HETEROGENEOUS INTEGRATION ROADMAP

2020 Edition

Chapter 17: Test Technology

http://eps.ieee.org/hir

The HIR is devised and intended for technology assessment only and is without regard to any commercial considerations pertaining to individual products or equipment.

We acknowledge with gratitude the use of material and figures in this Roadmap that are excerpted from original sources.
Figures & tables should be re-used only with the permission of the original source.

IEEE ELECTRONICS PACKAGING SOCIETY · ELECTRON DEVICES SOCIETY · IEEE Photonics Society · ASME SETTING THE STANDARD · semi

Table of Contents

To download additional chapters, please visit
http://eps.ieee.org/hir

CHAPTER 1: HETEROGENEOUS INTEGRATION ROADMAP: DRIVING FORCE AND ENABLING TECHNOLOGY FOR SYSTEMS OF THE FUTURE

CHAPTER 2: HIGH PERFORMANCE COMPUTING AND DATA CENTERS

CHAPTER 3: THE INTERNET OF THINGS (IOT)

CHAPTER 4: MEDICAL, HEALTH & WEARABLES

CHAPTER 5: AUTOMOTIVE

CHAPTER 6: AEROSPACE AND DEFENSE

CHAPTER 7: MOBILE

CHAPTER 8: SINGLE CHIP AND MULTI CHIP INTEGRATION

CHAPTER 9: INTEGRATED PHOTONICS

CHAPTER 10: INTEGRATED POWER ELECTRONICS

CHAPTER 11: MEMS AND SENSOR INTEGRATION

CHAPTER 12: 5G COMMUNICATIONS

CHAPTER 13: CO DESIGN FOR HETEROGENEOUS INTEGRATION

CHAPTER 14: MODELING AND SIMULATION

CHAPTER 15: MATERIALS AND EMERGING RESEARCH MATERIALS

CHAPTER 16: EMERGING RESEARCH DEVICES

CHAPTER 17: TEST TECHNOLOGY

EXECUTIVE SUMMARY AND SCOPE 1
SECTION 1: RF TEST 2
SECTION 2: TEST OF PHOTONIC DEVICES 8
SECTION 3: LOGIC DEVICE TESTING 30
SECTION 4: SPECIALTY DEVICE TESTING 35
SECTION 5: MEMORY TEST 40
SECTION 6: ANALOG AND MIXED SIGNAL TEST 42
SECTION 7: WAFER PROBE AND DEVICE HANDLING 49
SECTION 8: SYSTEM LEVEL TEST 60
SECTION 9: DATA ANALYTICS 68
SECTION 10: 2.5D & 3D DEVICE TESTING 89
SECTION 11: KEY DRIVERS AND TEST COSTS 93

CHAPTER 18: SUPPLY CHAIN

CHAPTER 19: SECURITY

CHAPTER 20: THERMAL

CHAPTER 21: SIP AND MODULE SYSTEM INTEGRATION

CHAPTER 22: INTERCONNECTS FOR 2D AND 3D ARCHITECTURES

CHAPTER 23: WAFER-LEVEL PACKAGING (WLP)

Chapter 17: Test Technology

Send corrections, comments and suggested updates to the TWG chair, using our HIR SmartSheet:

https://rebrand.ly/HIR-feedback

Executive Summary and Scope

In this 2020 update to the heterogeneous integration testing roadmap, we have simplified and focused our message. This was done by combining and connecting previously separate sections.

For example, with this update we introduce the **Data Analytics** section led by Ira Leventhal. This new section discusses the expanding role of AI in test while recognizing the close interconnection between this topic and previously separate topics of Test and Yield Learning, Adaptive Testing, and Reliability Testing.

Another combined topic is the new **Wafer Probe and Device Handling** section which combines four related topics of probes, probers, handlers and test sockets which were individually discussed in previous versions of this roadmap.

In addition to the major changes above, industry changes warranted significant updates to several sections as highlighted below.

Sections Updated in 2020 Revision

Analog/Mixed Signal Devices: Technology advances are expected to challenge mixed-signal devices in about five years as speeds and resolutions increase while voltage swings continue to drop.

Data Analytics: This new topic area brings together different sections in the previous revision where data generation tasks and data utilization tasks are separately discussed. With this change the team starts exploring the contribution of Artificial Intelligence (AI) in the industry. This section shares the challenges and roadmap for the industry to making data-driven decisions which improve both our product costs and quality.

RF Devices: As 5G devices ramp in high-volume manufacturing, the best-known-methods for testing these devices, and the challenges and opportunities for the market, have become clear. Additionally, higher-performance elements of the 5G solution set with frequencies up to 70GHz now challenge the industry for test solutions. The updated section describes in detail the expected solutions and their roll-out plans.

Photonic Devices: As data-center electrical data rates push beyond 112Gbps, the value of migrating to an optical interface is now clear. The challenge for the industry is how to effectively test these interfaces in a cost-effective fashion.

Logic Devices: This section highlights the challenges and solutions available to help the industry transition with a variety of new technologies. This discussion includes a focus around repairability as well as the critical need for Improved access and diagnosability.

Wafer Probe and Device Handling: It is critical that the industry have cost effective contacting solutions in both wafer probe and package test. The accelerated demand for higher pin counts and finer pitch, together with higher-power solutions, is being driven by the AI and data-center growth fueled by the world-wide response to COVID-19 and adds urgency to the tasks this team has mapped out in this latest revision.

Acknowledgments

This update to the test roadmap is output from a team of more than 120 test experts hailing from more than 50 companies world-wide. The leadership team for this effort is shown below:

Test Technology Working Group, Heterogeneous Integration Roadmap Leadership Team

Dave Armstrong & Ken Lanier: Co-Chairs

2.5D/3D device test	Zoe Conroy	Photonics Test	Dave Armstrong, Tom Brown, Sylwester Latkowski
Analog and Mixed-signal Test	Don Blair	RF Test	Don Blair
Cost-of-Test	Ken Lanier	System Level Test	Harry Chen
Data Analytics	Ira Leventhal	Specialty Device Test	Wendy Chen
Logic & SOC Device Test	Marc Hutner	Wafer Probe and Device Handling	Jerry Broz
Memory Test	John Caldwell		

Section 1: RF Test

Introduction:

In the last HIR RF Test roadmap in 2019, the key trends and difficult challenges were covered for the short term (<5 years), medium term (6-10 years) and long term (10 years +) in the areas of Mobility (LTE-4G, Wi-Gig and 802.11a/b/g/n to 802.11ac WiFi):

- IoT (The huge numbers of RF sensors that transfer data into a global data distribution center, and processing the large volume of data)
- Infrastructure/Automotive RADAR/Industrial (collision-avoidance detection systems and point-to-point backhaul)
- The trends were compiled into a format of test parameters that were numerically predicted for the different terms. The topic of 5G was touched on in 2019 but was in an infant stage at that point in time. In 2020, the total market share for 5G jumped to 9% and is projected to grow to 28% in 2023 as shown in Figure 1.

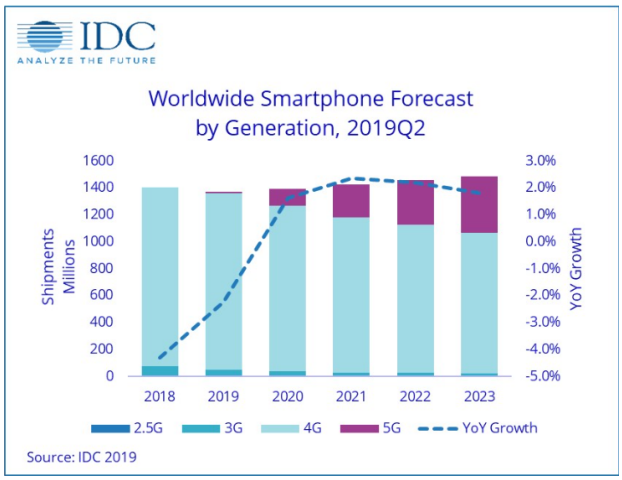


Figure 1. 5G Market Forecast [1]

- COVID-19 is a driving factor in the need for more bandwidth on the cellphone due to video conferencing taking the place of being there "in person" and is likely to hasten adoption and high-volume manufacturing of 5G solutions. For these reasons, this 2020 update will focus on 5G test.

Benefits of 5G

The benefits of 5G are shown in Figure 2. They are numerous and compelling factors for the new standard.

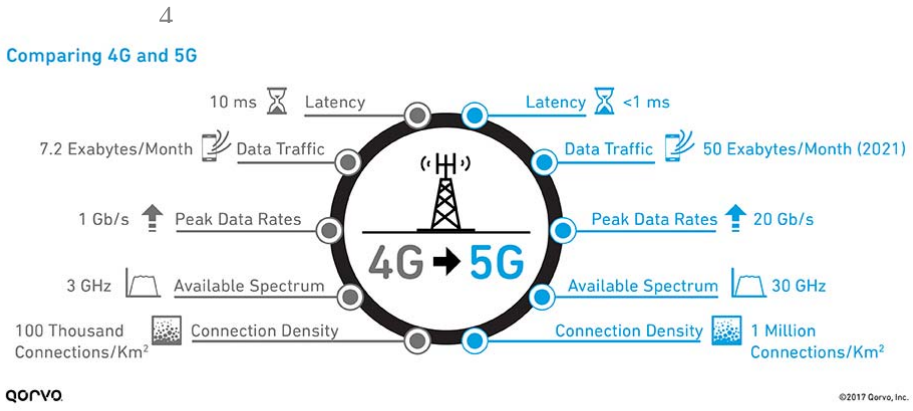


Figure 2. Benefits of 5G[2]

[1] https://informationmatters.net/5g-wireless-market-size-forecasts/

[2] https://www.qorvo.com/design-hub/blog/getting-to-5g-comparing-4g-and-5g-system-requirements

More details are shown in the Heterogeneous Integration Roadmap 2019 Edition in the 5G section.[3]

5G Topology

The 5G Topology shown in Figure 3 drives the need for:

- High density base stations for infrastructure
- Higher frequencies and increased bandwidth
- Cross-band MIMO transceiver tests
- Antenna tests with beamforming
- Low cost and high-volume handsets

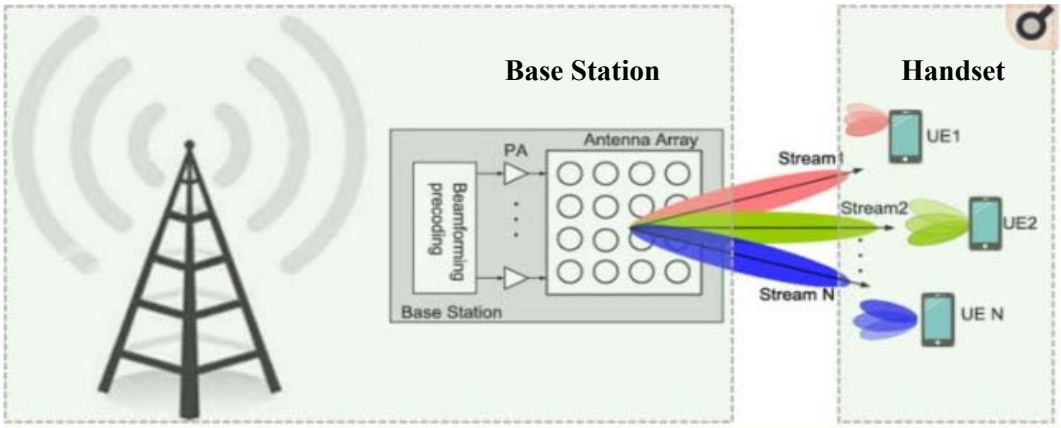


Figure 3. 5G Topology[4]

Overall 5G testing challenges

- Black type means manufacturing solutions are known and deployed widely
- Red type means manufacturing solutions are known but NOT deployed widely. Further investigation is needed.

5G transceivers have a huge dynamic range. Most ports are bidirectional with power ranging from -100 dB Rx up to +46 dB Tx. As many as 32 paths per device site involve these issues:

- Fixturing could be an issue with cabling and connectors
- Because of the resources needed, it could be difficult to test in parallel or multi-site parallel
- Loss and calibration (de-embedding the test fixturing)

Possible test interfaces could require non-contact testing (2 sides)

- Over-The-Air (OTA) test to the antenna in the package
- Anechoic chamber (high COT means this may be a characterization-only solution or perhaps a simplified chamber for production).

Socketing requires higher bandwidth, smaller pitch, tighter tolerances, and various package types.

Extending traditional transceiver tests into the FR2 (mmWave) bands are the most challenging and will be the focus of this section as shown in Figure 4. This includes:

- FR2 technologies are different than the current cell phone architecture.
- It is not simply a matter of scaling. High IF is up/down converted to mmWave FR2.
- FR2 has large frequency spans.
- Interfacing at FR2 frequencies is less stable. It is roughly 5X the speed of FR1 and the mechanics and electrical physics are challenges.

[3] https://eps.ieee.org/images/files/HIR_2019/HIR1_ch12_5G.pdf
[4] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6211090/

Band	ƒ (GHz)	Common name	Subset of band	Uplink / Downlink[B 1] (GHz)	Channel bandwidths[5] (MHz)
n257	28	LMDS		26.50 – 29.50	50, 100, 200, 400
n258	26	K-band		24.25 – 27.50	50, 100, 200, 400
n260	39	Ka-band		37.00 – 40.00	50, 100, 200, 400
n261	28	Ka-band	n257	27.50 – 28.35	50, 100, 200, 400

Figure 4. 5G FR2 frequency allocations[5]

5G mmWave bandwidth is 800 MHz with 8 carriers. As a result, there are carrier interference challenges:

- Test economics for the handset will be a challenge.
- Other bands may be used (unused bands from the FR2 table).
- New components haven't been tested in high-volume manufacturing before, such as beamformers and antenna arrays.
- 5G antennas are not omnidirectional like their predecessors (3G/4G), so the test methodologies need updating.

5G Antenna background

Beamforming

Beamforming consists of modifying the gain and phase of multiple antennas to focus and steer the transmitted or received radio energy in a particular direction. There are Analog and Digital techniques for beam forming. Some brief descriptions are:[6]

- **Analog** - A single signal is fed to each antenna element in the array by passing through analog phase-shifters where the signal is amplified and directed to the desired receiver. The amplitude/phase variation is applied to the analog signal at the transmit end where the signals from different antennas combine.
- **Digital** - the conversion of the RF signal at each antenna element into two streams of binary baseband signals – cosine and sine – are used to recover both the amplitudes and phases of the signals received at each element of the array. The amplitude/phase variation is applied to the digital signal before the DAC conversion at the transmit end. The received signals from antennas pass from ADC converters and DDC converters.

Beamforming is a special implementation of MIMO. Arrays of multiple antennas control the direction of a wave-front by sending the same signal from multiple antennas with different magnitude and phase. The signals are multiplied by complex beamforming weights; the phase determines the direction and the amplitude controls beam width and sidelobe level.

Multi-antenna techniques for beamforming

- Antenna arrays provide high antenna gain and side lobe attenuation.
- The higher the number of antenna elements in the phased-array system, the narrower the beam width and the better the beam can be focused.
- Beams can switch direction and other properties in microseconds; the range of possible directions is limited but can be extensive.
- Multiple beams can be transmitted at the same time.
- Defective single antenna elements do not lead to complete system failure because of redundancy.
- As carrier frequencies go up, antenna element dimensions shrink.

The 5G Antenna Test list includes[7]

- Gain
- Spurious emission tests
- MIMO multiple paths (8 or more)

[5] https://en.wikipedia.org/wiki/5G_NR_frequency_bands
[6] https://blog.pasternack.com/antennas/analog-vs-digital-beamforming/?gclid=EAIaIQobChMImZLGhc6Q6QIVJAnnCh0vDAeyEAAYAyAAEgJKmfD_BwE
[7] https://www.keysight.com/us/en/assets/7018-06369/brochures/5992-3413.pdf

  - Test the interference between paths
- The phased-array part of the antenna is the challenge (see next sections on Base Station and Hand Set test for more details)
- Test could be done at:
  - Package test – antenna integrated into the package
  - SLT – antenna integrated into the cell phone case
- Known-Good Die (KGD) could be a solution to reduce overall costs. If customers are driven to KGD, then test must cover KGD.
- New methodologies are needed to minimize test and cost of test.

5G Base Station Attributes and Testing Challenges

Base Stations need more digital, more power, multiple packages, multiple interfaces, with multiple antenna elements so more heat is generated. Since 5G infrastructure comes first, it is more urgent and will set the pace for the test strategy.

Antennas will not be integrated into the base station. They will be "standalone" due to the large size of the fixed antenna array.

Power requirements are higher, at 10W, than for the handset, and there are more thermal challenges on multiple antenna arrays (entire chip array).

The burden of negotiating the transceiver link is on the base station and there can be different frequencies for the links (one frequency for one type of data, a different one for other types. Example: voice and data). Much of the "intelligence" of the link is in the base station.

A higher "burden" of quality of test and more test coverage (quantity) is on the base station because the COT for the base station test can be "amortized" over the 5G service to consumers. Also, the number of IP blocks in the handset are fewer than thise in the base station.

More absolute accuracy for the antenna is needed on the base station side (beamforming - phase and gain)

5G Handset Attributes and Testing Challenges

The handset antenna is "omni-directional" in aggregate through the use of multiple antenna arrays; smaller/simpler arrays of antennas (4 as an example) than for the base station, and they consume less power.

Cost of test (COT) for the handset will be much more sensitive because:

- The consumer sees this cost directly in the purchase of the handset
- Elimination of tests can be considered as more test data/experience is gathered (yield learning)

Over-the-Air (OTA) testing is feasible, but it has specific impacts including modeling of the contactor; reflections and the alignment in the contactor shift the phase more at higher frequencies. Two potential options are placing the antenna in the package or in the cell phone case. System-level test (SLT) becomes more critical for designs with the antenna in the cell phone case because this may be the only practical way to access it, so interconnects need to be a part of this test.

Test coverage could depend on yield:

- Higher yield could allow some go/no go testing.
- A lower yield may require more qualitative tests, and trimming or other calibration adjustments may be needed.

Calibration must be done through the socket, not just to the socket. There is no single standard production solution right now. They are all non-standard.

5G Antenna Handset Architecture

There are some different antenna architectures for the handset antenna. These include arrays in the phone case or separate ICs distributed across the phone PCB. In Figure 5, the beam-steering antenna is in the case of the phone.

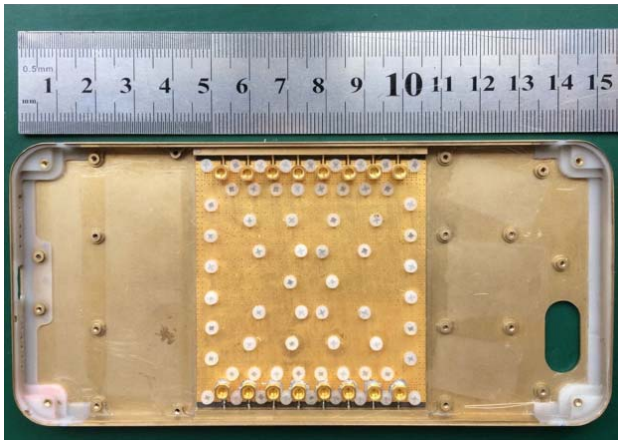


Figure 5. 5G Antenna in the cell phone case[8]

This may place a heavier burden on SLT if the antenna cannot be tested within the handset IC. It is possible the antenna could be tested before placement into the cell phone case, but interconnects would still need to be tested at SLT. This is being investigated further.

Another option is to put the beam steering antenna into separate ICs in the phone as shown in Figure 6.

Figure 6. 5G antenna as a standalone IC on the phone PCB[9]

In this case the antenna can be tested as an IC; however, SLT will continue to be needed to test the interconnects.

Testing Strategy

Test Strategies are evolving but will have the common elements of guaranteeing a working phone at "power up" and cutting test costs to the minimum.

Is this a Chipset or single SOC device?

The minimal chipset could be a front end (filters, amplifiers, switches), baseband modem, and antenna.

As happened in the previous 4G space, a very low-cost market demand could drive the 5G chipset to become a single IC with limited features. In the case of 5G, the physics may dictate the application design and usage (antenna, package, etc.).

Are there any testing "shortcuts"?

Loopback is a very low-cost shortcut, but it is not widely done in the RF lower bands today; it is not known if this will meet the testing goals.

DFT is another possibility but again, today this is not widely done on the RF lower bands. This needs more investigation in the 5G FR2 space.

[8] https://spectrum.ieee.org/tech-talk/telecom/wireless/a-beam-steering-antenna-for-real-world-mobile-phones

[9] https://www.theverge.com/2018/7/23/17596746/qualcomm-mmwave-5g-antenna-smartphones-qtm052-networking-speeds-size

Another strategy could be to do the simple tests on the IC and then do the rest at SLT. This could be implemented several ways including:

- Instead of doing the full range of frequency tests, measure at the worst-case frequency.
- Using a golden Device under Test (DUT) approach on the load board is a possibility that solves the OTA testing with the antenna in the golden DUT.

System Level Test (SLT) volumes will be high in order to:

- Ensure a good customer experience at first "power up" and guarantee the functionality based on the end use application.
- Ensure a minimal amount of lost/returned product.
- Learn what is the correlation between final (package) test and SLT. The experience from fails in SLT will then be integrated into the final test list.
- Provide feedback to design, test, and packaging.
- In the model described in the Digital logic test section,[10] this would be a "shift right" in the production cycle. The issues with this are the further right you go, the more a failure costs. The further left you go, the cheaper the cost is to find a failure; however, SLT is still necessary to ensure consumer "turn the phone on and it works" quality.

Package Test specific issues include catching dicing damage of the die, checking the interconnects, and doing the necessary AC/RF testing to ensure the part passes.

Wafer Tests are mainly DC and digital today. As much RF testing as is possible is also done here.

A probe card for 5G FR2 must have shorter needles because of the higher frequencies. Coaxial probes are not practical at the wafer pitches. Membrane probes and short needles need to be less than 1 mm long to have a discontinuity less than one tenth of the wavelength, which is expensive.

Having a single-site high-performance probe card for characterization and a less expensive high-parallelism card for production is a good alternative, but calibration/de-embedding still needs to be addressed.

Summary

5G is an upcoming and growing market with consumer demand that will require high volumes and reasonable price points. As a result, manufacturing solutions will be developing and evolving to meet these challenges. 5G base stations and handsets are out today in relatively small numbers so the technology is there to do the test. However, it needs to evolve in the ways depicted in Figure 7.

Now	Future
Test methodologies are established and available	These methodologies need to be adapted to a lower COT and to handle higher volumes.
Lots of testing done initially	Reduce/simplify tests as yield learning is achieved.
Base station infrastructure is rolling out	There is less cost pressure for test costs because the carriers can pass some of that cost along to many consumers.
5G handsets are introduced to the market	Test methodologies need to adapt to low cost/high volume for handsets (consumer).
Fixturing/Socketing/Contacting Challenges Physics/Mechanics of freq. and number of ports 28…39 GHz / 32 ports Number of ports/sites is the key	Need to fill the Technology and Cost Gaps through adapting/evolving current solutions and better/simpler methodology.
Antenna-In-Package (AIP) devices can be tested currently both at probe and package.	Need to improve the handling solution to deal with AIP devices/packaging and further evolve a solution for AIP at probe.

Figure 7 – What the future holds for 5G test

[10] See section 3

Section 2: Test of Photonic Devices

Executive Summary

In the electronic integrated circuit (EIC) industry, testing has become a mature process supported by practices and equipment that have been heavily optimized to drive down the cost and time spent on IC testing. In contrast, development of similar methods and tools for the photonic integrated circuit (PIC) community is still at an early stage, and the extra complexity that arises from having to measure both in the optical and the electrical domain poses many challenges. In this section. we define a number of key areas where development is needed, and in each of these areas we strive to leverage as much as possible the existing knowledge, practices and infrastructure from the EIC industry.

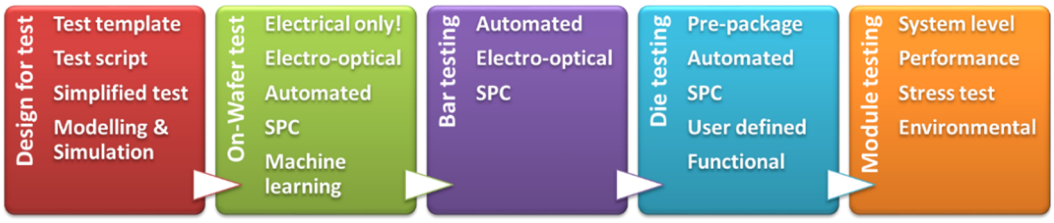


Figure 1. Overview of the test processes across the manufacturing chain of photonic integrated circuit based modules. Statistical process controls (SPC) require adequate test methods and data collection plans which should be accounted for already at the design phase.

The term PIC refers to an immensely diverse field of different implementations where we need to consider different (1) materials (InP, GaAs, Si, polymer, Triplex, glass), (2) integration schemes (monolithic, hybrid, etc.), (3) packaging (hermetic, non-hermetic, material) and (4) optical couplers (gratings, edge, mirrors coupled to single fibers, fiber arrays, lensed fibers, etc.). This leads to a first key development area: standardization of test metrics. New standardized testing methodologies and qualification parameters need to be devised that apply to all technologies, types of packages, and all relevant environmental conditions – leading to a truly platform-agnostic test solution.

A second focus area is to consolidate the design and test work flow (see example in Figure 1). A four-step method is proposed to enhance collaboration between designers, fab engineers and test engineers. Variations in dimensional and physical properties of materials and modules need to be understood and taken into account during design. This permits engineers to predict the influence of process variations on measurement results and allows them to design dedicated and improved test structures up front. By repeating these steps in combination with a careful analysis of the stored data, the number of devices to be measured and tracked can be reduced and the functional yield is expected to increase.

This targeted reduction in number of devices brings us to a third key area: test time reduction. There is a clear need for fully automated test systems. On the one hand this includes inline and where possible in-situ process testing at wafer level such as critical-dimension (CD) monitoring, defects counting, ellipsometry, etc. On the other hand, this includes the (out-of-line) automatic functional testing at wafer, bar/die and module level. For the functional test, a massively parallel test approach is envisaged in order to bring down measurement time and cost.

More specifically for wafer-level testing, this highly parallel test approach can be enabled by scalable and modular test equipment and an increase in the number of electrical and optical input-output (IO) ports per test site. For electrical measurement instrumentation, this modular approach is already quite well established; for optical instruments this is an emerging concept. In order to increase the number of optical IOs per test site from 10s to 100s of couplers in the next 10 years, multi-core fibers or fiber arrays will have to be used in combination with an optical interposer to reduce the pitch of optical IOs. Measuring optical signals indirectly using on-chip photodiodes is another interesting option to eliminate the need for optical alignment.

In contrast to silicon photonic (SiPh) chips, InP and GaAs chips often require bar/die level testing. The main reason is that often cleaved facets with ultra-low antireflection coating are needed for full device operation. In the future the target will be to replace cleaved facets by on-wafer etched facets and to replace facet coating by on-wafer coating. Then most of the bar/die testing can be avoided and wafer level testing can be used. In the case of SiN devices, it is expected that also in future, die testing will play an important role.

In addition to functional testing, also the area of reliability/lifetime testing needs to be addressed. To date, commercially available lifetime testing equipment is mainly based on fiber-attached mounted single dies or mounted single devices that are being tested with free space measurement setups (e.g. laser lifetime measurements using large

area photodiodes). Such investigations therefore cause high costs. Performing these tests at wafer level, simultaneously on multiple dies, is an option to be investigated. This will also require developments in terms of instrumentation, e.g. high-power laser sources for accelerated lifetime testing.

The IPSR-I Test TWG addresses overall lifetime test issues resulting from the inclusion of photonic capabilities into devices and products. Its emphasis is on wafers and dies with photonic functionality and assemblies and products that include these devices. Systems in Package (SiP), assemblies and bulk systems are addressed to the extent viable given the diversity of test needs that are specific to applications. As shown in Figure 2, the test issues for wafers, die, SiP, and systems will be addressed at the Design, Qualification, Validation, Production, and In-Use stages of product life cycles. Current and anticipated optical parameters to be tested and their value or level are considered along with the test access issue at each stage of the product life cycle.

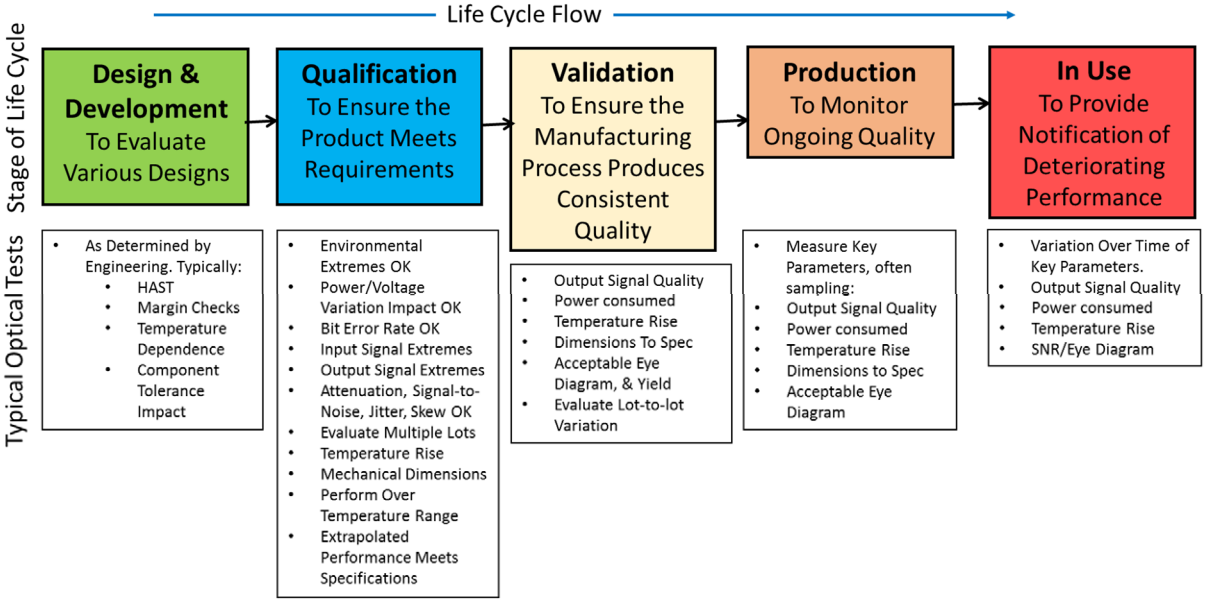


Figure 2. Test Needs During an Optical Product Life Cycle

Telecommunications test equipment, components, and methods were and are being adopted for optical testing of products used for non-long-haul applications. The traditional methods are being extended and new methods developed to address test needs for photonic wafers, photonic integrated circuits, and SiPs that utilize optics and complete systems. Utilizing these extended methods requires optical probing of both wafers and dies combined with electrical probing, resulting in a series of mechanical issues. The inclusion of optical probing, especially single-mode probing, requires gratings or other access points on the wafers. For individual die, dual mode (electrical and optical) probing is especially difficult due to the small size of die and difficulty of holding and locating probes accurately. At the SiP level, the problems are easier because the device is larger, not as fragile, and is often designed to facilitate dual-media probing. The wafer, die, and SiP probe fixtures tend to be expensive due to the complexity and accuracy required. System level test access is usually easier because at that level, electrical interfaces and optical connectors are included as part of the device under test (DUT).

In addition to probe access, optical test methods to simultaneously characterize and compare multiple optical lanes, channels and/or ports at the same time are needed. One need is comparative simultaneous testing of multiple signals from arrays of ribbon fibers, waveguides, sources or detectors for optical skew, jitter, etc. A related need is to simultaneously evaluate optical signals multiplexed on one fiber or waveguide. Applications with arrays exceeding 256 ports (fibers or waveguides) or >256 multiplexed wavelengths are forecast in the next ten years.

In addition to the standard telecom optical parameters such as power, wavelength, attenuation, jitter, signal-to-noise ration (SNR), etc., emerging applications aim to utilize virtually every parameter that light can have, potentially requiring the extension of test capability in multiple dimensions such as polarization, phase noise, amplitude noise, spatial modes, multiple fiber cores, etc. While these emerging needs are potentially very broad, the near-term emerging needs seem most likely to be extensions of data communications needs.

Optical communication applications are likely to utilize 650 nm to 2000 nm wavelengths, multiplexed wavelength spacing down of 25 GHz, detector responsivity of ~1 A/W, receiver sensitivity as great as -45 dBm, power levels of 1 Watt or less, symbol rates of 100 Gbaud per lane, modulation schemes utilizing up to 10 bits per symbol,

polarization multiplexing, BERs of 10^{-12}, etc. Over time, these parameters will improve so test capabilities will need to stay ahead of them. Data rates as high as 500 Tbps per fiber are likely to emerge in the next 10-15 years.

Sensor applications are likely to grow significantly in the next 10-15 years as remote fiber sensors are integrated into physical structures for strain and temperature sensing, and as chip-level chemical and biological sensors are introduced into the marketplace. While these applications will still require the same baseline test solutions as is required for telecom and datacom, the functional tests are likely to be quite different.

