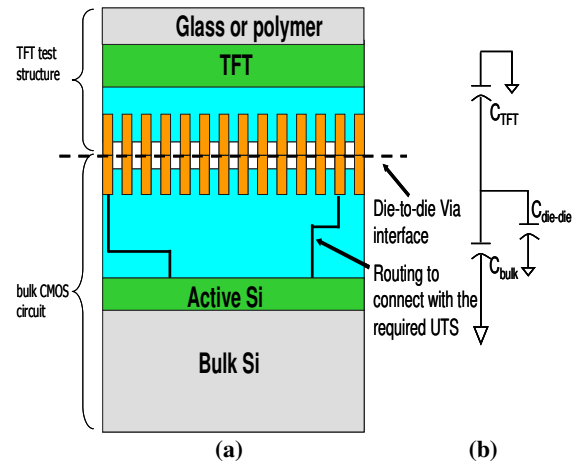
Fig. 4 shows the block diagram of the test architecture which constitutes of two blocks, namely (a) regular UTS (r-UTS), and (b) dedicated UTS (d-UTS). The *r-UTS* can be configured as *latch*, *xor*, *nand* and *nor* gates while the *d-UTS* can be configured as any test module e.g., MISR, LFSR, scan chain. The possible applications of r-UTS can be (a) collection of data from "hard-to-observe" test points for test/diagnosis during on- line/off-line test, (b) insertion of test data at "hard-to-control" test points, and (c) reduction of DFT overhead from the silicon die. The d-UTS, on the other hand, can be used to alleviate the BIST overhead from the silicon die. In Fig. 4, we have elucidated this concept by constructing a Built-In Logic Block Observer (BILBO) [2] with the TFT devices. *The length and characteristic polynomial of the BILBO can be configured through programmable wires* (using switch matrices as shown in Fig. 4). In this example, we have illustrated a 3-bit BILBO of characteristic polynomial $1 + x + x^2$. The dots in Fig. 4 represent connection between two crossing wires. The wires are programmed in such a way that a 3-BILBO (as illustrated by schematic in the inset of Fig. 4), is constructed. The major advantage of the BILBO is that it can be programmed as simple latch, scan chain, LFSR or MISR (by controlling inputs B1, B2, in Fig. 4) depending on the test requirements. Note that, one unit of the d-UTS may consist of many r-UTS's (Fig. 4). Although we demonstrated this example with two rows of d-UTS in Fig. 4, the actual implementation of the test architecture may contain several rows of d-UTS's which can be configured by the user to fully or partially construct the BIST. It can also be observed that d-UTS's are effective in testing the silicon die because they are fully configurable, however, they can be costly in terms of complexity and self-test. Therefore, only a minor area of the test architecture can be devoted to the d-UTS design.
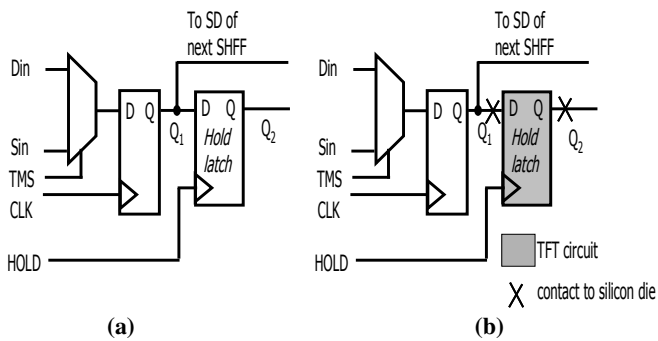
2.2 Testing of the test structure

Fig. 7 (a) Normal enhanced scan cell; and, (b) hybrid enhanced scan cell with TFT UTS integrated as hold latch

Fig. 8 Test point insertion. C_1-C_1 and C_n-C_n are die-to-die vias

To ensure the testability of the silicon die, the test structure should be *fault-free* itself. Since the test structure typically operates at slower clock, it can be tested by the external tester eliminating the need of built-in self-test (BIST). It can be noted that the r-UTS's of the proposed test structure are independent components. Therefore, they can be tested in parallel to reduce the test time. A possible test strategy can be as follows (Fig. 5 (a)): (a) shifting appropriate pattern in the configuration register serially to configure the r-UTS's to operate as a particular logic (say, *nand*, *nor*, *xor* and *inverter* etc); (b) shifting test patterns in the input registers so that a particular test vector is applied to the r-UTS's; and, (c) latching test response in the output register and shifting them out serially. These three steps can be repeated to exhaustively test all possible functions that can be generated by the r-UTS's. The serial output of a row of r-UTS's can either be sent to the tester for analysis or can be compacted by configuring the existing d-UTS as MISR. The MISR signature can be sent to the tester at the end of the test cycle. The integrity of d-UTS's can be tested by configuring them as scan chains, shifting test data and observing the serial output. Since the feedback path of d-UTS's is nothing but a row of r-UTS's, they can be tested similar to other r-UTS's.

