VLSI Term Project

Testing the Latent Delay Defects

Important Definitions:

Latent defects-hidden undetected flaws in testing which can't be identified by any user until some special set of operation are performed. These flaws can be detected only when a specific task is performed in unusual circumstances. Latent defects can cause undesirable delays in the circuit & affect circuit reliability. generally modeled as small-delay defects,

Aging is a technique of gradually increasing the priority of processes that wait in the system for a long time.

Hidden delay defects (HDDs) are small delay defects that pass all at-speed tests at nominal capture time. It is an important indicator of latent defects that lead to early-life failures and aging problems that are serious especially in autonomous and medical applications.

Automated test equipment (ATE) is a computer-operated machine used to test devices for performance and capabilities. A device that is being tested is known as device under test (DUT). ATE can include testing for electronics, hardware, software, semiconductors or avionics.The objective of ATE is to quickly confirm whether a DUT works and to find defects. This testing method saves on manufacturing costs and helps prevent a faulty device from entering the market.

Important Points:

- An effective way to screen out HDDs is to use **Faster-than-At-Speed Testing (FAST)** to observe outputs of sensitized non-critical paths which are expected to be stable earlier than nominal capture time.
- Normally Faster-than-At-Speed Testing (FAST) technique is an effective way to screen out HDDs, to observe outputs of sensitized non-critical paths which are expected to be stable earlier than nominal capture time.
- Even with aggressive FAST testing, HDDs generate only very few failing test response bits.

- Steps to overcome defects-presented the very first **logic fault diagnosis technique** that can identify HDD by analyzing fail logs produced by FAST.
- The technique uses **MRD back-propagation** to obtain initial HDD candidates which are then scored and ranked using simulation and variation-tolerant observation matching.
- The experimental results show a success rate of 94% even with very limited amount of failure data.
- FAST diagnosis established in this work provides the last missing component to rapidly improve yield and reliability by learning from otherwise unknown HDDs.

I. Existing Method (Previously followed Method):

Conventional At-Speed Testing

Process:

An at-speed delay testis applied using a single frequency matching the functional clock speed of the design.

A common type of fault model is called the “transition” or “at-speed” fault model, and is a dynamic fault model, i.e., it detects problems with timing. Compared to static testing with the stuck-at fault model (0/10), testing logic at-speed requires a test pattern with two parts classified as slow-to-rise and slow-to-fall faults.

The first part launches a logic transition value along a path, and the second part captures the response at a specified time determined by the system clock speed. If the captured response indicates that the logic involved did not transition as expected during the cycle time, the path fails the test and is considered to contain a defect. The time allowed for the transition is specified, so if the transition doesn’t happen, or happens outside the allotted time, a timing defect is presumed.

The transition fault model uses a test pattern that creates a transition stimulus to change the logic value from either 0-to-1 or from 1-to-0. For example, the pattern’s launch event may propagate a 0-to-1 (rising edge) transition along a specific path while holding all other conditions constant, as shown in Figure 1. Then the capture event pulses a functional clock to latch in the path’s response to the transition. If the “high”

value was not detected at the capture point in time, the path fails the test and is considered to have a “slow-to-rise” defect.

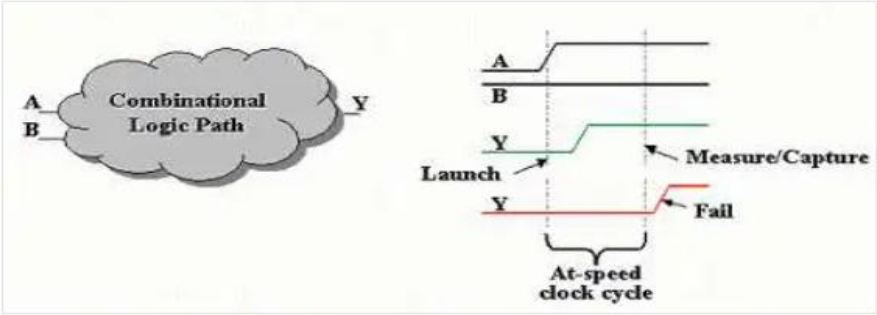


Figure 1: Pattern launch event propagates transition

Disadvantages:

- Early life Failures (ELF) may lead to economic losses for overall equipment manufacturers.
- Doesn’t capture the small delay defect/ hidden delay defects causing the performance of the circuit to become slow & when left unattended leads to aging

II. Proposed Method

Faster-than-At Speed Testing (FAST)

Aim/ Objective of the Project:-

To **improve the reliability** of current and future designs, this paper proposes **new backtracing and response matching methods** that yield high diagnostic success rates even with very limited amountof failure data.