Quantum technologies add yet a different dimension for testing. The use of single-photon and entangled-photon sources and circuits will yield its own complexity. There is currently no standardized test equipment for these applications; however, testing methods are currently being developed with University, Government, and Industrial research labs and will require consideration in future editions of this document.

REGARDLESS OF APPLICATION, THE PRINCIPLES OF *DESIGN FOR TEST* REMAIN THE SAME: THE USE OF OPTICAL TEST ACCESS POINTS, BUILT-IN SELF-TEST (BIST), REDUNDANCY FOR SELF-REPAIR, RE-PURPOSE AND PROGNOSTICS TO REPORT CHANGES AND DETERIORATION DURING OPERATION OVER THE LIFE CYCLE OF OPTICAL PRODUCTS ARE DESIRABLE AND OF VALUE IN AN INCREASING NUMBER OF APPLICATIONS. THESE TESTS SHOULD BE CONSIDERED FOR INCLUSION NOT ONLY IN DESIGNS, BUT ALSO IN SOFTWARE DESIGN TOOLS AS WELL.

INTRODUCTION

This Test section focuses on unique attributes of testing optical devices. No attempt has been made to duplicate required and typical electrical or mechanical testing. The chapter is open-ended on optical applications testing with much of the material broadly applicable. It does, however, concentrate primarily on testing data communications products.

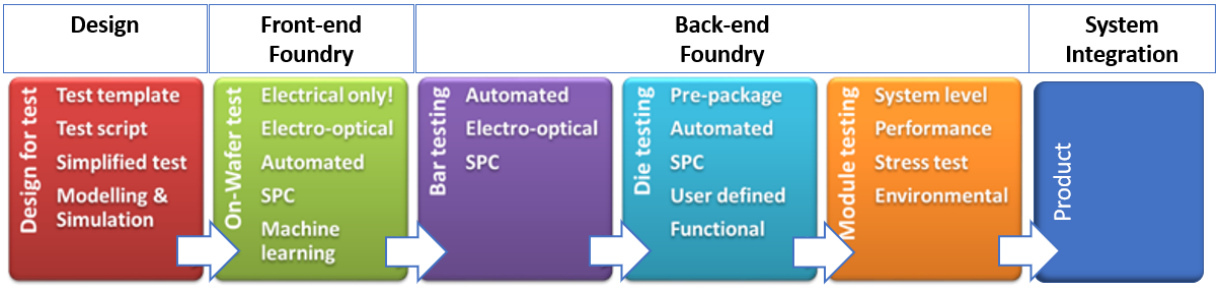


Figure 3. An overview of the PIC production chain for test.

In each step of the test chain that is followed by the components that will form an end product, different requirements and methods are used. This chapter will discuss both the separate steps, and the connection between those steps, regarding the product and data flow.

Areas of testing need during a product life cycle are:

- during development to prove functionality and de-bug devices
- qualification testing
- pre-production validation
- in-process production testing to assure product quality, reliability and to improve yield.

This section contains an overview of PICs made on InP, SiN, SiPh, GaAs, Polymers and CMOS platforms. Elements such as fiber couplers, fiber arrays, lenses, optical and electrical interconnects and the standardization of test port positions (optical, DC, RF) will also be discussed.

The kinds of testing required vary over the life cycle of a product (Figure 2). This figure lists typical optical device test activities and requirements during the life of a device from conception through the in-use and end-of-life phases.

Lanes vs Channels

An optical lane utilizes one wavelength traveling from one point to another in one fiber or waveguide. Information may be imposed on the beam utilizing any methods such as On-Off-Keying, PAM XX, dual polarization or any method that effects only that wavelength.

A Channel may consist of a single lane but often, even usually, has multiple lanes. The lanes may be on multiple parallel fibers or waveguides, or on the same fiber or waveguide utilizing wavelength division multiplexing. Many datacom standards utilize multiple lanes to achieve their data rates.

Evaluating technical capability is most easily done utilizing lanes because channels that combine many lanes make it difficult to understand the underlying technology. System designers, however, find the channel view more useful as the technology details are not important at their level.

ROADMAP OF QUANTIFIED KEY ATTRIBUTE NEEDS

ON-WAFER TESTING

Ideally, the estimation of the yield of a wafer for a particular PIC product occurs on-wafer at the final stage of the fabrication. This way, the number of handling steps can be reduced, and even eliminated (such as the handling of the bars during the bar-level tests). On-wafer electrical and optical testing is possible for PIC products by introducing the grating couplers and/or vertical mirrors into the PIC fabrication process. As a result, the grating couplers and/or vertical mirrors on the wafer need to be coated at the wafer level, and the electrical measurements have to be done in a pulsed-mode operation (preferably with standardized pulse widths). Therefore, for those PIC products that do not require endface coupling, this solution is ideal in that all the relevant testing from PIC performance to far-field measurements of the grating couplers/vertical mirrors can be done in a straightforward manner.

Most optical measurements of devices such as measurement of optical waveguide losses require the use of vertical in/out grating couplers or angled mirrors. Furthermore, on-wafer facet coating of all devices will be needed to significantly reduce all bar and die measurements. For electrical wafer testing, mature electrical tools will be used. The amount of I/O will increase in the future. Today, typically 4 and 16 I/O are used:

Table 1. On-wafer testing

	[unit]	2020	2025	2030	2035	2040
# of Optical I/O	per chip	16	32	64	128	256
Pitch of Optical I/O	µm	256	127	65	20	15
Optical I/O Geometry		Linear Array	Stacked Linear Array	Small diameter or multicore fiber	Multicore fiber	Multicore fiber

By use of fibrer arrays, multicore fibers or different multiple I/O optical probe technology, a parallelization of the measurements will be achieved in the future. To this end, standardized input/output grating couplers will be used to match the chosen optical probe technology. On the other hand, such measurements will be possible only if suitable modular and scalable measurement equipment is available, e.g. mutli-channel and fast high-power tunable laser sources, multi-port power meters, polarization controls, and optical switch matrixes.

The following tables summarize the current status and the expected development of device testing within the next 15 years. Laser diodes (LD), photo-detectors (PD) and Mach–Zehnder Modulators (MZM) are considered for different fabrication technologies as applicable, with Table 2 for silicon, Table 3 for InP and GaAs and Table 4 for SiN.

Table 2. Wafer level high throughput testing -Si-Photonics

Wafer level high throughput testing -Silicon-Photonics	[unit]	2020	2025	2030	2035	2040
Propagation loss ex. Si-wire, polymer, and rib waveguides (measure some waveguides per chip)	[dB/cm]	30 min/chip Semi-automated	5-20 min/chip* Fully automated	5-10 min/chip with 2-4 probes (1-5 min/2-4 chips)	2-5 min/chip multi 4-8 probes (0.5-2 min/4-8 chips)	
Insertion loss ex. modulators (measure some devices per chip)	[dB]					
Spectrum (Operating wavelength) ex. Grating couplers, Directional couplers, Ring resonators (measure some devices per chip)	[dBm/nm]					
Contact resistance	[Ω cm^2]	use current/mature electrical tools				
Sheet resistance	[Ω/sq]					
Vπ L or (Vpi) ex. MZI/Ring switch (heater), MZI/Ring modulators (pin/pn junction)	[V cm (or V)]	20 min per device or circuit semi-automated	10 min per device or circuit semi-automated	5 min per device or circuit fully automated	2 min per device or circuit fully automated	1 min per device or circuit fully automated
f_{3dB} (EO bandwidth) ex. MZI/Ring modulators	[GHz]					
Responsivity ex. Ge photodiode	[A/W]					
Dark Current ex. Ge photodiode	[nA]					
f_{3dB} (EO bandwidth) ex. Ge photodiode	[GHz]					
Eye pattern ex. MZI/Ring modulators, Ge photodiode						
BER (Bit Error Rate) ex. MZI/Ring modulators, Ge photodiode						
environmental testing ex. temperature -40 to +85 ℃.		Adapt current/mature environmental tools				

* not including wafer load/unload time (Front Opening Universal Pod (FOUP) or Front Opening Shipping Box (FOSB) to testing state to FOUP/FOSB.

Table 3. High throughput testing of InP and GaAs devices

Parameter	Unit	Test Level	Test Time Goal
output power – laser diode (LD)	[mW]	bar level	1 min/device
threshold current LD	[mW]	bar level	
SMSR LD	[dB]	bar level	
RIN LD	[dBc/Hz]	bar level	
Frequency noise spectrum LD	[Hz^2/Hz]	die level	
Frequency response LD	[GHz]	die level	
Extinction Ratio LD	[dB]	die level	
Sensitivity MZM	[$V_{2\pi}$]	bar level	2 min/device
Modulation bandwidth MZM	[GHz]	bar level	
Insertion loss MZM	[dB]	bar level	
Responsivity PD	[A/W]	wafer level	1 min/device
Bandwidth PD	[GHz]	wafer level	

Table 4. Wafer level Inspection- SiN devices

Parameter	**Unit**	**Test Level**	**Inspection Time Goal**
Layer stack quality	Particle density	Experienced eye	5 min
Lithography quality	Resolution test & defect count	manual	5 min
Etching quality	Compare with litho check	Microscope manual	5 min
Top cladding quality	Visual defect check	Microscope manual	10 min
Actuator check	Visual defect check	Microscope manual	5 min
Wafer level bump check	Visual defect check	Microscope manual	10 min

BAR/DIE TESTING

In contrast to Si-photonic devices, InP and GaAs chips often require on-bar/die testing (see Table 5). The main reason is that often a cleaved facet with ultra-low antireflection coating is needed for full device operation. In the future the target is to replace cleaved facets by on-wafer etched facets and to replace facet coating by on-wafer coating. Then most of the bar/die testing can be avoided and wafer-level testing can be used.

In the case of SiN devices, it is expected that, also in the future, die testing will play an important role.

Table 5. Bar/die testing InP and GaAs devices distinguishing serial (single device at a time) and parallel testing.

	[unit]	2020	2025	2030	2035	2040
output power LD	[mW]	serial	parallel	parallel wafer level test		
threshold current LD	[mW]	serial	parallel	parallel wafer level test		
SMSR LD	[dB]	serial	parallel	parallel wafer level test		
RIN LD	[dBc/Hz]	serial	parallel	parallel wafer level test		
Frequency noise spectrum LD	[Hz^2/Hz]	serial	parallel	parallel wafer level test		
Frequency response LD	[GHz]	serial	parallel	parallel wafer level test		
Extinction Ratio LD	[dB]	serial	parallel	parallel wafer level test		
Sensitivity MZM	[$V_{2\pi}$]	serial	parallel	parallel wafer level test		
Modulation bandwidth MZM	[GHz]	serial	parallel	parallel wafer level test		
Insertion loss MZM	[dB]	serial	parallel	parallel wafer level test		

Table 6. Die testing SiNx devices

Die testing	[unit]	2020	2025	2030	2035	2040
Waveguide defect check	# defects/chip % of defect chips yield	End facet incoupling of light and VIS & IR monitor on top of chip- Manual place, Auto align. Operator required.	Auto place & align and auto record. No operator. Manual analyze. Obtain statistics.	Full automation of acquisition and analysis		
Waveguide loss check	dB/ cm, loss spectrum	Manual place & Auto alignment to loss test structures. Operator required.	Auto place & align & auto measurement. No operator. Manual analyze. Obtain statistics.	Auto place Auto align & measure Auto analysis		
Building block test (DC, MMI, RR, MZI)	Insertion loss, splitting ratio, FSR,	Manual place & Auto alignment and measure. Operator required.	Auto place, align & auto measurement. No operator. Manual analyze. Obtain statistics.	Auto place Auto align & measure Auto analysis		
Actuators	# Defects / chip Sensitivity ($V_{2\pi}$)	Manual place & Auto alignment and measure Operator required	Auto place, align & auto measurement No operator. Manual analyze. Obtain statistics.	Full automation of acquisition and analysis		

Life cycle testing (Mounted dies but also earlier)

To date, commercially available lifetime testing equipment is mainly based on fiber-attached mounted single dies or mounted single devices that are being tested with free space measurement setups (e.g. laser lifetime measurements using large area photodiodes). Such investigations therefore result in high costs. For some of the test conditions (e.g. high temperature, high humidity) it is not clear how the materials (e.g. epoxies) used for fiber attachment will behave, so we actually need to study (1) the reliability of the fiber attachment and (2) the reliability of the opto-electrical components, and make sure one reliability aspect does not affect results for the other.

In the future, a parallelization of these lifetime measurements will be tackled, operating a large number of devices at wafer scale at the same time avoiding the mounting of single dies. As to the testing of e.g. photodiodes or waveguides, high-power laser sources are needed (especially considering parallel testing and thus splitting the laser output over multiple channels). However, for lifetime testing of the devices, including the mounting/bonding process, testing at the die level will be required.

Table 7. Life cycle testing InP and GaAs devices

	[unit]	2020	2025	2030	2035	2040
Optical output power	[mW]	established				
Threshold current	[mA]	established				
Optical wavelength	[nm]	none	started	established		
SMSR	[dB]	none	started	established		
Responsivity	[A/W]	none	started	established		

Testing parameters of production processes

To guarantee stable and high-yield production processes, it is a stringent requirement to test a large variety of different processing parameters while processing the wafers. These parameters to be checked still strongly depend on the devices fabricated. However, with a view to a steadily increasing monolithic integration depth, the parameters to be checked per wafer will converge in the future. The most important processing parameters to be checked during the processing are:

- etch depth uniformity: nm resolution, over 1 micrometer etch depth
- waveguide sidewall roughness: sub nm resolution, over 1 micrometer depth
- layer thickness and refractive index: $1*10^{-4}$ accuracy, versus speed of measurement
- waveguide width measurements: nm resolution
- resistance measurements of heaters
- conformal filling measurement
- photoluminescence: micrometer resolution
- particle analysis
- imbalance in coupler structures (trimming possible)
- atomic force microscope to determine grating depths (trimming possible)
- resist thickness and structure width
- doping levels

Furthermore, several components can be measured during wafer processing before finalizing it. This characterization can be done by electrical/optical probing:

- photodiodes
- lasers
- heaters
- separation / isolation resistance

Table 8. Testing production processes for polymer devices

	[unit]	2020	2025	2030	2035	2040
Dimensions (line width, hole diameter)	[nm]	1 nm (use mature CD SEM tool)				
LER (Line Edge Roughness) *1	[nm]					
Film thickness	[nm]					
Alignment precision of lithography (especially, pn ion implantation for optical modulator)	[nm]	Use mature optical overlay equipment				
Passive optical properties *2 Propagation loss	[dB/cm]		5-20 min/chip Fully automated	5-10 min/chip (parallel) 2-4 probes (1-5 min w/2-4 chips)	2-5 min/chip (parallel) 4-8 probes (0.5-2 min/4-8 chips)	
Insertion loss	[dB]					
Spectrum	[dBm/ nm]					
Qualification of polymer materials	°C	Manual	Manual	Automated		
Active and passive polymer curing, poling, Teng-Mann testing control		Manual	Manual	Automated		
Poling measurements		Manual	Manual	Automated		
Electrical probe and contact to polymer materials		Manual	Manual	Automated		
EO polymer r33 performance (EO activity in material)	pm/V	Established	Mature	Mature		

*1. LER: Si top LER can be measured, but side wall roughness cannot be measured directly. Correlation between CD SEM and optical properties should be investigated using image analysis and empirical approach.
*2. Contactless and non-destructive inline optical testing (with no particle pollution)

Table 9. Testing production processes for InP and GaAs devices

	[unit]	2020	2025	2030	2035	2040
etching depths	nm	Established				
waveguide width	nm					
resistance measurements	Ω					
on wafer facet coating reflectivity	%	No current capability	No planned capability	Planned Capability	Implemented	Established
thickness and refractive index of deposited dielectric	nm	Established				
thickness of overgrowth layers	nm					

A projection of the key industry needs is shown in Table 10

Table 10. Key challenges with respect to test between 2020 and 2040

	2020	2025	2030	2035	2040
Adopt semiconductor EIC industry test practices					
Test procedures from custom to standardized					
Standardization of test structures					
Test data exchangeability and analysis					
Technology agnostic testing					
Test automation					
Design for test					
Application agnostic testing					

Red: Not current industry practice; **Orange:** Partial industrial coordination; **Yellow:** Significant industrial coordination and compatibility; **Green:** Established Industry standard.

Each category is broken down in more specific subcategories in the following tables, following the same roadmap guidelines. Each table addresses areas such as key challenges, test practices, transition from custom to standardized procedures, transfer of data, adopting semiconductor test practices, and design for test both at the die level and the software level. The tables show competences going out beyond 5 years and emphasize relative strengths for each area.

Table 11. Adopt semiconductor EIC industry test practices

	2020	2025	2030	2035	2040
6 Sigma methodology					
Documenting and reporting					
The same metrics but methods may vary					
Optimized test at wafer-level					
DC testing in electrical – electrical domain					
Revised accept-reject methodology					

Table 12. Transition from custom to standardized procedures.

	2020	2025	2030	2035	2040
Standards instead of custom approaches					
Prioritize tests across full PIC value chain					
Testing metrics					
Relevance of a test					
Standardized test structures					

Table 13. Transfer of test data across the PIC value chain

	2020	2025	2030	2035	2040
Implementation in PDK **Improved design tools (EPDA)**					
Correlation of the test outcomes **Improved processes** **Identification of redundancies**					
Accessible scope – potential IP issues					

Table 14. Technology agnostic testing

	2020	2025	2030	2035	2040
Across (currently) major technologies InP, SiPH SiN, Electro-Optic (EO) polymers					
Open for emerging platforms polymer, diamond, rare earth ion doped, three-dimensional (3D) PICs, SoC (high temperature)					
Hybrid integration photonic cross platform electronic-photonic chip level (EPICs) electronic-photonic PCB-chip					
Testing PICs with CMOS circuits/testing					

Table 15. Automation of test at wafer, bar, die, module and system level testing

	2020	2025	2030	2035	2040
Wafer - level					
Bar and die – level testing					
Standard test interfaces (layout templates)					
Technology agnostic					
Scalability					
On-chip self-diagnostics (Utilizing electrical-to-electrical testing)					

Table 16. Design for test

	2020	2025	2030	2035	2040
Test oriented layout templates					
Implementation in PDKs					
Test scripts for generic die testing					
Training of PIC designers					

SITUATION ANALYSES

MANUFACTURING PROCESSES

Design

Testing is a time-consuming and complex task, and a significant contribution to the final cost of a device, but essential at different stages in the development of a photonic integrated circuit. Testing aims to verify fab process tolerances, validate foundry manufacturing, extract building block parameters, and analyze system-level performance of overall circuits. Since there is a sheer variety of PIC designs, it is very complicated to have a generic or application-independent testing procedure.

Testing starts at the design stage. On one hand, foundries design their own test cells for process control monitoring (PCM), containing specific structures to verify and validate the wafer fabrication, ensure basic building block performance, obtain statistics for PDK maturing, and estimate yield. On the other hand, designers should also include smart and proper test structures in their designs, not only to ease the testing task later on, but to verify complete design functionality, extract additional parameters, and to corroborate data provided by foundries. However, there are currently no standards for these test structures, so it usually depends on the designer's experience and proper communication with testing engineers.

In a view of the PIC production chain after the design stage, as depicted in Figure 3, actual testing can be divided into a number of categories: wafer level, bar level, bare die level and package level.

There are three main drivers to develop integrated photonic solutions, particularly in telecom and datacom applications. The first driver is that the cost of integrated photonics modules tracks much closer to the cost of traditional electronic ICs and are a fraction of the cost of the traditional optical solutions. The second driver is that

Integrated Photonic modules are also much lower power and smaller. The third driver is that integrated photonic modules offer much more in terms of scaling towards ever growing demand for bandwidth and speed.

Commercial companies are designing and manufacturing integrated photonics modules for all of the above reasons. Companies must be able to sell these modules while maintaining acceptable profit margins as expected by the financial markets. To obtain these margins, yields need to be high and the cost of test low. This applies to the cost of the needed test equipment for both validation and production.

The end-to-end cost to manufacture, assembly and test a complete integrated photonic module must be taken into account when designing the PIC, the electrical ICs and picking the composite components. Thought needs to be given to how much electrical testing can be done; that will predict the functioning of the optical waveguides, and whether there are optical wafer acceptance test parameters that will also predict functioning devices.

Automated test equipment (ATE) for optical solutions is difficult and expensive to implement and may not well represent the performance of the final integration of a laser and other components into the modules. For example, if the electrical testing of the PIC can be 95% accurate to predict good die, the cost calculation may show that it is actually cheaper to throw away bad modules built with the 5% bad die, than to pay for ATE for optical test. It is recommended that commercial photonic companies do an end-to-end cost and yield analysis of each test step vs the cost of test equipment to understand the most cost-effective test solutions.

It is very important to take the data collected across the production chain into account during the design phase and the complete process as depicted in Figure 4. Test results are collected and analyzed by designers and manufacturing engineers at each individual step. Although the nature of test at each step differs, a data flow allowing for open exchange between all stages is crucial for development of correlations and optimization of test processes.

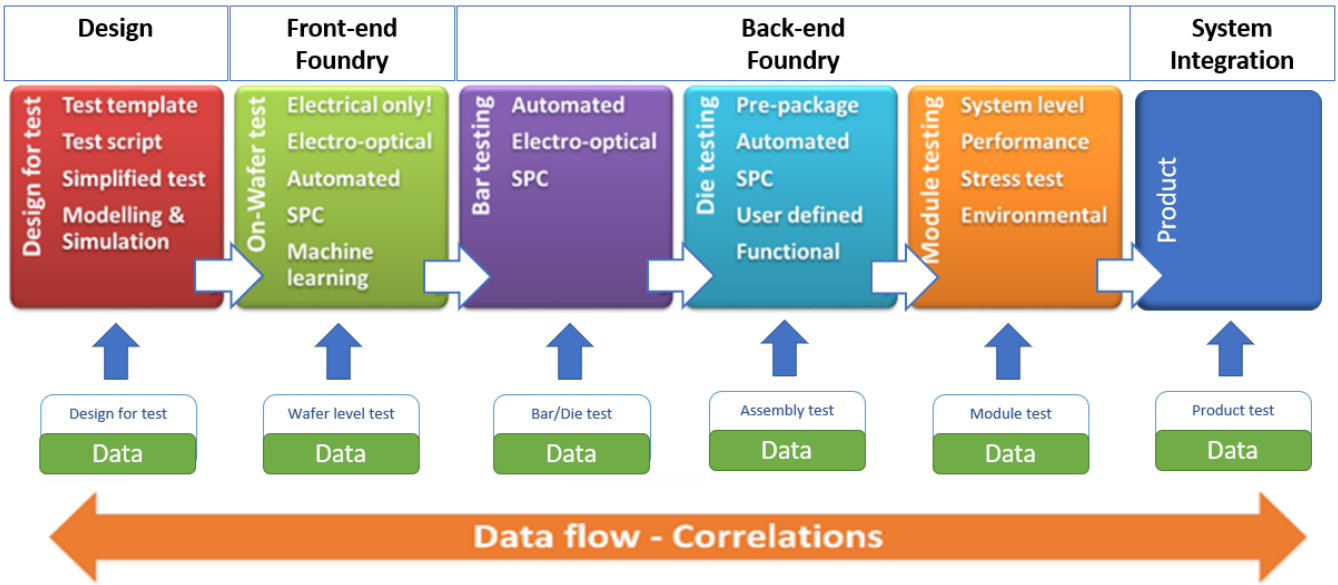


Figure 4. Test framework for data flow and exchange across production chain from a design to product (www.openepda.org).

Test frameworks (www.openepda.org) utilizing open standards assure data interoperability, exchangeability and traceability, and inherently enable such data flow. Such a test framework is hardware agnostic and can be deployed around any tool and equipment configuration. It decouples operators from low-level hardware control, allowing for dynamically adjusted test sequences and customized analysis modules. Machine learning techniques may further augment the capabilities of such an approach by increasing efficiency of data analysis and decision making.

The demand for test differs at different development stages of the manufacturing process, as suggested in Figure 5(a). It is particularly important at the initial start-up phase and for changes as the manufacturing process develops. The rate of this change is different at each phase and is inversely proportional to the yield. The cost of test is proportional to the product of the number of items tested, frequency of test events and the production volume. Also, since the cost of test must be borne by the passing devices, the cost of test per good-device goes up significantly as the yield drops (Figure 5 (b)). The long-term cost-of-test trend follows the demand for test and gradually decreases as the manufacturing process matures.

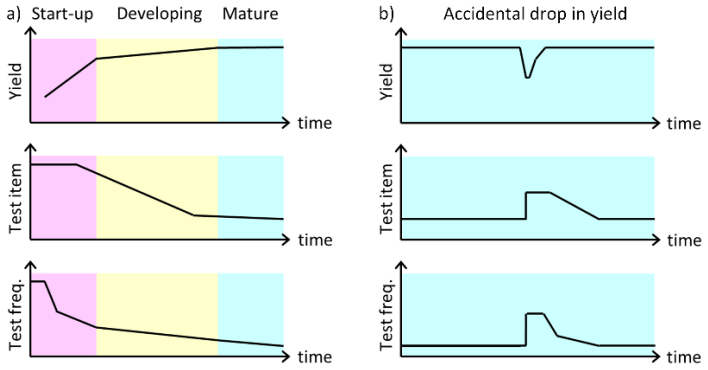


Figure 5. (a) Demand for test at different development stages of production process over time in perspective of yield, number of tested items and frequency/number of test cycles of test events. (b) Disruptive event in yield with an impact on number of tested items and test frequency.

In order to manufacture efficiently, a certain level of test for screening purposes has to be considered and put in place at different points of the production chain, as presented in the Figure 6.

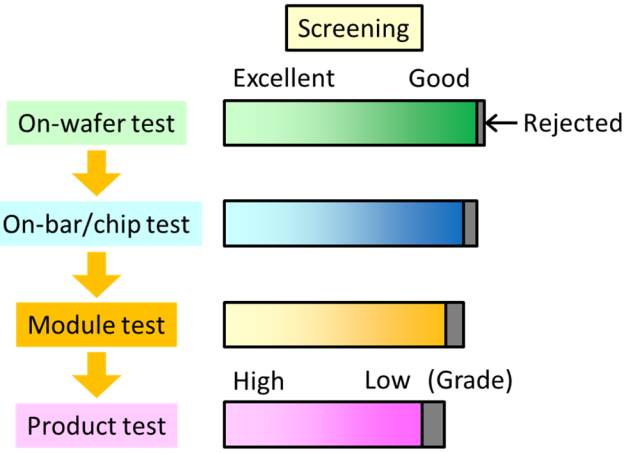


Figure 6. Test for screening is essential for manufacturing products efficiently.

The widespread use of test processes across the full production chain results in massive test data being collected in the manufacturing process. The test strategy for maintaining high yield, screening, etc. is determined based on such massive test data and it is very important to establish the best test strategy. This will remain challenging because the best test strategy varies in each situation.

Wafer level Testing

Wafer testing is mainly carried out by front-end foundries and test houses, and provides both visual and electrical and/or optical testing.

Within the foundry during wafer manufacturing, both visual inspection and quantitative measurements are performed at each step in the fabrication process. The primary goal here is to examine the wafer for irregularities in the process. For example, film thicknesses, etch depths, and line widths can all experience process variations. Most of these measurements are part of the standard suite of inspection methods for electronic ICs. Visual inspection is still manual much of the time and is usually executed only in limited regions of the wafer, requiring machine vision systems providing enough resolution, multi-axes positioners, and capability to inspect different wafer sizes and materials. Once the fabrication process is complete, and before singulation, final wafer metrology is carried out.

Optical microscopes are useful for obtaining physical dimensions of the fabricated structures, layer thicknesses, surface roughness, and (for III-V materials) can also create photoluminescence maps, giving information of material composition and quality, bandgap determination, etc. Modern digital microscopes have sophisticated stitching and profilometry algorithms that can provide an accurate picture of trench depths and waveguide path accuracy over relatively large areas.

Scanning electron microscopy (SEM) testing has always been an important technique at the wafer level. However, integrated photonics circuits add a level of difficulty to SEM testing because of the insulating oxide layers both above

and below the silicon and silicon nitride waveguides. These oxide layers result in charging; without a conductive coating on the wafer surface, an oxide-coated PIC may yield a distorted image.

Since defects/irregularities detection at this early stage of development will have a great impact on the final cost of PIC development, it is necessary to develop test pattern-recognition algorithms for the visual inspection and metrology and formulate a pass or fail criteria depending on technology (InP, SiN, SiPh, etc) and structures.

At the wafer level, along with visual inspection and taking advantage of the pre-existing equipment for CMOS technology, electrical testing can also be carried out. As a matter of fact, at this level, electrical measurements are easier than optical ones, because at the wafer level, coupling light to or extracting light from individual dies is a challenging task given the tight tolerances required. Electrical measurements can be carried out much faster and at lower equipment cost.

The straightforward approach for carrying out optical tests is by means of vertical grating couplers. However, these structures can only be implemented in a few technologies (e.g. SiPh, SiN) and have limited spectral bandwidth. Thus, different structures/approaches have been proposed to overcome these limitations, such as reflective 45º mirrors, evanescent coupling, etc., or by means of indirect electrical measurements (Contactless Integrated Photonic Probe, CLIPP) when possible.

Still, optical testing is quite challenging. Wafer prober stations are difficult to interface with optical probes, so both need to be adapted/designed according to the requirements, and this makes the resulting solutions not very flexible, increasing the throughput but limiting the re-usability for different applications.

Furthermore, optical probes require (semi-)automated alignment, along with very high resolution.

Bare die level Testing

Ideally, foundries should provide both parameter values and their expected statistical variation to end-users; these are usually built into the parametric models used to describe each building block in the photonics process design kits. However, due to all of the challenges associated with optical testing at wafer level, bare die testing is still carried out to extract most building block parameters.

Bare die testing typically has low levels of automation and is therefore costly and complex. Thus, only a small subset of tests and structures are measured by the foundry, and it is up to test houses or end-users to carry out more complex, extensive characterization of individual building blocks, or system-level functionalities of the overall circuit.

It is worth noting that many measurement procedures for PIC testing are not yet standardized. This is true both for testing of individual building blocks (sub-circuits and devices) as well as for full, functional circuits. This often leads to a wide variety of test setups (leading to different measurement results) between foundries, test houses, designers, and end-users. Defining standards/procedures at this early stage of the technology may seem too ambitious, but instead what is proposed, in the short term, is to define good practices.

Generic Photonic Device Testing

Figure 7 illustrates the general test requirement – the need for both electrical and optical test inputs, and then analysis of the electrical and optical outputs from the device under test (DUT). In addition, environmental parameters, such as temperature, humidity, vibration, etc. may be test inputs. Finally, in addition to the optical and electrical test responses, physical factors, such as temperature rise, may be outputs that are monitored during testing.

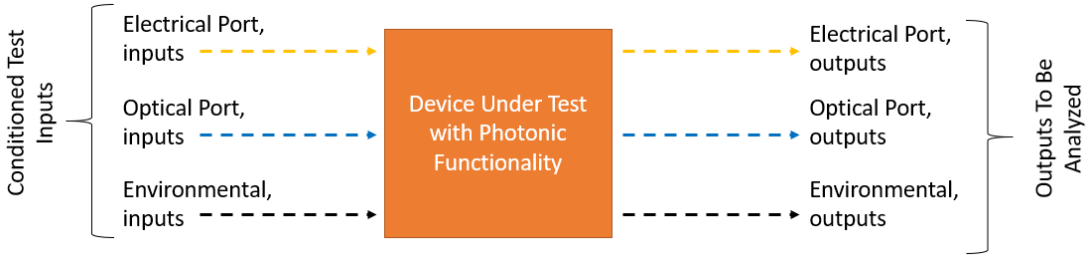


Figure 7. Generic Photonic Product Test Environment

The electrical ports are electrical contacts, or arrays of contacts, for power, control, monitoring of functionality, and, of course, data inputs and/or outputs.

The light directed to an optical input port may be in the form of one or more optical beams or it can be guided by one or more fibers. In either case, it is typical to arrange the parallel inputs in a regular array, the size of which must

be considered in the PIC design. The light can either be a constant source (that is then modulated by the device under test) or it can carry a modulated data stream (if the intended function of the device is to de-modulate the data). The optical interconnections may be in the form of butt/edge coupling, a grating coupler, or an evanescent coupling resulting from proximity of parallel waveguides. Making these test connections, especially the optical connections, is frequently a major project.

