Test Planning and Test Access Mechanism Design for Stacked Chips using ILP

Sengupta, Breeta; Larsson, Erik

Published in:
VLSI Test Symposium (VTS), 2014 IEEE 32nd

DOI:
10.1109/VTS.2014.6818764

2014

Link to publication

Citation for published version (APA):
Sengupta, B., & Larsson, E. (2014). Test Planning and Test Access Mechanism Design for Stacked Chips using ILP. In VLSI Test Symposium (VTS), 2014 IEEE 32nd (pp. 1-6). IEEE--Institute of Electrical and Electronics Engineers Inc.. https://doi.org/10.1109/VTS.2014.6818764

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Test Planning and Test Access Mechanism Design for Stacked Chips using ILP

Breeta SenGupta and Erik Larsson
Lund University
Lund, Sweden
Email: Breeta.Sengupta@eit.lth.se, Erik.Larsson@eit.lth.se

Abstract—In this paper we propose a scheme for test planning and test access mechanism (TAM) design for stacked integrated circuits (SICs) that are designed in a core-based manner. Our scheme minimizes the test cost, which is given as the weighted sum of the test time and the TAM width. The test cost is evaluated for a test flow that consists of a wafer sort test of each individual chip and a package test of the complete stack of chips. We use an Integer Linear Programming (ILP) model to find the optimal test cost. The ILP model is implemented on several designs constructed from ITC’02 benchmarks. The experimental results show significant reduction in test cost compared to when using schemes, which are optimized for non-stacked chips.

I. INTRODUCTION

The semiconductor technology development makes it possible to manufacture very complex integrated circuits (ICs). A single chip (die) can contain billions of transistors. To enable ICs with more transistors, it is possible with ICs where multiple chips are packaged in one single package. Examples of such ICs are multi-chip modules (MCMs), where the chips are placed laterally, and system in packages (SiPs), where the chips are stacked vertically and connected by bonding wires or solder bumps. The most recent advancement is ICs where multiple chips are stacked vertically and connected by interconnects known as Through-Silicon Vias (TSVs). These stacked ICs (SICs) have benefits such as increased performance, decreased power consumption and reduced form factor [1].

While manufacturing of ICs with a single chip per package, so called non-stacked ICs, is complex, manufacturing of SICs is more complex due to the need of additional manufacturing process steps related to the making and insertion of TSVs, thinning of chips, and alignment and bonding of the chips. A majority of IC manufacturers (approximately 85%) expect test cost to be the bottleneck in the production of SICs in the forthcoming years [2]. Test cost can be reduced by addressing problems related to fault modelling, wafer probing, design-for-test (DfT) architecture, and optimization of the test plan and test architecture [3].

In this paper we focus on test planning (test scheduling) and test access mechanism (TAM) design to reduce the test cost. We assume SICs developed using a core-based design methodology where each core is equipped with a core test wrapper. The test cost is highly dependent on the test time and the DfT hardware. The DfT hardware for core-based ICs consists of (1) the TAM for the on-chip test data transportation between the automatic test equipment (ATE) and the embedded cores of the IC, and (2) the logic for core wrapper at each core. As each core is equipped with a wrapper, the cost for implementing the core wrappers becomes fixed. Thus, the cost of the DfT hardware highly depends on the TAM width. The test time and the TAM width are related. A narrow TAM leads to low DfT hardware but long test times, while a wide TAM increases the DfT hardware but enables lower test times. For that reason, we minimize the test time and the TAM width. Both test time and the silicon for TAM are purchased; hence cost in $.