The performance and scalability of HDD diagnosis method is validated using fault injection campaigns with large benchmark circuits. Here the goal is to assess the severity of the reliability problems and to distinguish defective chipsfrom chips that are slow due to parameter variations.

Distinguishing the effects of parameter variations from small delaydefects is also crucial for avoiding unnecessary yield loss. The proposed solutions exploitthe observation that slow chips show a different behaviorthan defective chips for varying supply voltages.

If a failure can only be explained by multiple fault locations, the chip is probably a slow chip due to parameter variations.

Process:

In FAST, various subsets of delaytests are performed with multiple capture times.

Fig. 2 shows an example of a test where a single delay test is applied twotimes and failure information is collected at two capture timescnom and cFAST.

On the one hand, since cFAST is earlier thanthe nominal circuit delay, outputs of long paths (o2) need tobe masked. On the other hand, the delay defect is observableat output o0 at time cFAST, but its fault effect has alreadyvanished at time cnom.

Since a test has different outcomes at different clock frequencies, the information on capture times needs to be included in the fail log.

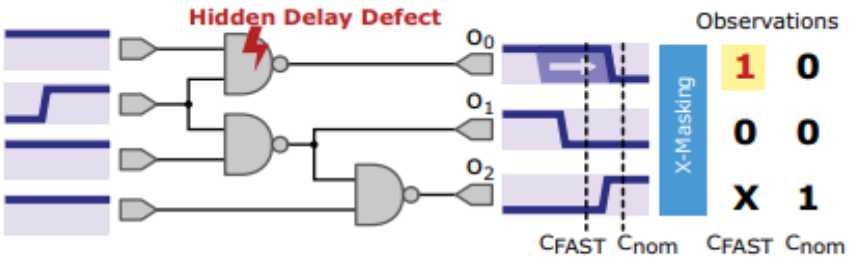


Fig. 2. Faster-than-at-speed testing with one test and two capture times discovering a hidden delay defect.

(1) HDD DIAGNOSIS

Two main phases: - effect-cause backtracingand subsequent cause-effect fault simulation

A. Phase 1: Backtracing FAST Observations

- used to identify the most suspicious circuit structures.
- Previous backtracing techniques usuallyoperate on the results of multi-valued logic simulation to determine sensitized paths.
- Backtracing in the proposed HDD diagnosis approach uses timing simulation data to better estimate sensitized paths

- The combinational portion of the circuit is extracted from the design and all scan cells are replaced by pairs of pseudoprimary inputs (PPIs) and pseudo-primary outputs (PPOs).
- The gates in the combinational portion of the circuit under diagnosis (CUD) are topologically ordered to facilitate the propagation of simulation data from PPIs to PPOs.
- The basic unit of computation is a waveform that contains the complete switching history of a signal line. First, the waveforms at PPIs are initialized with the launch transitions corresponding to a delay test pattern. The remaining waveforms in the combinational portion are calculated level by level and by using the nominal pin-to-pin and interconnect delays as illustrated in figure. As this computation progresses, the waveforms store all occurring transitions and glitches which are then available for all internal signals and PPOs.

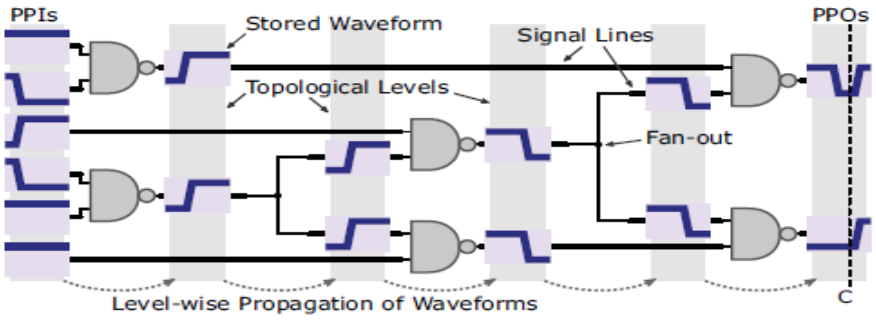


Fig. 3. Principle of waveform-based timing simulation of circuit.