The photonic input and output signals may have multiple parameters needing measurement. Specifically, optical signals may require measurement of the following properties of individual sources, lanes or channels:

- Power
- Polarization
- Direction
- Mode profile
- Wavelength
- Variation in any parameter (e.g. power, wavelength, polarization) over time
  - modulation (fast)
  - drift (slow)
- Skew between beams
- Component loss
- Return loss

System (functional) properties:

- Signal-to-Noise Ratio (SNR), Relative Intensity Noise (RIN), Phase-Noise, Crosstalk
- Bit error rate (BER)
- Variation of system figures of merit with temperature and other environmental inputs

A typical measurement requires a source with a means of controlling one or more of these parameters. The source will typically be coupled into a single-mode fiber that, for some systems, must be polarization-maintaining. The source can be tunable (in cases where the wavelength dependence is part of the measurement), narrowband (for monochromatic measurements) or broadband in certain sensors applications. The light is injected into waveguides via optical ports and routed (according to the PIC design) through one or more active or passive components. Active components typically require one or more electronic drive signals. These components (the devices under test) will modify the input optical signals in a manner determined by the magnitude, phase, and frequency of the electrical drive signals. The resulting electrical and optical signals are on the corresponding output ports.

Environmental inputs may include all of the usual variables: temperature, humidity, temperature cycling, Highly Accelerated Stress Test (HAST), vibration, shock, etc. Physical outputs include temperature rise, mechanical changes such as delamination, cracking, swelling, wire breaks or optical chain interruptions, etc.

Photonic test requirements vary by the test level (wafer, die, photonic SiP, system) and test need (access, sources, detectors, functions). Table 17 gives a generic view of the testing needs for items containing photonic elements at the various levels. A similar table can be developed for specific applications to provide some insight into the related requirements for each application.

Table 17. Photonic Test Requirements

	Test Need			
Test Level	**Optical Access**	**Sources**	**Detectors**	**Functions**
Wafer	45° mirrors, vertical grating couplers, etched factes, cleaved fiber, tapered fiber, lensed fiber, focused free-space beam, evanescent coupling	External sources injected via fiber or free-space access Integrated sources	External photodetectors, potentially in arrays; Imaging sensors Optics to collect and/or image light to be detected. Integrated photodetectors	Wide variety of device characterization and functional tests; media loss/cm, insertion loss, modulation depth/bandwidth, polarization control, wafer uniformity, detector sensitivity/responsivity, temperature sensitivity, die-to-die variation, skew between outputs
Chip	The same as wafer options plus edge coupling to embedded or surface waveguides.	The same as wafer options	The same as wafer options	The same as wafer options plus edge coupling impacts on loss, spectral bandwidth and polarization
Photonic SiP	Butt coupling or expanded beam connector, evanescent coupling, fiber splice, ribbon fiber splice	External or on-chip laser source to simulate application related requirements.	External detector or detectors, potentially in an array, gathering light from an edge emitting waveguide or vertical emitting 45° mirror, or vertical emitting grating	The same as wafer and chip options, plus characterize package connections, and application specific tests such as eye diagrams, BER, environmental sensitivity
System	Conventional optical connector, fiber splice	External or internal laser source or sources to simulate inputs.	As needed to measure and evaluate system outputs.	Intensity, skew between lanes, polarization, eye diagram, SNR,
In Use, Over Lifetime	Limited, if any. System dependent access.	Both self and remotely initiated data reporting	Primarily wireless or electronic	Monitor & report performance changes. Initiate self-repair.

Several types of product testing of devices, including those with photonic capability, usually are required:

- Test during development to ensure the design "works"
- Qualification testing, typically done before a product is committed to wide use
- In-process testing to monitor manufacturing process quality
- Final testing before each individual product is shipped to a customer

GENERAL TEST EQUIPMENT

Optical device test equipment is available from multiple suppliers. Historically, the telecommunications industry was the major consumer, but in recent years the use of optical communications for short distances, such as Local Area Networks (LANS), Fiber to the X (Home, Office, Curb, Node, etc. – FTTX), Active Optical Cable (AOC), and in Data Centers, as well as a variety of sensors, has broadened the demand. Much of the demand emerging for these new optical applications is filled by utilizing equipment developed for and derived from that used by the Telecom sector. As these applications grow in importance, specialized equipment is emerging and becoming available.

Devices under test can vary dramatically, so defining all the needed instruments for PIC testing is not the most appropriate approach. Since electrical and optical domains merge in PICs, optical and electrical equipment is needed. In the long run, testing should be as automated as possible, meaning that wafer-level testing should provide known good dies, reducing the fraction of dies that need to be tested at the package level. However, standard wafer probe stations are mainly intended for electrical testing, so the main challenge is to customize the probe station so that it is

equipped to carry out complete (optical and electrical) testing. After dicing, a pick-up element with accompanying alignment metrology is necessary to release and properly place dies for testing on a thermally controlled stage.

CRITICAL (INFRASTRUCTURE) ISSUES

In general, the ambition in manufacturing is to measure as little as is necessary to gain the relevant information on product quality. However, to reduce the assembly-related costs, it is essential to closely monitor the complete fabrication process of the photonic integrated circuits (PIC) or even of the final device. While this may appear to be a very expensive procedure, test and selection after fabrication and assembly is even more expensive. A critical point here is to select the good dies at a very early stage (i.e. the wafer stage) in the manufacturing chain. This is referred to as known-good-dies (KGD).

Beginning with wafer level production, each cleanroom step needs to be measured and validated. Since silicon-photonics manufacturing utilizes CMOS processes, it can benefit from such well-established characterization methods as ellipsometry, atomic-force microscopy (AFM), scanning-electron microscopy (SEM), white light surface profilometry and optical microscopy for layer thickness, critical dimension (CD) and line edge roughness measurements. These methods are well known and well established throughout the industry and, when used appropriately, can guarantee high throughput and high yield. For photonic integration technologies on other substrates and using other material systems, compatibility with CMOS optimized tooling is limited, hence transfer of those test and characterization methods may require additional R&D efforts for adaption to such technology processes. In addition, the electrical tests of the integrated circuits are quite well established. Automated test equipment, test heads, including specially adapted probe cards of different designs in combination with fully automated wafer probers, contribute as well to a high throughput for the electrical measurements in the manufacturing chain of integrated circuits (ICs). An exception to this is the need for electronic probes for very high bandwidth detectors, since it is not yet standard practice to automatically test electronic circuits at speeds higher than 40 GHz.

The equipment infrastructure for photonic measurements is not as mature as that for electrical test equipment. One major challenge for the coming years will be to assure availability of high volume and high throughput die and wafer level (WL) measurement equipment for photonic integrated circuits and electronic-photonic integrated circuits, so-called EPICs. A logical step would be to extend the electronic test equipment to enable photonic measurements using a similar physical architecture. For wafer level measurements, this would correspond to optical solutions within the:

- Wafer prober/wafer handler
- Probe Card
- Automated Test Equipment (Test head or test matrix, measurement instruments)

The wafer prober and wafer handler need to be similar in structure. They handle the wafers, do the x,y,z and theta alignment and step through the devices on the wafer. A feedback from the automated test equipment is beneficial but all current probers are capable of handling such input. However, the prober needs to be optimized for the particular optical input/output ports to be probed. For example, an edge input/output requires an optical fiber probe with a turning mirror, and must be aligned both in angle and position (typically better than 1 μm) to accomplish the coupling. A grating coupler has a more relaxed positional tolerance but requires precise angular control. In both cases, the mechanical requirements are more stringent than is required to probe electronic test pads. For very high-speed applications, the electrical bandwidth of the probe card may need to exceed 40 GHz. When probing detectors directly, it is then helpful to include a transimpedance amplifier in the probe head.

There are currently several suppliers of automated test equipment that provide tools such as high-speed optical power meters or optical network analyzers that allow engineers to develop a measurement setup very similar to the IC test setups. The real challenge is to feed and read out the optical and electrical signals together to and from the device under test. In the case of pure electrical probe cards, the electrical contact needles (e.g. cantilever, vertical or MEMS needles) provide physical contact to the DUT bond pads. In the case of standard IC probers, the alignment tolerances of several micrometers between electrical interface (needles) and device under test are quite moderate.

The optical interface will be realized by a grating coupler on the DUT side and a waveguide of some kind, e.g. an optical fiber, on the prober side. The lateral tolerances for such an optical interface are much tighter and lie in the sub-micrometer range. This gap between the two tolerance ranges will make it difficult to extend the electronic measurement equipment to the photonic measurements or even ensure a compatibility of common IC-Wafer level test equipment and procedures.

It is critical to correlate the physical measurements to both the specific functional measurements and the overall behavior of the photonic integrated circuit.

Some physical measurements (for example, loss measurements, spectral bandwidth, polarization dependence) are very straightforward. Others present a more significant challenge: for example, a suitable measurement of the effective index and group index of the waveguide mode and accurate measurements of the modal profile near couplers are critical measurements that are not yet standardized.

In the future, calibration facilities will be needed to provide calibration targets which can be traced back to a standard. For instance, test PICs calibrated for known component loss, waveguide loss, and dispersion could be made available to equipment manufacturers, foundries and packaging facilities. This is especially important when new materials are introduced; in such cases both the performance from a materials standpoint as well as a device performance standpoint will be vital. Calibration facilities could be e.g. NIST (USA), PTB (Germany) or NPL (U.K.).

TECHNOLOGY NEEDS (RELATED TO MILESTONES)

This section focuses on the technological needs regarding testing tools, methods, measurement time, capacity, speed, and accuracy. For an efficient workflow between the different production partners (e.g. design houses, wafer processing fab, backend and packaging groups and customer), standardized measurement techniques with defined measurement data transfer are a stringent requirement. These topics will be addressed in the following sections.

PRIORITIZED RESEARCH NEEDS

Prioritized research needs that require results in less than five years include the following:

- Ability to test photonic properties of wafers during fab to ensure that wafers are good.
- Processing ever-faster (100Gbps+) data streams. Test time is often determined and limited by memory I/O data rates, so increasing these will remove a barrier to lower cost. Developing test equipment with more capability than the devices to be tested is a continually moving target.
- There will be a need for test platforms that are compatible with the test needs of different applications. It will be therefore necessary to separate, where possible, baseline testing needs from application-specific functional tests. Some examples of baseline testing are: component and connection loss measurements; parametric measurements (e.g. signal vs. temperature, wavelength, polarization); wavelength and wavelength-spectrum measurements; detector efficiency, etc. Some examples of application specific functional tests are: Bit-Error-Rate-Test (BERT); modulation depth and modulation bandwidth; calibration of phase shifters; etc.

PRIORITIZED DEVELOPMENT AND IMPLEMENTATION NEEDS

Eventually, the ability to support 500 Tbps/fiber data transfer rates will be needed. An important issue is the nature of the data stream – how much parallelism, what modulation format, etc.

DESIGN

Standardization of test metrics and qualification

For any desired photonic-IC-based product, there can be numerous different implementations that will meet the same required result. Some of the options to consider are:

- material system (InP, Si, GaAs, polymer, Triplex, glass),
- integration scheme (monolithic, hybrid, heterogeneous, free-space component based),
- packaging (non-hermetic, hermetic, housing material),
- optical alignment (fibers, fiber arrays, lenses, grating couplers, vertical mirrors).

Therefore, standardized testing and qualification is an indispensable step towards reliable photonic-IC-based products independent of the particular implementation that was employed.

The measurement and qualification rules for the electronic ICs are not sufficient to qualify photonic ICs, since these do not consider the optical parameters. Generally, photonic signals are more susceptible to environmental conditions/changes than electronic signals. In addition, sub micrometer alignment accuracy of fibers/lenses to the photonic IC requires the use of mechanical/chemical-based fixtures with different materials with different properties, which makes the performance not only depend on the photonic IC itself, but also is heavily dependent on the type of packaging. As a result, new standardized testing methodologies and qualification parameters need to be devised that apply to all technologies, all types of packages, and all relevant environmental conditions.

In order to monitor the full supply chain of the PIC product, testing has to occur at the following multiple stages:

1. during PIC Design: design of all the structures needed for testing purposes mentioned below, plus a standardized test device to monitor the performance of the PIC along the optical path (such as using a directional coupler to tap a small portion of the light and couple that to a sensitive photodetector or a grating coupler or vertical mirror to see the spectral characteristics)
2. at Wafer Level: measurement of process-related parameters that are relevant to the performance of the devices on the PICs (waveguide width, etch depths, contact and series resistances, grating depths, etc.), wafer-release measurements at the end of fabrication (preferably, but not necessarily, fully electrical using the light sources and the detector devices on the wafer) to determine whether the devices meet the targeted criteria, electrical and optical on-wafer measurements using grating-coupler or vertical-mirror structures for determining/estimating the yield of the wafer (pulsed measurements with a standardized pulse width)
3. at Bar Level: testing of the coating performance, testing of the far-field pattern of the output waveguides, PIC testing using multiprobes/fiber arrays to select the known-good-dies for packaging afterwards
4. at Die Level: selection of the known-good-dies that were not damaged by the singulation process
5. on Packaged PICs: full testing of the PIC performance on the final package with the proper heatsink, burn-in of the PICs, final selection of the PICs that meet the targeted specs
6. during qualification: standardized qualification tests (such as temperature, humidity, shock, vibrations, etc.) for the environmental conditions relevant to the location where the PIC will be deployed

An important element of the standardized testing is the exchange of information between the various parties, and therefore, all data collected along the production line: –design, wafer processing, backend, packaging, system tests – should be stored in a standard data file format to enable an easy and simple exchange of the data between the partners. A centralized data base for all the data is needed to allow for a traceability of the full production processes. Cloud sharing of the data could be a possibility for sharing, provided that IP issues and safety regulation allow for this. The implementation of an industry standard regarding the data format is needed.

Depending on the material system used, standardized test protocols should be implemented that define the testing conditions such as e.g. temperature, moisture, surrounding atmosphere, maximum operation conditions, etc., in order to achieve a comparableness of measurements from different companies. This includes also a definition of the measurement equipment including the required minimum specifications. In some cases, e.g., for optical linewidth measurement, the measurement principle (e.g. homodyne or heterodyne approach) should be defined.

Special test structures must be defined and processed with all wafers (in particular for multiple-project wafer runs) to allow for the tracking of the processing quality. All measurements should be defined in a way that the results will be operator independent. The measurement itself should be as simple as possible with a strong analytical significance. The choice regarding the most suitable measurement method should be driven by cost-awareness and economic considerations. Furthermore, the measurements should allow a correlation between measurements at an early stage and a later stage in the production line.

All measurements/tests in principle should allow for an upscaling for volume and mass production.

Tools and methods

As mentioned above, in most cases special test structures that allow tracking and checking the processing and device quality have to be defined and processed. A generic four step method should be implemented leading to a steady improvement of the processing quality and the device yield:

Step 1. Investigation of many variations:

The processed test structures will be used to measure and evaluate the critical dimensions, such as

1. variations of processes (width, thickness, sidewall angle, etching depth, alignment, etc.)
2. variations of optical and electrical properties
3. variations of module properties

Step 2. Simulation using obtained variations:

If Step 1 is carried out for a large number of processed wafers (>100), designers can take the collected data and simulate the variations of some of the optical and electrical properties using the known variations of the processes, similar to Joint Test Action (JTA) in the electronic world. Thus a correlation between design, process variations and measurement results can be found. This can lead to a redesign of the respective components to increase the functional yield of the devices taking the processing variations into account.

Step 3. Decide on necessary and sufficient testing structures

When required, the designer can develop novel and improved test structures to be implemented on the wafers.

Step 4. Storage of test data, reduction of testing components

By repeating Step 1…3 in combination with a careful analysis of the stored data, the number of testing devices to be measured and tracked can be reduced and in parallel the yield of the functional devices will increase. A close collaboration between the designers and the Fab engineers is required here.

Examples of suitable tools for in-line testing are:

1. critical dimension SEM (line width, hole diameter, line edge roughness)
2. atomic force microscope (line width, etching depth)
3. film metrology equipment (film thickness)
4. step measurement setup (etching depth)
   - ellipsometer (absorption, refractive index, thickness)
   - mass spectrometer in etching systems
5. optical material detection systems in etching systems
6. opto-electrical wafer probe setups for on wafer characterization
7. optical measurement equipment for optical loss measurements
8. photoluminescence measurement equipment
9. x-ray system for strain measurements

Ideally, the goal for the measurement method should be to extract the scattering and transfer charactieristics (s-parameters) of all the devices on the chip. These allow creation of calibrated, compact models representing indicvidiual on-chip functions (building blocks) and enable more accurate simulations of the performance of the full functional chip. The extraction of the s-parameters can only be done by measurement methods such as the optical coherence reflectometry (OCR). InP chips should be able to integrate an OCR on the chip, but for the Si and Triplex technologies it is going to be more difficult.

High throughput testing (sub-second per chip)

For volume production of photonic chips, fully automated testing (e.g. waveguide defect testing, waveguide loss testing, laser performance testing, building block testing, actuator testing, etc.) is required. Automatic testing comprises, on the one hand, inline automated testing at the wafer level, such as e.g. critical dimension SEM, ellipsometric layer thickness measurement or step height measurements. On the other hand, off-line automatic testing at wafer, bar/die and module levels are included. Examples for such offline measurements are frequency response measurements, output power measurements, responsivity measurements, etc.

High throughput, automatic measurements require standardized measurement procedures in combination with a standardized data format to enable an efficient data exchange between designers, engineers and customers. For large volumes (mass production) simple automated measurements will not be sufficient. In this case a massive parallel testing will be required in order to cope with the large number of devices to be tested and to bring the costs for testing down.

Only with massive parallel testing can the average testing time per chip be reduced, which is the main cost driver here. However, in analogy to the electronic world, massive parallel testing will only be applied if the number of chips to be tested has a significant magnitude. Typically, prior to using massively parallel test equipment, test procedures will be developed and tested with a single-device test procedure. To enable massively parallel testing, suitable standardized layouts, tests and package templates have to be developed. Also sacrifice chip area has to be defined if needed. In most cases these areas will be located near the dicing line and at the wafer edges. Most of these measurements will be carried out on wafer level; however, part of the measurements will have to be carried still out on bar/die level.

WORKFORCE DEVELOPMENT

It is well known that testing requires not only appropriate facilities and dedicated measurement equipment, but also experienced personnel. Broadly speaking, PIC testing can be performed by three different agents: foundries, test houses, and finally end users, with each being carried out at different stages of the development process.

The design and test development of regular electrical ICs needs a highly technical skilled workforce. PICs need a similar skillset, with the addition of engineers who have a deep understanding of optical parameters, optical performance in photonic ICs and module design and integration, as well as the manufacturing processes to achieve a

working photonic module and manufacturing processes to achieve high yields. PhD-level engineers and technicians who have an understanding of technology, optical and electrical design and characterization and packaging technologies and the leverage into high volume manufacturing are desired. The constant communication and interaction of the engineers who have these skills is needed to achieve an optimal design that is highly testable in the bring-up lab and on ATE. During the new-product introduction phase, teams of these engineers need to be highly engaged in debug and any redesign required to help achieve manufacturing yield goals. Manufacturing test programs for the optical and electrical parts are highly customized for the application and are ideally done by test engineers who can work alongside the design team during the test development phase. Once done, the test programs can be handed to the high-volume manufacturing test houses, with a clear technical hand-off to make sure the testing is run correctly. Ideally the test house has equivalently skilled engineers who can take over the production testing and be the first-in-line resource for troubleshooting when any production test issues arise.

Testing is essentially manufacturing activity, and thus it requires education and training in a series of disciplines and skills. Table 3 and Table 4 provide some guidance on these needs.

Table 3. Academic Education Requirements

Knowledge Required	**Content**
AC and DC electricity & electronics	Voltage, current, frequency, power, electronics, transformers, capacitors, inductors, transistors, ions, conductors, semiconductors, non-conductors, electrical to optical and optical to electronic conversion.
Basics of Optics	Ray tracing, lenses, mirrors, prisms, wavelength, phase, polarization, intensity, beam divergence, beam focus, optical modes, E and H fields as related to the Poynting Vector, light in fibers, both single and multimode, etc.
Characteristics of Signals	Power, transmitting information, signal to noise ratio, modulation methods including OOK, orthogonal signals, multiplexing, demultiplexing, Shannon Limit, etc.
Basics of Statistics	Gathering data, maintaining integrity, managing data bases, standard deviation, mean, median, Parato charts, statistical process control, control limits, Cp, Cpk, etc.
Measurements	Basics of mechanical, electrical, optical metrology. Repeatability, gage studies, etc.
Financial basics	Basic business financial concepts; revenue, costs, elements of cost, product cost elements, overhead, cash, AR, AP, depreciation, equity, etc. “The $ in must be greater than the $ out”. “We make investments in order to make more money back utilizing the result of the investment,” etc.
Design for Test	Understand and implement desing for test (DfT) concepts. Transfer product and application requirements, boundaries of manufacturing processes into test protocols. Understand implcations resulting from such and account for at the design phase of a PIC product. Master design and simulation practices using software tools and aid the product design team

Table 4. Training Requirements

Skill Require	**Areas of Training***
Personal Behavior	Show up on time. Be prepared to perform your job.
Safety	Rules, behavior, precautions, etc., related to safety for machinery, chemicals, slips and falls, people related, spills, MSDSs, PPE, “see something, say something” rule.
Quality	Follow the rules. Ensure procedures are followed. Go beyond the formal requirements and propose improvements. Follow the Japanese “5S” rules. Follow “Deming’s 14 Rules for Management”. Use statistics to improve yield and minimize variation.
Cost	Why cost is important, sources of cost, minimizing cost, proposing cost reductions, minimizing waste, maximizing reuse and recycling.
Equipment operation	Safe operation, instrument setup, calibration, standard operations, maintaining records, impact of each process on cost, use of the operating manual, machine maintenance.
Metrology	Use of callipers, electronic and optical measurement methods, storage of data and analysis, ensuring accuracy.
Interpreting Instructions	Read what it says, ask question, make sure you understand, do not “assume”, eliminate and resolve ambiguities,
Completing Jobs On Time	Ensure you understand what is required; ensure all of the instructions, materials equipment and other resources are available. Start as soon as possible. Look for potential barriers ahead and ensue they are eliminated. Be prepared to revise your approach. Ask for help. When you make a mistake, admit it, learn from it, ensure you do not make it again. Do not hide your errors.
*While training is often highly specific to each job, basics apply to all jobs.	

GAPS AND SHOWSTOPPERS

STANDARIZATION

Standardized testing metrics and procedures are essential for developing PIC markets further. Some specific killer applications (interconnects, automotive, sensors, etc.) are needed to accelerate standardization. Necessary test items depend on a particular application, and a specific application makes them clear. A promising big market provides a powerful incentive for PIC companies such as PIC device companies, PIC foundries, and PIC testing equipment companies.

Necessary test items should be standardized across the full PIC value chain. Testing designs and procedures are then standardized. The design tools for testing should be implemented in EPDA and PDK. Testing should be accuracy and fast. On-chip self-diagnostics like that of EICs will be needed in the future.

PIC device engineers need to clarify testing equipment specifications (electrical and optical probes, functionalities, accuracy, speed, etc.). They should collaborate closely with PIC testing equipment engineers.

Standardization seems a difficult challenge in this field because it needs many people's efforts and some sufficiently attractive markets. If this challenge is achieved, we will be able to develop various kinds of PIC products with a minimum of effort.

PLATFORM AGNOSTIC TESTING

The basic testing setup is common in a variety of PIC technologies (SiPh, InP, GaAs, SiN, polymer, etc.). Technology-agnostic testing is very important. The standardized testing equipment should be used for a variety of PIC testing with minor modifications. Various PIC companies should cooperate with each other across technical boundaries. The PIC devices are tested at a variety of sample shapes (wafer, bar and die). Sample-shape agnostic testing is also very important.

AUTOMATION

Fully automated PIC testing equipment is essential for developing PIC markets further. Mature EIC industry test practices should be emulated, and original PIC industry test practices should be developed. Various types of fully automated transceiver testing (OOK, PAM4, QPSK, 16-64QAM, etc.) will be needed.

HIGH SPEED (RF BANDWIDTH) TESTING

PIC testing equipment must measure both low-speed and high-speed properties. Fully automated high-speed test (>10-100 Gbps) at wafer level are not easy. Probe contact becomes critical. We have to investigate which of the commonly used probe technologies for high-speed electrical wafer level testing are compatible with optical testing (e.g. some probe cards will make it difficult to probe optically).

OPTICAL TESTING FOR MANUFACTURE

Contactless and non-destructive inline optical testing equipment with no particle pollution, which is acceptable for a PIC fab, will be needed. The inline optical testing can improve a product yield.

USER SUPPORT

User-friendly GUIs and a variety of testing scripts are needed. PIC tests are generally difficult because electrical and photonic knowledge are needed. Helpful training manuals and courses are necessary.

ANALYSIS OF TESTING RESULTS

We have to research relationships between testing results at each level and product performance. A PIC accept-reject methodology should be established for each product. For example, one faulty sub-system does not necessarily disqualify functionality of the full circuit. In addition, statistics and analysis of testing data should effectively be transferred across the PIC value chain.

COST

Fully automated optical and electrical testing equipment will be very expensive. We should share expensive testing tools based on standardization and platform-agnostic testing. Testing time (including setup, calibration, wafer load and unload etc.) should be short enough because time is money. But testing should be accurate enough.

We have to make the best use of testing results to achieve a good product yield and high product performance. The testing results should also be used to revise a product design and develop new products with much higher performance.

HIGHER PIC TECHNOLOGIES

Some specific applications help to solve the above problems. Higher PIC technologies are necessary to realize such applications. For example, low-loss propagation, low power consumption and high-speed optical modulation, photo detection and amplification, high temperature stability, high r33 materials etc., which translate into high performance, will be expected in SiPh, InP, GaAs, SiN, polymer, etc.

- The 50 GHz barrier resulting from conventional CMOS capability forcing parallel solutions rather than higher baud rates.
  Low speed of suitable assembly, test and other process equipment resulting in high costs.
- Inability to overcome the cost-driving, rate-limiting step/bottleneck of manufacturing/testing such as the number of assembly steps or length of time to perform test, especially BER testing. "Time is money."
- Limits resulting from adapting existing equipment, materials and methods to optical test as more specific equipment is not available. Currently the demand for such specialized equipment is not sufficient to incentivize equipment manufactures to make it available due to high non-recurring engineering (NRE) costs and low return on investment.
- Designing for Manufacturing and test:
  - Maximizing output to reduce cost
  - Studying designs to trade off accuracy and speed
- Inability to utilize materials or processes due to environment-related constraints (RoHS, REACH, WEEE, etc.)

RECOMMENDATIONS ON POTENTIAL ALTERNATIVE TECHNOLOGIES

0. Silicon waveguides to 1D/2D photonic crystal waveguides or plasmonic waveguides. Some devices become much smaller (leading to higher-density photonic integrated circuits).
1. Combinations of active and passive polymers for alternative Silicon (and other) PIC designs and automated test, calibration and verification procedures.
2. Utilize laser processing to make optical waveguides in-situ for effective optical connections and optical structures.
3. Utilization of plasmons to minimize size and maximize functionality.

Contributors

Tom Brown, University of Rochester - chair
Dave Armstrong, Advantest - chair
Sylwester Latkowski, Photonic Integration Technology Center, Eindhoven University of Technology - chair

Robert Pfahl, iNEMI	Graham Reed, University of Southampton
Dan Evans, Palomar Tech.	Richard Otte, Promex Industries Inc.
Bill Bottoms, 3MTS	Lionel Kimerling, MIT
Makoto Okano, AIST	Martin Möhrle, Fraunhofer HHI
Tobias Gnausch, Jenoptik	Michael Lebby, Lightwave Logic Inc.
Jeroen de Coster, IMEC	Chris Roeloffzen, LioniX International
John MacWilliams, Consultant	Willem Vos, University Twente
Chris Coleman, Keysight	Jan Mink, VTEC
Keren Bergman, Columbia University	Gordonn Liu, Huawei
Sam Salloum, Tektronix	Jan Peters Weem, Tektronix
Scottie Wyatt, Tektronix	Iñigo Artundo, VLC Photonics
Rocio Banos, VLC Photonics	Michael Garner, Stanford University
Robert Polster, Columbia University	Philip Schonfield, RIT
Ignazio Piancentini, Ficontec	Shangjian Zhang, UEST
Eugene Atwood, IBM	Zhihua Li, IME
Yi Zhang, Teradyne	Zoe Conroy, Cisco
Fen Guan, Global Foundries	Carl Buck, Aehr Test Systems

Tsuyoshi Horikawa, Photonics Electronics Technology Research Association (PETRA), Japan

We would like to thank all who contributed to the Test TWG, both in workshops, on-line conferences and with written comments.

Section 3: Logic Device Testing

Executive Summary

Logic testing of complex digital functions is a key focus area during integrated circuit manufacturing. Since the introduction of scan in the 1970's we have faced many challenges to achieve our desired product quality. As an industry, we addressed the challenges of logic testing with multiple technologies such as LSSD and multiplexer-based logic scan, Automatic Test Pattern Generation (ATPG), scan compression, logic built-in self-test, boundary scan, core wrapping, testpoint insertion, and the recent addition of the cell-aware fault model. In the 2019 HI Roadmap[11] we highlighted that test times have continued to increase and that new opportunities for logic test innovations and optimizations are emerging. In this update, we will start to discuss these challenges and some of the technologies that are being considered to address them.

This update is organized into the following sections: Dynamics of pattern growth; Logic test no longer dominates total test time; AI devices and logic test; Functional test resurgence; In-field testing; Optimizing impact of test insertions; and Interface challenges. The intent is to give the reader a preview of the challenges and opportunities that will be discussed in the major 2021 update, which will also include updated modeling for memory and test time.

DYNAMICS OF PATTERN GROWTH

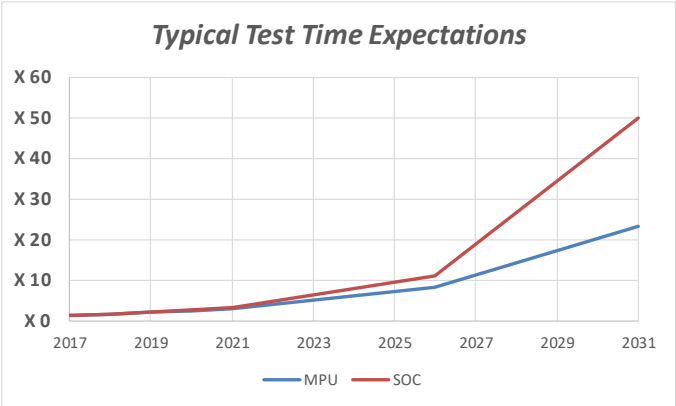


Figure 2: 2019 HIR prediction of test time growth

Previously in the HI Roadmap we have explicitly tackled the pattern growth problem first, and then translated that into a test time increase given a fixed test interface bandwidth. This worked well because the die interfaces were directly and synchronously coupled to the logic under test. Going forward, due to the increased adoption of DFT techniques (e.g., scan compression), this relationship between test access and logic under test is changing. As such, we will need to address test data volume separately from test time. It should also be noted that the trends outlined in the 2019 roadmap do not account for new fault models and will cause a 30-70% pattern count increase given the same design. In the roadmap tables we outline the growth of transistors and IP blocks, and how that maps to an increase in design complexity. Heterogeneous integration may drive an increase in the pattern count for logic test due to an increased gate count being accessed through a reduced test interface.

As the industry moves towards smaller technology nodes, the standard logical fault models (stuck-at and transition, for example) are no longer sufficient to detect some of the subtle defects that occur during manufacturing. In addition, the push towards zero defective parts-per-million, driven by some mission-critical applications such as automotive and medical, is driving faster adoption of new defect-oriented fault models, such as cell-aware, across the industry. Cell-aware fault models are based on an abstraction of the defects that happen inside the cell boundaries at the transistor level rather than focusing on the defects that can be modeled on the logical representation of the library cell. Since the number of faults based on an abstraction of defects at the transistor level is much larger compared to the logic level, the pattern count required to detect such defects is usually higher. The combination of different fault models along with higher test coverage requirements is expected to continue pushing pattern counts higher. Increased pattern count and test time have been outpacing improvements in scan compression and increases in the scan shift rate. This problem will worsen as new fault models are introduced to reach enhanced quality goals, and the data volume may increase even beyond what is projected in our model. This will be addressed in a future revision of the HI Roadmap. On the other hand, commensurate improvements in defect coverage and diagnostic accuracy can be expected with the additional pattern types.

The effective scan compression rate will continue to increase, but the increase is accounted for in a different way. The Logic Assumption sheet reflects the effectiveness of compression of a single IP block or subsystem. Effectiveness of compression at this level of the design is slowing. Several years ago, the concept of hierarchical scan was introduced by the EDA community. This is accounted for in the Logic Test Data Volume requirement

[11] https://eps.ieee.org/technology/heterogeneous-integration-roadmap/2019-edition/hir-test-chapter.html

sheets and realized with the number of identical IP blocks, where the same patterns could be applied to multiple identical instances of the same block concurrently.

Logic toggle activity, or power considerations during test pattern shift and capture events, will be an increasing concern. For low-power devices that are designed to support functional toggle activity in the 5-10% range, any increase in toggle activity during test application may lead to incorrect test results or even device failure due to excessive voltage droop or device heating. These concerns are valid for all devices since power and timing envelopes are optimized for mission mode rather than test mode.

