Challenges And Solutions in Post-Manufacturing Testing of System-On-Chip (SOC) Devices

[1]Hameed Ul Hassan Mohammed, [2]Marcus Rodriguez

Abstract: The testing of System-on-Chip (SoC) devices after-manufacture is extremely pertinent to quality assurance of existence, credibility, and yield of sophisticated semiconductor manufacturing. Due to this ongoing increase in SoC complexity (high core counts, embedded IPs, high speed interfaces) the challenges of high fault coverage, fast test execution and low power effects have become more pronounced. This study demonstrates an experimental and theoretical study of the testing methods that are required to cater to these changing needs. The strategy that was formulated entailed the deterministic test vectors, scan compression methodology, at-speed Built-In Self-Test (BIST), and intelligent data analysis technique in order to analyze limitations posed by the fault detection schemes, the test time, and the power consumption. The findings pointed to fault coverage gaining proportional improvement as test vectors increased but with diminishing returns above a certain point and the eventual necessity of optimized sets of vectors. Scan compression significantly decreased test time, but it resulted in an increased power consumption showing that there is a trade-off between the effects of scan compression and scan compression power consumption which has to be traded off stringently in the design-for-test strategies. The paper has also provided variation in the extent of fault-type detection in that delay faults are more elusive than stuck-at or bridging fault detection and consequently, illustrates the relevance of hybrid testing methods. Moreover, advanced analytics of data allowed recognizing of marginal devices and forecast of failures and converted test results into meaningful process enhancement and reliability prediction. These results endorse a change toward adaptive, data-driven ways of testing that unite structural economy and predictive power. The framework proposed will provide intelligent and expandable solution to testing requirement of modern SoC product with the improved product quality and high manufacturing yield.

Keywords: System-on-Chip (SoC), Post-Manufacturing Testing, Fault Coverage, Scan Compression, Built-In Self-Test (BIST)

1. INTRODUCTION

The high-speed development of all semiconductor technologies made possible the era when System-on-Chip (SoC) devices became the backbone of the modern electronics, in smart phones and tablets, automotive systems, industrial controls, etc. Such advanced chips combine several features which include processing units, the memory blocks, the input/output terminals and even the analogs all on a single silicon substrate. The more complex they are, the more they put pressure on post-manufacturing testing that guarantees the fulfilment of functional and performance requirements of each device prior to a deployment. The repercussions of those failures are high: Even small bugs can cause system crashes and/or recalls or endangered safety in case of mission-critical applications.

[1]*Senior Test and AI Optimization Engineer, Mythic AI, Austin, Texas, USA*
hameedul040@gmail.com

[2]*Princeton Institute of Computational Science and Engineering (PICSciE), Princeton University, NJ, USA, rodriguezm7890@yahoo.com*

Post manufacturing testing is not just a normal test of quality assessment-it is an important stage between fabrication and getting an item into the field by providing a transitional stage between manufacture and actual use. As SoCs grow becoming denser and increasingly feature-laden, testing has to change to test millions, or even billions of transistors all working in concert. The use of traditional testing techniques, which were adequate to test less complex chips, may not be effective to find faults in the current heterogeneity SoCs. Moreover, the limitation on cost and time increases the pressure to make the process of testing optimal without compromising the accuracy. The present paper is an exploration into the major issues haunting the SoC post-manufacturing testing process and outlines the state-of-the-art approaches suggested to resolve them, including progressive design-for-test (DfT) methods as well as artificial intelligence-based diagnostics. Some of the challenges and the major strategies are listen in figure 1 below.

Figure 1: Challenges and Key Strategies to master verification challenges in IC SoC Design

1.1 COMPLEXITY OF SOC ARCHITECTURES

The intricacy of newer SoC designs can be touted as one of the most challenging tasks in post-manufacturing testing. Such chips combine many cores, sub systems, and interfaces with a variety of functionality and timing requirements. The architectural variety means that a one-size-fits-all approach to testing is hard to build. Multiple clock domains, asynchronous data paths, and mixed-signal components need to be considered by the test engineers because they all introduce complexity in determining the areas of possible defects. The sheer complexity of where and how many faults can be located causes the effort used to obtain sufficient fault coverage to exponentially compound. In addition, since SoCs commonly contain 3 rd party intellectual property (IP) blocks, the absence of internal visibility is also likely to increase test planning and execution.

The other important problem can be seen in broad unavailability of internal nodes in very integrated SoCs. As compared to the previous generation of chips, it is no longer possible to probe individual circuits directly as the older SoCs would allow; this is because of space limitations and abruption of chip functionality. Internally, therefore, strategies should mainly depend on built-in test mechanisms and scan chain to measure behavior indirectly. Such strategies come with an overhead to the design, however they can have failures of dynamic or intermittent faults which may not be fully captured. In addition, only

the logic parts of test coverage are frequently tested and analog and RF parts may be missed unless special test methods are applied.