- The simulation model calculates and stores separate waveforms for fan-out stems and all fan-out branches. This is used to facilitate fine-grained backtracing and fault injection. Every stored waveform corresponds to a potential HDD candidate location. We hereafter refer to these locations also as signal lines $l \in L$.

B. Phase 2: HDD Candidate Simulation

- Cause-effect inject-and-validate phase uses diagnostic fault-simulation to find the candidateswhose behaviors best match the observations in the fail log
- Each HDD candidate is injected into the simulation model one-by-one and simulated. HDD diagnosis uses full timing simulation for each candidate to predict its fault effects at PPOs.
- These predictions therefore include the effects of all re-convergencies and hazards within the circuit. The fault simulation results are compared with the observed failures in the fail log to calculate matching scores.
- Since the simulation is performed with nominal timing and the observations in the fail log are potentially altered by process variations, the variation-tolerant scoring is adapted.
- Finally, the HDD candidates are sorted by their scores to generate a ranked list as the final HDD diagnosis result.

(2) BACKTRACING FAST OBSERVATIONS

timing simulation of all considered delay tests.

Let $t \in T$ be a delay test that shows some observed failures in the given fail log. Test t is simulated with full (nominal) timing, yielding for each signal line $l \in L$ a waveform with all expected transitions over time. In particular, we obtain the expected $\mathrm{LST}(t, o)$ for all PPOs $o \in O \subset L$ as these are just the times of the latest transition in each of their waveforms.

Let $f = (t, o, c)$ be an observed failure for test t at PPO o at time c. Such an observed failure implies that the actual $\mathrm{LST}^*(t, o)$ of PPO o is later than c, while the expected $\mathrm{LST}(t, o)$ is earlier than c. For this failure to occur, the latest transition of PPO o must have been delayed by a HDD by at least $c - \mathrm{LST}(t, o)$. As mentioned previously, failures might have been observed at multiple capture times at the same PPO o for the same test t.

Definition 2. *Let T be a set of delay tests, O the set of all PPOs, and F a fail log. The* minimum required delay $\mathrm{MRD}(t, o)$ *for a test $t \in T$ and PPO $o \in O$ is defined as:*

$$\mathrm{MRD}(t, o) = \max\{c - \mathrm{LST}(t, o) | (t, o, c) \in F\}.$$

If there is no observed failure involving t and o, then $\mathrm{MRD}(t, o) = 0$.

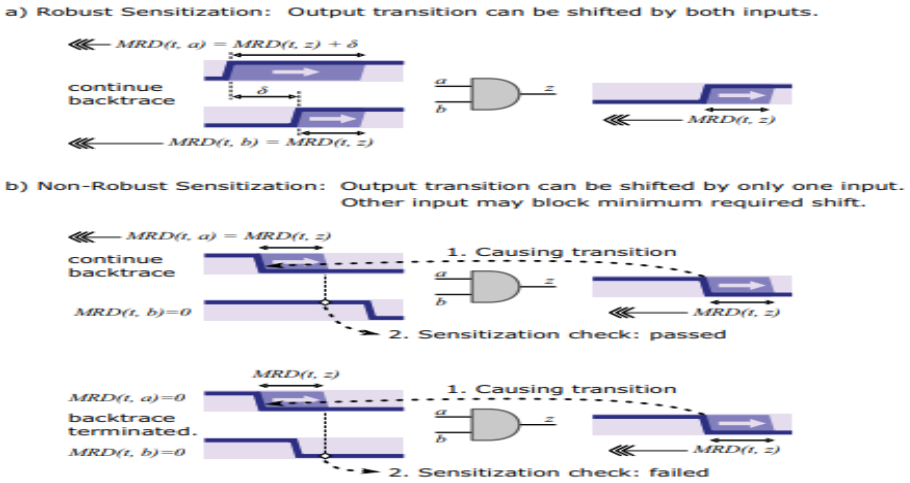
The fail log F is converted to a set of MRDs for each failing PPO and its test. If multiple observed failures exist for a PPO o and a test t, only one MRD is included since it covers all observations at o for test t for the purpose of backtracing.