2.3 3-D integration of the test structure

The TFT test structure can be integrated by face-to-face bonding using die-to-die via interconnects with the silicon wafer as shown in Fig. 6(a). The die-to-die vias are placed on the top of the metal stack of both dies and are heat bonded after alignment [10]. The die-to-die vias on TFT are connected to the inputs and outputs of the UTS by using local interconnects and vias. On the other hand, die-to-die vias on silicon can be connected to the *node of interest* through local wires and metal-to-metal vias. Fig. 6(b) shows the model of die-to-die vias.

2.4 Test and diagnosis by using the test structure

Testing of the silicon die basically requires appropriate configuration of the UTS's. The d-UTS's can be configured to build a portion of BIST circuitry to assist during testing (manufacturing test or periodic field test). The r-UTS's can be configured as required by the user to reduce the DFT overhead from silicon design and/or to
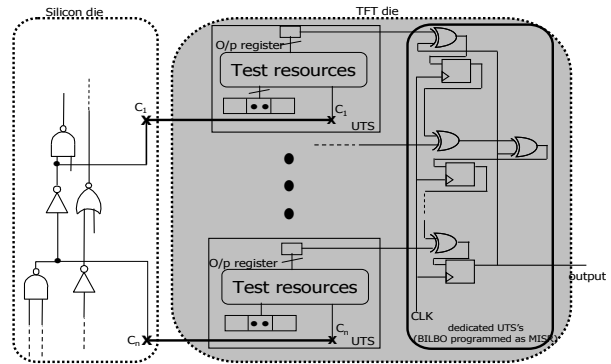
improve the controllability and observability. As discussed in Section 2.3, the inputs/outputs of the UTS's are connected to the appropriate nodes of silicon die through local interconnects and die-to-die vias. A potential application of the test structure is illustrated in Fig. 7. In this example, we have demonstrated how the r-UTS's can be used as hold latches for constructing the enhanced scan cell. Fig. 7(a) shows the normal enhanced scan cell while Fig. 7(b) shows the hybrid enhanced scan cell with TFT UTS configured as hold latch and integrated with the scan cell in bulk-Si. Node 'Q1' of the scan cell is connected to the UTS input and the output of UTS is connected to node 'Q2' in bulk-Si. The configuration register is programmed so that the UTS can be configured as latch. This eliminates the need of designing the hold latches in silicon; reducing the DFT overhead significantly while having the same flexibility in delay test. Note that, the latch should be bypassed in normal mode to reduce delay overhead (by using a switch).

Another example of utilization of the test structure would be to monitor *hard-to-observe* nodes during structural or functional test. An entire column of r-UTS can be reserved for application as observability points. All hard-to-observe nodes should be connected to these UTS's which can be configured to act as a *buffer*. The buffered outputs can either be scanned out to the external tester or can be compacted using the d-UTS's (by configuring them as MISR). We have illustrated a conceptual diagram in Fig. 8 where wire C_1-C_1 and C_n-C_n represent die-to-die vias corresponding to 1st and nth probe point. The test structures can also be utilized for on-line monitoring of certain nodes for diagnosis purpose. Since clock frequency of test structure is slower than the bulk-Si counterpart (due to slow TFT devices), the node outputs can be sampled occasionally. The sampled outputs can either be compacted using the TFT MISR or can be stored for further analysis/debug.