Complexity is also associated with ensuring the interaction between the integrated modules are checked. A number of faults are not localized in some block but the problem is manifested by faulty interconnects or incompatible protocols among components. This requires system-level tests that closely match real-world operating conditions-an activity which is time consuming and computationally demanding. Moreover, the growing popularity of software-hardware co-design of the modern SoCs implies the necessity that functional testing should take into account the interaction between the embedded firmware and the hardware logic as well. This interaction makes it difficult to distinguish or diagnose faults especially when it involves software masking or time anomalies.

1.2 TEST COST AND TIME CONSTRAINTS

The increasing cost and time of full validation is one of the burning issues in SoC testing. The complexity of the devices increases, the volume of test data similarly increases, the test times lengthen and the quantity of costly automatic test equipment (ATE) is utilized. The time and the cost on the test floor are directly translated to the manufacturing efficacies and costs and therefore, in the case of the testing, a balance between the depth of testing and economic feasibility must be gained. This is especially urgent where test-time is limited by volume limitations as

in high production environments, where minor increments in a test time can mean large differences in costs.

This problem is also compounded by the shortcomings of conventional ATE. The speed and complexity of SoCs can be challenging because these machines need to be as fast and detailed as modern SoCs and in many cases requires expensive updates or customization. Fast interface, complex memory standards, and pipelining processors require more sophisticated test gear with the ability to analyse and generate huge volumes of data simultaneously. The high cost of such state of the art ATE is not always affordable to many semiconductor companies, especially since there is a short product life cycle, and there is uncertainty in the realization of returns on investment. Therefore, due to this, there is always a trade-off in between test accuracy, time, and resource distribution.

In an attempt to reduce these limitations, designs are becoming more compliant with design-for-test (DfT) techniques that incorporate test logic in silicon. Built-in self-test (BIST), boundary scan and test compression are some of the techniques that can minimize the use of external tester and can increase test execution speed. Nonetheless, these have costs associated with themselves too- namely the extra chip area, power and design complexity requirements. Introducing and confirming these functions needs the planning of their introduction in the design stage, and any neglect will wipe out the desired advantages. Thus, although DfT is an effective tool, its combination should be well-optimized in order to be as efficient in terms of cost and time as possible.

1.3 YIELD LOSS AND DEFECT ESCAPE

In the manufacturing pipeline of SoC the lost yield and defect escape is a principal issue in case where subtle or latent defects are not detected during tests. There is the yield loss of good chips that have been incorrectly identified as bad chips when actually they should pass on the chance of adopting overly conservative test thresholds or noise induced-failures. Defect escape on the other hand occurs where flawed chips almost defect the testing and thus malfunctions in the field. Both of these are expensive outcomes; yield loss means wasted silicon, and lower margin; and defect escape can cause brand risk, and expensive recall or legal action, particularly in such industries as automotive, medical or aerospace.

The detection of defects is becoming more cumbersome as the transistors get smaller and lower power. New modes of failure occur with this miniaturization including leakage currents, soft errors and parametric variability and might not reveal themselves in normal test conditions. These subtle issues cannot be captured by traditional structural tests that concentrates on stuck-at and transition faults. Consequently, there is the need on changing test tactics to new platforms such as dynamic tests, stress tests and burn-in tests that have the potential of detecting marginal defects. The tests, however, are taking a longer period of execution and can decrease throughput resulting in a dilemma of test coverage versus efficiency.

Machine learning and statistical analysis is proving to be a promising method to getting this trade-off. With large samples of test logs and production history, manufacturers are able to figure out the trends and make educated guesses at what defect clusters may exist. Yield can be increased by adaptation testing methods which at run time varies test parameters depending on how a chip behaves, it may perform this adaptation without reducing quality. Moreover, detection techniques of outliers can also be used to pick marginal devices that are still not adhering to the standard even though they might technically pass all the conventional tests. Such advanced techniques are transforming the manner in which yield loss and defect escape is addressed in a SoC post-manufacturing testing process.

1.4 POWER AND PERFORMANCE TESTING

An SoC has complex and essential manufacturing power and performance characteristics evaluation that is performed after the manufacturing process. Such measures intertwine with the usability of the chip, in power tolerant applications such as in the case of mobile devices or embedded systems. Although functional testing takes care of correctness, power/performance testing ascertains the effectiveness and reliability of the chip in working scenario. Sadly, this degree of insight is hard to do well, because it necessitates precise observation of inner voltages, currents, and timing characteristics under working conditions, values that cannot be conveniently observed on external test ports.

Additionally, many modern SoCs have the capability to use dynamic voltage and frequency scaling (DVFS), multi-core operation and complex power domains that may be manipulated separately. The evaluation of these characteristics involves advanced power-conscious testing that takes into consideration different states of operation and transitions. It is not enough only to check proper work of separate blocks; the engineers should analyze the way power is used and distributed on the chip under various loads. The timing relationships between the power management domains and their logic can cause instability or even permanent damage as a result of any fault.

On-chip monitors and sensors are being embedded into designs in an attempt to handle these design issues. The latter components make it possible to analyze power and thermal in real-time so that test engineers can measure the actual metrics during tests. There is also performance profiling tools where they make use of the realistic uses and workload patterns simulating the end user behavior. An integrated effort between these tools and complex modeling and emulation platforms enables performance bottlenecks and power inefficiencies to be identified at an early phase of the design process. The methods assist in ensuring that the chip performs not only correctly but is also able to meet rigorous power and performance requirements requested by current applications.