The test flow impacts the test cost. For non-stacked ICs, the test flow is well-defined and consists usually of two test instances; wafer sort (testing of the unpackaged chip (bare die)) followed by package test (final test) (testing the packaged chip). At the two test instances, the TAM design is the same and, typically, the test plan is the same. It is therefore sufficient to find one optimized TAM design and one single test plan; an order in which the cores of the IC are tested. The test plan is first applied at wafer sort and then at package test using the same TAM architecture. For SICs it is very different. Testing can be applied at the following instances: wafer sort (testing individual chips prior to integration into the stack), intermediate test (testing a partially constructed stack), post-bond test (testing the complete chip stack), and package test (testing all chips in the packaged IC). We make two observations. First, the more test instances that are used, the higher is the chance to detect manufacturing defects. However, more test instances lead to higher test cost, especially in terms of test time. Reducing the number of test instances, decreases the test time but may lead to a higher manufacturing cost. If, for example, only package test is performed, all chips in the stack are wasted if a single chip is defective. Second, the testable units differs between the test instances. As will be detailed in Section II, optimizing the TAM design and the test plan considering the cores at each chip separately, leads to a sub-optimal solution when all cores are jointly tested during package test. In the worst case, it leads to a TAM design that is infeasible for package test.

In our previous work [4], we studied the test flow problem and showed that it is most suitable to use a test flow which
consists of wafer sort of each individual chip and package test of the complete stack of chips. In this paper, we assume such test flow, and address the problem of finding the most suitable test plan and TAM design for testing the cores of each chip during wafer sort and for testing all cores of the complete chip stack during package test such that the overall test cost, given as a function the test time and the TAM width, is minimal. As will be shown in the paper, efficient optimization of the test plan and the TAM must jointly consider testing of each individual chip at wafer sort and the testing of the complete stack at package test. We make use of an Integer Linear Programming (ILP) model to minimize the overall test cost. We applied our scheme on several designs constructed from ITC’02 benchmarks and compared the results against two schemes for non-stacked ICs. The experiments show that proposed scheme results in significant lower test cost.

The rest of the paper is organized as follows. In Section II, related work is reviewed and the work in current paper is motivated. The proposed ILP formulation is detailed in Section III and the experimental results are presented in Section IV. The paper is concluded in Section V.

II. RELATED WORK AND MOTIVATION

In this section we first detail previous related work on test planning and TAM design for non-stacked ICs. To motivate the need of current paper, we then demonstrate shortcomings when using schemes that are developed for non-stacked ICs when developing TAM design and test planning for SICs. Finally, we review related works on SICs to demonstrate the novelty of current paper.

For non-stacked ICs, several works addressed test planning to minimize the test cost for testing core-based systems [5, 6]. Zorian proposed for systems where each core is a testable unit that is tested with built-in self-test (BIST), a scheme to find a test plan where the test cost in test time and DFT hardware is minimized while power constraints are met [5]. The DFT hardware is given by the number of BIST control lines. As all cores employ BIST, the cost of BIST circuitry is not included in the cost function. Chou et al. proposed for the same problem a systematic approach where test time is minimized while resource constraints and power constraints are not violated [6].

IEEE 1500 [7] was developed to enable core isolation so that each core can be tested as an independent unit. Several works addressed the co-optimization of test planning and TAM design for core-based systems with IEEE 1500 wrapped cores where the test cost in terms of test time is minimized at given TAM width constraints [8, 9]. In these works, the problem at core-level is to form so called wrapper-chains such that test time and the need of useless test data bits are minimized for a given core with scan elements, that is scan-chains, inputs, outputs, and bi-directionals. The problem is illustrated in Figure 1. The core has 4 scan-chains, each 100 flip-flops long. The scan-chains can form 1, 2, 3, or more wrapper-chains. The test time of the core depends on the number of test patterns, which is fixed, and the number of wrapper-chains, which are to be defined from the scan elements. A low number of wrapper-chains leads to long scan-in and scan-out times, which increases the test time. A higher number of wrapper-chains reduces scan-in and scan-out times but requires higher TAM width to connect the wrapper-chains. The need of additional but useless test data bits depends on how well balanced the wrapper-chains are. Figure 1 shows two wrapper-chain configurations with 2 and 3 wrapper-chains, respectively. The test times are the same as the longest wrapper-chain in each configuration has a length of 200 flip-flops. However, in the case with 3 wrapper-chains, 200 bits of extra useless bits must be added for each pattern to ensure that all scan-chains can capture test data at the same time. Iyengar et al. showed that the problem of forming wrapper-chains at core-level is \( NP \)-hard and made use of a best fit decreasing (BFD) algorithm to form the scan elements into a given number of wrapper-chains such that test time is minimized [9].