3. Optimized TFT for Proposed Application

In previous section, we elaborated the test architecture and test/diagnosis of the silicon die by using TFT based UTS. In this section, first we present a brief introduction of the conventional TFT devices. Then we
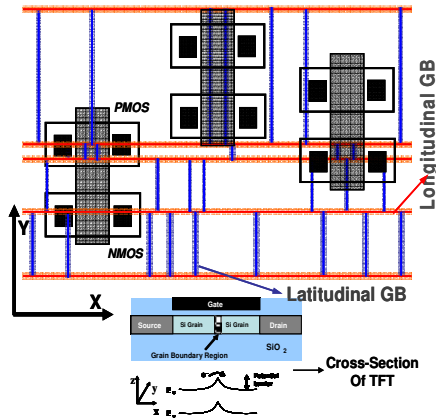
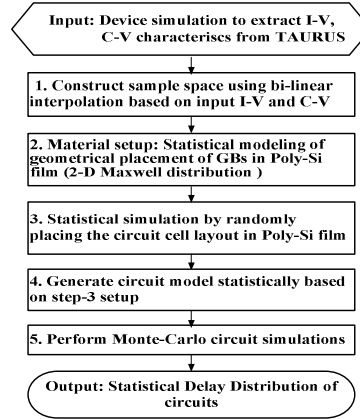**Fig. 9 Process variation due to placement of layouts in polycrystalline film.**

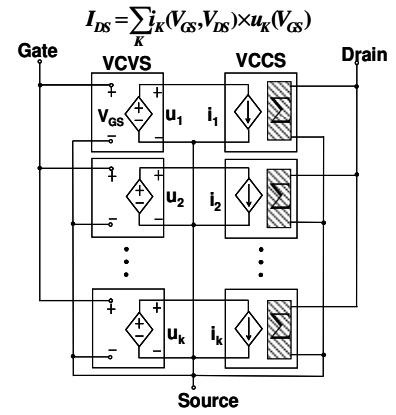**Fig. 10 Flowchart of statistical simulation of TFT circuits**

Fig.11 HSPICE compatible DC model of the TFT device

briefly discuss the optimization methodology [6] to design the TFTs for robustness, low power and high performance.

3.1 Basics of TFT

LTPS TFTs have been widely used in LCD as pixel-switching-elements. Compared to conventional bulk-Si device, the fabrication costs of LTPS TFT is very low. Moreover, it can be built on flexible substrates (glass, polymer etc) making them a promising candidate for 3-D integration as an add-on to the Silicon. Clearly, for such application, it is essential for LTPS TFTs to have CMOS-compatible performance, low power consumption and less variability. State-of-art TFT technology cannot provide the desired performance with low power consumption due to the extremely low current drivability. Poor device characteristics result in high supply voltages (10~20V) to meet the performance target. In contrast to bulk MOSFET, the polycrystalline Silicon channel material in LTPS TFT consists of a number of single crystal grains with highly defective regions in between, called grain boundaries (GBs) that limits the performance.

3.2 Device Optimization for Power and Performance

Conventionally, there are two possible methods to improve the performance of Poly-Si TFTs, (a) increase the grain size, and; (b) scale down the device size so that the number of GBs is reduced in the channel and I_{ON} is improved. However, obtaining very large grain size not only increases fabrication cost but also increase thermal budget. In keeping with the general trend of CMOS technology, the channel lengths of Poly-Si TFTs can also be scaled down to the submicron regime. However, aggressively scaling channel length induces short channel effect and increases device-to-device variation. It has been demonstrated in [6], that in scaled TFT device, T_{si} (body thickness) can be scaled instead of scaling T_{ox}, improving sub-threshold slope without inducing extra thermal budget and cost.

3.3. Process Variation Estimation and Compensation

Although the overall performance of the device can be significantly improved by device optimization as described above, the statistical variation induced by GBs can be a major concern. To estimate the GB induced variations and to evaluate its impact on designing low power and robust digital circuits, we present a statistical simulation technique. The GB induced variation mainly arises from the statistical GB distribution on a die and can be modeled using this Quasi-Manhattan Structure (QMS) as shown in Fig. 9. In QMS, the grain size distribution follows the Maxwell distribution [11] along X and Y directions. The key distribution parameter is the average grain size.

Fig. 10 illustrates the flowchart of our proposed statistical simulation methodology. In Step-1, device simulator Taurus [12] is used to design and optimize the device, from which I-V and C-V device characteristics are extracted. Then we construct a sample space (including all possible GB distributions) by using bi-linear interpolation based on the I-V and C-V characteristics obtained from device simulation (Step-2). This approximation is reasonable since we observed less than 3% difference between interpolated and simulated value using Taurus. In Step-3, the Maxwell distribution of GBs in 2-D poly-Si film is modeled. Note that, device simulation using Taurus is highly accurate but the simulation efficiency is very slow. Therefore, we develop Hspice-compatible behavioral model (Fig. 11). In this model, the current I_{DS} can be expressed as follows

$$I_{DS} = \sum_K i_K(V_{GS}, V_{DS}) \times u_K(V_{GS})$$

where i_k is the voltage-controlled current source and u_k is the interpolation coefficient which denotes the voltage-controlled voltage source (the index k selected by V_{GS}). In addition, for transient analysis, the gate-to-source (C_{GS}) and gate-to-drain capacitance (C_{GD}) is modeled by polynomial fittings.