A. MRD Back-Propagation

For the sake of clarity, we will only discuss this calculation for fan-outs and primitive one- and two-input gates in this paper:

For each non-zero MRD (t, o), an individual back-propagation is conducted to obtain a MRD (t, l) for each internal signal line l ∈ L. First, we initialize MRD(t, l)=0 for all signal lines $l \neq o$. All cells and all fan-outs in the circuit are processed in reverse topological order. Whenever a non-zero MRD (t, l_o) is present at an output line lo of a cell or a fanout, new MRD (t, l_i) values are calculated at the respective input lines l_i.

1) **Fan-Outs:** A fan-out is a signal that connects a single driver to multiple receiving gates. During back-propagation, each branch of a fan-out as well as its stem are regarded as separate internal signal lines with separate MRD values.
2) **Interconnects and One-Input Cells:** Interconnects with a specified transport delay as well as one-input cells such as buffers and inverters are transparent with respect to MRD back-propagation. Any MRD at their output is simply copied to their input, because the required delay-shift of the last transition is independent of the transport or the pin-to-pin delay of the respective cell.
3) **Two-Input Cells:** The back-propagation at multi-input gates is more complex because of possible non-robust sensitizations. If the signal at the off-path input of a two-input gate is stable and has no transitions, the MRD is propagated along the sensitized path in the same way as with one-input cells. If both inputs have transitions, robust and non-robust sensitizations need to be distinguished as shown in figure:



ig. 4. MRD back-propagation at an AND-gate for robust sensitization (a)

B. Initial Candidate Set Generation

B. Initial Candidate Set Generation

The results of all individual back-propagations are now combined to generate an ordered list of the most probable HDD candidates. Each HDD candidate consists of a location l, a polarity (slow-to-rise or slow-to-fall), and a delay or size. First we describe how to obtain likely candidate locations, and then we explain how to calculate polarities and delays to form the HDD candidate list.

For each $\mathrm{MRD}(t, o)$, we define two sets of candidate signal lines as:

$$L_{cand}(t, o) = \{l \in L | \mathrm{MRD}(t, l) = \mathrm{MRD}(t, o)\},$$
$$L'_{cand}(t, o) = \{l \in L | \mathrm{MRD}(t, l) > \mathrm{MRD}(t, o)\},$$

with all $\mathrm{MRD}(t, l)$ being obtained by back-propagating $\mathrm{MRD}(t, o)$. $L_{cand}(t, o)$ contains all signal lines l whose $\mathrm{MRD}(t, l)$ does not contain any slack, and $L'_{cand}(t, o)$ contains all signal lines with additional slack in their $\mathrm{MRD}(t, l)$ originating from robust sensitizations (input a in Fig. 4a).

For each signal line $l \in L$ in the circuit, we further define hit-counts $h(l)$ and $h'(l)$ as the numbers of appearances of l in all $L_{cand}(t, o)$ and $L'_{cand}(t, o)$, respectively:

$$h(l) = \sum_{(t,o,c) \in F} |\{l\} \cap L_{cand}(t, o)|,$$
$$h'(l) = \sum_{(t,o,c) \in F} |\{l\} \cap L'_{cand}(t, o)|.$$

The list of candidate signal lines is composed of two parts. First, all $l \in L$ with $h(l) > 0$ are added in descending order of h. Then, all $l \in L$ with $h(l) = 0$ and $h'(l) > 0$ are appended in descending order of h'.

Let l be a candidate signal line. The fault size $\delta(l)$ of the HDD candidate on l is determined by taking the maximum of all $\mathrm{MRD}(t, l)$ obtained during backtracing. This is a lower bound for the actual HDD fault size and ensures that the fault effect of the HDD candidate can propagate to all observation points reported in the fail log.

points reported in the fail log.

In addition to the calculation of $\mathrm{MRD}(t, l)$, back-propagation also records the polarity of the shifted transitions. As a signal line l may be reached by back-propagation multiple times, back-propagation may record only falling transitions, only rising transitions or both falling and rising transitions being shifted by the minimum required delay. If only falling transitions were shifted, the corresponding defect is likely of *slow-to-fall* (STF) polarity. In this case, an HDD candidate $\mathrm{STF}[l, \delta(l)]$ is added to the candidate list. If only rising transitions were shifted, a *slow-to-rise* HDD candidate $\mathrm{STR}[l, \delta(l)]$ is added to the candidate list. If both falling and rising transitions were shifted at signal line l while back-propagating from different observed failures, both $\mathrm{STF}[l, \delta(l)]$ and $\mathrm{STR}[l, \delta(l)]$ are added to the candidate list.