The problem at system-level is given a TAM width \( W \) to find the most suitable number of TAM groups, their widths, and assign the cores to the TAM groups such that test time is minimized. Figure 2 shows an example of a TAM design (left) and a test plan with test time \( T_{sch} \) (right). As the test time at wafer sort \( T_{ws} \) and the test time at package test \( T_{pt} \) are equal to \( T_{sch} \), the total test time \( T \) is \( 2 \times T_{sch} \). During the optimization, the following two problems must be addressed, which will be illustrated with the help of Figure 2. First, consider Core A (details in Figure 1), which is assigned to \( w_1 \) with a width 3. This is sub-optimal when considering Core A alone as 2 wrapper-chains results in equal test time but requires less extra useless test data bits (discussed above). However, taking the overall perspective, the alternative is suitable. Second, the usage of TAMs are not perfectly balanced as \( w_2 \) is not fully utilised (not used until \( T_{sch} \)). As with unbalanced wrapper-chains, unbalanced TAMs adds extra useless test data.

While the above approaches are suitable for non-stacked ICs, they are not suitable for SICs. Assume a SIC with two
identical chips where the TAM and the test plan are optimized for the testing of the individual chips (see Figure 2). At package test, the setup for these two chips connected to an ATE is shown in Figure 3. Note that neither the test plan nor the TAM are optimized for this setup. In this case, due to identical chips, the TAM of the two chips becomes the same. Hence, connecting the chips becomes straightforward. However, in general, the chips in a SIC are not identical. The TAM can differ if each chip is optimized individually. If the widths of the TAMS in Figure 3 differs, it becomes cumbersome or even impossible to transport test data. Assume a stack of chips where the lowest chip is connected to the ATE. If this chip has the lowest TAM width of all chips, there is no bandwidth to send test data to chips higher in the stack designed with higher TAM width. Alternatively, if all chips have the same TAM width but the number and widths of TAMS differs between chips, package test might be possible if the chips are tested in a sequence one after the other.

Test planning to reduce test cost given by test time and DfT hardware has for SICs been addressed in [10, 11]. The schemes do however not assume any particular DfT architecture and make it difficult to reuse DfT hardware added for wafer sort test at package testing. We proposed a scheme to minimize test cost given by test time and DfT hardware assuming a JTAG 1149.1 architecture for test data transportation [12]. The main drawback with the JTAG 1149.1 architecture is the low bandwidth for test data transportation. Marinissen et al. proposed an architecture with IEEE 1500 compatible wrappers to enable core-based testing of SICs [13]. While, the architecture enables core-based test of SICs, there is no TAM design and test planning scheme proposed for SICs, which is the topic of current paper.

III. ILP FORMULATION

In this section, we detail the ILP formulation. The used notations are collected in Table I. Figure 4, which will be used to illustrate the notations, shows a SIC with \( I = 2 \) chips (chip 1 and chip 2). Chip 1 contains \( C_1 = 3 \) cores denoted \( c_{11}, c_{12}, \) and \( c_{13} \). Chip 2 contains \( C_2 = 4 \) cores denoted \( c_{21}, c_{22}, c_{23}, \) and \( c_{24} \). Figure 4 shows that the TAM width \( W = 5 \) is distributed into \( K = 2 \) TAM groups, \( w_1 \) and \( w_2 \). The width of \( w_1 \), given by \(|w_1|\), is 3. The scan elements at core \( c_{11} \) are formed into \( wc_{11} \) wrapper-chains. As the width of \( w_1 \) is 3, the number of wrapper-chains at \( c_{11} \) can be no more than 3.