After developing the Hspice model, we randomly place the circuit layout in the Poly-Si film (Fig.9) (Step-4) and Monte-Carlo circuit simulations are performed based on the Hspice model (Step-5). As indicated before, the
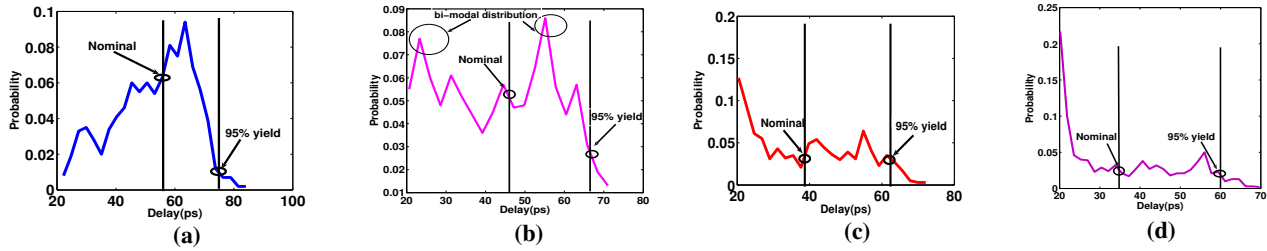
Fig 12 Statistical delay distribution of an inverter with average grain size of (a) 200nm, (b) 300nm, (c) 400nm, (d) 500nm

average grain size is not constant but process dependent. In this work, we have considered laser-annealed re-crystallization process where the average grain size is typically in the range of 200nm to 500nm. Fig. 12 (a)-(d) illustrates the delay distribution of an inverter (FO=1) obtained by following the simulation procedure shown in flowchart (Fig. 10) for different average grain sizes. It depicts that the discrete and randomly located GBs produce multi-modal delay distribution (contrary to uni-modal delay distribution in conventional bulk-Si) for average grain sizes above 300nm. We observed that the overall performance gets improved with the increase of average grain size but at the cost of large variation (σ/μ). To ensure a robust and stable functionality of TFT circuits, Multi Finger (MF) design technique proposed in [13] can be used.

For the circuit simulations in the following sections we use the device models for the nominal (i.e., *mean*) as well as 95% yield (which we refer as *worst* for simplicity) targets for various average grain sizes as shown by bars in Fig. 12.

4. Evaluation of Unit Test Structures

In Section 2, we proposed COMT, MUX-1 and MUX-3 as three possible implementations of UTS. In this section, we will evaluate them in terms of power, delay, variations, complexity etc when they are configured to realize a combinational or sequential logic functions. The proper choice of UTS can be made based on design target.

4.1 Simulation setup

For simulation of the unit test structures, we use the generated Hspice model of the TFT devices from Section 3. To observe the impact of technology, we have performed simulations for average grain sizes ranging from 200nm to 500nm. For each average grain size, we use both nominal and worst case devices to design the UTSs (i.e., COMT, MUX-1 or MUX-3). The same procedure is repeated to simulate all UTSs for different grain sizes using nominal and worst case device to estimate the variability induced by GBs. A supply voltage of 1.5V is used for simulations.

4.2 Combinational logic

The propagation delays of the three primitive gates (with fan-out=1) realized using the COMT, MUX-1 and MUX-3 UTSs are shown in Fig. 13. It can be noted that the worst case corners always show larger delay than that of the

nominal case for a particular grain size. This is mainly due to the considerably degraded current drive ability (i.e., less I_{ON}) induced by GB variation in terms of number, location and orientations in worst case devices. However, the delays of all three gates reduce with increasing grain sizes. In addition, Fig. 14 indicates that enlarging the grain size has a minor impact on the power consumption (which is mainly dominated by the dynamic power). Therefore, the performance loss in the UTS circuits due to GBs can be compensated by enlarging the grain size (i.e., by using better fabrication process) without increasing the power dissipation.

Fig 13 also illustrates that the MUX-1 UTS results in smallest delay, followed by COMT and MUX-3 UTS. The increased delay of MUX-3 arises from more number of stages. Furthermore, MUX-3 suffers from worst process variation which can be observed from the spread of delay (i.e., the difference between mean and worst case delays) as shown in Fig. 13. Also, the power consumption of MUX-3 structure is larger than the other two structures for *nand* and *nor* gates due to the presence of larger numbers of transistors and hence increased switching energy. For instance, MUX-3 uses 12 transistors while COMT uses 4-10 transistors and MUX-1 uses only 4 transistors. Therefore, MUX-1 has the smallest power consumption, delay, area and process variation among all these three UTS configurations.