The list of initial HDD candidates, sorted by $h(l)$, is given to a fault simulator for final scoring. Our experiments show that among all successful candidate list generations (i. e., the real culprit is indeed among the candidates), the real culprit is almost always among the first few hundred candidates.

(3) HDD CANDIDATE TIMING SIMULATION

1) Confidence Estimation for Variation Tolerance

Fig. 5 illustrates the basic idea of evaluating the values of and the confidences in predicted responses. The waveform at a PPO o is a step-function that alternates between 1 and 0 according to the logic values of the signal over time. This step-function is multiplied with a Gaussian probability density function with its mean at a capture time c and a chosen standard deviation σ. The area A of the resulting product gives the probability that the scan cell at PPO o in the chip captured a logic 1. The predicted logic value at the PPO o is 1 if $A > 0.5$ and 0 otherwise. The confidence in the predicted logic value is given by $|2 \cdot (A - 0.5)|$. For example, the confidence is 0.0 if $A = 0.5$ and 1.0 if $A = 0.0$ or 1.0.

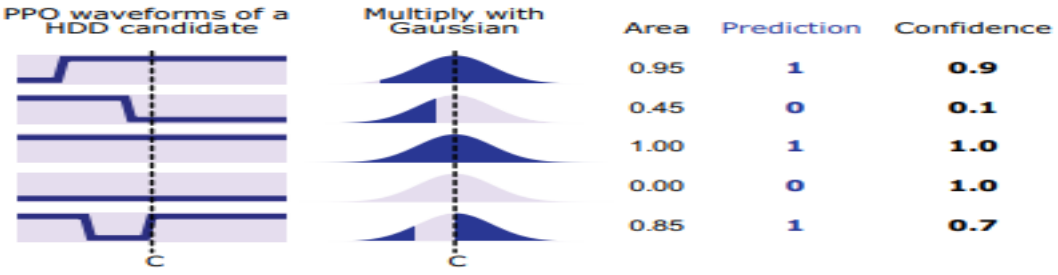


Fig. 5. Prediction of a captured value at time c and the estimation of simulation confidence [26].

The standard deviation σ determines the sensitivity of the confidence prediction to nearby transitions. A reasonable estimation of σ is given by the expected standard deviation of the LST at the PPOs of the circuit under variation. This standard deviation can be determined by running about 100 monte-carlo simulations and fitting a normal distribution to a histogram of

2) HDD Candidate Scoring

For a response bit, there are four possible situations, commonly known as TPSP, TPSF, TFSP, and TFSF.

• TFSF (Tester-Fail Simulation-Fail): The fail log contains an observed failure and the simulation also predicts a failure.

• TFSP (Tester-Fail Simulation-Pass): The fail log contains an observed failure but the simulation predicts a correct value

• TPSF (Tester-Pass Simulation-Fail): No failure was observed but the simulation predicts a failure.

• TPSP (Tester-Pass Simulation-Pass): No failure was observed, and the simulation also predicts a correct value

are added to s_{TFSP} or s_{TPSP}, respectively.

The goal is to order all HDD candidates by their estimated probabilities of being the actual culprit. This order is defined by a single final score s for each candidate that can be calculated with a formula of the general form:

$$s = s_{\mathrm{TFSF}} + \alpha \cdot s_{\mathrm{TFSP}} + \beta \cdot s_{\mathrm{TPSF}} + \gamma \cdot s_{\mathrm{TPSP}}.$$

In practice, $s_{\mathrm{TF}} = s_{\mathrm{TFSF}} + s_{\mathrm{TFSP}}$ is always very close to the number of observed failing outputs, and $s_{\mathrm{TP}} = s_{\mathrm{TPSF}} + s_{\mathrm{TPSP}}$ is close to the number of passing test response bits. Therefore, both s_{TF} and s_{TP} do not vary significantly between the different HDD candidates and are not very useful for ranking. If we treat s_{TF} and s_{TP} as constants within a ranking of candidates, the general score formula can be simplified as follows. First, the constants are subtracted from all scores without affecting the ranking:

$$s = s_{\mathrm{TFSF}} + \alpha s_{\mathrm{TFSP}} + \beta s_{\mathrm{TPSF}} + \gamma s_{\mathrm{TPSP}} - \alpha s_{\mathrm{TF}} - \gamma s_{\mathrm{TP}}.$$