Given as input to the ILP is a core-based SIC with \( I \) chips where all cores are wrapped with IEEE 1500 wrappers. As all cores are wrapped, the cost of wrappers is fixed and the additional DfT hardware is only depending on the TAM width \( W \). The output of the ILP is:

- a TAM design where defined is the TAM width \( W \), the \( K \) number of TAM groups, and the width of each TAM group \( w_k \) is \(|w_k|\), and
- a test plan where the scan elements at each core are formed into wrapper-chains and the cores are assigned to TAM groups.

The objective is to find a TAM architecture and a test plan with minimal cost in TAM width and test time. It should be noted that both test time and silicon for DfT are to be purchased, which means the cost function defines the cost in $.

The general form of the ILP model is as follows [14]:

\[
\begin{align*}
\text{minimize} & \quad \sum_{i=1}^{I} T_{\text{cost}}(w_{c_{ij}}) \\
\text{subject to} & \quad T_{c_{ij}}(w_{c_{ij}}) \leq T_{ws_i} + T_{pt} \\
& \quad \alpha \cdot T + \beta \cdot W \\
& \quad \forall i \in I, \quad \forall j \in C_i, \quad \forall k \in 1..K \\
& \quad y_{ijk} = 1 \quad \text{if core } c_{ij} \text{ is assigned to TAM } w_k; \quad \text{otherwise } 0
\end{align*}
\]

Table I

<table>
<thead>
<tr>
<th>Notation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( i,I )</td>
<td>Chip ( i ) in a SIC with ( I ) chips where ( i \in 1..I )</td>
</tr>
<tr>
<td>( c_{ij} )</td>
<td>Core ( c_{ij} ) is core ( j ) at chip ( i ) where ( j \in 1..C_i ) and ( C_i ) is the number of cores at chip ( i )</td>
</tr>
<tr>
<td>( wc_{ij} )</td>
<td>Number of wrapper-chains at core ( c_{ij} )</td>
</tr>
<tr>
<td>( W )</td>
<td>Total TAM width</td>
</tr>
<tr>
<td>( w_k, K )</td>
<td>TAM group ( w_k ) where ( k \in 1..K ) and ( K ) is the total number of TAM groups</td>
</tr>
<tr>
<td>( T_{c_{ij}}(w_{c_{ij}}) )</td>
<td>Test time of core ( c_{ij} ) when the scan elements are formed into ( wc_{ij} ) wrapper-chains</td>
</tr>
<tr>
<td>( T_{ws_i} )</td>
<td>Test time of all cores assigned to TAM ( w_k )</td>
</tr>
<tr>
<td>( T_{ws_i} )</td>
<td>Test time for wafer sort of chip ( i )</td>
</tr>
<tr>
<td>( T_{pt} )</td>
<td>Test time for package test of the SIC (all chips jointly)</td>
</tr>
<tr>
<td>( T )</td>
<td>Total test time which is the sum of all times at wafer sort and package test, given as: ( \sum_{i=1}^{I} T_{ws_i} + T_{pt} )</td>
</tr>
<tr>
<td>( \text{test cost} )</td>
<td>Total test cost given as: ( \alpha \times T + \beta \times W )</td>
</tr>
<tr>
<td>( \alpha, \beta )</td>
<td>( \alpha ) and ( \beta ) are designer specified constants to capture the importance of test time and TAM, respectively</td>
</tr>
<tr>
<td>( y_{ijk} )</td>
<td>1 if core ( c_{ij} ) is assigned to TAM ( w_k ); otherwise 0</td>
</tr>
</tbody>
</table>
The test time of cores assigned to TAM group $w_i$ is given as the sum of test time of the cores $c_{ij}$ where 

$T_{w_i} = \sum_{j \in C_i} T_{c_{ij}}(w_k) \times y_{ijk}$ \hspace{1cm} (2)

where $y_{ijk}$ is 1 only if core $c_{ij}$ is assigned to TAM group $w_k$. The test time of cores assigned to TAM group $w_k$ is given as the sum of test times of the cores $c_{11}$ plus $c_{12}$.