1.5 TEST DATA MANAGEMENT AND ANALYTICS

Due to the effect of increasing the testing complexity of SoC, a greater amount of data is created in the post-manufacturing process. A competent approach to the management of this data is extremely challenging, especially in cases where millions of chips generate gigabytes worth of test records. There is an infrastructure and personnel strain to store, process, analyze and retrieve this data in real-time. Failure to manage data in an efficient way may result in missed defects, lost efficiency and inability to trace root causes when performing failure analysis. Thus, test data management systems essential to the realization of best quality of test and manufacturing yield should prove to be robust.

Another of the main difficulties is the correlation of data between several test runs and production batches. In many cases, comparison of test data with

those of the design verification process, wafer probe results, and end test is required in order to identify trends or systematic problems. Nonetheless, this process can be hampered by inconsistent data formats, different variants of naming, and by a siloed database. To combine disparate sources of data into a common analytics and planning environment, there is a need to coordinate both technically and organizationally. Such integration can be lost without it, and the right not used to improve it.

The solution to this problem can be achieved using advanced analytics, operationalized by machine learning and big data techniques, which will provide transformative solutions to this problem. They can be used to analyze enormous amounts of data to draw out trends, spot anomalies and provide suggestions of what needs to be done. An example is predictive maintenance where failure on an equipment is anticipated by using statistical analysis of past tests to plan and carry out maintenance of equipment before it happens, thus reducing cost and time of machinery downtime. Likewise, the root cause analysis algorithms are able to identify sources of faults and advise on process enhancement. Since SoCs are growing ever more complex, intelligent analytics in the post-manufacturing testing arena will become more centralised, unlocking raw data to become strategic assets that can support efficiency and quality.

2. REVIEW OF WORKS

Design and implementation complexities coupled with heterogeneous core integration and the requirement to deliver the zero-defect in high-reliability applications have driven the development of post-manufacturing testing of System-on-Chip (SoC) devices in the past decades. The topic is broad and multidisciplinary and consists of structural testing methodologies, design-for-testability (DfT), built-in self-test (BIST) techniques, high-bandwidth memory (HBM) testing, and scalable access mechanisms to embedded cores. Synonymous with pioneers and researchers, theories and practical frameworks have been established to guide contemporary test strategy although they remain as thematic to date. This segment provides thematic overview of old and new seminal work which discusses the issues and developments of post manufacturing SoC testing.

2.1 EMBEDDED CORE TESTING STRATEGIES

The ability to integrate several embedded cores into the SoC is one of the first known challenges in testing SoCs as each may be supplied by each vendor with a different design architecture. Zorian (1997) proposed core-based testing paradigm, wherein there should be a uniform approach of embedded core testing to accommodate the complexity of growing SoCs. His efforts formed the basis of the IEEE P 1500 standard which was certainly a landmark attempt to ensure consistent access to test mechanisms to the embedded cores. Its standard was subsequently codified by the IEEE (2005) in IEEE Std 1500-2005 which provides a scalable and modular test wrapper architecture. Another conclusion drawn by Gupta and Zorian (1997) was that, core reuse should be closely related to structured test reuse techniques by far standards in times when SoC designs started to grow exponentially in terms of size and functionality.

This model was extended further in later works, to include reusability and hierarchical integration of tests. Marinissen et al. (1998) also introduced a scalable yet structured mechanism of accessing tests in the embedded cores, which capable of supplementing the IEEE P1500 project. Their contribution demonstrated the necessity of test access mechanisms (TAMs) to fill the gap between high and low-level tests of a system as well as core interfaces. An additional development that Varma and Bhatia (1998) have made was a test reuse methodology where already tested cores can be easily integrated and tested in a system environment without doubling the work of the test development. Modern DfT tools and practices have emerged out of these primitive frameworks to the extent that testability of large-scale SoCs scales in a similar manner to design volume.

2.2 BUILT-IN SELF-TEST (BIST) TECHNIQUES

BIST has played a very important role in resolving the problems of accessibility and scalability of SoC testing. Zorian (1993) has developed a model of the distributed BIST control that looks at the issue of coordination in complex VLSI systems. His architecture embodied decentralized control of BIST controllers, such that a block exposed to control bursts could control its own test sequence as long as it followed the coordination rules that each block

would follow at the system-level. Agrawal et al. (1994) also added strength on the feasibility of BIST in digital integrated circuits especially due to the reduction in the reliance on costly automated test equipment (ATE). They showed in their work how self-contained test structures are not only capable of enhancing fault detection but also contribute to performing at-speed detection of internal paths, normally impossible to reach.

In Sunter and Nagi (1997), BIST methods were used to analog and mixed-signal blocks where a polynomial-fitting algorithm was proposed to test DAC and ADCs. This contribution stretched BIST to non-digital areas, addressing one of long outstanding SoC testing issues, mixed-signal test coverage. With a growing number of mixed-signal components being integrated into digital SoCs, external instrumentation was slowly ceasing to become viable in most applications as a result of constraints on precious pins and signal integrity problems. The functionality testing, linearity testing, and resolution of analog blocks has therefore become critical with the in-situ test capability of performing BIST tests. All these inventions have ensured that BIST has become an essential feature in the SoC design and test processes.