Through simple re-arrangements, two partial scores can be eliminated:

$$\begin{aligned} s &= s_{\mathrm{TFSF}} - \alpha \cdot s_{\mathrm{TFSF}} + \beta \cdot s_{\mathrm{TPSF}} - \gamma \cdot s_{\mathrm{TPSF}} \\ &= (1-\alpha) \cdot s_{\mathrm{TFSF}} + (\beta - \gamma) \cdot s_{\mathrm{TPSF}}. \end{aligned}$$

Without changing the ranking, we can divide all scores by the constant $(1-\alpha)$. Furthermore, we can write the constant $-(\beta-\gamma)/(1-\alpha)$ simply as ω to yield:

$$s = s_{\mathrm{TFSF}} - \omega \cdot s_{\mathrm{TPSF}}.$$

The weight ω determines by how much simulation mispredic-

3) Candidate Simulation and Ranking

C. Candidate Simulation and Ranking

Phase 1 of the diagnosis procedure generated a list of HDD candidates that are sorted by the number of times its location was visited by backtracing. These HDD candidates are fault-simulated one by one, starting from the candidate with the highest hit-count. Each candidate is simulated with all failing tests that are present in the fail log. With the simulation results, $s = s_{\mathrm{TFSF}} - \omega^* \cdot \frac{s_{\mathrm{TPSF}}}{s^{\max}_{\mathrm{TPSF}}}$ for each candidate is calculated. In the calculation of the scores, both passing and failing response bits of the failing delay tests are considered.

Simulating all passing tests will yield a more accurate s_{TPSF}, which may further improve the ranking. However, this adds considerable fault-simulation runtime that may become infeasible for large designs.

Advantages:

- **Improve the reliability** of current and future designs**.**
- Backtracing prunes the vast search space of all possible HDDs down to an initial set of HDD candidates for further evaluation by diagnostic fault simulation**.**
- Does not require costly fault simulation, avoids any explicit enumeration of sensitization or propagation paths. Therefore, the computational resources needed for backtracing are only slightly more than those required for a fault-free timing simulation of all considered delay tests.
- HDD Candidate simulation & Ranking method yields more accurate results.

CONCLUSION

The goals of our experiments were to evaluate the performance in terms of results and runtime of the proposed HDD diagnosis algorithm. They were conducted on large ITC'99 benchmark circuits [44] that were synthesized using the SAED 32 nm technology library and a standard commercial tool flow. For each circuit, the FAST-ATPG approach from [12] was used to generate test pattern sets, observation times as well as a list of all HDD detected by the generated test sets. To generate a diagnostic test case, we randomly picked an HDD from the set of all detected HDD, injected it into the circuit, and simulated the FAST procedure to generate the fail log. Note that all culprits are actual hidden delay defects that cannot be detected with traditional at-speed delay testing as they do not break the timing of the CUD. To the best of our knowledge, no logic diagnosis approach in the literature targets diagnosis of HDDs. Therefore, it is not possible to compare our results to any other diagnosis algorithms.

Limitations of MRD Back-Propagation

There are cases of limitations in which MRD back-propagation may fail to include the actual fault-site.

(1) When a fault-effect propagates through the circuit as a hazard or glitch. Since backpropagation is based on altering transitions that already exist in the fault-free case, any additional hazards caused by the HDD in the CUD are not considered.

(2) If a fault-effect re-converges at a gate. In many cases (e. g., if the reconversion point is robustly sensitized), back-propagation will still follow some of the propagation paths and still reach the actual fault-site. However, if the reconversion point is non-robustly sensitized and a fault-effect is required on the off-path signal to obtain the output MRD, back-propagation may stop prematurely.

(3) For large designs candidate simulation and Ranking may become infeasible since the fault simulation runtime increases in the process.

Referred Papers:

Logic Fault Diagnosis of Hidden Delay Defects -Stefan Holst1, Matthias Kampmann2, Alexander Sprenger2, Jan Dennis Reimer2,Sybille Hellebrand2, Hans-Joachim Wunderlich3, and Xiaoqing Wen1

References

https://semiengineering.com/knowledge_centers/test/scan-test-2/

https://www.techopedia.com/definition/2148/automatic-test-equipment-ate