The wafer sort test time $T_{w_{s_i}}$ of chip $i$ is given by:

$T_{w_{s_i}} = \max_{k \in K}\{T_{w_k}\}$ \hspace{1cm} (3)

The wafer sort test time $T_{w_{s_i}}$ of the chip in Figure 2 is given as the sum of the test times of Core A+Core C.

The test time of all cores in all $I$ chips in the SIC assigned to TAM group $w_k$ is given by:

$\sum_{i \in I} \sum_{j \in C_i} T_{c_{ij}}(w_k) \times y_{ijk}$ \hspace{1cm} (4)

The objective function to be minimized is the test cost given as:

$\alpha \times T + \beta \times W$ \hspace{1cm} (1)

where $W$ is the TAM width, $T$ is the total test time of wafer sort of each chip $i$ plus the test time of the complete SIC, and $\alpha$ and $\beta$ are user-defined constants set to define the importance of test time and DfT hardware.

Below we detail the computation of the total test time $T$. $T_{c_{ij}}(w_k)$ denotes the test time of a core $c_{ij}$ when the scan elements are formed in $|w_k|$ wrapper-chains and assigned to TAM group $w_k$ is given by the Best Fit Decreasing (BFD) algorithm for wrapper-chaining [9]. For example, the test time of Core A in Figure 1 when the four scan chains of length 100 flip-flops are formed as three wrapper-chains is given as: $(200 + 1) \times p + 200$ where $p$ is the number of patterns (+1 is for the capture cycle).

The test time of all cores at chip $i$ assigned to TAM group $w_k$ is given by:

$T_{w_k} = \max_{i \in I}\{\sum_{j \in C_i} T_{c_{ij}}(w_k) \times y_{ijk}\}$ \hspace{1cm} (5)

where $T_{w_k}$ denotes the test time of all cores assigned to TAM $w_k$.

While Eq. 3 and Eq. 5 have similarities, note that the former only considers the cores at one given chip $i$ while the latter considers all $I$ chips.

Finally, the total test time $T$ is the sum of the wafer sort test time $T_{w_{s_i}}$ of chip $i$ and the package test time $T_{pt}$ of the complete SIC:

$T = \sum_{i \in I} T_{w_{s_i}} + T_{pt}$ \hspace{1cm} (6)

The constraints are as follows:

- The sum of the TAM widths $w_k$ at each chip is:

$\sum_{k \in K} |w_k| \leq W$ \hspace{1cm} (7)

- Each core $c_{ij}$ is assigned to exactly one TAM group $w_k$ where the number of wrapper-chains ($w_{c_{ij}}$) at core $c_{ij}$ satisfies $w_{c_{ij}} \leq w_k$ (all $I$ chips). The number of wrapper-chains at a core cannot be higher than the width of the TAM group to which the core is assigned. The number of wrapper-chains can, however, be less. For example, while a core with only one scan-chain can only form one wrapper-chain, this wrapper-chain can be assigned to a TAM group that has a width higher than one. It is not optimal as it enforces extra test bits.

- Each wrapper-chain is assigned to exactly one TAM wire.

### IV. Experimental Results

The objective of the experiments is to demonstrate that the proposed ILP scheme results in a lower test cost compared to when making use of schemes developed for non-stacked ICs. The proposed ILP scheme for SICs is detailed above and the following two schemes for non-stacked ICs were used.

- Scheme 1, the TAM for each chip is optimized independently of all other chips in the SIC. It means that each chip gets the TAM that is most suitable for its wafer sort.