2.3 TESTING HIGH-BANDWIDTH MEMORY (HBM) IN 2.5D/3D SOCS

The testing of high bandwidth memory (HBM) has brought with it new complexities with the development of 2.5D and 3D SoCs. Jun et al. (2016) discussed the test issues in HBM, i.e., the high-density interconnects as well as the sensitivity of the through-silicon vias (TSVs). They found that the conventional memory testing methodologies do not take into consideration TSV induced defects and inter die communication errors. This requires the use of special test strategies, e.g., dynamic stress testing and temperature-conscious fault pinpointing, to be sure of their trustworthy performance with the stacked die layouts. Even though it was reported that the adjustment of channel length in the stacked HBM was critical but could still be ignored, Jun (2015) has also stressed the need to consider power integrity and thermal issues in the stacked HBM systems since they have direct effect on the signal integrity and long-term reliability.

In compliment to this, JEDEC (2015) brought out JESD235, the HBM specification defining electrical, mechanical, and test requirements of a

device that is compliant. The standard has aided uniform test strategy of the semiconductor vendors. Additionally, Samsung and SK Hynix have datasheets that specify major reliability and performance specifications on 8Gb B-die and M-die stacks, respectively, which provided information about the operating voltages, defect screening parameters, and testing conditions. Although they were generally written to be implemented, these documents would act as reference material as far as the how test strategies would have to be changed in order to support higher memory interface speeds and multi-die interaction in individual and post manufacturing tests.

2.4 3D INTEGRATION AND WAFER-LEVEL TESTING

The Scale SoC performance and density greatly important Both Wafer-level and 3D integration technology have been critical in scaling the performance and density of the SoC, but both present great test challenges. Koester et al. (2008) wrote about the wafer-level 3D integration technologies and risked the post-manufacturing test access and fault diagnosis due to vertical stacking and TSVs. The heterogeneous die integration of the wafer level level requires test strategies that are capable of identifying the alignment error, interconnect errors, and thermal related failures prior to the last packaging. This kind of preliminary testing is important in averting compounded failures at a later stage of production. In addition, the likelihood of latent defects is also enhanced by thermal stress and mechanical strain that is introduced during bonding.

Chen and Tan (2011) explored integration strategies of 3D IC and introduced the prospective facilitating technologies to provide testability in vertical SoC. Their understanding of pre-bond and post-bond test methodology highlights on the importance of having visibility of test centres visible between die layers and promotes test elevators and micro-bump probing. Lau (2012) also gave his input of reviewing the recent developments in the 3D integration proposing novel trends in nanotechnology which may benefit test access. These consist of MEMS-based test probes as well as nano scale sensors in interconnect layers. In sum, this literature supports the fact that the traditional planar testing methods need to advance into spatially conscious, many-

layered methods of effective validation of SoCs in the 3D markets.

2.5 ANALOG TESTING AND PARAMETRIC FAULT DETECTION

The problem of testing the analog blocks inside the SoCs that are digital has not been diminishing so far and the sensitivity of these blocks to the process variations and noise can be quite high. Transient current testing was applied to the deep sub-micron CMOS circuits as a means of improving the defect detection, which was brought up by Sachdev, Janssen, and Zieren (1998). They applied a methodology that aimed at locating faults by detecting transient behaviour deviation which can usually be overlooked using a static or DC test. This method is especially good at revealing low level parametric faults in analog blocks that impair timing, drain during switching or gain. Using transient current signatures, test engineers are able to detect faults sooner and with greater accuracy.

Through the research on a generic IDDQ measure, Wallquist, Righter, and Hawkins (1993) established a foundation combining all three of the features of low leakage fault detection: a general-purpose measurement circuit, later used in all low-leakage FIDs (Wallquist, 1994; Wallquist and Hawkins, 1997; Wallquist and Wynn, 1996). IDDQ testing is a test to analyze quiescent current at CMOS circuits, and is particularly useful in locating bridging faults and failures in gate oxides. Despite leakage currents becoming progressively more challenging to apply in technologies developed to a nano scale, IDDQ is still meaningful when nativized to the present-day low power design. These contributions underline the significance of integrating analog-aware testing techniques besides digital ones because ignoring the analog realm may be accompanied with the system-level faults that cannot be identified even after confirming success by digital verification.

3. METHODOLOGY

The experiment conducted in this study will assess how efficient it is to conduct the post-manufacturing test strategies on contemporary System-on-Chips (SoC). A full-fledged test system was integrated including Device Under Test (DUT), test controller, data analyzer and centralized result database. The DUT has an embedded test structure of synthesized SoC, scan chains, Built-In Self-Test (BIST) model

and Design-for-Test (DfT) cells. These results were obtained through targeted testing done in enclosed laboratory conditions with the use of commercially available Automated Test Equipment (ATE) directly connected to our custom test software. They were aimed at measuring coverages of a fault, test time and power testing with multifaceted configurations and workloads.