Power and thermal considerations during shift and capture have continued to apply pressure to pattern counts. Ideally, the more portions of a design participating in a parallel test application, the faster one can achieve desired quality levels and reduction in the cost of test. However, this usually involves a higher percentage of flops switching during shift and capture, so causes larger power draws and heat generation. To fix the problem, various techniques have been developed which can be applied to adjust activity, during both shift and capture. For example, serially testing components inside a device can silence blocks not being tested, shift channel inputs can be held static for portions of a shift cycle, capture activity can be modulated per pattern, etc. DFT hardware can also be inserted to adjust activity, adjust the compression levels, constrain the clock gating logic, add gating logic at the head of scan chains to limit the number of chains that are allowed to toggle simultaneously, etc.

A high-speed serial interface, leveraging IEEE 1149.10 or a functional protocol such as USB, enables an optimal scan architecture to be applied for each subsystem, because it decouples the number of scan IO between tester and internal scan chains. These high-speed interfaces allow large volumes of test data to be applied to the chip at a very high transfer rate. Inside the device, the data is converted to parallel scan chain patterns which can be applied at an appropriate speed to keep thermal issues and local IR-drop issues from affecting test outcomes.

LOGIC TEST NO LONGER DOMINATES TOTAL TEST TIME

Although the absolute number of digital logic gates relentlessly increases and demands test time accordingly, the relative proportion of test time per device that is spent on traditional scan tests has been falling as other test types consume more time and prove harder to scale. These other tests are necessary for other circuit types which are now commonly integrated alongside the digital logic on the same die or in the same package, including embedded memories with BIST and repair features, as well as analog and mixed-signal (AMS) circuits requiring specification testing, tuning, and trimming. In addition, functional tests are being applied to exercise the chip in mission mode for a more realistic determination of performance limits. Volume data collection is often performed during testing to provide feedback to the foundry, or to provide feedforward data for subsequent test insertions. The result is that traditional logic tests based on scan patterns from ATPG (to which significant optimizations like scan compression have been applied) are no longer the dominant component of test time. Even for “big-D little-A” chips with a small amount of analog circuitry, the disproportionate time it takes to stimulate and measure those analog circuits can exceed the time taken for high-throughput scan tests. The industry is only beginning to tackle the problem of scalability for AMS circuit testing through standardization of defect-oriented tests which are applied through DFT access points.

In addition to test time being consumed for these new circuit test types enabled by integration, more time is also being devoted to other actions as the objectives of the manufacturing test flow itself have steadily grown beyond simply distinguishing good parts from defective parts. The tester has become the focal point for identifying repairable defects (as in memory arrays with spare rows and/or columns) and harvestable chips whose defective circuits can be selectively disconnected (“down-cored”) and sold as lesser-capable products. Likewise, the tester is used as the platform to determine the performance bin for each unit; this process may involve multiple dimensions (maximum frequency, minimum power, number of defect-free units) as well as the determination of the fuse settings or laser trimming corresponding to the final configuration on a per-part basis. The tester is also used to gather extended data (like memory bitmaps or continue-on-fail test results for an entire test set) on a 1-in-N sample of the parts it tests for use in continuous process improvement efforts. These extra steps compete with the original purpose of screening out bad material but presumably provide a positive return on investment for the test time spent.

Another trend which is upsetting the traditional pareto of test time usage is the increasing incorporation of firmware-based architectures into chip designs. Many SOCs utilize dozens of embedded microcontrollers to perform and manage operations, and each of these requires an associated embedded memory to be loaded with firmware before it can execute its mission-mode duties. If structural-only tests are used (as in traditional scan-based digital testing), then this architecture is immaterial, but if a portion of the testing is to be functional, then this firmware-based architecture implies that the tester must spend time loading the firmware and launching the microcontrollers to

perform their tasks while orchestrating that activity with the application of tests through other available interfaces. This is complex work which can consume considerable test time and impose requirements on tester architectures that go beyond traditional ATE designs. A uniquely challenging example is when high-speed serial I/O interfaces (which must be tuned and calibrated by firmware running on an embedded microcontroller) are re-purposed to deliver packetized scan data to test the rest of the chip: these I/Os must be tested first through some other means (e.g. structural test), then placed in mission mode for training, then used to deliver structural tests to the rest of the SOC.

AI DEVICES AND LOGIC TEST

As the industry is grappling with how to improve AI performance, new architectures are being deployed for modeling cutting-edge neural networks for AI devices targeted for data centers, end-point applications, or IoT devices. These designs pose new challenges as well as opportunities for testing of logic devices. One of the trending applications for such AI devices uses a massive amount of parallelism involving compute engines and memories to mimic complex neural networks. This results in huge designs with cores that are repeated thousands of times interconnected via an intelligent network. Another technique that is being adopted for easier handling of placement and routing of such large devices is a tile-based design flow. Tile-based designs are connected by abutment, so all interconnects are routed through the individual tiles. Test implementation of such devices must be based on a divide-and-conquer approach. The test can be implemented at the core level, but if the number of cores is in the thousands, a cluster of cores (or tiles) is grouped together and DFT is implemented at that level. Due to the identical nature of these designs, the DFT implementation can be replicated across all clusters and test data can be broadcasted to all of them. However, there should be a mechanism in place to observe the test responses individually so that faulty tiles can be easily identified so appropriate action can be taken. Figure 2 illustrates various place and routing techniques for AI devices. Various factors such as pin availability for scan test, power dissipation during test, test verification time at the chip-level, repair, etc. are some of the major challenges with regards to DFT implementation for AI designs that must be addressed.

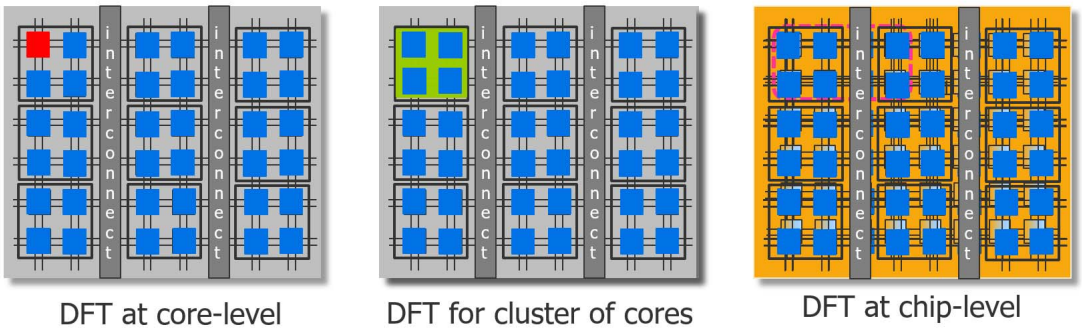


Figure 2: Typical AI device with multiplicity of identical cores

FUNCTIONAL TEST RESURGENCE

Why are we seeing an increased interest in functional test, after decades of focus on structural test? The question comes back to effectiveness of coverage and efficiency of exercising each of the fault models. Over the past three decades we have improved the structural fault models as well as the way we deliver the tests. We continue to innovate, as can be seen in the earlier discussion on high-speed scan. The issue we are seeing is that emerging defects are not caught at the structural level, as they involve a combination of blocks, functional states, and environment that leads to a system error. System on chip (SOC) or System in Package (SIP) devices may contain billions of standard cells and thousands of IP blocks, so what is the best way to identify the critical tests? In addition, these errors may be tolerable and only in a small number of cases activated by the application software. Given that the number of possible defects that are detectable are much smaller than the number of potential defects, one must prune the test set to focus on the likely defects. The alternative to this is to develop tests for all cases, which would cause a significant increase in test time that would not be acceptable because test costs would significantly increase.

Functional test is not new to our industry. We have a rich history of running industry benchmarks on processor devices during system level testing. Processor companies with evolutionary architectures (e.g. x86) have identified a series of tests representing application software that could be written in the future and mimic the activity at the device level. Such tests are difficult to generate because the sequences are non-deterministic. ATE has evolved to support some non-determinism (e.g. Serdes), but there is still work to be done in order to mimic the behavior of the system in which the device under test would normally reside. For high volume devices, many companies have

developed application boards with a socket, device handler, and a library of software functions used for a system level test (see the 2019 HI Roadmap for more information).

For the next-generation applications that do not have a long product history, how do we get the benefit from this test approach? There are several technologies that will enable additional functional coverage. ATE has developed instruments that are protocol-aware and can mimic non-deterministic environment stimuli and response for the device under test. The second is leveraged from logic verification: a new approach called Portable Stimulus Standard was approved in 2018 by Accellera[12]. The concept is that the HW features of each sub-block are described, as well as the accessible interfaces and test instruments. The portable stimulus solver would figure out how to create tests for each of the described features given a reduced set of access interfaces. This will enable a test to be targeted to a specific ATE insertion with different test instrumentation access. Many of these tests are software that is run on the device under test. The benefit is that appropriate tests can be selectively applied at different insertions targeting the best economic point.

IN-FIELD TESTING

Mission-critical applications such as automotive, data centers, and medical devices have resulted in a resurgence of the need for in-system (or in-field) testing of ICs. It is not only important to test such devices post-manufacturing to eliminate any defective parts (also termed as time T0 testing), the new application domains mandate that the electronic parts are monitored throughout the life-cycle of the device to detect any reliability issues related to aging or wear-out. Traditionally, semiconductor ICs have very low wear-out rates early in their life-cycle but start failing as time elapses. Numerous practices have been adopted in the industry to design, manufacture, and test for high reliability, but as these devices are being deployed for safety-critical applications, the importance of monitoring has become a necessity. Built-in Self-Test (BIST) has been around for a long time, but in practice it was primarily deployed for testing certain IC components such as embedded memories during manufacturing test. BIST techniques are now being used for time T>0 testing (in-field) for memories, digital logic, clock controllers, PLLs, mixed-signal components, etc. In addition, there is now a need for BIST and repair techniques that can be run on-chip very efficiently. Factors such as test application time and test quality are of great importance for in-field testing, although to a varying degree. For example, on-chip test application time is critical for automotive ICs as the tests have to be completed within a very small time window – whereas, for industrial applications such as data centers, there is relatively more time for achieving target test quality. One application of in-field testing is to monitor performance degradation and adjust the operating point of the device based on the test results. Some systems have the ability to notify the operator for preventive maintenance.

OPTIMIZING IMPACT OF TEST INSERTIONS

The number of test insertions as well as the test intensity has increased from the traditional wafer and package test insertions of a decade ago. Wafer test is the lowest cost from a product scrap point of view, but has the highest interface cost (probe card, prober), and it is more difficult to achieve a high site count. With System Level Test (SLT) the interface cost per site is lower with a capability of high site count, but the scrap cost is much higher. The questions become where is the best economic point to apply a test, and can a test be retargeted to any test insertion? Traditionally each insertion has been viewed as a unique event with different test flows. Looking forward to the future, with test interfaces leveraging functional interfaces, it becomes possible to apply a test at any point in the product life-cycle including in-field. If tests are designed for portability, if one test changes from a low yield impact to a high yield impact, it can be moved from SLT to wafer probe. The instrumentation and test access are different at each phase of production, which is why a test retargeter is required. Having multiple test insertion points enables more tests to be run, driving a higher end-product quality with an optimal coverage cost.

INTERFACE CHALLENGES

A challenge going forward with heterogeneous integration will be IO interfaces such as die-to-die communication or direct access mechanisms. If a die is designed to be placed in a stack, with no communication off-package, the IOs may be designed for limited drive strength, minimal physical interface, and reduced ESD protection. The largest of these challenges is with the IO strength for getting diagnostic data from the die to the ATE at wafer probe. To mitigate the inability to reuse functional IOs as test IOs, some DFT will be required to utilize sacrificial test pads or a dedicated test interface. Post-integration test access has recently been addressed by IEEE Std 1838. Some of the structures may be leveraged for singulated die test access, but parametric and physical test issues remain.

[12] https://www.accellera.org/downloads/standards/portable-stimulus

New interfacing alternatives may be required which enable communication between ATE and low-drive die-to-die interfaces. There are also challenges for singulated die handling which will need to be addressed toward the goal of achieving known good die. Probing microbumps is an ongoing problem which is discussed in the HIR probe section and optical interfaces in the HIR photonic testing section.

SUMMARY

The preceding sections highlighted the changes that are emerging in the logic test area. We have provided details of the dynamics of each area that will undergo significant changes. To highlight the major themes, we have developed a table to describe the overarching themes.

Trend	Short term 0-5 years	Long term >5 years	Challenges
Scan pattern growth of >30%/year	Multiple test modes which reassign common IOs to test different core partitions	High-speed serial interfaces carrying packetized scan data: more scan bandwidth	Rate of adoption of new scan interfaces
Functional test resurgence	Beyond SLT, further adoption focused on Portable stimulus	Functional test on ATE and SLT using Software Test Libraries	Establishing coverage metrics
Demand for in-field testing growing due to functional safety	Re-use of DFT-based instruments at power-up e.g. MBIST and LBIST	BIST + Software Test Libraries at power-up and on-line Safety critical requirements driving new functions	Integration of DFT-based BIST with mission mode control and reaction
Increasing IO interface challenges	Sacrificial pads and dedicated DFT interfaces	ATE infrastructure to contact advance interfaces	Electrical, optical, and mechanical interface sensitivities

Team Leader: Marc Hutner

Section 4: Specialty Device Testing

A classification of specialty devices was defined in industry roadmaps beginning in 2006, driven by strong high-volume market demand, but having odd test requirements. Examples are CMOS image sensors, LCD drivers, MEMS devices (including multimode sensors), actuators, bio-MEMS, and similar non-standard devices.

Trends Impacting this Technology Area

The applications of Mobile personal devices, IOT, Healthcare, Automotive/ADAS and Robotics are key drivers of specialty devices to motivate innovative technologies and the high growth rate of volume.

The trends of technologies: (Near Term < 5 years)

The mobile and wearable devices for IOT and healthcare applications are major drivers, impacting technology trends in the near term.

- The trends for multi-mode MEMS sensors is fusing multiple functionalities together in one device and also reducing the size of the package to be smaller and thinner for adding value in a compact unit. See Figure 1.

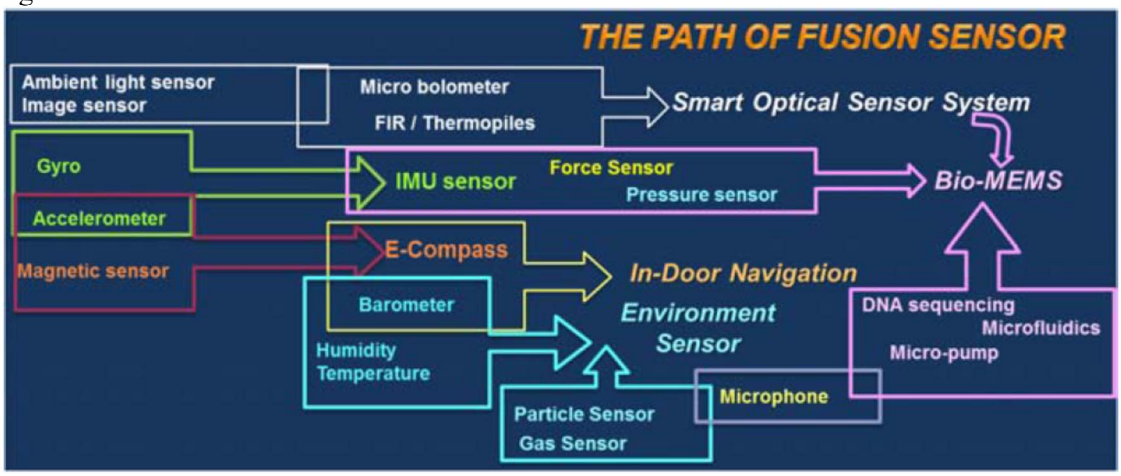


Figure 1: The path of MEMS sensor fusion

- The technology trends for image sensors lead to highly integrated multiple wafers using a 3DS (three-dimensional stacking) process and novel packaging technologies for enhancing image performance with cost-effective mass production solutions. The first successful step involving 3DS wafer processing of image sensors was the BSI (Back Side Illumination) process which bonded a photo-sensor wafer together with a back-side mixed-signal data processing wafer, connecting their signals through TSVs. When image-sensor pixel numbers increased and sensor size shrank below 1.4um, the BSI process had the advantage of a lower signal to noise ratio than the traditional FSI (Front Side Illumination) process. The next step in the image-sensor wafer-integration process adds a memory-cell wafer between the photo sensor wafer and the mixed-signal data processing wafer, which could enhance image performance and the speed of data processing in a variety of imaging applications such as 3D imaging, face recognition, and image capture, with frame rates over 1000 frames/second. See Figure 2.
- Image-sensor packaging innovation began using WLCSP (Wafer Level Chip Scale Package) to develop the glass-based stacking wafer-level package with TSV or Shallcase type processing, as shown in Figure 3.
- The trends for new WLP for image sensors are WLO (Wafer Level Optics) and WLCM (Wafer Level Camera Module) which stack optical systems on the image-sensor wafer using a wafer-level packaging process to reduce the size of optical systems and increase the efficiency of mass production. See Figure 4.

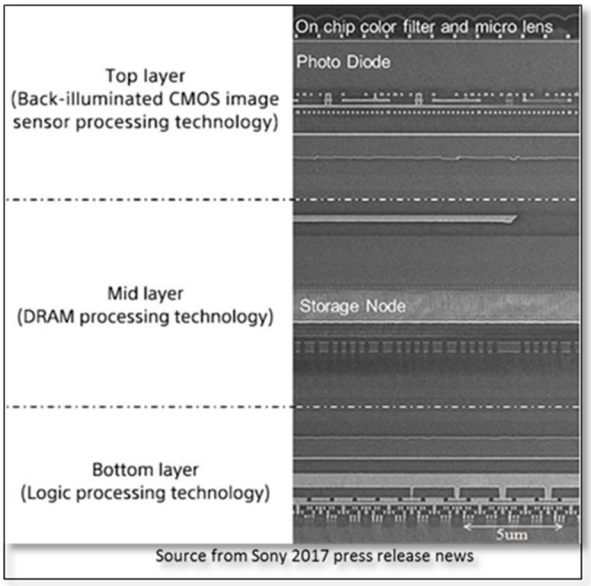


Figure 2: Image sensor 3D stacking wafers
(Backside illumination photo sensor wafer + memory cell wafer+ data processing wafer)

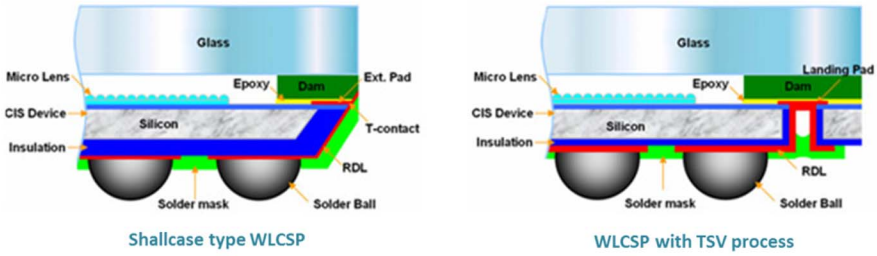


Figure 3: The structure of Image sensor WLCSP(wafer level chip size package)

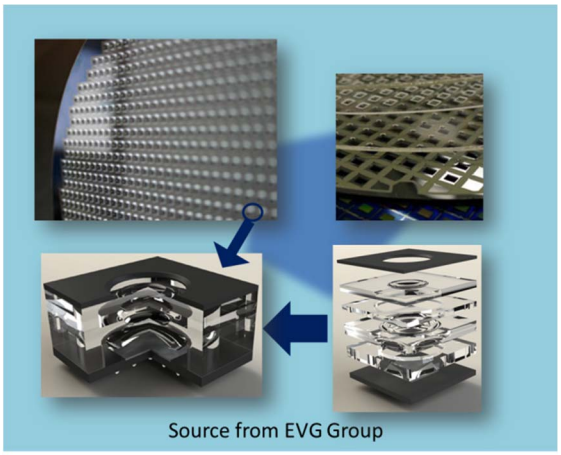


Figure 4: Image sensor WLO (Wafer Level Optics) packaging

The trends in technologies: (Medium to Long Term < 15 years)

The automotive, robotic, medical and intelligent artificial organ fields are the next wave of drivers for specialty devices which impact technologies in the medium and longer term.

- Reliability will become a very important subject for specialty devices. Burn-in and tri-temperature testing will become necessary test procedures during mass production.
- Built In Self-Diagnostic, Self-Calibration & Compensation and Self-Repair technologies will become important design skills to apply on specialty devices for enhancing reliability performance.

Concerns: Test Challenges

LCD display drivers:

LCD display drivers are unique because of their die form factor, which can have larger than a 10:1 aspect ratio and thousands of very narrow gold-bump pads requiring contact for test. In 2017, in-line probing pad pitch for LCD display drivers already was down to 18µm, and stager pad pitch was 13µm. Right now, only the cantilever probe card is a major cost-effective solution for achieving probing of LCD drivers with such narrow and fine pitch pad with gold bump in mass production.

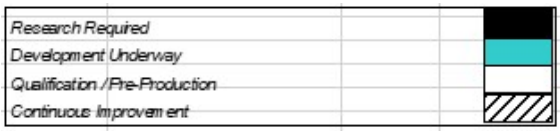
An upcoming test challenge is the data transfer speed of I/O, which will increase to 2.5 Gbps in 2019 and is predicted to be up to 6.5 Gbps within 10 years. We need to overcome the challenges of probing fine-pitch bumping pads with high-speed signals with economical probing solutions.

Image sensor devices:

The testing of image sensor devices needs to consider special test requirements for optical systems and huge image data processing. The innovative technical trends of highly integrated 3DS CMOS image sensors increase the difficulties of special test requirements and challenges.

		Year of Production	2017	2018	2019	2020	2021	2026	2031
Process Integration	**Test Method**	**Challenges**							
Wafer level probe (BSI process + Memory+ ASIC , 3 layer W to W)	CP	Wafer probe with multi-sites	Multi- insertion						
			Single insertion (ATE)						
		Wafer probe by full wafer contact	Optical system (Visible light)						
			Technology of optical asserroy & probe card						
			(ATE) / challenges test system resource						
WLCSP (BSI process + Memory +ASIC , 3 layer W to W)	FT	Test after singular (Pkg form)	Multi- insertion						
			Single insertion (ATE)						
		Test after dicing (wafer form)	Optical system (Visible light)						
			Probing methods and accessories, multi-sites						
		Test after dicing with full wafer contact (wafer form)	Optical system (Visible light)						
			Probing methods and accessories						
			(ATE) / challenges test system resource						
WLO-P (Multi-layer W to W)	FT	Test after singular (Pkg form)	Single insertion (SLT)						
		Test after dicing (wafer form)	Optical system (Visible light)						
			Probing methods and accessories						
		Test after dicing with full wafer contact (Full functional test under wafer form)	Optical system (Visible light)						
			Probing method and accessory						
			Challenges ATE or SLT test system resource						

Table 1: Specialty device odd test potential solutions table – Image sensor device



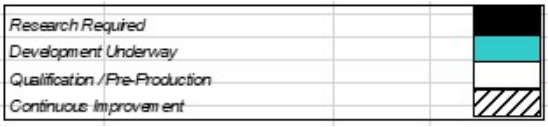
Automotive ADAS applications and intelligent machine vision especially need the functionalities of image sensors with a wide spectrum (from UV to FIR), high dynamic range, good S/N (Signal to Noise) ratio, fast data frame rate, better quality and reliability, which challenges test system design. Burn In solutions also need to include optical stress for sorting out defects in the coating process on the photo sensor surface.

MEMS devices (Sensor, Actuator and Biological):

MEMS were successfully applied for various sensors for sensing motion, magnetic field, optics, sound, air pressure and vibration, flow, chemical composition of air, DNA sequencing, and other characteristics, and the market volume is increasing rapidly due to IoT, healthcare and automotive applications. Testing MEMS sensor devices with suitable physical stimulus and cost-effective solutions for the various types of sensors is difficult and tricky. Especially testing the expanding kinds of fusion sensors will bring many test challenges.

			Year of Production	2017	2018	2019	2020	2021	2026	2031
Process integration		**Test Method**	**Challenges**							
IMU sensor (Accelerometer + Gyro)	CP / wafer probe	Probing MEMS wafer (DC only)	Probe card technology/multi-sites vs. cost							
		Probing each wafer with full functions (Multi-insertion)	Motion Prober system and probe card technology/multi-sites vs. cost							
		Probing 3DS wafer with full functions (Single-insertion)	DFT design and implement							
	WLP	Test before dicing (Wafer form)	DFT design and implement							
		Test after singular (Pkg form)	Handling small size package							
	FT	Final test with full functions (Multi- insertion)	Test cost is high							
		Final test with full functions (Single- insertion)	Reduce test coverage rate							
			DFT design and implment							
		Burn In Test	SLT							
			BISX (Build In Selft X) X= (Test, Diag-nostic, Correlation,Compensation/Repair)							
Navigation (G-sensor+ Gyro+ Magnetic sensor + Barometer)	CP / wafer probe	Probing MEMS wafer (DC only)	Probe card technology/multi-sites vs. cost							
		Probing each wafer with full functions (Multi-insertion)	Probing system with situmlus & probe card technology							
	FT	Final test with full functions (Multi- insertion)	Test cost is high							
		Final test with full functions (Single- insertion)	Reduce test coverage rate							
			DFT design and implment							
		Burn In Test	SLT							
			BISX (Build In Selft X) X= (Test, Diag-nostic, Correlation,Compensation/Repair)							
Enviromential Sensor (Pressure + Humidity + Gas) Sensor	CP / wafer probe	Probing MEMS wafer (DC only)	Probe card technology/multi-sites vs. cost							
		Probing each wafer with full functions (Multi-insertion)	Probing system with situmlus & probe card technology							
	FT	Final test with full functions (Multi- insertion)	Test cost is high							
		Final test with full functions (Single- insertion)	Reduce test coverage rate							
			DFT design and implment							
		Burn In Test	SLT							
			BISX (Build In Selft X) X= (Test, Diag-nostic, Correlation,Compensation/Repair)							

Table 2: Specialty device odd test potential solutions table – MEMS Fusion Sensor



DFT for MEMS sensor devices is a new technology and needs research and innovative development for different kinds of sensor structures. MEMS sensors DFT needs to develop the stimulus source and the sensor together in the MEMS structure as a BIST (Build-In-Self-Test) cell. When testing, the cloned control signal of physical stimulus is generated from the MEMS ASIC to enable the MEMS BIST cell to imitate physical stimulus for testing the sensor cell to achieve the DFT concept. This concept could also implement the technologies for BISD (Build-In-Self-Diagnostic), BISC (Build-In-Self-Correlation/Compensation) and BISR (Build-In-Self-Repair) to enhance reliability of MEMS sensors for automotive and medical applications. The key during testing is to make sure the BIST cell works well.

Beyond MEMS sensors, there are also actuator and biological applications such as micro-mirrors, MEMS speakers, RF switches, energy harvesting, microfluidics, micro-dispensers and artificial organs, plus others. The

challenges of testing MEMS actuators and biological devices are that test methods are hard to standardize and depend on the structure for each different kind of MEMS device. Especially for the testing of biological devices, the test environment can be severe and needs to pass safety certification based on the laws of different grades and countries.

Summary

Specialty devices as defined need odd test requirements and are driven by strong high-volume market demand. Under these two conditions, the trends of specialty devices will be drive toward highly integrated multi-functions in one smaller unit to overcome ASP (Average Sale Price) erosion, and testing procedures will move toward high parallelism to reduce test cost. Test challenges will follow the same trends to overcome testing evolutionary HI (Heterogeneous Integration) specialty new product through cost effective solutions.

References

"Sony Develops the Industry's First 3-Layer Stacked CMOS Image Sensor with DRAM for Smartphone", Sony press release, February 7 ,2017

"Xintec 12-inch WLCSP Line Ready for Mass Production in 2H 2015" , DIGITIMES, March 2015

"EVG's wafer-level optics (WLO) manufacturing solutions", EVG press release, September 11 ,2017

Section 5: Memory Test

There have been significant changes in the Memory environment in recent years. Up until 2003, DRAM bits comprised approximately 90% of bits shipped per month. By 2009, NAND had become the dominant form of memory and comprised 85% of monthly bits shipped. During the same period, NOR Flash has largely migrated from parallel interface devices to serial interface devices with extremely small form factors in order to reduce PCB size, complexity and power. With the introduction of mobile smart phones in 2007 and tablets in 2010, the traditional dominance of PC DRAM has been eroded by lower-power LPDRAM which is required for longer battery life. New generation memory types such as PRAM, RRAM, STT RAM, and CBRAM along with 3D multilayer instantiations of NAND will further change the memory environment over the next decade. The continuing shift to a battery-powered wireless environment will continue to drive changes in memory usage. Existing memory types will likely not be replaced, but will be used in joint solutions with newer memory types.

Memory density has kept pace with Moore's Law since the first DRAM was manufactured in 1969, but lithography cycles as well as the performance roadmap are expected to push out for litho nodes less than 20nm, so the typical 2-year technology pace is forecasted to stretch to initially 3 years, and then 5 years later in the roadmap. Multi-layer 3D NAND technologies allowed the use of longer gate lengths, so the 2-year technology pace should continue for the near term.

From a test perspective, most memory will be structurally tested at wafer test utilizing low-pin-count interfaces of 4 to 10 pins per device. Structural test, when used, will likely include a self-test performance validation.

Table 1: Memory Test Requirements

	2018	2019	2020	2021	2026	2031
DRAM Characteristics						
Capacity (bits) [1]	16G	16G	32G	32G	64G	128G
I/O data rate (Gb/s)						
PC DDRx	3.2	4.4	4.4	6.4	8.4	8.4
GDDRx	10.0	12.0	16.0	16.0	16.0	16.0
LPDDRx	4.0	4.8	4.8	4.8	8.4	8.4
Hybrid Memory Cube (HMC [SerDes])	15.0	15.0	15.0	15.0	25.0	30.0
Wide IO	0.3	1.1	1.1	1.1	2.2	2.2
High Bandwidth Memory (HBM)	2.0	2.6	2.6	2.6	3.0	3.6
NAND Characteristics						
Capacity (bits) [2][3][4]	1T	1T	2T	2T	16T	64T
Maximum I/O data rate (Gb/s)	0.53	0.53	0.53	0.67	1.07	1.60

Manufacturable solutions exist, and are being optimized

Manufacturable solutions are known

Interim solutions are known

Manufacturable solutions are NOT known

Notes:

1. DRAM bit capacity per die
2. 3 bits per cell introduced in 2009
3. 4 bits per cell introduced in 2012
4. 3D multi-layer introduced in 2014

DRAM

Historical trends show that PC DRAM has doubled its performance every 5 years and will reach a data rate of 8.4 Gb/s per I/O in 2022 with DDR6. However, DDR4 appears to be the end of the DDRx era as there are no further enhancements of the DDRx architecture beyond DDR4 in definition or development. DDR4 itself has severe PCB design restrictions in order to meet the I/O performance requirements, so an enhancement of the current DDRx single-ended interface will present additional challenges and constraints. The Wide IO memory architecture is targeted to mobile applications such as phones and tablets and is an evolution of the DDRx that decreases I/O bit rates while expanding the number of I/Os up to 512. Hybrid Memory Cube (HMC) is targeted to servers, where it provides very

high performance over a SERDES interface, though at increased cost. High Bandwidth Memory (HBM) is similarly targeted to graphics video cards and applications. Based on existing roadmaps, DDR3/4 will continue to serve the PC market for the foreseeable future. LPDDRx and GDDRx DRAM families will continue to be a driver for the near future.

DRAM will become increasingly difficult to scale in sub-20nm nodes, and transistor wear-out will increase the frequency of errors. On-chip error correction and memory management will likely become a requirement before 2020. Dynamic failure detection, analysis, and repair will become necessary over the product life. To enhance test productivity, new test-oriented architectures will be required. On-chip correction may also change the DRAM test paradigm.

Maintaining high ATE test parallelism is required over the roadmap period to manage test cost. However, probe card performance test at high parallelism may be a challenge at high GT/s due to the interface routing complexity required. These challenges will ultimately drive the need for die self-test.

Flash

NAND will double in density ever year in the short term and slow to a doubling every 2 years. The doubling every 2 years will be faster than the projected lithography node migration due to the increase in the number of bits stored in a single memory cell from one and two to four on some cell types. Use of error correction allows greater uncorrected error rates and enables increased bits per memory cell. NAND bus width has continued to be dominantly 8-bit with a decreasing number of products at 16-bit I/O. As large amounts of NAND are being consumed in Solid State Drives (SSD), a new NAND interface may emerge that is more optimized for SSD use.

The need for internal voltages that are 3-8 times the external supply requirements is expected to continue in the test process, driven by the hot-electron and Fowler-Nordheim charge transport mechanisms. Increased absolute accuracy of supply voltages will be required in the future due to the trend toward lower voltages, but percentage accuracy relative to the supply voltage will remain a constant. I/O voltage decreases are pushing the operational limits of standard tester load circuits; new methods will be required in the future.

Wafer test generally does not require the performance of package test, but error detection, error analysis, and redundancy processing are required.

NOR memory density is expected to increase slowly over the roadmap period and remain flat toward the end of the roadmap. NOR has been transitioning from a parallel to a serial interface since 2007 to reduce package size and power. Further increases in NOR performance along with increasing test requirements are not expected.