4.3 Sequential logic

Besides combinational logic, we also simulated the COMT, MUX-1 and MUX-3 to realize sequential logic (i.e., latch). As depicted in Fig. 15(a), MUX-based latches consume less power compared to COMT (contrary to the primitive gates). This is mainly due to reduced short circuit power and switching energy in the transmission gate based multiplexers used in MUX-1 and MUX-3 UTS. Similar to the combinational logic discussed before, the power remains fairly constant with increasing average grain sizes.

We have also plotted the propagation delay of the latch for both nominal and worst case corners (Fig. 15(b)). The results show that COMT-based latch always have larger delay and variations than MUX-1 and MUX-3 due to its increased number of stages.

From above discussions we can conclude that, (a) MUX-1 is the best choice in terms of area, delay, power and
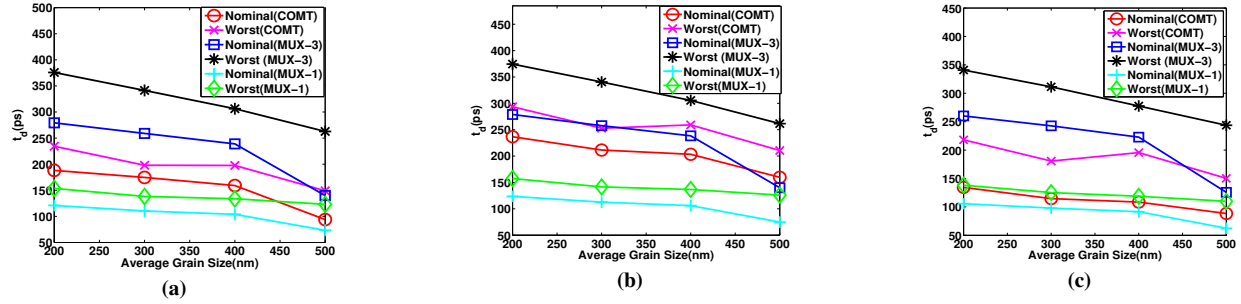
Fig 13 Delay vs average grain size of COMT- and MUX-based (a) NAND gate, (b) NOR gate, and (c) XOR gate for nominal and worst case
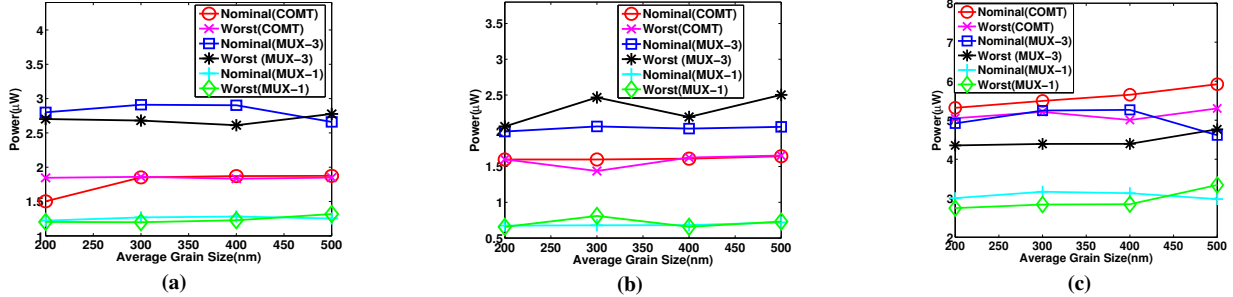

Fig 14 Power vs average grain size of COMT- and MUX-based (a) NAND gate, (b) NOR gate, and (c) XOR gate for nominal and worst case

variations, (b) COMT is simple and only requires a small number of configuration registers and can be a good choice in terms of power. However, the area overhead and variations can be a concern if complicated functionalities are required from a single UTS, (c) MUX-3 is more generic and flexible. It can also provide more numbers of functionalities with acceptable power, performance penalty and variations. However, it requires large size of switch matrix and configuration register.

5. Potential Applications in Test

In this section, we discuss some potential applications of the proposed test architecture in both off-line and on-line test for assisting BIST and reducing the DFT overhead from the silicon die.