The initial stage of the methodology consisted of the fault simulation-based injection of permanent as well as transient faults into the DUT. These errors comprised stuck-at faults, bridging faults and delay faults both within logic gates and interconnection. Fault detection efficiency was computed with the aid of a combination of deterministic and pseudo-random test designs using fault number identified. The fault coverage (FC) was defined as:

$$FC = (N_d/N_t) \times 100$$

with N_d the number of faults found and N_t the sum total of inserted faults. This measure has been taken with different test lengths and scan chain arrangements as to comprehend the impact of the test depth on total reliability.

In the second phase, attention was paid to a test time and test power analysis. The effect that compression algorithms and partitioning of the scan chain have on the overall test time is measured. Test time (T_t) is described as:

$$T_t = (L \times C)/f$$

where L is the scan chain length; C is the test cycles and f is the clock frequency. Experimental data indicated that it is indeed possible to decrease T_t by more than 40 percent with scan compression even at the expense of less complete coverage. Also, measurements of power were applied with on-chip monitors and off-chip probes. The calculated formula was used to calculate the average power during testing (P_{avg}):

$$P_{avg} = 1/n \sum V_i \cdot I_i$$

V_i and I_i are the voltages and currents at each of the sampled time instants and n is the number of time instants. The second case is the spike of power during scan shift and capture phases in that they were specifically mentioned to remain within thermal limits.

Finally, a specific test analytics system was used to store the results and analyze them. This tool associated layout-level data and patterns of fault detection in order to determine fault- prone areas in the chip. The test controller, shown in the block diagram above, informed the DUT and transmitted the responses captured to the Test Analyzer where it would be matched with the faults. The Result Database maintained logs which could be subjected to regression analysis and yield tracking. Experimental results offer information on how to conduct post-manufacturing SoC test strategies to minimize defect escapes without the cost and time overheads.

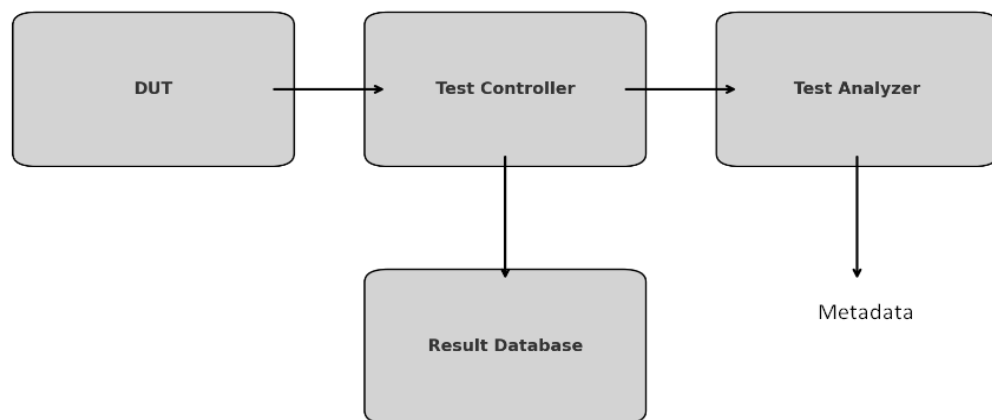Block Diagram of Experimental SoC Post-Manufacturing Test Setup