Stacking of various types of Flash and other memory and/or logic components in a single package has become standard and is expected to continue. Multiple die within a package has complicated the package test requirements and has increased pin count and number of DUT power supplies required. Data and clock rates for flash will increase, but there is expected to be a wide variability in the requirements based upon the end application.

Embedded Memory

Embedded memory consumes greater than 80% of transistors in many MPU and SoC designs and will scale with the increase of transistors in these devices. Embedded Flash and DRAM bits will not match the density of standard DRAM and NAND. Newer memory types such as RRAM or STT RAM may become embedded over the course of the roadmap.

To enhance test productivity, new test-oriented architectures and/or interfaces will be required. Built-in self-test (BIST) and built-in self-repair (BISR) will be essential to test embedded DRAM and Flash memories cost effectively. The primary test algorithms for Flash memories will continue to be Read-disturb, Program-disturb, and Erase-disturb while March tests with all data backgrounds will be essential for embedded DRAM.

Considerable parallelism in test will be required to maintain test throughput in the face of rising memory densities. In some cases, test is made cost-effective by double insertion of devices rather than testing both logic and embedded memories on the same tester. In double insertion, embedded Flash and DRAM are tested and repaired on a memory tester, while the logic blocks are tested on a logic tester.

Section 6: Analog and Mixed Signal Test

Executive Summary

The economic benefit of monolithic integration (SoC) and system in package (SiP) is well established and continues. This integration has combined digital logic with processing, analog, power management, and mixed signal routinely in a single package and often on the same die. This trend has increased the breadth of interface types on a single part and given rise to test equipment that mirrors this range with a corresponding breadth of instruments. Now this trend has again escalated with the emergence of through silicon via (TSV) packaging technology driving the challenge in a 3rd dimension.

An important trend impacting mixed signal and analog testing is the compelling economics of multi-site testing for devices manufactured in extremely high volumes, also called parallel test. To support parallel test, many more instrument channels of each interface type are required to keep test cell throughput and Parallel Test Efficiency (PTE), also known as Multi-Site Efficiency (MSE), high; this is of increasing importance to avoid severely impacting Units Per Hour (UPH).

A similar concept but in a dimension relating to the single device itself is testing multiple IP cores within the device in parallel (concurrent test). This has many of the requirements and challenges of parallel test, but also includes some unique ones. A key one is having the ability in the design of the IC to test IP cores independently, in parallel. Test Access Mechanisms (TAMs) are the ability of IP cores to be accessed and controlled independently from other IP cores. The most powerful economic advantage results when being able to test multiple IP cores in parallel, while at the same time testing multiple devices in parallel.

The increasing number of interfaces per device and the increasing number of devices tested simultaneously raise the need to process an increasing amount of data in real time. The data from the mixed signal and analog circuitry is typically non-deterministic and must be post processed to determine device quality. This processing must be done in real time or done in parallel with other testing operations to keep test cell throughput high. In fact, as site count increases, overall throughput can decrease if good PTE is not maintained.

Looking forward, the breadth, performance, density, and data processing capability of ATE instrumentation will need to improve significantly to provide the needed economics. The area undergoing the most change is RF/microwave and so it is covered in its own separate section. The digital and high-speed serial requirements for mixed signal devices are equivalent to logic and are covered in that section. The requirements for the TAM are covered in the DFT SOC Device Testing section. The requirements for DC trim accuracy are included in the Mixed Signal tables (see Table 1).

DC Accuracy updates for 2020

The 2020 update for DC accuracy includes ever-increasing low-end accuracy requirements driven by lower VDD values and more fuse blowing and servo techniques being used to cost effectively make the DUT more accurate and improve the specifications and yields.

COT is always important and more parallelism in terms of IP blocks within a device (IP block) and multi-site parallelism is key to this.

Quality also needs to be improved with these accuracy improvements. Pre and post inline checking and the comparison of lot runs looking for common tests that always pass or fail will be aided by using Artificial Intelligence (AI) and Machine Learning (ML) to handle and simplify large volumes of data. Other quality improvements include inventorying the tests that have been run and having more quantitative (actual value) versus qualitative (pass/fail) testing. There is always a cost trade-off balance.

Power updates for 2020

The other end of the spectrum for 2020 is high power (current and voltages) being driven primarily by server farm power needs and automotive and battery management systems as shown in table 1 in the Note 8 section.

Because of the higher power, some tests that run a device at full power must be run very quickly and then turned off so as not to damage the parts that require special cooling. In these cases, precision pulses are required on tests like RDSon which pulses a high current at a very short pulse width to test the on-resistance of a switch.

Quality improvements here would include thermal testing and management throughout the test flow. For example, high power tests which would generate a lot of heat could be interleaved with low power test to allow the device to cool down.

Handlers with built-in cooling for the device is another option to be looked at for devices requiring the cooling.

Some process technologies once considered niche are gaining mainstream acceptance, including GaN (gallium nitride) and SiC (silicon carbide) devices.

SiC is projected to hit $1.5B by 2023 for these types of applications[13]:

- Electric Vehicle
- Train
- Charging Infrastructure
- Motor Drivers
- Photo Voltaic (PV)
- Wind Power

GaN is projected to hit $500M by 2023 for these types of applications[14]:

- Data Centers
- Fast Charger
- LiDar
- Wireless Charging
- Electric Vehicle

Power devices using GaN and SiC have higher band gaps compared to their silicon counterparts. The benefits are[15].

- Higher power density
- Smaller size (smaller wafer & die)
- Better high temperature performance because their band gap is higher than silicon
- Higher frequency response
- Lower ON-resistance
- Lower leakage, so there is a need for sourcing higher test voltages, as well as appropriate low current measurement sensitivity.

The test requirements to test GaN and SiC devices are

- Breakdown voltages up to 3000 V or even higher
- More than 100 A
- Junction capacitances for dc biases up to 3000 V
- High SiC and GaN voltages and fast switching speeds
- Testing these devices at their specified voltage, current and power rating
- Test fixturing:
  - A proper test fixture solution is extremely important to ensure safety (due to the high voltages and currents used)
- Supporting the wide variety of power device package types.

The breakdown voltage test has special techniques being investigated involving Paschen's Law. To summarize: above a certain pressure, increasing the pressure raises the breakdown voltage or allows a narrower gap without breakdown at a set voltage.

Analog Mixed Signal Updates for 2020

Pulse Amplitude Modulation – 4 levels (PAM4) (Note: Optical PAM4 is not addressed in this update)

The attributes of PAM4 include:

- 4 amplitude levels
- 2 bits of information in every symbol: ~ 2x throughput for the same Baud rate, ie, 28 GBaud PAM4 = 56 Gb/s
- Lower SNR, more susceptible to noise
- More complex Tx/Rx design, higher cost

It is used extensively in the JESD 204B/C standard.

[13] https://www.systemplus.fr/wp-content/uploads/2018/07/YD18027_Power_SiC_2018_Materials_Devices_Applications_July2018_Yole_Sample-1.pdf

[14] https://compoundsemiconductor.net/article/106038/Would_Apple_Change_The_Power_GaN_World%7BfeatureExtra%7D

[15] https://www.powerelectronics.com/technologies/power-electronics-systems/article/21860727/testing-gan-and-sic-devices-faqs

The transmitter (Tx) can be measured with high-speed digitizers, samplers, digital oscilloscopes or even a digital comparator. The receiver (Rx) signal is generated by RF DACs. RF design rules come into play at these high frequencies.

DSP is required to get an optimal eye opening which entails equalization for both PRE and POST processing. PRE processing is used to clean up the stimulus to the Rx, and POST processing is used to clean up the measured data from Tx. Amplitude accuracy is important because of the 4-level algorithm of PAM4. High-accuracy timing and low jitter are important to get a good eye opening.

Challenges in Analyzing PAM4 signals include:

- Sampling Point: Finite rise times and different transition amplitudes create inherent ISI and make clock recovery more difficult (TransImpedance Amplifiers have CDR integrated into them).
- Quantization error plays a role when you take PAM4 measurements versus NRZ. Transition times of the PAM4 data signal can create significant horizontal eye closure due to the higher transition density.
- Noise Tolerance: Instead of having the full amplitude range, there is only 33% of the amplitude because the voltage range is divided into four levels (refer to the Figure 1). Lower PAM4 insertion loss compensates for the 9.5-dB loss in SNR because the eye height for PAM4 is 1/3 of the eye height for NRZ, SNR loss = 20* log10 (1/3) = ~9.5 dB. When other non-linearity is included, it is approximately 11 dB.
- Non-Linear Eyes: The system-margin bottleneck lies with the worst eye. Nonlinearity starts right at the Tx output, and is composed of RLM loss + SNDR loss + other losses like SNDR (ISI).
- Clock Recovery is used on the Rx side to minimize low frequency jitter.
- Fixturing – getting the signal to the DUT
    - Integrated resources are difficult to design at these speeds but are sometimes easier to fixture. External Boxes are available but then are more effort and expense to route to the device. Line loss and jitter are a challenge.

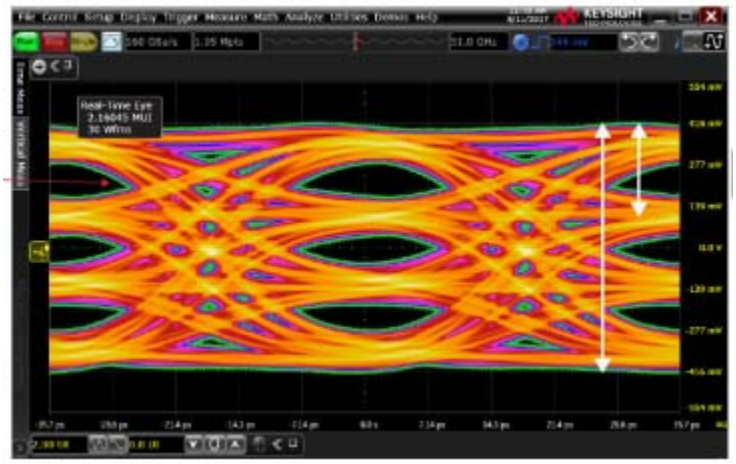


Figure 1. Scope Capture of PAM4 Signal

A typical test list for PAM4 looks like this:

Tx using a PRBS13 waveform:

- Output waveform
- Level Separation Mismatch Ratio
- Eye Symmetry
- Eye Height (amplitude) and Width (timing)
- Transition Time
- Signal-to-noise-and-distortion ratio (SNDR)
- Output Jitter
    - Jrms
    - Even-Odd Jitter (EOJ)
- Spacing of the PAM4 levels
- Eye Linearity: ratio of min to max PAM4 eye amplitudes as shown in Figure 2
    - Eye linearity = min(AVupp, AVmid, AVlow) / max(AVupp, AVmid, AVlow)

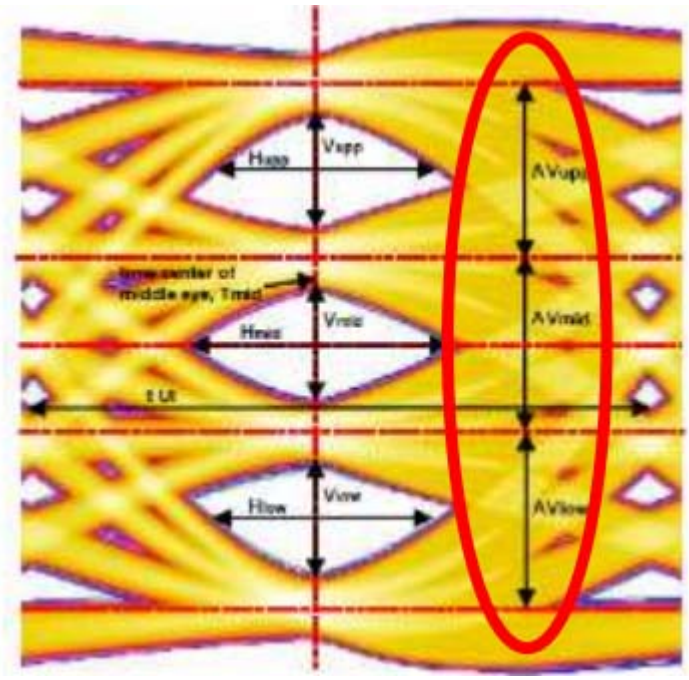


Figure 2: Eye Linearity

Rx tests

These tests involve how much distortion and jitter can be placed on the incoming signal to the receiver and still "read" the correct data stream.

- Jitter tolerance defined as how much jitter the receiver can tolerate
- Other potential receiver "stress tests"
  - Eye Skew (Timing)
  - Eye non-linearity (Amplitude between levels)

Key Test Trends

Short-Term Trends (< 5 Years)

There are three important trends. The first is to deliver adequate quality of test. Most analog/mixed-signal testing is done through performance-based testing. This includes functional testing of the device and then analyzing the quality of the output(s). This requires instrumentation capable of accurately generating and analyzing signals in the bandwidths and resolutions of the device's end-market application. Both of these parameters are trending upwards as more information is communicated between devices and/or devices and the physical environment. See the Mixed Signal Test tables (Table 1) for updates and future needs.

The second key trend is the need for higher DC accuracy. Many of the converters and precision references are made more accurate by doing a measure and trim step. The trim can be accomplished through several means; one of the more recent and cost-effective ways is through register programming of the device. The trim takes a relatively lower performance device and adds high accuracy to it through a DC test and register programming. In the past, this was done for medium performance devices, but now the test methodology has matured, and it is being applied to high accuracy/resolution devices. The change is that in this class of devices, much higher DC accuracy is required to make a valid test.

The third key trend is to enable the economics of test through instrumentation density and Parallel Test Efficiency (PTE). The level of parallelism requires an increase in instrumentation density.

These trends of increasing ATE instrument channel count, complexity, and performance are expected to continue, but at the same time the cost of test must be driven lower (see the areas of concern listed below).

Analog/mixed-signal DFT and BIST techniques continue to lag. No proven alternative to performance-based analog testing has been widely adopted and more research in this area is needed. Analog BIST has been suggested as a possible solution and an area for more research. Fundamental research is needed to identify techniques that

enable reduction of test instrument complexity, partial BIST, or elimination of the need for external instrumentation altogether.

The Ethernet trends are continuing into higher speeds – 28, 40 Gbps per channel and even beyond.[1] There continues to be the need for backwards compatibility to the many existing digital communication standards.

Table 1: Mixed-signal and DC Test Requirements

	2020	2021	2026	2031
Low Frequency Waveform [Note 1]				
SFDR	145	145	145	145
SNR	120	120	120	120
THD	140	140	140	140
BW-Minimum (kHz)	50	50	50	50
BW-Maximum (kHz) [Note 2]	500	500	500	500
High Frequency Waveform Source / Measure [Note 3]				
Level V (pk–pk)	<4	<4	<2.5	<2.5
BW (MHz)	250	250	500	500
Sample rate (MS/s) [Note 5]	500	500	1000	1000
Resolution (bits) AWG/Sine	16	16	18	18
Noise floor (dB/RT Hz)	-140	-140	-150	-150
Very High Frequency Waveform Source / Measure [Note 4]				
Level V (pk–pk)	<4	<4	<4	<4
Accuracy (±)	0.50%	0.50%	0.50%	0.50%
Measure BW (GHz) (under sampled)	9.6	9.6	15	15
Capture Depth Mwords	4	4	4	4
Min resolution (bits)	8-10	8-10	8-10	8-10
DC Accuracy (Note 6)				
DC force (uV)	50	50	50	50
DC measure (uV)	50	50	50	50
DC force (nA) (Note 7)	5	5	1	1
DC measure (nA) (Note 7)	5	5	1	1
DC Power (Note 8)				
DC force V Constant	120	120	140	140
DC measure V Constant	120	120	140	140
DC force A Constant	80	80	100	100
DC measure A Constant	80	80	100	100
DC force V Pulse	80	80	100	100
DC measure V Pulse	80	80	100	100
DC force A Pulse	30	30	50	50
DC measure A Pulse	30	30	50	50
Ethernet				
Speeds (Gbps)	40	40	100	400

Manufacturable solutions exist, and are being optimized	
Manufacturable solutions are known	
Interim solutions are known	
Manufacturable solutions are NOT known	

NOTES:

1) Audio / Precision; Source & Measure specifications (22 KHz BW)

2) Major testing condition

3) Target Devices are Wireless Baseband, xDSL, ODD, Digital TV (Track Mobile Baseband)

4) Target Devices are HDD, Radar, WiGig

5) For Measure Sample Rate: Dependent on method, tracking or Front End filter.

6) The purpose of DC accuracy for this table is for high resolution force/measure and trim

7) Devices may also need high current with the less accuracy

8) Markets include Automotive, Battery Management and Power. This does not include high voltage breakdown test.

Difficult Challenges in the Short Term

- As reflected in the tables, manufacturing solutions exist for the immediate future testing needs. However, high DC accuracy for sourcing, measuring and for trim/fuse blowing/register-setting in a manufacturing environment could be at issue depending on how high a resolution/accuracy the DUT is. Also 40 Gbps Ethernet has known manufacturing solutions, but none are optimized.
- Time-to-market and time-to-revenue issues are driving test to be fully ready at first silicon. The analog/mixed-signal test environment can seriously complicate the test fixtures and test methodologies. Noise, crosstalk on signal traces, added circuitry, load board design complexity, and debug currently dominate the test development process and schedule. The test development process must become shorter and more automated to keep up with design. In addition, the ability to re-use analog/mixed-signal test IP is needed.
- Increased use of multi-site parallel and concurrent test of all analog/mixed-signal chips is needed to reduce test time, in order to increase manufacturing cell throughput, and to reduce test cost. All ATE instrument types, including DC, require multiple channels capable of concurrent/parallel operation and, where appropriate, fast parallel execution of DSP algorithms (FFTs, etc.) to process results. In addition, the cost per channel must continue to drop on these instruments as the density continues to increase in support of parallel test drivers.
- Improvements in analog/mixed-signal DFT and BIST are needed to support the items above.

Medium-term Trends (6 to 10 years out)

- For Wireless Baseband, xDSL, ODD, and Digital TV (Track Mobile Baseband) devices, the source and measure bandwidths, sampling rates and resolutions increase, while the noise floors are decreasing.
- Additionally, DC force and measure accuracies get more challenging.
- Ethernet speeds trending to 100 Gbps [2] have only interim solutions identified.
- Higher speeds and modulation will necessitate PAM to handle the increased data bandwidth – for example, PAM4, 8 or 16 at speeds of 32 GBPS. [3], [4]

Difficult Challenges in the Medium Term

- As the capability requirements increase, there are solutions available, but they do not lend themselves easily to high volume manufacturing.
- Basic physical and electrical properties come more into play. For example, a -150 dB noise floor is possible, but special fixturing is required that is difficult to deploy into a manufacturing environment.
- Ethernet speeds of 100 Gbps [2] have only interim solutions identified.

Long-term Trends (10 years+ out)

- Ethernet speeds trending to 400 Gbps [5], [6]

Difficult Challenges in the Long Term

- Ethernet speeds of 400 Gbps do not have known manufacturing solutions identified.

SUMMARY

Cost continues to be the most critical pressure and concern for analog mixed signal because much of the volume for this is consumer oriented. However, in the medium and long term, performance starts becoming an issue for high-volume manufacturing in terms of bandwidth, sample rate, resolution and noise floor to keep up with the newer devices on the horizon. Ethernet in the medium and long term has manufacturing challenges both in optimization and known solutions.

References

1. https://www.cisco.com/c/dam/en/us/products/collateral/switches/catalyst-6500-series-switches/white-paper-c11-737238.pdf

2. https://www.networkcomputing.com/data-centers/25-50-and-100-gigabit-ethernet-data-center/1422885308
3. http://www.ieee802.org/3/100GNGOPTX/public/mar12/plenary/tremblay_01_0312_NG100GOPTX.pdf
4. http://www.ieee802.org/3/bm/public/sep12/ghiasi_01a_0912_optx.pdf
5. http://searchdatacenter.techtarget.com/feature/Understanding-tomorrows-Ethernet-speeds
6. https://www.keysight.com/main/application.jspx?ckey=2716254&id=2716254&nid=-32176.0.00&lc=eng&cc=US
7. https://en.wikipedia.org/wiki/Effective_number_of_bits

Section 7: Wafer Probe and Device Handling

Wafer probe and component test handling equipment face significant technical challenges in each market segment. Common issues on both platforms include higher parallelism and increasing capital equipment and interface cost.

Device Handling Trends

Increased parallelism at wafer probe drives a greater span of probes across the wafer surface and significantly increased probe card complexity. Prober and probe card architecture should evolve to simplify the interface, however just the opposite is happening: ATE tester complexity is decreasing and more technology and complexity is built into the probe card interface. A better thermal solution is a very important parameter along with performance for better yield management. Memory applications are increasing the total power across a 300mm wafer, and wafer probe needs to dissipate this total power to sustain the set-temperature during test. Power density per DUT is increasing and it's very challenging to manage a stable wafer-level test temperature. 3D integration technology requires very precise probing technology in X, Y and Z, as micro-bumps may be easily damaged during the probing process. MEMS applications require a variety of testing environments such as pressure, magnetic, and vacuum environments; also, wafer shape and package style are becoming very unique depending on the application type.

Reducing the cost of wafer-level and package-level test in the face of more challenging technology and performance requirements is a constant goal. The demand for higher throughput must be met by either increased parallelism (even with reduced test times), faster handler speed, or process improvements such as asynchronous test or continuous-lot processing. 3D integration technology requires new contact technology for the intermediate test insertion which will be added between conventional front-end process and back-end process. New contact technology to probe on the singulated and possibly thinned die's micro-bumps or C4 bumps after the die is mounted on an interposer is needed. For the die-level handler, the main tasks are the alignment accuracy to enable fine pitch contact, die level handling without damaging the die, and the tray design that supplies/receives the die.

Packages continue to shrink, substrates are getting thinner, and the package areas available for handling are getting smaller at the same time that the lead/ball/pad count is increasing. In the future, die-level handlers as well as package handlers will need the capability to very accurately pick and place small, fragile parts, yet apply similar or increasing insertion force without inducing damage.

Temperature ranges are expanding to meet more stringent end-use conditions, and there is a need for better control of the junction temperature, immediate heat control technology, and temperature control to enable stable DUT temperature at the start of test. Power dissipation overall appears to be increasing, but multi-core technology is offering relief in some areas.

It is unlikely that there will be one handler that is all things to all users. Integration of all of the technology to meet wide temperature range, high temperature accuracy, high throughput, placement accuracy, parallelism, and special handling needs while still being cost effective in a competitive environment is a significant challenge.

Gravity feed, turret, and strip handlers have been added to the table while retaining the pick and place type handler. The gravity feed handler is used on SOP, QFN, and DIP packages. Turret handlers are widely used on discrete-type QFN devices. Strip handlers are used on the frame before singulation. Strip test enables high parallelism with fewer interface resources, which enables cheaper test cost. These additional three types of handlers are widely used on relatively low-end or low-cost devices. Evolution of these handlers is quite different but important for various type of LSI.

Table 1: Test Handler and Prober Difficult Challenges

Pick and Place Handlers (High Performance)	Temperature control and temperature rise control due to high power densities
	Continuous lot processing (lot cascading), auto-retest, asynchronous device socketing with low-conversion times
	Better ESD controls as products are more sensitive to ESD. On-die protection circuitry increases cost.
	Lower stress socketing, low-cost change kits, higher I/O count for new package technologies
	Package heat lids change thermal characteristics of device and hander
	Multi-site handling capability for short test time devices (1–7 seconds)
	Force balancing control for System in Package and Multi-Chip Module
Pick and Place Handlers (Consumer SoC/ Automotive)	Support for stacked die packaging and thin die packaging
	Wide range tri-temperature soak requirements (-55ºC to 175ºC) increases system complexity for automotive devices
	Device junction temperature control and temperature accuracy +/-1.0ºC
	Fine Pitch top and bottom side one shot contact for Package on Package
	Continuous lot processing (lot cascading), auto-retest, low conversion times, asynchronous operation
Pick and Place Handlers (Memory)	Thin die capable kit-less handlers for a wide variety of package sizes, thicknesses, and ball pitches < 0.3mm
	Package ball-to-package edge gap decreases from 0.6 mm to 0 mm require new handling and socketing methods
	Parallelism at greater than x128 drives thermal control +/-1.0ºC accuracy and alignment challenges <0.30mm pin pitch
Prober	Consistent and low thermal resistance across the chuck is required to improve temperature control of the device under test. There is a new requirement of active/dynamic thermal control, which can control junction temperature(ΔT) during test
	Both Logic and Memory wafer generates more wattage/heat, demand of Heat dissipation performance improvement is expected. Especially Heat Dissipation at Hot temperature is challenging technology for wafer prober.
	There are wafer handling requirements of non-SEMI standard such as 3DI, MEMS, WLCSP and PsP applications. Those are thin, thick, unique shape so customized wafer handling technique/technology is needed. Wafer cassette is needed to be customized to meet the request as well.
	Probing on micro-bump is technically proven but there are many challenges "parallelism/multi-site", "Thermal conduction" and "bump damages/reliability"
	Advances in probe card technology require a new optical alignment methodology.
	Dicing frame probers can cover a wide temperature range, but a dicing sheet cannot cover the full range.
	Greater parallelism/multi-site, and higher pin counts require higher chuck rigidity and a robust Probe Card changer.
	Power Device application requires very thin wafer which drive need for 'Taiko Wafer' and 'Ring attached wafer' handling and more high voltage chuck technologies.
	Enhanced Probe Z control is needed to prevent damage to pads, there are solution in the market but those must be optimized to integrate onto wafer prober to meet needs of test cost requirement.
Gravity Feed Handlers	Thinner packages and wafer will require a reduction in the impact load to prevent device damage
	Test head size increase due to higher test parallelism may alter handler roadmap
	Reduction of static electricity friction and surface tension moisture friction on very small packages (<1 x 1 mm)
Turret Handlers	Test contactor support for > 100A current forcing on power devices
	Kelvin contact support (2 probes) to very small area (0.2 x 0.2mm) contacts on small signal devices
Strip L/F Handlers	Testing process infrastructure configuration
	Accuracy of the contact position for high temperature testing environment

Table 2 (part 1): Wafer Probe Technology Requirements

Year of Production	*2019*		*2020*		*2021*	
MPU, ASIC, SOC and Mixed Signal Products						
Wirebond - inline pad pitch	40		35		35	
Wirebond - stagger pad pitch	45		30		30	
Bump - array bump pitch	30		30		30	
Sacrifical pad pitch in a field of bumps	100		100		100	
I/O Pad Size (µm)	X	Y	X	Y	X	Y
Pad Materials						
Wirebond	30	30	30	30	30	30
Bump	30		30		30	
Sacrifical pad in a field of bumps	45	45	42	42	42	42
Wafer Test Frequency (Hz)	2.4G		2.4G		2-10 GHz	
Wafer Test Frequency (Hz) for HSIO	25Gbps/12.5GHz		56Gbps PAM4 28Gbps NRZ @ 14GHz		100Gbps PAM4 @ 28GHz	
Probe Tip Diameter Wirebond	7.5		6.5		6.5	
Probe Tip Diameter Bump	25		25		25	
Probe Force Bump(gf) - at recommended overdrive	1.5		1.5		1.2	
Size of Probed Area (mm^2)	20000		20000		20000	
Number of Probe Points / Touchdown	180000		200000		200000	
Maximum current per probe >130um pitch	2A		2A		2A	
Maximum current per probe <130um pitch	1A		1A		1A	
Maximum contact resistance	<0.5		<0.5		<0.5	
Probe test temperature range	-55	200	-55	200	-55	200
Automotive Radar						
Wafer Test Frequency (GHz)	80GHz		80GHz		80GHz	
RF I/O Geometry	Solder Ball		Solder Ball		Solder Ball	
I/O Size (um)	100um Cu Pillar SB		100um Cu Pillar SB		100um Cu Pillar SB	
I/O Pitch (um)	300um		300um		300um	
RF Ports per Site	14		14		13	
Sites being probed together	2		4		4	
Total Number of RF Ports	28		56		52	
High Speed Digitial (TIAm CDR, VCSEL, etc.)						
Wafer Test Frequency (GHz)	67GHz		67GHz		67GHz	
RF I/O Geometry	X	Y	X	Y	X	Y
I/O Size (um)	50	50	50	50	50	50
I/O Pitch (um)	80um		80um		80um	
RF Ports per Site	24		24			
Sites being probed together	2		8		8	
Total Number of RF Ports	48		96		96	
802.11ad						
Wafer Test Frequency (GHz)	64GHz		64GHz		64GHz	
RF I/O Geometry	Solder Ball		Solder Ball		Solder Ball	
I/O Size (um)	80um		70um		70um	
I/O Pitch (um)	150um		125um		125um	
RF Ports per Site	32		32		32	
Sites being probed together	8		8		8	
Total Number of RF Ports	256		256		256	
5G						
Wafer Test Frequency (GHz)	45GHz		73GHz		50-60GHz	
RF I/O Geometry	Solder Ball		Solder Ball		Cu Pillar w/o cap	
I/O Size (um)	100um		70um			
I/O Pitch (um)	150um		130um		130um	
RF Ports per Site	34+		38+			
Sites being probed together	8		8			
Total Number of RF Ports	64		>100			

Table 2 (part 2): Wafer Probe Technology Requirements. NOTE VCSEL and PIC have different requirements

Year of Production	***2019***		***2020***		***2021***	
Optical Probe - ***NOTE VCSEL and PIC have different requirements***						
Minimum pitch between fibers (um)	**127**		**120**		**120**	
Fiber optical alignment accuracy (Multi-Mode)	**< 5um**		**< 10um**		**<10um**	
Fiber optical alignment accuracy (Single-Mode)	**< 0.1um**		**< 0.1um**			
DRAM						
Wirebond - inline pad pitch	**50**		**50**		**50**	
I/O Pad Size (µm)	**X**	**Y**	**X**	**Y**	**X**	**Y**
Wirebond	**40**	**50**	**35**	**40**	**35**	**40**
Sacrificial Pads	**45**	**50**	**40**	**40**	**40**	**40**
Wafer Test Frequency for Sort(Hz)						
Test Frequency(Hz)	**250M**		**400M**		**400M**	
Shared Signal Line Test Frequency (Hz)	**125M**		**200M**		**250M**	
Minimum pulse width	**2.0nS**		**2.0nS**		**2.0nS**	
At Speed Wafer Test						
Test Frequency(Hz)	**3.2G**		**3.2G**		**3.2G**	
Probe Tip Diameter	**8.5**		**8.5**		**8.5**	
Probe Force(gf) - at recommended overdrive	**2.5**		**2.5**		**2.5**	
Size of Probed Area (mm^2)	**100% of wafer**		**100% of wafer**		**100% of wafer**	
Number of Probe Points / Touchdown - Memory	**130000**		**150000**		**150000**	
Maximum Current (mA)/pin	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**
	250	**<.001**	**250**	**<.001**	**250**	**<.001**
Maximum Resistance (Ohm)	**Contact**	**Series**	**Contact**	**Series**	**Contact**	**Series**
	<0.5	**<3**	**<0.5**	**<3**	**<0.5**	**<3**
Probe test temperature range	**-45**	**150**	**-45**	**175**	**-45**	**175**
NAND						
Wirebond - inline pad pitch	**80**		**80**		**65**	
I/O Pad Size (µm)	**X**	**Y**	**X**	**Y**		
Wirebond	**50**	**60**	**50**	**60**	**50**	**60**
Wafer Test Frequency for Sort (Hz)						
Wafer Test Frequency(Hz)	**100M**		**133M**		**133M**	
At Speed Wafer Test						
Test Frequency(Hz)	**600M**		**600M**		**2.4G**	
Probe Tip Diameter	**10**		**10**		**10**	
Probe Force(gf) - at recommended overdrive	**3**		**3**		**3**	
Size of Probed Area (mm^2)	**100% of wafer**		**100% of wafer**		**100% of wafer**	
Number of Probe Points / Touchdown - Memory	**80000**		**80000**		**80000**	
Maximum Current (mA)/pin	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**
	250	**<.001**	**250**	**<.001**	**250**	**<.001**
Maximum Resistance (Ohm)	**Contact**	**Series**	**Contact**	**Series**	**Contact**	**Series**
	<0.5	**<3**	**<0.5**	**<3**	**<0.5**	**<3**
LCD driver Products						
Bump - inline pad pitch	**18**		**16**		**16**	
Bump - stagger pad pitch	**10**		**8**		**8**	
I/O Pad Size (µm)	**X**	**Y**	**X**	**Y**	**X**	**Y**
Inline	**11**	**50**	**11**	**50**	**11**	**50**
Stagger	**15**	**30**	**12**	**40**	**12**	**40**
High speed I/O pin freq (Mobile/TV)	**4.5Gbps / 6.5Gbps**		**4.5Gbps / 6.5Gbps**		**4.5Gbps / 6.5Gbps**	
Probe needle structure	**Cantilever / Vertical**		**Cantilever / Vertical**		**Cantilever / Vertical**	
Probe Tip Diameter (um)	**8**		**8**		**8**	
Probe Force(gf)	**2**		**2**		**2**	
Size of Probed Area (mm^2)	**5600**		**6800**		**6800**	
Number of Probe Points / Touchdown	**12000**		**12000**		**12000**	
Maximum Current (mA)/pin	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**
	300	**<.001**	**300**	**<.001**	**300**	**<.001**
Maximum Resistance (Ohm)	**Contact**	**Series**	**Contact**	**Series**	**Contact**	**Series**
	<0.5	**<3**	**<0.5**	**<3**	**<0.5**	**<3**