### TABLE II

<table>
<thead>
<tr>
<th>Label</th>
<th>Design</th>
<th>Cores</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>d665</td>
<td>11</td>
</tr>
<tr>
<td>G</td>
<td>g1023</td>
<td>15</td>
</tr>
<tr>
<td>P</td>
<td>p34392</td>
<td>20</td>
</tr>
<tr>
<td>T</td>
<td>t512505</td>
<td>32</td>
</tr>
</tbody>
</table>

Fig. 4. Illustration of notations.
Note that after the optimization additional TAM wires can be added to a chip. For example, if the top chip requires a very wide TAM while all other chips only need a narrow TAM, the wide TAM is added to all chips to make testing of the top chip possible at package test.

- Scheme 2, the TAM for the lowest chip is optimized and the same test architecture is used for all chips in the SIC. In this case, all chips use the TAM optimized for wafer sort test of the lowest chip.

For the experiments, core-based SICs were constructed using four ITC’02 benchmarks: d695 (D), g1023 (G), p34392 (P), and t512505 (T), see Table II. To give an indication of the complexity of the designs, Table II also details the number of cores. Each of the ITC’02 benchmarks form a chip and by combining the four benchmarks in various ways, SICs with 2, 3 and 4 chips were constructed. For example, the DP design in Table III is a SIC with 2 chips consisting of d695 and p34392 where d695 is the lowest chip.

For the experiments, core-based SICs were constructed using four ITC’02 benchmarks: d695 (D), g1023 (G), p34392 (P), and t512505 (T), see Table II. To give an indication of the complexity of the designs, Table II also details the number of cores. Each of the ITC’02 benchmarks form a chip and by combining the four benchmarks in various ways, SICs with 2, 3 and 4 chips were constructed. For example, the DP design in Table III is a SIC with 2 chips consisting of d695 and p34392 where d695 is the lowest chip.

For the constants $\alpha$ and $\beta$ in the test cost (Eq. 1), we performed some experiments and found that suitable are the following settings: $\alpha = 1$ and $\beta = T_{max}/(0.5 \times (TAM_{max} - TAM_{min}))^2$ where $T_{max}$ is the test time when all chips are tested assuming one wrapper-chain, $TAM_{max}$ is set to $T_{max}/s$ where $s$ is the length of the longest scan-chain, and $TAM_{min} = 1$. By setting $\alpha$ and $\beta$ in this way, we have a general scheme for all design. Obviously, a designer can set $\alpha$ and $\beta$ in the most appropriate way for a given design. For the search space with respect to TAM width $W$, we have a general limit that has been used for all experiments. The highest allowed TAM width is limited by computing the sum of the test times of all cores assuming a single wrapper-chain. To obtain the highest TAM limit, this value is divided with the test time of the core with highest test time.

The results from the non-stacked schemes are in Tables III and IV. The results from the proposed SIC scheme are in Table V. The comparison between the three schemes is in Table VI. Tables III, IV and V are constructed in the same way. The results from the non-stacked schemes are in Tables III and IV. The comparison between the two non-stacked schemes against the SIC scheme is in Table VI. Table VI is organised as follows. Column one lists the designs. Columns two, three and four are organised in the same way with the test time, the TAM cost and the test cost for the two non-stacked schemes and the proposed SIC scheme. Column five reports the comparison between the non-stacked schemes and the SIC scheme. Column five reports the comparison between the non-stacked schemes and the SIC scheme. The results show that the SIC scheme produces the lowest test cost in all cases. For the benchmark DP, the SIC scheme is 125% better than Scheme 1 and 14% better than Scheme 2. At an average over all designs, the SIC scheme is 40% better than Scheme 1 and 63% better than Scheme 2.
V. CONCLUSION

In this paper we propose a scheme for test planning and test architecture design for core-based SICs based on ILP where the test cost, given as the test time and TAM cost, is minimized. We assume a test flow where each chip is individually tested at wafer sort and all chips (the complete SIC) are jointly tested at package test. In the experiments we compared the proposed scheme against two schemes for non-stacked ICs. The results show that proposed scheme results in a test cost at an average of 40% and 63%, lower than the two schemes for non-stacked ICs, respectively. As future works, we will not only address TAM constraints but also general test conflicts and test power constraints.

REFERENCES