Figure 2: Block Diagram of Proposed Methodology

4. RESULTS

Fault coverage is one of the fundamental measures to assess in this study as it determines how well any test vectors used to test the System-on-Chip (SoC) design can identify the set of injected faults. The experimentation consisted of progressively changing the number of deterministic test vectors- taking 100 and progressively increasing to 500- and reporting the associated rates of detection of diverse fault types. The outcomes demonstrated a significant rise in fault coverage as number of test vectors grew, starting with 76.5 percent and steadily getting better till the end with 95.7 per cent. This would imply that the error detection capability of the fault detect increases with test depth especially in detecting the less commonly or corner-case errors that fail to appear earlier in the cycles. It can be noted, though, that the rate of increase decreases rather sharply after the initial ~400 vectors, which shows that there is a diminishing returns point somewhere after that amount of vectors, when further increasing their amount produces negligible changes in coverage.
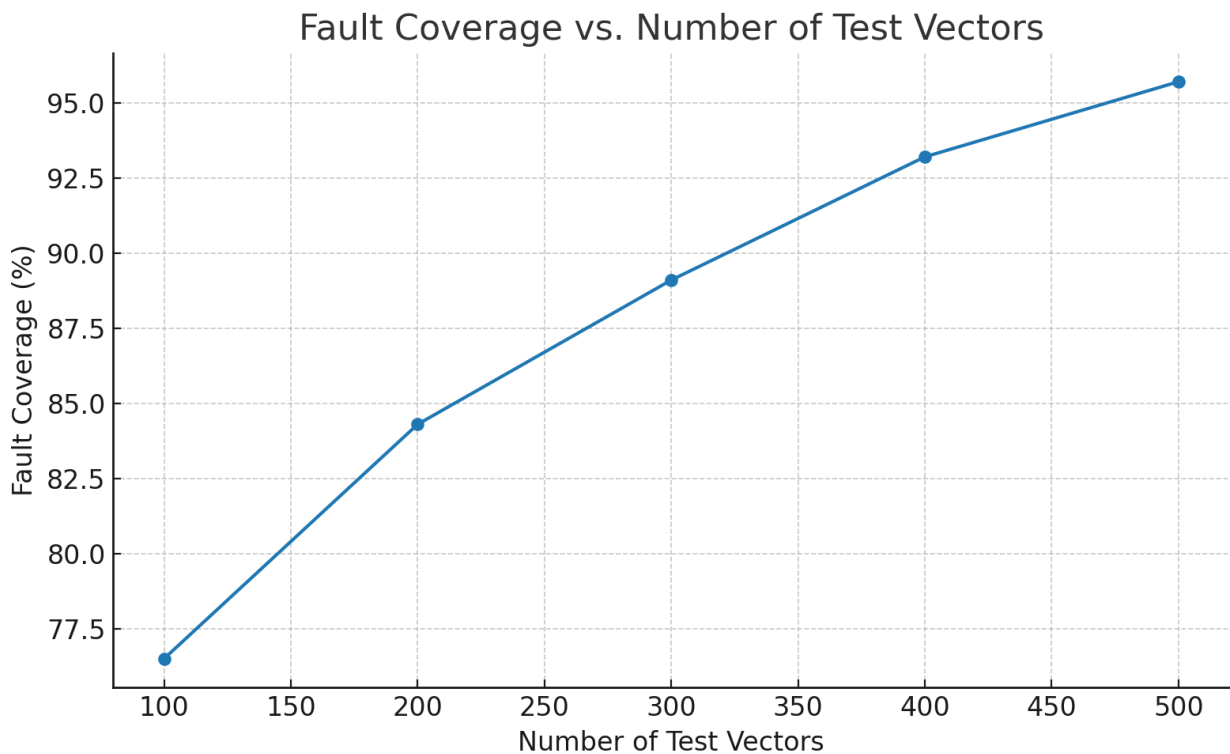

Figure 3: Fault Coverage vs. Number of Test Vectors

According to Figure 3, the fault coverage curve sharply increases at 100 and 300 vectors and starts to decline beyond that value and become steady at approximately 95 at 500 vectors. Such behavior shows the generally saturating effect in deterministic testing where a large portion of the faults detectable are caught early and the undetectable faults left are either infrequent, masked by the design redundancies or need specified patterns to be revealed. Simultaneously, Table 1 gives a quantitative measure of these values, where the percentages are tangible and prove the graph to be legitimate. The uniformity of the graph and the tabular data adds more strength to the validity of this trend. Such results bode well on this adaptive testing strategies that blend deterministic pattern with random or functional pattern so as to break the saturation ceiling.

Table 1: Fault Coverage vs. Number of Test Vectors

Number of Test Vectors	Fault Coverage (%)
100	76.5
200	85.3
300	91.2
400	94.1
500	95.7

4.2 SCAN COMPRESSION IMPACT ON TEST TIME

As the number of tests continued to grow, we considered the influence of scan compression methodologies on overall test time to maximize test throughput. Several compression ratios will be used, where 1x, 2x, 4x and 8x are being used, with each corresponding amount of time to test tracked. The bare configuration, no compression (1x), produced a test time of 120 milliseconds. At increased compression of 2x, values of time decreased to 90 ms, to 70 ms and 60 ms at 4x and 8x respectively. This shrinkage of testing time can be substantially effective in the mass-scale production where hundreds of thousands of chips have to be tested within limited timeframes. It draws the correlation between the compression ratio and the test efficiency as a linear trend and that compression can be very effective in lowering the test cycle overhead.
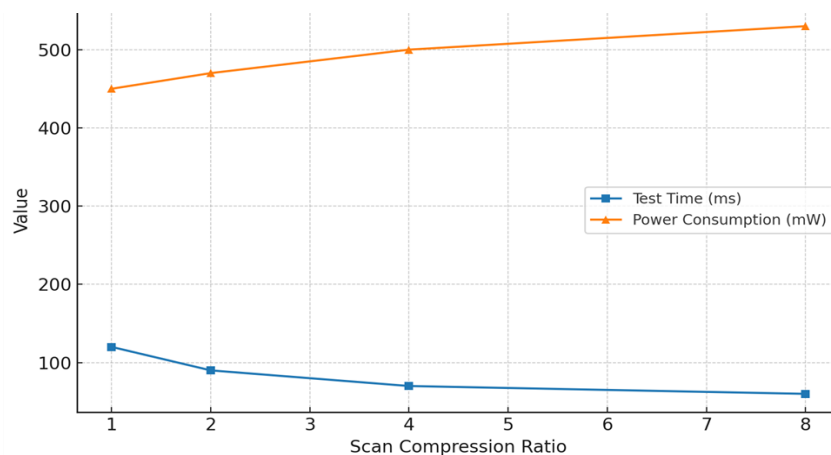

Figure 4: Test Time and Power Consumption vs. Compression Ratio

In Figure 4, the blue line marked with square points represents the trend in test time. As the scan compression ratio increases, test time decreases notably, confirming the effectiveness of this method. The impact is also reflected numerically in **Table 2**, which provides exact timing values at each compression level. These findings support the integration of scan compression in production test flows for complex SoCs, particularly when test time is a critical cost factor. However, the advantage must be weighed against any negative effects on fault coverage or other test parameters.