Table 2 (part 3): Wafer Probe Technology Requirements

Year of Production	***2019***		***2020***		***2021***	
CMOS Image Sensor						
Wirebond - inline pad pitch	**90**		**80**		**70**	
I/O Pad Size (µm)	**X**	**Y**	**X**	**Y**	**X**	**Y**
Wirebond	**60**	**70**	**60**	**65**	**60**	**60**
WLCSP	**46**	**100**				
WLCSP (TSV construction)	**55**	**55**	**40**	**40**	**40**	**40**
	200M		**200M**			
High speed I/O pin freq (Hz)	**2.5G**		**3G**			
Probe needle structure	**Vertical / MEMS**		**Vertical / MEMS**		**Vertical / MEMS**	
Probe Tip Diameter Wirebond (um)	**12**		**10**		**7**	
Probe Force Wirebond(gf)	**2**		**2**		**2**	
Size of Probed Area (mm^2) [3]- Visible light	**300x300**		**300x300**		**300x300**	
Number of Probe Points / Touchdown - IR [4]	**5000**		**10000**		**10000**	
Maximum Current (mA)/pin	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**
Visible light sensor	**250**	**<.001**	**250**	**<.001**	**250**	**<.001**
IR sensor [5]	**1000**	**<.001**	**1200**	**<.001**	**1200**	**<.001**
Visible Light Sensor / Optical Fiberoptic Transmission						
Maximum Current (mA)/pin	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**
Visible light sensor	**250**	**<.001**	**250**	**<.001**	**250**	**<.001**
Parametric (Process monitor)						
Inline pad pitch	**40**		**40**		**40**	
Inter-row pad pitch	**35**		**35**		**35**	
Pad Size (µm)	**X**	**Y**	**X**	**Y**	**X**	**Y**
In line pads	**20**	**20**	**20**	**20**	**20**	**20**
Probe Tip Diameter	**6**		**6**		**6**	
Number of pad rows	**2**		**2**		**2**	
Probe Force(gf) - at recommended overdrive	**2**		**2**		**2**	
Number of Structures /Touchdown	**8**		**8**		**8**	
Maximum Capaciance (pF pin to pin)			**1**		**1**	
Maximum Leakage (pA)/pin (10V / 1 Sec test)	**0.2**		**0.2**		**0.2**	
Maximum Contact resistance (Ohms)/pin	**0.3**		**0.3**		**0.3**	
Maximum Path resistance (Ohms)/pin	**3**		**3**		**3**	
Maximum Probe temperature Range (degrees C)	**-50**	**200**	**-50**	**200**	**-55**	**200**
Maximum test Frequency (GHz)	**3**		**6**		**6**	

Table 3: Wafer Prober Requirements

Year of Production	***2019***	***2020***	***2021***
Wafer Handling			
Wafer Size [inch]			
200mm Prober	6, 8	6, 8	6, 8
300mm Prober	8, 12	8, 12	8, 12
Min Bump Size[um]	15	15	15
Min Wafer Thickness[um]	200	100	100
Max Wafer Thickness[um]	3000	3000	3000
Max Wafer Weight[g]	350	350	350
Min Wafer Exchange Time (sec)	30	30	30
Tester Docking			
Test Head Weight[Kg]	1500	1500	1500
Probe Card			
Probe Card diameter[mm]	580	725	725
Probe Card PCB Thickness[mm]	10	18	18
Probecard Total Height [mm]			
Prober			
XY Accuracy (Probe to Pad) [±um]			
200mm Prober	2.0	2.0	2.0
300mm Prober	2.0	1.0	0.8
Z Accuracy (Probe to Pad) [±um]			
200mm Prober	5.0	3.0	2.0
300mm Prober	5.0	2.0	2.0
Chuck Planarity [±um]			
200mm Prober	7.5	7.5	7.5
300mm Prober	7.5	5.0	5.0
Chuck Maximum Force [Kg]			
200mm Prober	60	60	60
300mm Prober	450	450	500
Set temperature range [ºC]			
200mm Prober	-55 to +300	-55 to +300	-55 to +300
300mm Prober	-55 to +250	-55 to +250	-55 to +250
Chuck Temp. Accuracy [±℃]			
200mm Prober	1.0	1.0	1.0
300mm Prober	1.0	1.0	1.0
Chuck Leakage [pA]			
200mm Prober	0.1	0.1	0.1
300mm Prober	0.1	0.1	0.1
Total Power Logic (W/Die)			
300mm Prober	200	200	200
Total Power Memory (Watts Per Die)			
300mm Prober	0.75	0.80	0.80
Max Voltage [V]			
200mm Prober	10000	10000	15000
300mm Prober	10000	15000	15000
Max Electrical current [A]			
200mm Prober	300	300	300
300mm Prober	300	300	300

Table 4 (part 1): Test Handler Requirements

Year of Production	***2019***	***2020***	***2021***
Pick and Place Handlers (High Performance)			
Temperature set point range (°C)	-20 to 125	-20 to 125	-20 to 125
Temperature accuracy at DUT (°C)	±1.0	±0.5	±0.5
Number of pins/device	2500	4000	5000
Throughput (devices per hour)	2-10K	2-10K	2-10K
Sorting Categories	3-6	3-6	3-8
Maximum Power Dissipation (W/DUT)	400	500	700
Maximum socket load per unit (kg)	80	120	200
Maximum Package Size(mm)	50x50	75x75	90x90
Minimum Package Thickness (mm)			
Pick and Place Handlers (Consumer SoC/Automotive)			
Temperature set point range (°C)	-55 to 190	-60 to 200	-75 to 200
Temperature accuracy at DUT (°C)	±1.0	±1.0	±1.0
Number of pins/device	1000	1200	1200
Throughput (devices per hour)	2-30k	5-30k	5-30k
Sorting Categories	3-6	3-8	3-8
Maximum Power Dissipation (W/DUT)	40	40	40
Maximum socket load per unit (kg)	80	80	80
Minimum Package Size(mm)	2x2	2x2	2x2
Minimum Package Thickness (mm)	0.2-1.8	0.2-1.8	0.2-1.8
Pin/land pitch (mm)	0.3	0.3	0.3
Pick and Place Handlers (Memory)			
Temperature set point range (°C)	-55 to 155	-55 to 155	-55 to 155
Temperature accuracy at DUT (°C)	±1.0	±1.0	±1.0
Number of pins/device	50-1000	50-1000	50-1000
Throughput (devices per hour)	20-75K	20-75K	20-75K
Index time (sec)	2-3	2-3	2-3
Sorting Categories	5-9	5-9	5-9
Minimum Package Size(mm)	4x6	3x5	3x5
Minimum Package Thickness (mm)	0.2-1.8	0.2-1.8	0.2-1.8
Pin/land pitch (mm)	0.2	0.2	0.2
Ball edge to package edge clearance (mm)	>0.1	>0.1	>0.1
Gravity Feed Handlers			
Temperature set point range (°C)	-55 to 175	-55 to 200	-55 to 200
Temperature accuracy at DUT (°C)	±2.0	±1.0	±1.0
Parallel testing:	8 (2x4)	16 (2x8)	16 (2x8)
Throughput (devices per hour)	50k	50k	50k
Index time (sec)	0.6-0.8	0.6-0.8	0.6-0.8
Sorting Categories	3-10	3-10	3-10
Minimum Package Size(mm)			
Minimum Package Thickness (mm)			
Conformity tube type (mm)	280-580	280-580	280-580
Turret Handlers			
Serial testing	2-4	2-4	2-4
Index time (sec)	0.072	0.072	0.072
Throughput (devices per hour)	50k	50k	50k
Minimum Package Size(mm)			
Minimum Package Thickness (mm)			
Sorting Categories	5-9	5-9	5-9
Impact load to PKG (N)	3	3	3

Table 4 (part 2): Test Handler Requirements

Year of Production	***2019***	***2020***	***2021***
Strip L/F Handlers			
Temperature set point range (°C)	-55 to 155	-55 to 155	-55 to 155
Temperature accuracy at DUT (°C)	±1.0	±1.0	±1.0
Number of pins/device	6-250	6-250	6-250
Parallel testing:	1-256	1-256	1-256
Throughput (devices per hour) 1-16 parallel	20-120K	20-120K	20-120K
Index time (sec)	0.15	0.15	0.15
Sorting Categories	32	32	32
Min. Pkg. Size(mm)	0.8x0.8	0.8x0.8	0.8x0.8
Max. Strip Size(mm)	300x100	300x100	300x100

Test Sockets

The test socket is an electrical and mechanical interface responsible for good electrical connection and transference of high-integrity signals between the DUT and the PCB/tester through a mechanical contact mechanism in order to determine the electrical characteristics of the DUT. As semiconductor design and manufacturing capabilities have progressed in recent years, the testing process keeps raising the electrical and mechanical requirements of test sockets. Therefore, the socket technologies have been rapidly driven by significantly enhanced electrical and mechanical requirements, both of which are instigated by higher power/voltage/current, reduced package size, tighter pitches, higher pin counts, smaller solder resist opening, and so on. It has been indicated that electrical properties are determined by not only the electrical but also by the mechanical requirements. The multi-physics problems have made socket designs progressively challenging for these higher requirements. Current models show difficulty in making sockets for high ball count devices and achieving I/O bandwidths of > 20GHz.

Socket Trends

Table 3 contains the test socket technology requirements. The requirements have been divided into contacting NAND, DRAM, and SoC devices that are contained in TSOP, BGA, and BGA SoC packages respectively. The TSOP package is assumed to be contacted using a blade; the DRAM BGA is contacted with a spring probe, and the SoC BGA is contacted with a 50-Ohm spring probe. The test socket performance capability is driven by the pitch between balls or leads, so the lead spacing of the assembly and packaging roadmap was used to determine the pitch.

Contact blades are generally used for testing TSOP NAND Flash and contain a spring function in their structure, which is loaded by compressing the DUT into the socket. The structure is very simple and suitable for HVM; however, the contactor blade must be long to maintain the specified contact force and stroke, and to achieve a long mechanical lifetime. A weak point is that the blade contactor is not suitable for fine pitch devices due to the need to have isolation walls between adjacent pins. The thickness of the isolation wall must be thinner for finer pitches, which makes fabrication of the isolation wall more difficult. At the same time, the contactor blade thickness needs to be thinner for finer pitch, which complicates achieving the specified contact force, stroke requirement, and mechanical lifetime.

Spring probes, mainly used for testing BGA-DRAM devices, are formed by use of small-diameter cylindrical parts (probe and socket) and coil springs. Compression of the spring probe creates the contact load. In order to guarantee sufficient mechanical life, the probe diameter should be large enough to guarantee strength and durability and the length should be long enough to maintain sufficient travel under compression. The spring probe structure is relatively simple and easy to maintain and it is also easy to design a DUT loadboard.

According to the BGA-DRAM roadmap, the spring probe diameter will need to be smaller over time, driven by the finer pitch of the package ball roadmap. In addition, the spring probe will need to be shorter to meet the lower inductance values required to support the high frequencies of the roadmap I/O data rate.

Spring 50-Ohm probes required for BGA-SoC high frequency devices have coaxial structures that can reduce probe length transmission issues through impedance matching. However, advances in the package ball pitch through the roadmap will create restrictions to the coaxial pin arrangement structure (0.5 mm pitch in year 2016). The data rate will increase to 20GT/s in 2016, but the spring 50-Ohm probe will not have good electrical performance due to its multiple parts structure having higher contact resistance than other contactors. To support 50milli-Ohms of contact resistance starting in 2016, advances will be required in materials, plating, and structure.

Table 4: Test Socket Technology Requirements

Year of Production	***2019***	***2020***	***2021***
TSOP – Flash (NAND) – Contact blade			
Commodity NAND Memory			
Lead Pitch (mm)	0.3	0.3	0.3
Data rate (MT/s)	133	133	266
Contact blade			
Inductance (nH)	5-10	5-10	5-10
Contact Stroke (mm)	0.2-0.3	0.2-0.3	0.2-0.3
Contact force (N)	0.2-0.3	0.2-0.3	0.2-0.3
Contact resistance (m ohm)	30	30	30
Slit width (mm)	0.17	0.17	0.17
BGA – DRAM – Spring Probe			
Commodity DRAM (Mass production)			
Lead Pitch (mm)	0.25	0.25	0.2
DRAM RM GT/S	5.3	5.4	6.4
Spring Probe			
Inductance (nH)	0.2	0.2	0.15
Contact Stroke (mm)	0.2	0.2	0.2
Contact force (N)	<0.2	<0.2	<0.2
Contact resistance (m ohm)	100	100	100
BGA – SoC – Spring Probe (50 ohm)			
Lead Pitch (mm)	0,3 mm	0,25 mm	0,25 mm
I/O data (GT/s)	56 G/s	56 G/s	112 G/s
Spring Probe (50 ohm)			
Contact force (N)	0,3 (N)	0,2 (N)	0,2 (N)
Contact resistance (m ohm)	28 mOhm	28 mOhm	15 mOhm
BGA – SoC – Conductive Rubber			
Lead Pitch (mm)	0,3 mm	0,25 mm	0,25 mm
I/O data (GT/s)	56 G/s	56 G/s	112 G/s
Conductive Rubber			
Inductance (nH)	0,1 nH	0,1 nH	0,05 nH
Contact Stroke (mm)	0,1 mm	0,1 mm	0,05 mm
Contact force (N)	0.1	0.1	
Contact resistance (m ohm)	20 mOhm	20 mOhm	10 mOhm
Thickness (mm)	0.5	0.5	
QFP/QFN –SoC – Contact blade + Rubber			
QFP/QFN –SoC			
Lead Pitch (mm)	0.3	0.3	0.3
Data rate (GT/s)	20	40	40
Contact blade + Rubber			
Inductance (nH)	0.15	＜0.1	＜0.1
Contact Stroke (mm)	0.2	0.2	0.2
Contact force (N)	0.2-0.3	0.2-0.3	0.2-0.3
Contact resistance (m ohm)	30	30	30

Conductive rubber type contactors are used for BGA high frequency SoC devices. Conductive metal particles are aligned vertically in insulating silicone rubber which enables vertical contact and adjacent conductor isolation. Compared to other contacts, it is superior for uses with high frequency device test due to its low inductance and low contact height, but compression travel is limited. Conductive rubber will meet the fine-pitch requirement in the roadmap, but it is difficult to reduce contact force without decreasing the compression travel.

Contact blade + Rubber, generally used for testing QFP/QFN high frequency SoCs, is a combined structure of a short-length metal contact and compression rubber that makes contact thru force and travel. The required

compression force can be varied by changing the rubber material, but the life cycle is normally shorter than for a Contact Blade type contact.

Socket lifetime has not been pursued in this roadmap, but the lifetime problem will become more important in the near future as lead, ball and pad pitch becomes finer and pin counts get higher, which drives lower contact force to avoid lead/ball damage. Pb-free devices require higher contact forces than are required for non Pb-free packages.

Electrical Requirements

Socket electrical requirements include current carrying capacity (CCC) per pin, contact resistance, inductance, impedance, and signal integrity parameters such as insertion loss, return loss, and cross-talk. The higher the power and bandwidth the packages are designed for, the higher the CCC, the lower the resistance, and the better matched the impedance of the pins and/or sockets need to be. Data rate requirements over the roadmap timeframe are expected to exceed 20 GHz, which will greatly challenge impedance matching and potential signal loss. As package size, solder resist opening, and pitches become smaller and pin counts higher, the smaller pins required to fit within tighter mechanical constraints will greatly increase contact resistance and signal integrity issues. One of the critical parameters to stabilize the electrical contact and ensure low contact resistance is the contact force per pin, which generally ranges from 20 ~ 30 grams. As pitches get finer, smaller and more slender pins will be required, which may not be able to sustain a high enough contact force to have reasonable contact resistance. Due to the negative impact of mechanical requirements on electrical properties, it will be necessary to have improved electrical contact technologies or socketing innovations, in which the electrical properties and signal integrity will not be significantly impacted by or will be independent from stringent mechanical requirements. To handle these high-frequency signals, the user has to carefully consider the signal integrity of the overall test system including board design/components/socket.

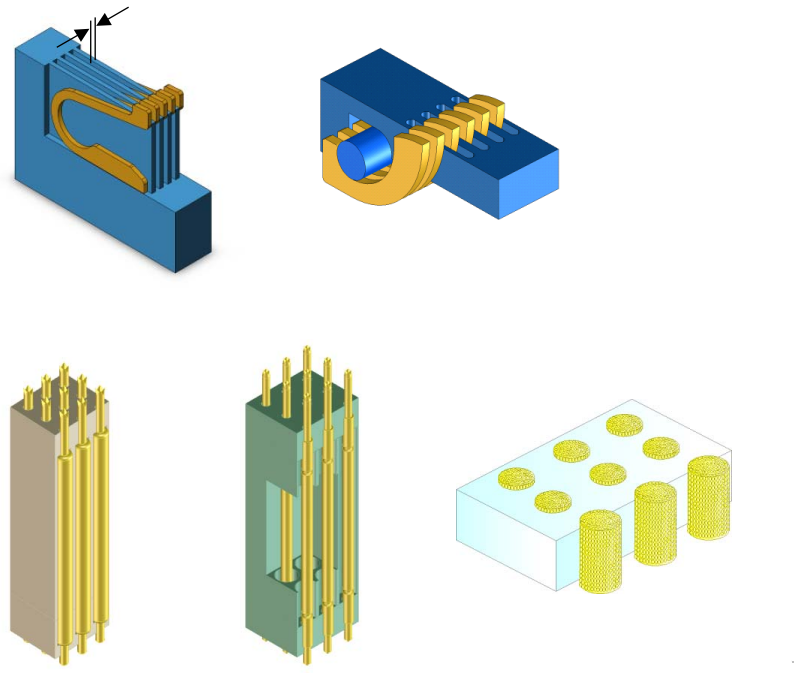


Figure 1: Contactor Types

Mechanical Requirements

The mechanical requirements include mechanical alignment, compliance, and pin reliability. Mechanical alignment has been greatly challenged by higher pin counts and smaller solder resist openings, particularly in land grid array (LGA) applications. Currently, the majority of test sockets use passive alignment control in which the contact accuracy between pin and solder resist opening is determined by the tolerance stack-up of mechanical guiding mechanisms. The limit of passive alignment capability is quickly being reached because manufacturing tolerance control is approximately a few microns. The employment of active alignment or an optical handling system is one

of the options to enable continuous size reduction of package and solder resist opening, smaller pitches, and higher pin counts.

Compliance is considered as the mechanical contact accuracy in the third dimension (Z-direction), in which the total contact stroke should take into account both the co-planarity of operating pin height and the non-flatness of the DUT pins, in addition to a minimum required pin compression. In general, the total stroke of the contact is between 0.3 mm and 0.5 mm. However, as required pin sizes get smaller, it may not be feasible to maintain the same stroke and thus the compression issue may become the bottleneck of electrical contact performance.

Contactor pin reliability and pin tip wear-out have also experienced challenges because tight geometric constraints prevent adding redundant strength to the pins. The testing environment becomes more difficult with higher temperatures, higher currents, smaller pin tip contacts, etc.

Section 8: System Level Test

System level test (SLT) refers to exercising the components of a system as an integrated whole to validate correct system operation for its intended end-use applications. We confine our notion of systems to be primarily electronics-based. As a whole, a system is comprised of both physical hardware and software programs. Electronic hardware scales from single chips, multi-die integrated packages, printed circuit boards (PCB), on up to racks of PCBs. For interacting with the physical world, some system hardware components can fall into non-electronic domains such as mechanical, optical, chemical, and biological. Software includes firmware, device drivers, operating system, and applications. The steady march of Moore's Law in semiconductors has enabled the creation of ever more complex systems. These systems are finding new applications and winding their way ever deeper into our daily lives. This trend is expected to continue at an accelerated pace as exemplified by smart cars, homes, cities, factories, healthcare, agriculture, etc. loosely aggregated under the umbrella term internet-of-things (IoT), enabled by high-data-rate 5G wireless communications.

The purpose and use of SLT differs depending on the perspective of the adopter: a component supplier versus a system integrator. Running SLT is a way for the component supplier to minimize defect returns from the system customer. Having SLT enables earlier time-to-market for the component while buying more time to improve ATE-based production test. The system integrator uses SLT to assess the quality of incoming components from multiple sources as part of an overall system validation process. Complex systems can have many configuration parameters which can be tuned by software to achieve optimal performance. Variability of components and their interactions may require the system integrator to run SLT to tune each system individually.

This section addresses a number of topics addressing SLT:

1. What's driving the trend toward more use of SLT?
2. What are the various SLT approaches and flows being used today?
3. What are the challenges and issues in using SLT?
4. What are the opportunities for improving SLT?
5. How will future system applications affect SLT development?

TREND TOWARD SLT ADOPTION

In the early days of simpler electronics, functional test predominated. As designs grew in complexity, propelled along the trajectory guided by Moore's Law, functional testing became untenable due to its high development cost and inadequate fault coverage to meet quality requirements. Scan-based structural test gained prominence since it enabled efficient automatic test pattern generation (ATPG) to achieve high fault coverage. As process dimensions shrank, fault models that underlie structural test kept pace with increasingly complex defect behaviors by evolving from stuck-at to transition-delay to cell-aware, delivering the required quality levels albeit with significantly higher test pattern size. However, the requirements are shifting yet again and the benefits of structural test may be hitting a limit. There is much anecdotal evidence that functional testing in the form of SLT is gaining more usage to augment structural test, in order to catch so-called "marginal" defects missed by ATE-based production test.

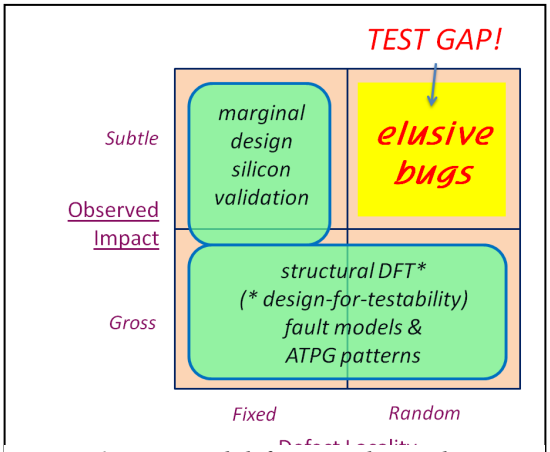


Figure 1. Marginal defects in advanced process nodes escape through test gap left open by existing test methods.

The notion of marginal defects arose in connection with inherent limits of optical lithography and process variation control in the manufacturing of devices at advanced nanometer nodes. Traditional test and debug methods are successful at catching defects with a gross observable impact that occur at random locations, or more subtle defects that are systematically localized. Left uncovered is a test gap for marginal defects that are both subtle in impact and randomly occurring (Figure 1). Exacerbating the issue is the growing divergence between test-mode versus in-system operation for today's multi-core system-on-chip (SoC) designs. For production test efficiency, complex clock and power domain interactions found during mission-mode are typically not exercised on ATE. Production test conditions can miss marginal defects which escape and cause failures under system operating conditions. These defects are difficult to resolve between field and factory, earning them the label of no-trouble-found (NTF).

Scan-based structural test is only applicable to synchronous digital logic. Specialized test schemes, usually executed via some form of BIST, exist for memory, high-speed I/O, and some analog functions. However, at the system level, defects that affect interactions among the constituent components are not well-covered by isolated

device/IP-level tests. Adding software to the mix further creates the possibility of soft failures when marginal devices are exercised under certain application scenarios that induce stress, triggering "application-enabled" defects. The trend toward product miniaturization in many end-use applications has driven packaging innovations to create multi-die system-in-package (SiP) solutions. For these advanced packages, testing to assure the structural and functional integrity of individual block-level parts may not be sufficiently comprehensive.

The test gaps identified above form the driving impetus towards SLT adoption. For the system integrator, performing SLT is par for the course, since the system is the end product. However, for the component supplier, SLT has not always been a necessity. To some extent, component functional test (in addition to structural test) performed on the ATE partially fulfills the purpose of SLT and is often good enough to ensure end-user system quality. But when both component and system complexities reach beyond a certain threshold, typical ATE-based functional test can no longer mimic the compendium of system-level interactions. For example, it is difficult to boot the entire Android operating system on a mobile phone SoC on the ATE. For the component supplier, it can be more cost-effective to build specialized SLT environments that can better mimic the end-user system and run a multitude of application scenarios to catch potential system-level defects. Thus, providers of sophisticated SoC and SiP components are seen to be early adopters of SLT.

SLT APPROACHES AND FLOWS

SLT practices vary widely, dictated by factors such as:

- System integrator vs. component supplier role
- Product market segment, unit volume, and lifetime
- Product quality and reliability requirements
- Competitive pressure in terms of time-to-market (TTM) and time-to-volume (TTV)
- Product profit margin

These factors determine how a company develops its own unique SLT approach, which can be characterized in a number of aspects:

- ATE vs. special SLT equipment
- Structural and functional SLT content
- SLT run time
- Full vs. sampled SLT
- SLT thermal, voltage, and cycling conditions
- SLT failure diagnosis

As shown in the example illustration (Figure 2) for a high-volume consumer mobile phone SoC, SLT today is inserted as the last step in the production test flow after wafer probe die test (WP), and packaged chip final test (FT). Typically, SLT is performed on special equipment distinct from WP and FT ATE. In this case, the SoC test board is comprised of the entire mobile phone system where the software stack from firmware to user applications can be run. The SLT flow is controlled by a PC-based test station with handler to load/unload multiple boards for parallel testing. If system components employ various scan-based test access standards such as IEEE 1149.x, 1500, and 1687 in a consistent and compatible manner, structural aspects of SLT to assess system assembly and connectivity can be executed more efficiently than relying purely on functional exercises. Reuse of modular BIST capabilities at the system level is another benefit.

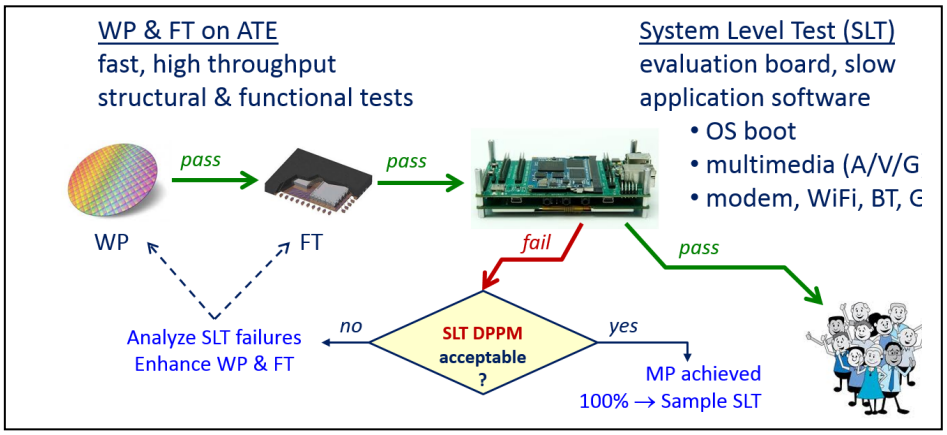


Figure 2. SLT flow for high-volume mobile phone SoC, initially 100% then reduced to sampling mode once WP/FT test quality has been enhanced to achieve acceptable SLT DPPM.

Typical SLT run times can vary from seconds to tens of minutes for high-volume products. For low-volume and long-lifetime products requiring high reliability, the system integrator may run SLT for hours

or even days. A component supplier can utilize SLT in the production test flow either as a permanent or a temporary step. Due to the relatively long run time of SLT, including it fully in production requires higher-cost SLT test equipment that can support very high concurrency in order to meet volume throughput demands. Full SLT typically applies to higher-end products selling at a premium, accompanied by higher user expectations. The cost structure of high-volume but lower-end products cannot support full SLT. Instead, to meet early TTM, full SLT is only utilized during initial production to deliver products of good-enough quality. Effort is spent on analyzing SLT failures to enhance WP & FT functional tests with the goal of reducing SLT DPPM. Once SLT DPPM reaches a level corresponding to acceptable customer quality expectations, SLT is shifted from full mode to occasional sampled monitoring. Quickly removing SLT as a throughput bottleneck enables the production TTV goal to be met.

Based on the experience of product deployment and in-field failures encountered, the system integrator may require the supplier to augment SLT with further preventive measures. For example, if power-up failure is noticeably significant, adding multiple cycles of OS boot under thermal stress to SLT may help flag problematic components. For safety-critical applications, it may require a serious and continuous effort in failure diagnosis to identify root-causes, leading to corrective actions including potential design and manufacturing changes.

There is increasing interest to employ big data analytics across the WP-FT-SLT flow to improve process efficiency and accuracy. These usually involve finding signature correlations among test measurements and failing behaviors which can be turned into predictive decisions as part of an adaptive test flow. Another approach being explored by industry is to move SLT upstream to WP, and for SLT to be ATE-based. This approach has two aspects: (1) replicate the full SLT environment and (2) add more SLT-like tests for the component. Being able to perform SLT at the wafer and/or die level will help provide Known-Good-Die (KGD) prior to subsequent multi-die integration steps. However, full replication of a highly complex SLT environment can be very expensive and is rarely justified except in special cases such as critical warfare electronics in a fighter jet. If system complexity is not high, it is possible to create the full system on the ATE load-board and run SLT using a low-cost tester (Figure 3). Besides the device-under-test (DUT), the load-board system includes flash memory to store SLT program components, a simple CPU as test controller, DRAM for system memory, and SLT self-checking circuitry to send DUT results to ATE. Such a "system functional test" approach has been used successfully to achieve exceptional quality for a line of low-cost consumer optical drive ICs.

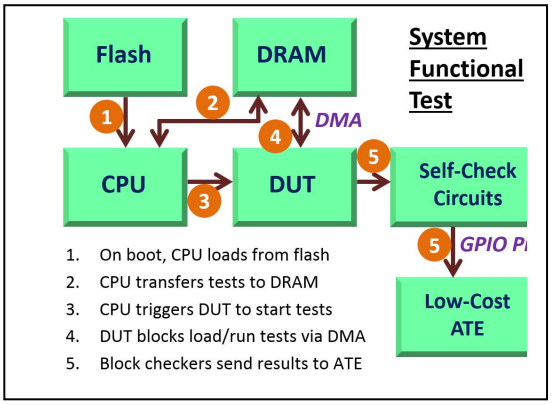


Figure 3. ATE-based system functional test.

Stand-alone component functional test executed on ATE can take advantage of on-chip processor and memory to run light-weight system programs to partially achieve the aims of SLT. Protocol-Aware (PA) testing is a recent development to ease the testing of modern device-to-device high-speed I/O interfaces that have complex non-deterministic protocols. It overcomes severe productivity challenges associated with traditional low-level fixed-timing ATE programming, and enables the test engineer to develop patterns at a higher level of system abstraction. Such patterns can cover more realistic and dynamic I/O protocol transactions to achieve better coverage.

SLT CHALLENGES AND ISSUES

Though its ability to catch defect escapes from WP/FT can be readily demonstrated, SLT's primary issue is the lack of quantifiable metrics to assess its effectiveness. Unlike structural test, SLT does not have well-developed fault models and an established statistical framework to link fault coverage to quality level. The fundamental reason underlying this widely-recognized issue has to do with the much more complex nature of SLT failures. Clearly, defects that manage to escape ATE test are harder to catch by definition. Reasons for failure in SLT are more likely to be murky and involve multiple contributing factors. For example, a compound failure may involve a piece of firmware code operating two interacting marginal devices in a manner that creates an unanticipated scenario that pushes system operation into an unsafe power-thermal zone. When running SLT, a common failure response is for the program to hang and time out without any indication of location and cause. Because SLT lacks the controllability and observability advantages of scan, there can be a long lag as in millions of clock cycles between when an error first occurs to when it propagates to a distant location where system failure is eventually observed.

Needed are methods of systematic SLT failure analysis to understand in more detail the defect mechanisms, error sequences, and associated statistical properties. That understanding then drives the development of effective test and

defect coverage strategies. Many elements of an SLT program are naturally derived from design verification where pattern development is guided by coverage of more abstract entities such as application scenarios, transactions and functions. Missing is a link from the higher-level abstraction to the underlying hardware in terms of covering where defects are more likely to occur, and de-emphasizing areas that are not problematic or already covered by WP/FT tests. Fault grading using traditional gate-level simulation is ineffectual due to severe run-time/capacity limitations and inability to model the entire system encompassing software, behavioral constructs, and non-digital hardware. Uniform treatment of low-level structural fault models without system context may waste computing resources and convey a false sense of adequate coverage. SLT program development, failure analysis, and enhancing ATE functional test remain largely a manual and inherently inefficient endeavor, contributing to its high cost of adoption.