5.1 Application in Off-line Test

5.1.1 Enhanced Scan cell

In this subsection, we demonstrate a potential application where the UTS can be configured as the hold latch of enhanced scan cell to reduce the DFT overhead from the silicon die using hybrid 3-D integration. As discussed in Section 2.4 (Fig.7), the inputs/outputs of the hold latch are connected to the nodes of silicon die through local interconnects and die-to-die via. For Hspice simulation, we have modeled the die-to-die via and local interconnects as lumped capacitors as shown in Fig 6(b). The capacitance of the local interconnects have been extracted for 130nm technology node [14] using Virtuoso [15]. The die-to-die via capacitance is assumed to be 0.82pF [16]. Based on this model, we compare the hybrid enhanced scan cell (i.e., the bulk-Si scan flip-flop integrated with the TFT hold latch) with the conventional bulk-Si enhanced scan cell in terms of delay and power (with supply voltage of 1.5V). A supply voltage of 1.5V is used for the simulations. We
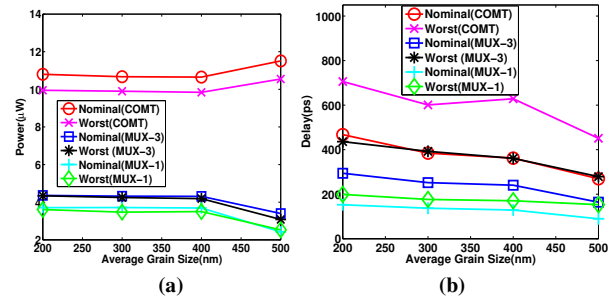

Fig 15 COMT- and MUX-based latch: (a) power and, (b) delay

observed that the delay of hybrid cell is slightly higher (~10%) than conventional cell due to the larger capacitance induced by the die-to-die vias. Note that the hold latch can be bypassed from scan flip-flop during the normal mode of operation. Therefore, in normal mode the enhance scan cells suffer from performance penalty only due to the die-to-die via capacitance. However, this performance penalty is negligible compared to considerable area savings (~15-20% due to absence of hold latch from the silicon die [17]). Furthermore, the power consumption of the hybrid cell is less than the conventional cell due to reduced gate capacitance (T_{ox}~10nm for TFT and T_{ox}~2nm for bulk) and hence, reduced power consumption of the TFT hold latch. We have plotted the power of all three UTS implementations to evaluate the hybrid enhance scan cell. Fig. 16 shows that MUX-1 and MUX-3 consumes less power than COMT due to the reduced switching energy, (same trend as the sequential logic in Section 4.3). Note that, more power savings can be achieved by operating the UTS's at further reduced supply (at the cost of performance penalty).
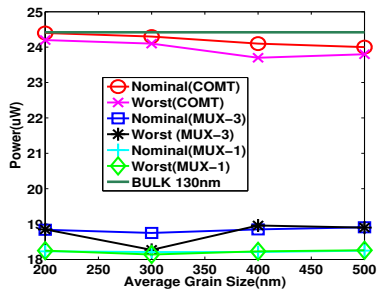
5.1.2 Built-in Logic Block Observer

Fig 16 Simulation results of enhanced scan cell

We have simulated a 3-bit BILBO to operate in all four possible modes as discussed in Section 2.1. The simulation results for the scan chain and LFSR modes are shown in Fig. 17(a) and (b), respectively. The figure demonstrates the functional correctness of the BILBO that is designed entirely using TFT devices. In Fig. 18(a), we report the worst case propagation delay of the BILBO. The results show similar trend as discussed in the previous subsections. That is, the propagation delay reduces with enlarging grain sizes and propagation delay of worst case corner is always larger than that of the nominal corner. Fig 18 (b) shows the maximum operating frequency of the BILBO. For comparison, we have also plotted the delay and maximum frequency of BILBO designed using bulk-Si. It can be concluded from this figure that by proper control of grain size and process variation, it may be possible to obtain a BIST circuit that can operate at a reasonable speed (i.e., 0.8-4.3 GHz compared to 8.4 GHz in conventional bulk). Note that, our frequency estimations are slightly optimistic because we have not considered the interconnect parasitics. We also show the power consumption of BILBO implementation using COMT, MUX-1 and MUX-3 for the grain size of 200nm (Fig 18(c)). It has been shown that the MUX-1 implementation consumes lowest power and operates with highest speed. From these results, it can be concluded that the test circuits designed using TFTs can operate at reasonable speed with low power consumptions by judicious choice of the unit test structure. Since the test circuits are re-configurable, it can be useful in testing a family of processors or DSPs in high volume manufacturing process.