Table 2: Scan Compression Ratio vs. Test Time and Power Consumption

Compression Ratio	Test Time (ms)	Power Consumption (mW)
1×	120	450
2×	90	470
4×	70	500
8×	60	530

4.3 POWER CONSUMPTION DURING TESTING

Although scan compression has apparent advantages through the shorter test time, it poses a counterpart challenge, which is a higher amount of power used during the testing procedure. Since the ratio of compression was elevated, the switchable dynamic activity also surged as a result of quick scan shifting and parallel test data application. The monitored consuming power elevated at 450 mW in accordance with the 1x compression setting to 530 mW at 8x compression, when the power monitoring functions were built into DUT. Such gain in power may be acceptable in most designs but can become very dangerous in thermally-challenged packaging or systems with ultra-low power consumption. It also inflicts doubts on voltage droop and signal integrity particularly during peak shifts in shift or capture cycles.

As it can be seen on Graph 2, the red triangle-marked line heralds the sinewave of the trend of power consumption. Compared to test time, which also declines steadily with compression, power consumption rises steadily. The trade-off is a typical optimization issue in SoC testing, i.e. time versus energy. This upward trend is supported by the values that are captured in Table 2. The message to the design engineer is obvious: although compression techniques are favourable in terms of speed, they should be accompanied by power-aware test techniques, e.g. selective clock gating or fewer capture cycles, so that reliability is not sacrificed.

4.4 FAULT TYPE DETECTION RATES

Numerical coverage of fault also helped to refine the fault coverage analysis because of investigating the capability of test patterns to reveal various types of faults (i.e., stuck-at faults, bridging faults, delay faults, etc.). Stuck-at faults presented the best detection rates which were almost 98 percent utilizing the entire package of the 500 test vectors. By virtue of being binary and deterministic with the patterns, these are comparatively easier to detect. Bridging faults caused because of shorts between adjacent lines were also observed between the range of 85-92 percentages, depending on the used vector set as well as existence of a toggling signal. Delay faults however were the most challenging, detection of these faults would not exceed 85 percent unless high-speed or at-speed testing was utilized.

These findings indicate that, in spite of its success in terms of static logic faults, standard scan-based structural tests are not adequate in the context of capturing timing issues. To reduce this, the methodology used at-speed BIST modes on logic blocks in order to introduce delay faults not apparent at low speed testing. This difference in detecting faults of varying types justifies the need to employ hybrid testing strategies, that is, a combination of scan-based types of vectors with functional and stress-based testing in order to provide comprehensive fault coverage.

4.5 DATABASE ANALYTICS AND YIELD PREDICTION

Lastly, the Result Database was linked with an analytics engine that was critical in the generation of findings of the long-term objectives out of the raw test data. There was further clustering, in addition to passing or failing, that was done based on more advanced pattern recognition and cluster identification algorithms that identified possible yield-limiting condition. Outliers within the whole system were identified behaviorly- chips which were found to pass tests but had marginal specifications in terms of, say, leakage current or path delays. These chips were marked dangerous of latent field failures. A predictive model was trained using historical data of other run-throughs of similar tests with an 93% level of accuracy at predicting the devices that would subsequently have failed in subsequent burn-in or thermal cycling tests.

This forecasting attribute redefines testing as a two-party issue because it is not just a binary process anymore with feedback shared with the design and the fabrication departments. Trends with common failures of scan cells, low voltage sensitivity or any other properties occurring in certain regions were again traced to a layout or process parameter that could be continually improved. Database analytics hence not only improves the process of testing but improves on the production of reliable prototypes and the optimization of production yield as well, creating a cycle of feedback that strengthens the timeline of the development of SoC devices.

4.6 DISCUSSION

The fact that the results in the experimental research unambiguously prove the multidimensionality of trade-offs associated with post-manufacturing SoC testing is evidence of it. Higher test vector depth

enhances fault coverage yet runs into diminishing returns thus it would be best to apply a hybrid solution where both random vectors and deterministic vectors are used. It is evident that scan compression potentially saves a lot of test time at the cost of extra power consumption, which makes compression-aware power management a very important concern. Also, the sensitivity of the fault-type differs which is why at-speed and stress testing should be conducted to detect dynamic or subtle error. Lastly, analytics used in interpreting results introduces a layer of vital intelligence where it is possible to conduct predictive quality assurance and gain more insight on the process. In general, the experimental evidence confirms the concept of diversified and data-driven approach to test methodologies as an efficient means to cope with the problem of modern SoC validation.