In terms of test equipment, building an ATE-based full SLT environment in CP/FT faces ever higher barriers as system complexity increases. Even if it could be done, the cost is likely to be prohibitive in most situations. Thus, component manufacturers are more likely to choose dedicated SLT equipment used in an additional test insertion. Throughput then becomes the primary issue. A typical component SLT test station today is configured to handle four to six units in parallel. In order to meet high throughput demands under the constraint of long SLT run times, the manufacturer is forced to add more testers, more operators, and more factory floor footprint. Massively parallel SLT architectures are needed to raise throughput measured in unit-per-hour (UPH) by an order of magnitude or more. Being able to perform thermal stress testing is also gaining importance with more SoC devices implemented in FinFET nodes where self-heating is a concern as well as managing heat dissipation in tightly packed multi-die SiP components . Consequently, the SLT environment must support the evaluation of system thermal integrity.

OPPORTUNITIES TO IMPROVE SLT

Though SLT methods and practices are developed in a somewhat ad hoc fashion based on circumstances unique to each organization, many opportunities exist to introduce more formal and systematic approaches that can be shared across the industry to create commonality. Over time, an ecosystem can be built up based on standards, EDA tools, and equipment to enable more purpose-driven and efficient SLT flows.

As groundwork, research should be conducted, both theoretical and experimental, to develop fault models that embody emergent properties of complex component interactions. Complex systems are inherently hazardous since it's nearly impossible to anticipate all scenarios of failure, especially when software (rarely defect-free) and human behavior are integral to the system. Even when anticipated, failure mitigation measures constrained by cost may imply the acceptance of calculated risks during system design, calculations that may contain faulty assumptions. A system-level fault model should encompass, as a minimum, physical mechanisms of variation, noise, and aging, all the way up to application-induced stress scenarios that expose marginalities leading to heightened probability of failure. Having such a fault model will enable further development of system-level capabilities for fault grading, ATPG, and statistical correlation of coverage vs. quality. Unlike traditional device-level fault models, system-level faults may be difficult to explicate by single or simple root-causes. Thus, new approaches to fault diagnosis that consider compound causal relationships need to be investigated.

Since SLT is primarily function-oriented (as opposed to structural), much of the test patterns are procured from design verification (DV) which has also been facing growing design complexity as a serious challenge, to the point that DV now dominates the cost of SoC development. Starting from manually written tests to verify functionality, impressive progress has been made in the development of coverage metrics to assess quality, automated test generation such as constrained-random to improve coverage, and the standardized Universal Verification Methodology (UVM). Though successfully deployed widely, UVM has run into reuse scalability limitations: first in moving up the scope of integration beyond the level of verifying IP blocks to verifying complete systems with embedded software; and second in extending to non-simulation verification platforms such as in-circuit emulation, FPGA prototyping, and post-silicon bring-up.

The latest push by the DV community to overcome these limitations is to develop a common abstract model of verification intent named the Portable Test and Stimulus Standard (PSS). As shown in the PSS diagram (Figure 4), block-level abstract models can be combined directly to construct higher-level system verification tests. The same abstract stimulus model can drive test generation tools to target multiple verification platforms. PSS offers an opportunity for SLT to capitalize on major investments being made by EDA tool providers and users to realize a broad vision of verification reuse. As illustrated, the PSS abstract model can be extended to include system-level fault model and stress scenarios. Test generation tools can be enhanced to traverse the system activity graph, solving a set of constraints derived from SLT perspective to generate test patterns which can be run on production SLT

platforms. The test community should participate in the PSS Working Group (PS-WG) to promote the inclusion of SLT in its scope.

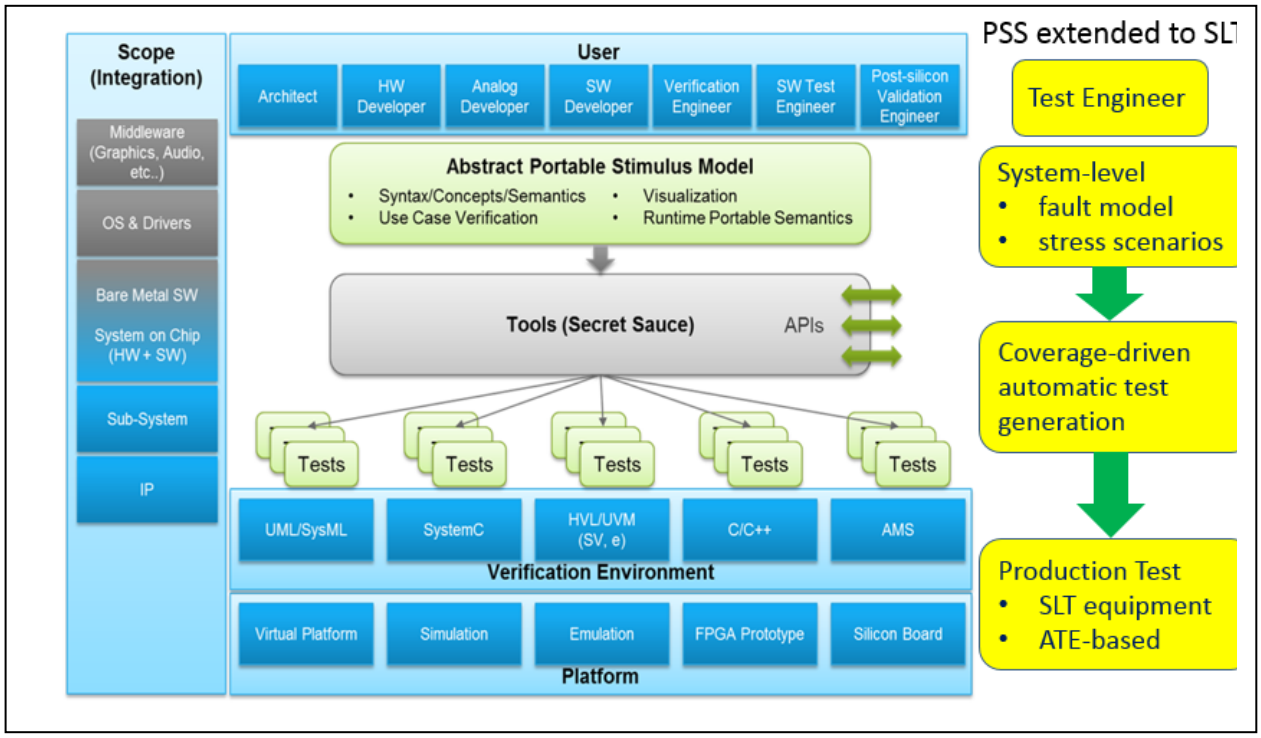


Figure 4. PSS serves as a framework to construct SLT flow that includes fault modeling and test pattern generation.

As with structural test, design modifications to increase controllability and observability should enable more effective and efficient SLT. Due to the close relationship between SLT and post-silicon validation, SLT design-for-testability (SLT-DFT) can draw on the latter's design-for-debug (DFD) practices as well as on-going industry and academic research to advance the state-of-the-art. In SoC designs, it is common to find on-chip instrumentation to support (1) profiling of performance aspects such as power consumption and speed, (2) monitoring of power grid noise and thermal excursions, and (3) debug and diagnosis of silicon failures by signal tracing. Instruments in hardware tend to favor observation so as to not disrupt normal system operation. Software also plays a crucial role in configuring initial silicon states, setting triggers for tracing, transporting logged data off-chip, and analyzing data for high-level debug. SLT can utilize data captured by embedded instruments during system operation to identify vulnerabilities in the silicon and improve failure diagnostic resolution. Of note is the recent appearance of commercial IP for embedded analytics (UltraSoC) and the growing popularity of USB as the device-under-debug to host interface. USB has the advantages of ubiquitous presence in SoC designs and high data transfer rate when compared to IEEE 1149.1 JTAG or IEEE 1687 IJTAG.

Another promising approach called Quick Error Detection (QED) aims to drastically reduce error latency via software transformations applied to existing validation tests. It has the option to insert hardware checkers (incurring a small overhead) to accelerate test execution time. Reduced error latency allows more precise localization of both functional and electrical bugs. QED transformations target bug scenarios abstracted from analysis of a diverse set of actual failures found in commercial multi-core SoCs. Thus, bug scenarios can be viewed as a kind of system-level fault model and the goal of QED is to generate tests to achieve high coverage of bug scenarios.

Embedded instruments provide a wealth of internal device data which can be exploited to optimize the entire production test flow. Adhering to the general concept of adaptive test (Fig. 5), upstream WP/FT data can drive downstream SLT decisions. For example, typical SLT failure rates (so called SLT DPPM) are quite low. Even a high SLT DPPM of 10,000 means 99% of devices will pass SLT. If one can predict based on WP/FT data even a modest portion, say 60%, of devices that are very likely to pass SLT, then limiting SLT to run only on the other 40% will realize large test cost savings. Furthermore, if WP/FT data can indicate potential weak spots in each device, SLT programs can be uniquely customized to target the weak spots and achieve more effective SLT screening, thus realizing higher quality. Customized SLT does require new flexibility for on-the-fly test program construction. To enable adaptive SLT, a learning phase is needed to find strong data correlations between WP/FT and SLT stages. The kinds of test data and correlation methods are active topics of investigation. For example, typical WP/FT test data may lack sufficient information. One approach applies delay-test scan patterns under non-destructive stress conditions on ATE to generate fine-grain localized internal health signatures which can be correlated against SLT pass/fail data to obtain predictive rules.

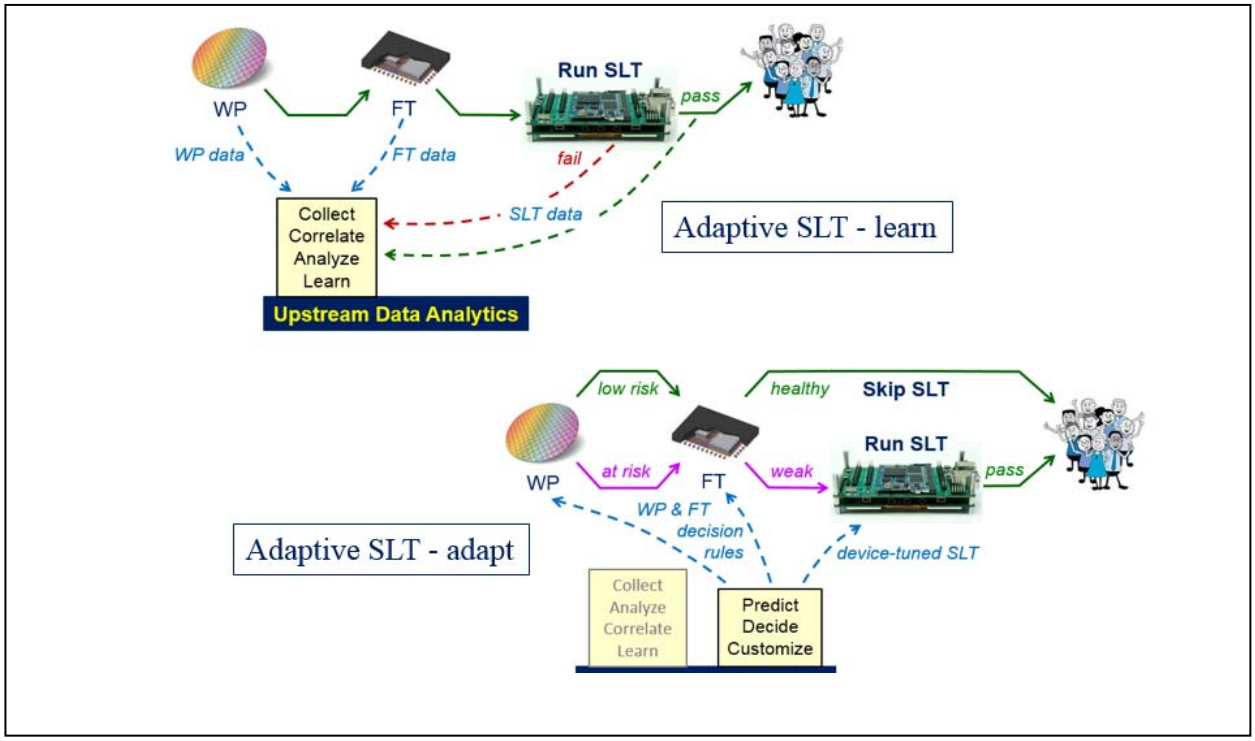


Figure 5. Adapt SLT flow using data analytics to optimize for cost efficiency and test effectiveness.

Sophisticated fault models such as cell-aware are being used today to target subtle system-level faults in WP/FT at great expense in test pattern count. Perhaps a more effective SLT can help off-load WP/FT to achieve an optimal and balanced production test flow. Creating a data loop between component suppliers' WP/FT and system integrators' SLT will require secure data exchange standards and means to protect sensitive proprietary information. Real-time data analytics also create opportunities for test equipment to add value by adopting machine learning accelerators.

IMPLICATIONS FOR SLT FROM FUTURE APPLICATIONS

The modern high-end car is a marvel of systems engineering with hundred-plus electronic control unit (ECU) sub-systems communicating over multi-level networks, all coordinated by software that can reach over 100M lines of code. Yet this level of complexity pales in comparison to the transformation currently underway that will take us into the future of self-driving electric vehicles. Withal the complexity lay the uncompromising requirements for reliability, safety, and resilience over an expected lifetime lasting ten years or more. This future robot on wheels brings together all of the key technological advances in 5G communications, artificial intelligence (AI) computing, and IoT applications. Needless to say, SLT will have to undergo a similar transformation to keep pace. The implications for SLT as applied to the robot car also extend to systems in other domains undergoing similar versions of the "smart" evolution.

Achieving the full potential of self-driving vehicles will require them to be part of an intelligent transportation system where wireless vehicle-to-vehicle/infrastructure (V2X) networking coordinates dynamic local operations. The key enabler is 5G millimeter wave (mmWave) beamforming which provides the fast network acquisition, short latency, and high data rate necessary for ad hoc network formation. Furthermore, due to space and weight constraints, the plethora of additional sensors dictate a portion of in-vehicle networks to also go wireless in support of various IoT-connected services. For 5G mmWave, the traditional cabled and controlled-environment testing approach no longer suffices. Instead, SLT has to be over-the-air (OTA) to check each individual antenna array element as well as the entire array's overall adaptive beamforming performance under realistic noisy conditions.

AI computing introduces software that is, in part, not written by humans, but instead trained by data in the form of deep-learning neural network (DNN) connection weights. The inexact nature of AI computation introduces a testing conundrum: Is a system failure due to defects in the DNN hardware implementation, or to insufficient DNN training that missed rare corner cases in the input data domain? Heavy AI computation also poses a power consumption challenge, especially if power comes mainly from stored energy, e.g., battery. Already inefficient conventional von-Neumann processors are being overtaken by highly parallel arrays of dedicated processing cores with local memory and data-flow accelerators. On the horizon are even more energy-efficient computation techniques spanning the spectrum from devices to algorithms such as analog, memristive, asynchronous, neuromorphic, and approximate. As structural testing techniques for synchronous digital designs do not easily extend to the above, it may be even more incumbent on SLT to fill the testing gap.

Future smart and energy-efficient systems may be architected from the ground up to be inherently redundant and error-tolerant much like the human brain. System construction will be based on complexity management principles of hierarchy, modularity, and abstraction. Such systems may have to operate continuously under extreme conditions with natural aging and degradation over time, thus requiring self-monitoring/testing/healing capabilities beyond T0 (defined as time of product shipment). Though component SLT may still be done as part of a less complex subsystem, SLT itself will evolve into the more encompassing built-in capability of self-aware systems.

Summary

SLT's rise in importance is inevitable as advances in semiconductor process and heterogeneous integration technology lead to ever more complex products. Its practice extends from system integrators to system component suppliers. Subtle software-hardware component interactions defy current production test screens, resulting in end-system application-dependent failures. Though SLT can help as an additional screen, there is much opportunity to improve its efficiency and effectiveness in methodology, automation, and test equipment. SLT also fits naturally within the overall adaptive test flow by utilizing cross-test data analytics. The test community should take advantage of the similarity between SLT and DV (particularly post-silicon validation) to push for PSS extensions to include SLT. Driven by upcoming 5G/IoT/AI-enabled transformations touching every aspect of our lives, SLT will most likely evolve into the self-monitoring/testing/healing capability of future smart systems.

References

S. Biswas and B. Cory, "An industrial study of system-level test," IEEE Design & Test, vol. 29, no. 1, pp. 19–27, February 2012.

P. G. Ryan et al., "Process Defect Trends and Strategic Test Gaps," International Test Conference, October 2014.

Z. Conroy et al., "A practical perspective on reducing ASIC NTFs," International Test Conference, November 2005.

A. Jutman, M. Sonza Reorda and H.-J. Wunderlich, "High Quality System Level Test and Diagnosis," Asian Test Symposium, November 2014.

H. H. Chen et al., "Predicting System-Level Test and In-Field Customer Failures using Data Mining," International Test Conference, September 2013.

Astronics Test Systems, "Massively Parallel System-Level Testing – The new Paradigm for Semiconductor Quality Control," https://www.astronics.com/test-systems/whitepaper---system-level-testing-for-semiconductors

Accellera Systems Initiative, Portable Stimulus Specification Working Group, http://www.accellera.org/activities/working-groups/portable-stimulus

P. Mishra et al., "Post-Silicon Validation in the SoC Era: A Tutorial Introduction," IEEE Design & Test, vol. 34, no. 3, pp. 68–92, May 2017.

M. Sadi et al., "An All Digital Distributed Sensor Network Based Framework for Continuous Noise Monitoring and Timing Failure Analysis in SoCs," Asian Test Symposium, November 2014.

D. Lin et al., "Effective Post-Silicon Validation of System-on-Chips Using Quick Error Detection," IEEE Trans. on Comp.-Aided Design of Integrated Circuits and Systems, vol. 33, no. 10, pp. 1573–1590, October 2014.

C. G. Shirley et al., “Board Manufacturing Test Correlation to IC Manufacturing Test,” International Test Conference, October 2014.
H. H. Chen et al., “Statistical Techniques for Predicting System-Level Failure using Stress-Test Data,” VLSI Test Symposium, April 2015.
H. H. Chen, “Data analytics to aid detection of marginal defects in system-level test,” Int. Symposium on VLSI Design, Automation and Test, April 2016.
A. Smith, “5G and IoT create a market for mmWave,” https://www.electronicsweekly.com/news/design/communications/5g-and-iot-create-a-market-for-mmwave-2017-06/

Section 9: Data Analytics

Data Analytics is a new section of the Test chapter in 2020, combining the *Adaptive Test* and *Test and Yield Learning* sections of the 2019 roadmap. Combining these sections provides a more cohesive roadmap for test applications. These applications have become increasingly interdependent and enabled by advances in Data Analytics, Big Data, and Machine Learning technologies.

Background

An IEEE Xplore® database search yields publications on Adaptive Test and Yield Learning dating back 30+ years. While advances in these areas have been achieved over the last several decades, it's challenging to apply the techniques holistically across the full semiconductor value chain. Limitations on our ability to efficiently collect, store, and analyze the massive amounts of available data have limited adoption to well-defined and self-contained applications. During the past few years, multiple technological advances have combined to change this landscape significantly:

- Internet of Things has facilitated the efficient collection of massive amounts of data
- Cloud Computing and Big Data technologies have turned data silos into Data Lakes and Data Meshes
- Tremendous advances in computational power and parallel processing have facilitated the adoption of advanced Data Analytics and machine learning models
- The combination of all the above has enabled rapid advancements in algorithm design and implementation

The foundation is now in place to strategically improve Adaptive Test and Yield Learning, by implementing Data Analytics, Big Data, and Machine Learning techniques.

Why is Data Analytics Important for Semiconductor Manufacturing and Test?

Today's challenges of increased design complexity and functionality, shrinking process nodes, increased quality and reliability requirements, and shortened time to market have combined to drive an exponential level of pressure to improve the semiconductor value chain. A massive amount of data – we conservatively estimate 1 terabyte (TB) data per day for a fully-loaded high volume fab – is collected across the semiconductor manufacturing supply chain and test flow. That data contains a wealth of information that can help optimize the overall test flow and discover hidden issues and relationships across process steps. For example, if the correlation between process drifts and yield is fully understood, immediate actions can be taken to maximize profit and ensure supply. Key insights can be unlocked by using advanced Data Analytics. Fabs can achieve further gains by strategically designing data collection to take full advantage of new and different analytic techniques.

Failure or delay in applying modern Data Analytics holistically across the semiconductor value chain will lead to increased costs and risks as fab complexities increase, and stop-gap measures are implemented to reach quality targets. Attempts to optimize manufacturing process steps individually, without full consideration of the interactions and dependencies across the entire process flow, will lead to diminishing returns. A test escape anywhere in the process flow reduces the quality level of the overall flow.

Tactical, localized solutions to manufacturing challenges are usually costly. One example is adding a System-Level Test (SLT) insertion as a back-end quality screen. By the time issues are detected, the manufacturing process is typically so far downstream that the effort required to truly root-cause and resolve them is only justified for the most major and systemic issues. With Data Analytics that relate SLT fails to other process and test data, issues can be caught sooner and SLT becomes one of a series of test insertions rather than a backstop.

The multiple benefits of applying advanced Data Analytics are described in the sections below.

Transforming the Backend to an Industry 4.0 Smart Factory

In essence, Smart Manufacturing is the trend towards automation and data exchange in manufacturing technologies. Representative processes include cyber-physical systems (CPS), the Internet of Things (IoT), Industrial Internet of Things (IIoT), cloud computing, cognitive computing, and artificial intelligence.

Today, there is a big push for automation and data sharing in the backend with expected improvements such as:

- Reduced costs
- Improved overall equipment effectiveness (increased uptime, accelerated output, decreased faults)
- Improved quality
- Data sharing across the value chain (although security concerns are significant)

The key requirement is the collection, sharing and acting on the data. For Smart Manufacturing there are three dominant data perspectives:

- Historic state - Review, analyze and model historical performance
- Current state - Monitor current state to enable real-time control
- Future state - Use history and current data to identify and plan for future improvements

It is crucial in Smart Manufacturing to measure and characterize every aspect of the manufacturing process, including material movement systems, products, machines and processes, without scrambling to consolidate data. A consistent and detailed strategy for collecting, analyzing, and categorizing data is essential.

Optimizing Cost of Test

The benefit of Cost of Test optimization is clear. As with any per-device costs involved in semiconductor manufacturing and test, savings drop directly to the bottom line as increased profit. However, this optimization process is challenging, given the tradeoffs that must be balanced:

- Adding tests to reduce the risk of test escapes will increase cost, and if other tests are removed to compensate, that may simply shift the problem to new test escapes. Any attempt to remove tests should be based on data-driven insights, within the context of achieving the more valuable objectives of optimizing quality and yield.
- HIR-related technologies such as chiplets drive a "shift-left" of testing to earlier insertions to guarantee known good die (KGD). However, overall test cell cost would significantly increase if tests are pushed to wafer probe (vs. package test or SLT) without consideration of the natural limits driven by interfacing and thermal testing challenges.
- SLT and burn-in insertions use equipment that is significantly lower cost-per-DUT than wafer probe and package test, but are both limited in the types of tests that can be performed.
- Incorporating additional design-for-test (DFT) circuitry uses valuable chip real estate but can reduce test system requirements.

Advanced Data Analytics are key to meeting these optimization challenges, through their ability to identify complex effects and interdependencies, both within a specific test insertion and across the entire test flow. This ability to optimize across the entire flow becomes especially important to keep test costs under control for complex devices requiring an added System-Level Test insertion, or devices for automotive applications that have rigorous multi-temperature testing and burn-in requirements. However, care must be taken not to over-focus on Cost of Test reduction, which can lead to missed opportunities for improving yield, quality, and reliability. Those factors have the potential for far greater financial and competitive benefits.

Please refer to section 11: Cost of Test in this chapter of the HIR for a detailed analysis of the Cost of Test roadmap.

Improving Quality, Eliminating Test Escapes, Getting to Zero Defective Parts-per-billion

The primary objective of production test is quality assurance. And the test process itself is an exercise in Data Analytics. The device under test is stimulated, response data is collected, and the data is analyzed to determine if the test passed or failed or if the device performs within specifications. Beyond this foundational notion, other analytical methods are used to further improve quality, especially in very high-reliability market segments such as automotive, which strive for "zero defects" outcomes. It is easy to simply add more test content and/or insertions, but that also adds cost. Instead, Data Analytics need to be used to optimize the test flow with the most effective content at each insertion.

Also, in the past 20 years or more, the industry has witnessed a tremendous growth in the application of advanced statistical methods in outlier analysis to detect and remove subtle time-zero or reliability at-risk devices. Data feed-forward methods are used to analyze upstream test data, to adaptively determine the appropriate downstream test content and minimize test escape rates. Data feed-backward methods are also used to adjust the manufacturing process and shorten the time to achieve entitlement yield and quality. Correlations across test insertions can identify drift or other issues. When available, historical data should be mined to set baselines, screening limits, and guardbands. Relatively new test methods have been deployed, which require analyses to measure the impact. This is required to ensure no new test escapes are created when displacing other forms of testing.

Finally, quality issues can originate in the test equipment instead of the silicon, and therefore test cell data must be analyzed to identify when that has occurred and how to mitigate it. These are just a few examples of how Data Analytics plays an ever-larger role in production test's primary mission of maximizing quality and reliability.

Improving Yield

According to the Integrated Circuit Engineering Corporation, yield is "the single most important factor in overall wafer processing costs," as incremental increases in yield significantly reduce average per-wafer manufacturing costs. In this regard, yield can be viewed as being closely tied to equipment performance (process capability), operator capability, and technological design and complexity.

Over the years, advances in fab technology such as more efficient air circulation systems and better operator capabilities, as well as efforts to lessen direct human contact with the production process through the use of automation, have led to a decline in particulate problems. Yet many semiconductor companies struggle to implement sustainable yield improvements due to resistance to changing established processes, an insufficient view of data, and isolated efforts, as well as a lack of advanced analytics capabilities.

Yield correlates to profitability. The traditional method to calculate yield is: (good devices per wafer)÷(total devices per wafer). A more holistic view of yield based on currency rather than simple counts provides greater insight into the value of yield improvement. Consider the following equation for yield: (# of sellable devices per wafer×average selling price)÷(manufacturing and test costs per wafer). Viewing yield in this way provides direct feedback on the ROI per wafer, and clearly shows that every marginal device added to the numerator, or every device that can be moved to a higher-tier bin, adds directly to the fab's bottom line.

Based on discussions with industry insiders, even a few tenths of a percent incremental improvement in yield may result in millions of dollars in incremental revenue. Because the marginal cost of each incremental device is effectively zero, every incremental good device is virtually pure profit.

Performance Grading/Binning, Build to Order, Tune to Actual Sales

To extract the most margin from a device it's increasingly common for one design to be used across a variety of applications and segments. Along with that, the ability to control how a device operates becomes more reliant on information that can be modified internal to the device. The more testing we apply to a device, the greater our ability to identify the optimal operating parameters per device. To leverage this knowledge we need feed-forward data systems to communicate part-specific information. Additionally, traditional binning during a single test insertion can no longer account for all the information needed to determine the grading of a device. Combining data from all insertions allows for greater modelling and control over device parameters.

An innovative approach could combine financial, sales and device data to tailor deliverables that exactly match customer requirements. This tuning optimizes manufacturing and test processes, which increases margins, improves lead time and increases supply elasticity.

Assembly Information Tracking

Increasing complexity of assembled devices and an increased need for device security drives a requirement to know where every component of a device originated and where it was placed on the device. For all components, we need the support of upstream supply chain sources and assembly facilities. These entities can provide data streams identifying date codes, supplier, assembly equipment, and lot information about each component. Analytics on this data stream can increase trust in the components, to ensure only validated devices are used. Further yield improvement is made by analyzing the data to identify faulty equipment, or suppliers that don't meet required quality levels.

Support of advanced packaging techniques, introducing heterogeneous devices packaging of high costs, such as SOCs (System on a Chip devices) with multiple HBMs (High Bandwidth Memories), or multiple "Tiled" SOCs. Data Analytics could provide the optimal mix and match for performance, quality and reliability of these very complex packaging solutions.

Data Analytics for Test - Building the Foundation

The long-term vision for Data Analytics is bold, where data from design, manufacturing, automated test, and system-level test will all be coherently aggregated and analyzed with advanced machine learning or other analytic algorithms, providing valuable insights that facilitate optimized process efficiency and yield. It will be especially important to lay the groundwork for successful application of Data Analytics approaches by embracing and implementing foundational elements such as full traceability throughout the manufacturing and test process flow, and adopting standards that promote the efficient collection, management, access, and sharing of data across multiple process steps and organizations. To successfully apply Data Analytics, the right data needs to be available when and where it is needed. This will require a holistic approach that includes effective negotiations for data access from suppliers and contract manufacturers, efficient and cost-effective methods for management and access of the data

across multiple process steps, and analytics tools that are accessible and usable by semiconductor test domain experts. It will be important to maintain separation between the data management technologies and the Data Analytics technologies so that rapid advances in the analytics technologies can be quickly adopted.

Data Analytics Roadmap

The color scheme of the *Enablers* column in the table below is:

- Yellow – solutions are known – but there are still adoption hurdles
  - Solutions may be widely accepted or solutions may only be company-specific.
- Red – Research & development is needed to develop solutions
  - Or some solutions may only be proprietary and not generally available.

Challenge	Current Status	3-to-5-year Projection	Enablers	Benefits
Transforming the Backend to an Industry 4.0 Smart Factory	Key building blocks for automation are in place but lacking some key requirements to enable autonomous decision-making (for example, standards).	Definition of key standards completed and widely adopted. Interoperability within a floor in place. Targeted deployments in place with interoperability across different floors/locations/organizations.	Machine-to-machine IoT communications infrastructure with native publishing and standard format generation. Equipment usage (including legacy) of standard data formats and messaging. Critical mass of integrated equipment and systems. Provision sufficient edge computing and storage. New paradigms in secure and trusted data sharing. 5G-based communications infrastructure.	Requisite foundational step to enable significant return-on-investment from implementing Smart Manufacturing.
Optimizing Cost of Test	Efforts have historically been on Cost of Test reduction at specific insertions, based on statistically- determined removal of tests during TTV phase. General agreement that more powerful capabilities are required.	Focus will shift from reduction to optimization, where the overall test budget is smartly applied across all test insertions based on advanced data analysis and Machine Learning algorithms.	Equipment usage (including legacy) of standard data formats and messaging. DFT and embedded sensor circuitry for cost-effective performance measurement. Proven data-driven approaches that provide significantly greater level of information vs. individual test insertion results.	Significant opportunities for test to become an increasingly value-add step: higher quality from better understanding of failure modes, higher yield from more intelligent binning, higher reliability from smart monitoring throughout device life cycle.
Improving Quality, Eliminating Test Escapes, Getting to Zero Defective Parts-per-billion	Post-test analytics, e. g., outlier analysis, are becoming commonplace. Still a struggle to put the right combination of tests in place during test execution to satisfy zero defects requirement.	More pervasive use of machine learning across all elements of test flow. Better data sharing across disaggregated manufacturing flows.	Equipment usage (including legacy) of standard data formats and messaging. Controlled and secure data sharing methods. Advanced embedded sensors and monitor circuits with dedicated machine learning and analytics for population baseline selection and meaningful parametric data for accurate outlier detection. Improved analytics and algorithms to identify at-risk product.	Higher quality products, faster and more successful product ramps to volume production, faster reaction to quality issues and more proactive responses in some cases. Ability to simultaneously optimize quality, yield, and cost.

Improving Yield	Yield analysis tends to be limited to one variable at a time commonality analysis. Data from different process and test domains tends to be stored with unique formats and access. The limiting issue is combining data from different processes in a format and system that understands them.	Data analysis tools will become more capable of reading disparate formats from multiple process steps. Access to greater "raw-material" data enables more valuable and capable systems to answer more complex questions. Transition from diagnostic to predictive, and ultimately prescriptive analytics.	Widely accepted data description, transmission, and definition standards. Data storage, query, and connectivity solutions to provide identified data to the chosen Machine Learning algorithm. An integrated data management infrastructure that enables cross-domain Data Analytics.	Predictive and prescriptive analytic solutions that utilize well -understood models to infer mult- variable causality downstream and make processing and metrology decisions upstream. The benefits will be reduced cost, increased manufacturing capacity, and higher yield and quality.
Performance Grading/Binning, Build to Order, Tune to Actual Sales	Commonplace requirement in order to fine tune design in process for optimum yield.	Tuning will continue to be required as margins are decreasing with smaller geometry technologies. Shift-left of volume/grading/binning decisions to earlier manufacturing and test steps. Analytics for correlating performance to fab and packaging process fluctuations.	Test Data Analytics to show optimal performance range. Can be home grown or industry available tool. Cloud-based solutions are becoming commonplace. Comprehensive Digital Twin capability for assembly and test equipment.	Increased ability to predict trends and plan accordingly rather than reacting to issues when they arise.
Assembly Information Tracking	First party components typically well tracked due to margin and quality requirements. 3rd party components offer limited visibility into source or device characteristics. Equipment on the manufacturing floor provides inconsistent data and formats.	Continued integration of 3rd party components and focus on performance will drive higher complexity in the assembly processes and greater effects on margin. Adoption of standard data and formats will be underway, likely starting as point solutions in the areas of greater complexity. Yield and quality excursion root-cause analysis will increasingly use BOM and equipment-related information.	Standards for all tools on a manufacturing floor to generate traceability of device and equipment data. Equipment trace data access and analysis capabilities. New paradigms in secure and trusted data sharing.	Enhanced security to ensure a device contains only trusted components. Increase or maintain trust and quality levels as assembly flows grow more complex and introduce new failure mechanisms. Faster root cause analysis and more effective excursion management and inventory management.