5.2 Other Possible Applications (On-line Test/ Verification/Repair)

(a)　　　　　　　(b)
Fig. 17 Hspice simulation of 3-bit BILBO in (a) scan chain mode; (b) LFSR mode of operation. Q_1, Q_2 and Q_3 are latch outputs

Besides off-line test, the proposed test architecture can be an effective tool for the system designers and software developers for on-line test and verification of the silicon die. The test structure can sample the outputs of the test points at low speed which can either be (a) compacted by configuring d-UTS to act as MISR and the signature can be used later for verification; (b) provided to the software layer for checking and verification. Since the UTSs can be configured to function as combinational and sequential logic, they also have the potential to be used for repair purposes (at the cost of speed). However, it would require extra effort for proper connection of input/outputs of defect-prone logic blocks to the UTS through die-to-die vias. The UTSs can be configured by the software during run-time if repair is required.

6. Discussion on Practical Challenges

In this paper, we proposed a TFT based test paradigm and provided different examples of DFT circuits (e.g., BILBO, enhanced scan cell) to prove the feasibility of the proposed test methodology. In this section, we address some of the challenges and issues as discussed below:

- *Scalability/speed and voltage mismatch between CMOS and TFT*: Conventional CMOS has been scaled down over generations making the transistors smaller, faster and allowing them to operate at lower supply. On the contrary, for traditional LCD applications, TFTs with large device dimension (~10 µm) usually operates
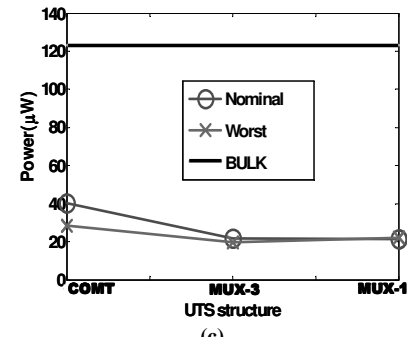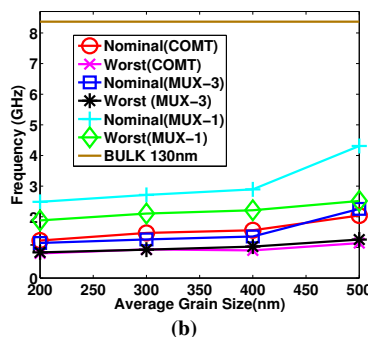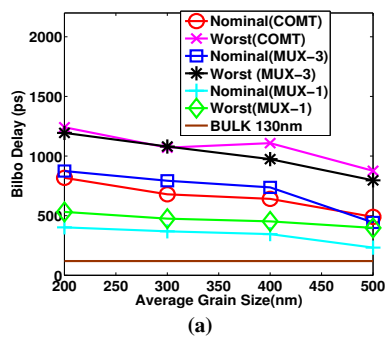

(a)　　　　　　　(b)　　　　　　　(c)
Fig 18 Simulation results of BILBO (a) propagation delay, (b) possible frequency of operation and (c) power for different UTS (i.e., COMT, MUX-1 and MUX-3) when the average grain size is 200nm
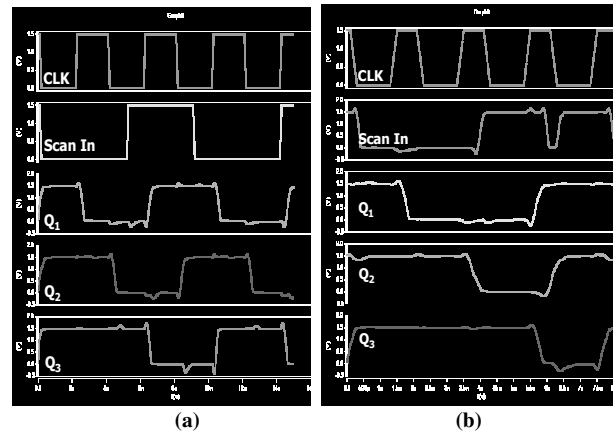
at high supply voltage (~20V) in order to achieve sufficient current drive ability. However in the proposed test paradigm, the supply voltages of the TFTs and the bulk-CMOS should be compatible in order to avoid level converters when applying 3-D integration. Proper device optimization and circuit level solution can be utilized to scale down the power supply of TFTs while maintaining sufficient current drivability and less variability. However, for proper integration of future scaled silicon technologies (with further reduced supply) the proposed TFT should be operated in sub-threshold regime (because $V_{T\_bulk}<V_{DD}<=V_{T\_TFT}$). Therefore, it is imperative to devise new device-level and circuit-level design techniques to allow the TFTs to operate in sub-threshold region.