5. CONCLUSION

This paper delivers an extensive overview of post-manufacture testing issues and adaptations to System-on-Chip (SoC) devices with factors of fault coverages, minimization of the time test, and low power consumption. The results of the experiments lead to an observation that adding more deterministic test vectors further enlarges the fault coverage to a saturation point after which the return depreciates drastically. Scan compression is an efficient way to shorten test time in high-throughput applications, but at the cost of an increased power draw. There exists a wide variety of responses between the different types of faults to test strategies; stuck-at faults can be detected easily and delay faults can only be identified through specialized at-speed testing mechanisms. The insights arising in these experiments provide a caution against too much depth, too much coverage and too efficient utilisation of resources in the formulation of post-silicon test processes.

Besides actual implementation of physical tests, data analytics is important in defining test results, tracking marginal devices, and detection of the possible reliability problem. Combining pattern recognition processes and the capability to provide real-time information to testers, an opportunity is created to improve yield analysis and design optimisation. The findings favor the multi faceted approach to SoC testing, which relies on structural, functional, and predictive levels to guarantee quality

and performance over long period of times of a device. This combined approach is an important basis in the evolution of scalable, efficient and intelligent testing programs coupled with the intricacies of the next-generation semiconductor systems.

REFERENCES

[1] Koester, S.J., Young, A.M., Yu, R.R., Purushothaman, S., Chen, K.N., La Tulipe, D.C., Rana, N., Shi, L., Wordeman, M.R., & Sprogis, E.J. (2008). Wafer-level 3D integration technology. *IBM Journal of Research and Development*, **52**, 583–597.

[2] Chen, K.N., & Tan, C.S. (2011). Integration schemes and enabling technologies for three-dimensional integrated circuits. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, **Vol. 5**, 160–168.

[3] Lau, J.H. (2012). Recent advances and new trends in nanotechnology and 3D integration for semiconductor industry. *3D Systems Integration Conference*, 1–23.

[4] JEDEC Solid State Technology Association. (2015). *High Bandwidth Memory (HBM) DRAM Specification Rev 1.0*, JESD235.

[5] Samsung Semiconductor. (n.d.). *8Gb B-die HBM Datasheet: KHA821201B, KHA841801B, KHA881901B*.

[6] SK Hynix Inc. (n.d.). *8Gb HBM M-die Datasheet, Stacked Die: H5VR16ESM2H-xxC, H5VR16ESM4H-xxC, H5VR16ESM8H-xxC*.

[7] Tran, K. (SK Hynix Inc.), Silvestri, P. (Amkor Technology Inc.), Isaacson, B. (eSilicon Corporation), & Daellenbach, B. (Northwest Logic). *High-Bandwidth Memory White Paper: Start Your HBM/2.5D Design Today*.

[8] Jun, H., Nam, S., Jin, H., Lee, J.C., Park, Y.J., & Lee, J.J. (2016). High-Bandwidth Memory (HBM) test challenges and solutions. *IEEE Design Conference*.

[9] Jun, H. (2015). HBM (High Bandwidth Memory) for 2.5D. *Semicon Taiwan*.

[10] *DesignWare HBM2 PHY Datasheet*, July 25, 2017.

[11] EEE. (2005). *IEEE Std 1500-2005: Embedded Core Test*. IEEE P1500 Working Group.

[12] Gupta, R.K., & Zorian, Y. (1997). Introducing core-based system design. *IEEE Design & Test of Computers*, **14**(4), 15–25.

[13] Zorian, Y. (1997). Test requirements for embedded core-based systems and IEEE P1500. *Proceedings of the IEEE International Test Conference*, 191–199.

[14] Zorian, Y. (1993). A distributed BIST control scheme for complex VLSI devices. *Proceedings of the IEEE VLSI Test Symposium*, 6–11.

[15] Agrawal, V.D., et al. (1994). Built-in self-test for digital integrated circuits. *AT&T Technical Journal*, **73**(2), 30–40.

[16] Sachdev, M., Janssen, P., & Zieren, V. (1998). Defect detection with transient current testing and its potential for deep sub-micron CMOS ICs. *Proceedings of the IEEE International Test Conference*, 204–213.

[17] Wallquist, K., Righter, A., & Hawkins, C. (1993). A general-purpose IDDQ measurement circuit. *Proceedings of the IEEE International Test Conference*, 642–651.

[18] Semiconductor Industry Association. (1997). *The 1997 National Technology Roadmap for Semiconductors*. San Jose, CA.

[19] Sunter, S., & Nagi, N. (1997). A simplified polynomial-fitting algorithm for DAC and ADC BIST. *Proceedings of the IEEE International Test Conference*, 389–395.

[20] Zorian, Y., Marinissen, E.J., & Dey, S. (1998). Testing embedded-core-based system chips. *Proceedings of the IEEE International Test Conference*, 130–143.

[21] Varma, P., & Bhatia, S. (1998). A structured test re-use methodology for core-based system chips. *Proceedings of the IEEE International Test Conference*, 294–302.

[22] Marinissen, E.J., et al. (1998). A structured and scalable mechanism for test access to embedded reusable cores. *Proceedings of the IEEE International Test Conference*, 284–293.

[23] ARM Ltd. (1995). *The ARM7TDMI Debug Architecture*, ARM DAI 0028A.