- *3-D integration*: Proper operation of the proposed test architecture requires 3-D integration. In this technology, some of the issues need to be further explored and resolved. For example, accurate alignment of the die-to-die via, bonding of the vias, heat dissipation through the 3-D structure etc are the challenges that should be addressed. Moreover, the assembling process for 3-D hybrid integration may introduce new defects. The yield loss due to test circuit assembly can be crucial.

- *At-speed test*: The proposed test structures operate in the range of hundreds of MHz to few GHz. Therefore, they may not be applicable for at-speed test of very high performance circuits. However, the test structures can be efficiently used for at-speed test of medium performance circuits– a large number of DSP applications require medium performance.

- *Design efforts for routing*: For gaining more benefits from the proposed test structures, the designers should route the *nodes of interest* to the die-to-die vias (as discussed in Section 2.4). The task has to be done in a close cooperation with DFT engineers.

- *Loading of internal nodes due to vias*: Routing of wires required to connect the *nodes of interest* to die-to-die vias induces extra loading on the internal nodes. Therefore, such nodes should be chosen carefully so that the critical paths are not affected.

7. Summary and Conclusions

We proposed a generic and reconfigurable test paradigm using low-cost LTPS TFTs for application in off-line as well as on-line test of complex VLSI systems. The test structure designed with TFTs, can be fabricated independently on a flexible substrate and integrated with the silicon die using 3-D die-to-die vias. Traditionally, TFTs operate at high supply voltages (~10-20V) due to presence of highly defective grain boundaries in the channel. In this work, we designed the test structure using optimized TFT devices that operate at CMOS compatible lower supply voltage for low-power consumption and medium performance. We also developed an efficient simulation environment for circuit implementation of our proposed test structure using TFT. We proposed three different circuit implementations of the unit test structures and evaluated them in terms of area, power, variations and performance. Further, we presented test and diagnosis scheme of the silicon die using the proposed unit test structures. Simulation results show that the test structures can operate at reasonable speed while achieving low power dissipation. The proposed test architecture provides an additional tool to the system designers and software developers for improving the test coverage, DFT overhead, on-line test/verification and correction as well as for reducing the time-to-market. We believe that the proposed test paradigm can provide maximum benefit if used judiciously by the hardware and software for on-line test/monitoring/verification purposes.

8. Acknowledgements

9. References

[1] JEDEC Solid State Technology Association, Failure mechanisms and models for semiconductor devices, JEDEC Publication JEP122-B, 2003.

[2] M. Breuer, Digital system testing and testable design, IEEE Press, 1995.

[3] S. Mukhopadhyay, "Design of reliable and self-repairing SRAM in nano-scale technologies using leakage and delay monitoring", *ITC*, 2005.

[4] S. Ghosh, A novel delay fault testing methodology using low overhead built-in delay sensor, *TCAD*, 2006.

[5] G. Swan et al., Crosscheck- a practical solution for ASIC testability, *ITC*, 1989.

[6] J. Li, "Exploring low temperature poly-silicon for low cost and low-power sub-micron digital operation," *DRC*, 2006.

[7] M. O. Thompson, "Laser processing of Si-TFT's on plastic: technology and lessons from FlexICs", lecture given at Cornell University, April, 2006.

[8] J. M. Rabaey, "Digital integrated circuits", Prentice Hall, 1996.

[9] L. Lavagno, Timed shannon circuits: a power-efficient design style and synthesis tool, *DAC*, 1995.

[10] B. Black, "3-D processing technology and its impact on iA32 microprocessor," *ICCD, 2004*.

[11] S. J. Cheng, "A statistical model to predict the performance variation of polysilicon TFTs formed by grain-enhancement technology", *TED*, 2004.

[12] Taurus Device Simulator, Synopsys Inc.

[13] J. Li, "Novel variation-aware circuit design of scaled LTPS TFT for ultra low power, low-cost applications," *ICICDT*, 2007.

[14] BPTM 130nm: Berkeley predictive technology model.

[15] Virtuoso design manual, Cadence Inc.

[16] W. R. Davis, "Demystifying 3D ICs: the pros and cons of going vertical", *D&T*, 2005.

[17] S. Bhunia, "A novel low-overhead delay testing technique for arbitrary two-pattern test application", *DATE*, 2005.