Impact of COVID-19 on Data Analytics Roadmap (Special Section for 2020)

We expect COVID-19 to be an accelerating force in the adoption of advanced Data Analytics. As activities migrate to a remote model, Data Analytics improves efficiency and facilitates increasingly decentralized operational models. Key examples where we expect accelerated adoption are:

- "Remote everything" is essential, across companies and continents. This drives better architectures for data management and delivery systems, to quickly access meaningful data and better anticipate where and when it is required. Tools will help disposition data between expensive quick-access and low-cost slower-access repositories.
- A bigger push to cloud-based services, and increased efficiency requirements for their use. While security concerns have fueled resistance to the cloud in the semiconductor industry, the requirement for

a more decentralized operational model, including substantial edge compute capability, may be a key catalyst for change.

- Predictive analytics and predictive diagnostics will become more important to maximize the effectiveness of support-related activities, especially where on-site travel is required.
- Smart Manufacturing becomes more important to help locate parts and ensure a reliable supply chain.
- Increased need for augmented reality-based remote support of factory floor equipment, new product ramp-ups, assembly monitoring, and supplier audits (need to deal with security risks/concerns).

To support the above, organizations need to increase focus on implementing modern Data Analytics tools. The operating model has been that test engineers know test, and data scientists/analysts know data, but we need to quickly bridge the gap between the two worlds.

Note: The remainder of this section includes the full text of *the Adaptive Test* and *Test and Yield Learning* sections of the *Test Technology* chapter of the 2019 HIR. This information has been retained in the 2020 HIR update for completeness, and for those who are interested in in-depth analyses on these topics. This entire section will be rewritten and streamlined for the 2021 full HIR revision.

Adaptive Test In-Depth Analysis

Adaptive Test is a set of IC manufacturing test methods enabling real-time optimization of the value of production test (in a fully automated way). These methods include the practice of using production test data to reduce/optimize test cost, improve product quality and reliability, drive yield improvements and improve product performance.

To effectively use Adaptive Test, additional development is needed in data infrastructure and data analytics, production test cell design, die traceability, and the coordination between IC manufacturers, IC test and assembly. The top challenges facing industry application of Adaptive Testing are listed in the portion of this section called "Adaptive Test Challenges and Directions."

There is a diverse set of Adaptive Test applications aimed at different product markets and requirements. Companies should target specific Adaptive Test applications aimed at their product requirements and target markets. Thus, the benefits of implementing Adaptive Test will vary based on type of product, manufacturing flow, volumes and technology.

This section provides:

1. A description of Adaptive Test and terminology used by its practitioners
2. Example applications of Adaptive Test as of 2019 and future opportunities
3. A list of challenges for the development and deployment of Adaptive Test.

To ensure the industry can fully exploit the benefits of Adaptive Test, this section describes:

1. The infrastructure requirements for test cells, data systems and device and system designs
2. A description of Adaptive Test challenges and the coordination needed between IC manufacturers, OSATs (Outsource Assembly and Test providers) and fabless system integrators.

Definition

Adaptive test comprises a set of methods for automatically changing manufacturing test conditions, manufacturing flow, test content, test limits, or test outcome to reduce test cost, increase outgoing quality and reliability, reconfigure parts, or collect data to further improve test and manufacturing. Adaptive Test strives to make these changes in a manner that does not significantly increase testing time or increase human involvement in test operations. The decisions on when and how to adapt the tests are made algorithmically by the tester, other test cell equipment, or an automatic data analysis system (given automation – the analysis time is significantly reduced compared with the traditional, engineering-intensive approach).

Adaptive Test Description

Adaptive Test is generally accepted as an advanced test strategy that can be used to achieve quality, yield, and cost goals that might not be reached by normal test methods. For example, Adaptive Test may modify a production test process in a number of ways:

1. **Test Conditions** (modifying test setup conditions or limits – such as voltage or clock frequency)
2. **Manufacturing Flows** (adding or deleting test insertions such as burn-in)

3. **Test Content** (modifying specific patterns or tests such as transition fault or IDDQ, respectively)
4. **Test Limits** (changing the pass/fail limits such as DC power or Vdd-min test specifications)
5. **Test Outcomes** (changing the binning or configuration of some die based on post-test analysis of the die's test results)

Adaptive Test applications are organized by when decisions are made to modify the test flow and to which device(s) the modified test flow are applied. The four most common categories are in-situ, feed-forward, feed-back and post-test. Figure 1 is a flow diagram depicting the relationships among these four categories.

1. **In-situ:** Data collected from the part being tested is used to modify the testing of the same device during the same test insertion. These methods include not only speed-binning and trimming calibration (which are not new), but using data from any test to modify test conditions or device settings for other tests for this specific device. For example, parametric data taken from on-product process monitoring structures may be analyzed – and the results used to drive subsequent test limits or test conditions or test-driven device reconfiguration.
2. **Feed-forward:** Data collected from a previous test step stage (e.g. probe, hot probe, burn-in) is used to change how the same parts are tested at a future stage. An example of the Feed-forward category are statistical methods which identify 'risky' dice or wafers and selects these components (only) for burn-in, or "clean" dice that may be candidates for reduced testing.
3. **Feed-back:** Data collected from a previous part (or parts) is used to modify the tests or limits of different devices yet to be tested. Skipping some test patterns on high-yield wafers, adding more tests to low-yield wafers or refining statistical models used for die classification are examples of this category.
4. **Post-Test:** Data sample statistics or other analysis is performed between test steps and is used to reclassify certain devices or to change future manufacturing flow and test conditions for these devices. Part Average Testing and outlier identification methods are examples of the Post-Test category.

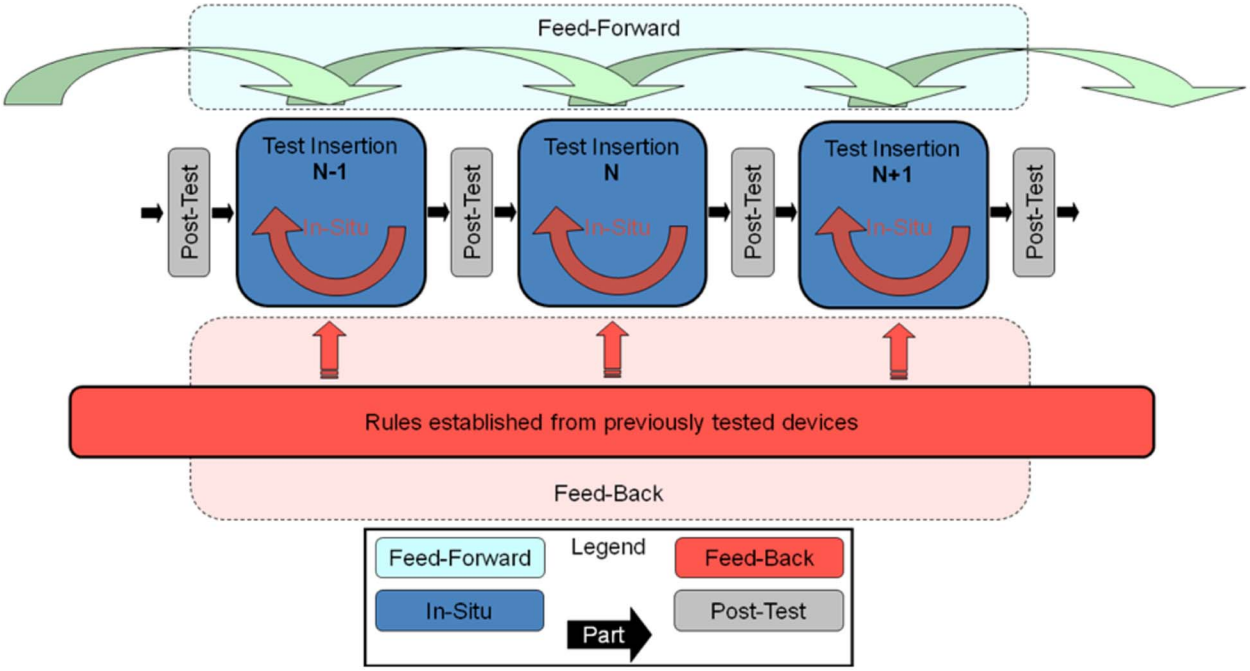


Figure 1: Adaptive Test supports feed-forward and feed-back data flows. Adaptive Test provides for data use for test decisions either in-situ (i.e. during a given test step) or post-test.

Example Applications

Below is a list of example Adaptive Test applications. Each example is labeled by one or two categories outlined earlier. In addition to clarifying the categories, the examples demonstrate the shift from manual and static methods to automatic methods with little or no human intervention during test execution. Note that the use of die ID (e.g., a die-specific identifier such as wafer/XY coordinate and lot information, or a unique identifier that is fused on each

die) is a key enabler for many of these applications. There is a list of references that include many example applications at the end of this section.

Some of these applications are more widely used today than others. Common methods include:

- Data feed-forward
- Good die, bad neighborhood
- Post wafer test statistical analysis – e.g., pickmap updates
- Device trimming or reconfiguration
- PAT (Parts Average Testing) and DPAT (Dynamic Parts Average Testing) – for outlier detection
- Burn-in optimization

1. **Dynamic test flow changes (In-situ, Feed-forward):** Die production data is monitored within the test program to add or remove tests, to selectively perform per-die characterization for yield learning, or to collect data for later diagnosis. This application supports many common real-time Statistical Process Control (SPC) methods.
2. **Statistical screening (Post-Test, Feed-forward):** After wafer or lot data collection, identify die which are outliers or mavericks as possible sources of test-escapes spikes or reliability failures. Statistical screening is Feed-forward because results can be used to route target dies through test flows different from the main flow.
   - PAT – Part Average Testing is a statistical technique in which a die is tested against static or dynamic test limits derived from other die in a common subgroup
   - NNR – Nearest Neighbor Residuals is a statistical technique relating a univariate or multivariate test result to a model derived from a local region on wafer of the device under test.
3. **Single-step flow control (Feed-forward):** Data from one test step is used to optimize testing at the next test step to focus subsequent screening on issues observed in the manufactured parts.
   - For example, inline test modifies wafer test; wafer test modifies package test; burn-in modifies final test; or package test modifies card/system test.
   - One method for doing data feed-forward is to store results in on-chip memory (e.g., Flash) that will be read and used at subsequent test steps.
4. **Off-tester optimization of test flows (Feed-back):** Off-tester data analysis drives test flow changes for future devices (fully automated).
   - For example, off-line analysis could optimize test flows, test content and test measurement routines using input from many sources including historical data, test capacity, required turn-around times, DPM (defects per million) requirements, expected yields and parametric data.
5. **Production monitors and alerts (In-situ, Feed-forward, Feed-back):** Data from multiple sources is merged for statistical analysis to control production test optimization beyond what has historically been possible.
   - For example, subtle parametric shifts from marginal wafer probe contacting can be automatically identified and action taken during production testing.
6. **Die matching (Feed-forward, Post-test):** Production data from various sources is used to support the build/test process for multi-chip applications and many of today's board build process to match specific die combinations during assembly.
   - Note die-matching data transfer may require world-wide data sharing, across multiple companies and throughout the entire supply chain.
7. **On-chip test structures and sensors (In-Situ, Feed-forward, Feed-back):** Data collected from auxiliary on-chip test structures such as ring oscillators, selected critical paths, on-chip voltage and thermal sensors, or on-chip reliability monitors is used to modify the die's test content, test limits or future test flow.
   - Sensor measurements can be used at all levels of assembly and test (including system operation) to monitor and adjust functionality.
8. **On-chip configuration (In-situ, Feed-forward):** Production test data (including test structure data) is used to adjust features in the design to improve a die's performance, power, margin, yield or reliability.
   - Emerging ICs have more on-chip configuration and adaptability such as clock tuning, partial goods (redundant spare cores), and voltage and frequency adjustments (including per core).
9. **Component System Level Test (SLT) test optimization (In-Situ, Feed-forward, Feed-back, Post-test):** Test results from current or prior operations are used to customize the test flow or enable test sampling.

10. **Card/System configuration and test (Feed-forward, Post-test):** Component test results (such as parametric data, yield characteristics or partial good data) are used to customize the card/system test flow or customize card/system test conditions.
    - Data feed-forward from die testing are fed-forward and used by the board test program to make decisions on whether to add specific content to test for marginality.
    - Data feedback is used to deliver card/system test results to IC suppliers – who use this data to adjust test content, test condition or test limits.
    - In-situ card/system test measurements (such as on-chip sensors for voltage or temperature) are used to modify board testing (e.g., adjust margins and/or performance). On-chip sensors can also be read during field usage to monitor aging and perform in-field adjustments and/or send results back to IC suppliers to adjust their test limits.
11. **Adaptive Diagnostics (In-situ, Feed-forward):** Test results drive advanced diagnostic data collection.
    - For example, on-chip BIST (built-in self-test) circuitry can be programmed on-the-fly to localize and characterize specific types of failures. But these methods must only be selectively applied to ensure reasonable test time/cost.
    - Many emerging chips have programmable on-chip test controllers that can interpret test results on-chip – and take action (test, diagnostics, characterization) without requiring extensive data collection being transmitted to/from the test equipment.

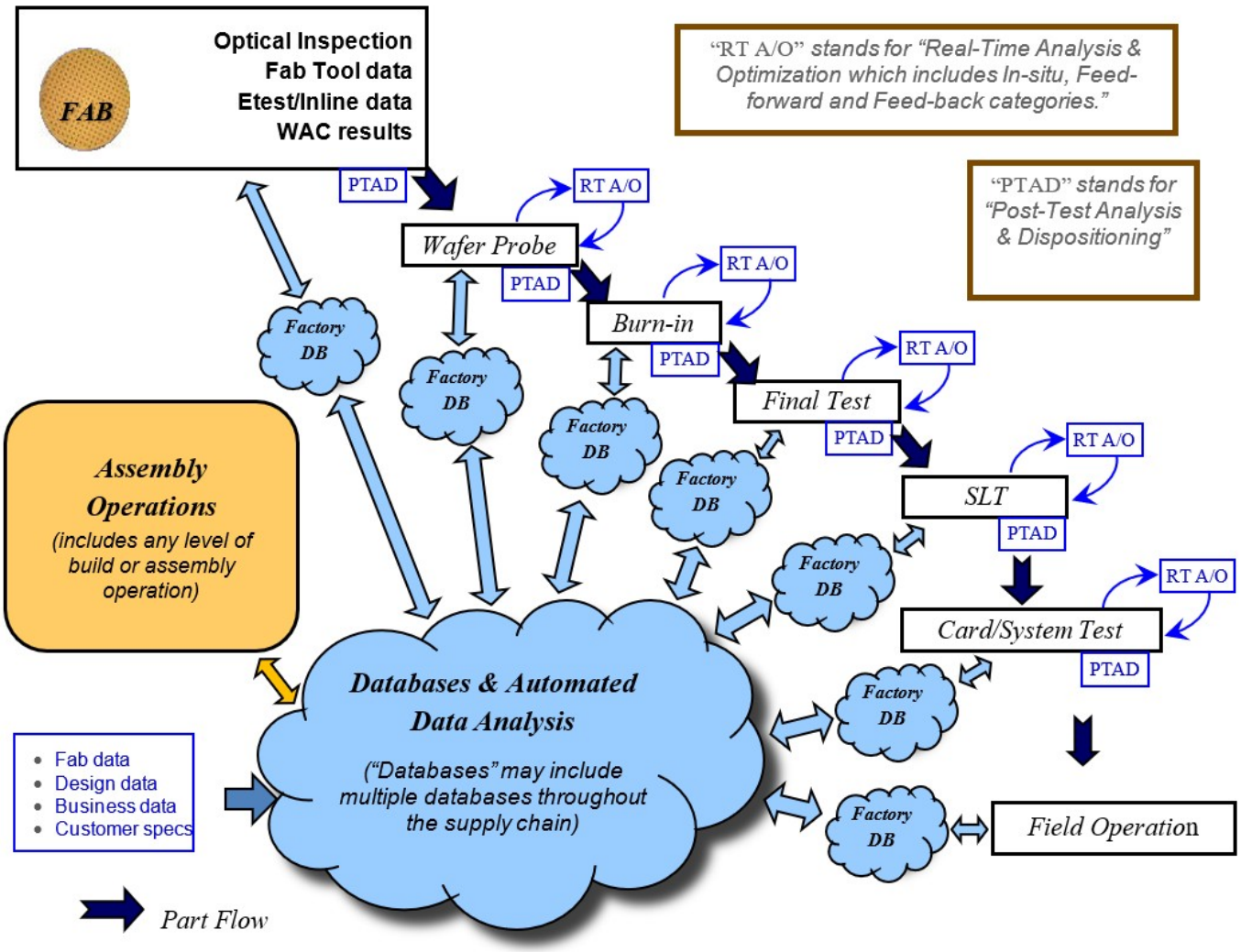


Figure 2: The architecture of Adaptive Test organizes each insertion's test data into one or more databases. A waterfall of manufactured parts may insert, join or query databases for test flow decision-making.

Adaptive Test Architecture and Infrastructure

Adaptive Test makes decisions throughout the manufacturing process based upon data from multiple sources and using data of varying detail and completeness. Before actionable decisions can be made with multiple-sourced data, new data integration requirements are needed. Some integration requirements are unique to Adaptive Test and are different from data requirements used at any of the originating sources.

Figure 2 displays a model of the entire End-to-End flow of parts under test and Adaptive Test applications. Note that in the flows shown in Figure 2; feed-forward, in-situ, feed-back and post-test dispositioning opportunities can occur at each test step. Although Figure 2 shows the total data store as a single database, the actual data structure would probably consist of multiple distributed database hierarchies each with unique capacity, latency and accessibility characteristics.

Overall Data Model Requirements for Adaptive Test

Making the decisions to adapt the test attributes listed above first involves collecting the proper data and then organizing the data into a structured data model so that the right data can be accessed when and where it is needed. At the appropriate time, data of the proper scope – that is data from a particular test run or data from a particular part, wafer, or lot – is accessed from the data model and processed by the applicable decision algorithms. Similarly, the test variables, such as limits, conditions, flow and content, must be changed at the right time to complete the adaptation decision.

The data model can exist entirely in an off-line database apart from the tester, or be distributed between servers and the tester depending on the latency requirements and convenience. To branch a test flow for a particular part (a real-time decision) latency must be short, i.e. there can be no significant impact to test time. To support low-latency requirements, the data needs to either be stored on the tester or be rapidly pulled into the tester. To make an outlier decision such as to re-bin already tested parts, longer latencies are tolerated such as from the time of test until wafer maps are uploaded or the parts are dispositioned. Longer latencies mean an off-line database can be used.

Decisions to adapt a test are often based on comparing the variation observed on the sample of parts in question to a model of the expected variation. In outlier detection, parametric test limits are adapted to track expected variation so as to only discard parts with unexpected variation. Tests can be temporarily dropped from a test flow when their control charts show the manufactured material and test process is in control and within specification. If and when monitoring based on sample testing shows the material or test process has changed and gone out of control, the tests are reinstated on every part. Similarly, diagnostic data collection tests can be added to a test flow when certain types of failures occur more frequently than expected.

Generally, more adaptability means more frequent decision-making in the test flow with the goal of improving the trade-off between shipped product defect level and yield/cost. Adaptability follows a bottom-up progression from the conventional static limit, to a static parameter variance model (static PAT), to a variance model with variable parameters (dynamic PAT), to choosing variance model equations based upon well-grounded principles. Moving up this progression requires not only more data but also a better understanding of the processes that cause the test response to vary. This progression also means the decision-making generally moves from an off-line human activity to an on-line machine activity.

Data requirements unique to Adaptive Test center on the database policies of latency, access and retention period. Latency measures the time between the request for a data item and the availability of the requested item. Access refers to the scope of the user community that can store, retrieve, and act on a data item. Retention period measures the time the data item is electronically available within the required access time period.

- **Local processing in the test cell requires low latency.** For example, access latency should be in a few milliseconds if data is to be retrieved on a per device level – Real-Time Analysis & Optimization (RT A/O). Post-Test Analysis & Dispositioning (PTAD) applications may have latency requirements of a few seconds to a few minutes. Normally data volumes for these steps would be relatively small. Processing in a central database (e.g., "The Cloud") has more relaxed timing constraints (minutes, hours), but typically deals with much larger data volumes.
- **Data access requirements are influenced by the diversity of users "touching" the data.** Applications that bridge assembly or test companies, fabless design companies, and test developers require robust access mechanisms to ensure the correct people and processes access only the appropriate data.
- **Data retention requirements depend on the specific market requirement.** Some requirements drive data retention of all shipped products to 10 years or more.

Many areas of IC manufacturing are increasingly more comfortable with using data from the cloud, but a notable exception is the test cell. Test cell integration of Adaptive Test algorithms is one of the most challenging applications. For example, local test cell actions (such as "clean probe-card now") were the sole responsibility of the specific test floor and were designed to guarantee the test cell integrity and test cell-to-test cell correlation. Adaptive Test changes this paradigm in a number of ways:

- **Algorithms will be owned by multiple involved parties,** including wafer fab, design house and test floor. Some algorithms may originate from commercial providers, others from the involved parties themselves. They all need be executed synchronously in a real-time environment.
- **Data collection as well as data access** (e.g., to upstream data in case of data feed-forward) **becomes a mission-critical task of a test operation as well as the entire supply chain.** This challenges the reliability of the underlying infrastructure, which likely spans multiple companies, geographic areas and cultures.
- **It is required to simulate the impact of algorithms on historical databases** to understand how to maximize the value of Adaptive Test without creating adverse side effects. This requires the exact same algorithm to be executed in as diverse environments as a cloud database and a test cell measuring real-time data.

As a consequence, the industry needs to develop:

- **Provenance models** to allow disparate data sources and users to access and trust data.
- **Data exchange formats** which are flexible, compact and standardized so that only minimal extraction and translation effort is required. A common set of indices is required, such that data remains identifiable and traceable even across heterogeneous supply chains.
- **Recipe management systems** which can handle a diverse set of recipe origins, check for (likely unintended) interactions and maintain consistency across non-synchronized update cycles from the various origins. Version control systems for these recipes are also required.
- **Monitoring and rule-checking systems** must be enabled to monitor the health of Adaptive Test algorithms (are basic assumptions met?) and the health of the distributed data, and escalate errors to the right entities in an actionable format.
- **Distributed data distribution infrastructure** which can handle a diverse set of data origins, monitor data changes, and maintain consistency across non-synchronized updates from the various sources. Data version control systems are also required, as are logging systems able to trace changes.
- **Shared analysis** may be the only way to jointly discover problems and opportunities. The large quantities of proprietary data may require that analysis be performed in a distributed manner with intermediate results shared.

Global Adaptive Test Infrastructure Requirements *(data exchange and archiving, reference Figure 2)*

This section describes the data storage and exchange requirements across the supply chain.

Data requirements that are different but not unique to Adaptive Test include date and time stamping, test naming, and date recording methods. For example, Adaptive Test data stamps should be consistent across all insertions and between companies. Current date stamping practices are ad-hoc with some companies using different date formats and date references at different test insertions. Database standards exist for date stamping such as Coordinated Universal Time. A time and date stamp requirement policy eases integrating test data when some units are retested and simplifies merging two (or more) data sets in an unambiguous time order. Similar issues arise in recording floating point results of the same test from different insertions with different precisions and formats.

The following are important attributes of a global infrastructure:

- **Latency.** Global latencies can be quite long, only needing to satisfy the needs of the typical uses. Data at this level is packaged in some lot-related group so access time in the order of lot processing times are considered adequate. These times (for non-local data) could be in days.
- **Volume.** While data volumes have increased and storage times extended so has the ability to store this data. The real issue is not the storage but the locating and use of this data. Data volumes are in the order of 2-20 Gbytes per tester-week with history storage spanning 4 to 20 years with the longer times being for automotive, medical and some industrial applications.
- **Provenance.** All characteristics of data must be trusted. Who created it, has it been changed, and is it accurate are all facets which must be known.

- **Security.** Adaptive test requires the sharing of data but it must control that sharing to those who have the need and right to use it.
- **Reliability.** Testing cannot stop due to infrastructure equipment or communications failures. Checks are also required to ensure that the required data is collected and stored.

Production Test Floor Infrastructure and ATE (and component/socketed SLT) Requirements

This section describes the data infrastructure requirements local to the Test Floor.

The test cell is expected to deliver a cost-effective means to screen defects for quality, classify devices for performance and collect data for learning. The rate of product complexity is increasing with more clock domains, voltage planes, IOs and configurable fuses followed by the introduction of parallel testing of multiple, dissimilar devices with 2.5D and 3D stack packaging. In parallel, the business demand for higher quality and reduced product cost severely challenges the ability of the test cell to continue to provide an effective test solution and still continue to reduce the cost of test. Adaptive Test methods provide levers to address these additional demands but not without disruptions of their own.

Adaptive Test requires the test cell to be able to accept input from external and internal sources, apply device-specific models to determine test conditions and to evaluate results for flow control and device binning. This materially changes the setup, execution and categorization requirements of the test cell and affects both low-level software capability such as firmware as well as high-level software such as the executable test program. Of particular challenge is the relationship of flow control and binning when the test flow becomes a function of non-deterministic evaluations influenced by dynamic test points, limits and device configuration.

The following are important attributes of a test floor infrastructure:

- **Latency.** Adding Adaptive Test should not impact the throughput for a test cell. Data collection must have minimal impact. In addition, the various equipment in a cell must respond quickly to changes driven by an adaptive test rule. For real-time control, the response time should be in the low milliseconds for each part. Some applications require require that data taken from a part can be used for binning that same part in real time. There is also lot feed-forward which is more of a lot-based time scale, usually hours.
- **Volume.** Required data volume to enable Adaptive Test varies by Adaptive Test method. Incremental data for real time decisions is small whereas data for feed-back applications can be much larger. Data volumes in the range of 2-20 Gbytes of data per tester per week (while archiving this data monthly) are not unusual. The increase in data volume also brings challenges to the network infrastructure supporting the tester with increased demand on both reliability and bandwidth.
- **Provenance.** All data generated from a test cell must be fully trusted. Data should be signed and possibly encrypted.
- **Security.** Within a test floor, this is not usually a concern. With the advent of keys and other sensitive data being stored into parts, it is becoming necessary to comprehend security methods such as selective encryption of data from specific parts.
- **Reliability.** Testing cannot stop due to infrastructure equipment or communications failures. Each test cell must continue at least the current lot without floor communications. This implies that each cell has sufficient data resources (data bases and storage) to continue.
- **Legacy Support.** Support for legacy testers, handlers, probers and products.

In addition, future test cells will have to support the following:

- Per-device test flows (and per-device limits, test conditions and content) based on external inputs, the device itself, and dynamic business rules.
- A move from being the entire cell controller to a test engine with a standard API.
- Asynchronous and distributed (multi-insertion) test flows.

Much work is being done today outside of the semiconductor industry to address the needs for machines to communicate and for the data generated by these interactions to be collected, analyzed and acted on. The semiconductor industry should take advantage of these as appropriate. Groups like SEMI CAST (www.semi.org/en/collaborative-alliance-semiconductor-test) have been organized to stimulate discussion within the industry to standardize around available technologies.

- **IoT (Internet of Things) and MTM (Machine-to-Machine).** Communications protocols, e.g., MQTT (Message Queue Telemetry Transport) must be both secure and extendable.

- **Distributed file systems.** Systems like IPFS (Interplanetary File System) are proposed to handle the distribution of asynchronous data while maintaining trust.
- **Streaming data analysis.** While data mining of large collections of data is a popular topic today, the concept of treating data as a stream is more appropriate to the needs of the semiconductor industry.
- **Replay/simulation capability.** Data systems must have the capability to evaluate the impact of different test rules and flow. For example, there must be a "replay" capability where a user can change test rules – and evaluate the quality, cost and other impacts.

Test results driving reconfiguration of "Adaptive Designs"

More and more designs are being reconfigured during testing. Examples include partial goods (on-chip redundancy), VDD/frequency adjustment per die, and local clock tuning. In most cases this product personalization will be based on either test measurements or data feed-forward from other operations. In some cases, this reconfiguration will be based on "application demand".

Testing resilient features

Resilient features are on-chip structures that can be used to configure the product to work around hard defects or to tolerate latent defects. These structures span a wide range of circuits and architectures, including fuses, redundant memory elements, architectures capable of operating on reduced numbers of logic elements like CPU cores or graphics components, error-detection and retry schemes for hard and soft errors, and the sensing and control circuitry for adaptive designs. Like every other circuit, these structures must be themselves tested and characterized, though these circuits present unique testing challenges beyond standard logic and memory elements, including temporary configurations (for fuses), soft-repair vs. hard-repair validation (for memories), combinatorial explosion of down-cored variants of redundant features, the need for error injection to test recovery circuits, and analog stimulus for sensors (such as voltage or aging monitors).

While resilient features are widely used for memories, it is currently less frequently applied for logic cores. This could change if an automated approach were available to help chip designers employ partial-good die on their chips; this will need to include power-off means for bad cores and functional operation in the presence of non-functional core instances. This is in addition to DFT to isolate such cores and fuses to disable the bad ones once identified by testing. EDA companies need to pursue means to help designers add this to their chips.

Non-deterministic device behavior: Test and run-time availability

Non-determinism is incompatible with traditional cycle-accurate automated test equipment, but is nonetheless becoming typical on modern SOCs. Several new I/O protocols are non-deterministic, as are the standard methods to avoid metastability in clock-domain crossings (which are commonplace in highly integrated devices). Fine-grained power gating and power-state management can change the configuration of a device and its execution profile during test and normal operation. Adaptive designs take this notion even further with architectural features which can perform state rollback and pipeline retry based on events at arbitrary times during execution. The result is that test patterns, particularly functional patterns which execute in mission mode, must either prevent or be tolerant of non-deterministic response. The former raises coverage questions, the latter pattern and ATE interface challenges.

Testing Adaptive Designs

Adaptive designs bring the complexity of dealing with advanced power management such as power gating, variable configuration of IP (such as IO and arrays), self-defining performance bucketing and part-specific reconfiguration (such as redundancy, repair and harvesting) to a test environment traditionally characterized by a linear test flow measuring to fixed corners to verify device operability. Instead, on-chip sensors are used to detect the workload, voltage, temperature, and timing margin of the chip as it operates and dynamically adjusts power supplies, clock frequencies, thermal control, and even the instruction stream. The adaptive features of a design make it much harder to define (and thus characterize) both typical and worst-case use models, which in turn makes it more difficult to test appropriately. Additionally, the removal of excess margin represented by traditional guard-banding increases the risk of exposure to subtle defects, necessitating both higher coverage and better correlation between structural and functional test modes.

An emerging direction is to apply Adaptive Test techniques (which modify the parameters or content of a test program based on information collected about the device under test from the current or previous test insertions) to adaptive designs (which modify their own operating point or execution flows based on internally generated feedback). The proclivity of an adaptive design to compensate for the environment in which it is being (functionally) tested will

present challenges for data gathering by the Adaptive Test process beyond opening control loops to test at fixed conditions. A means to record and store the conditions to which the device is tested, organized in a manner for ease of retrieval and consumption, is required.

Online Testing for Automotive ICs

Automotive standards such as ISO26262 are driving the need to perform testing in real-time while the system is in use. Common online testing methods include Logic Built-in Self Test (LBIST) and memory BIST and repair. Results from these online tests will drive reconfiguration – for example, to work around logic blocks that fail the online LBIST. It is clear that, as online testing is used more widely, the industry will need to develop methods to validate these BIST methods and ensure they are defect free. Also, it will be required to verify reconfiguration capability.

Adaptive Test for Fab Process Feedback & Control

There is a long history of defect-related fault information identified through test for use in yield improvement by the fab, but parametric tuning feedback related to performance has been much more limited. The use of data collection and statistical models applied through Adaptive Test methods is showing promise and an increasing level of interest in providing product performance related feedback to the wafer fabs. Example applications include optimizing the "sweet spot" for N and P device targets to meet customer demand distributions as well as providing direction in optimizing process modules for transistor-limited or R/C-limited performance improvement.

Adaptive Manufacturing (Post Fab)

An emerging direction is using test results to drive other IC production steps such as packaging multi-chip products. For example, the card/board assembly operation may require that specific dies or types of dies be used on specific boards based on previous test results. Given the emergence of multi-chip packages (such as 3D ICs) and power constraints, specific bare dies will need to be selected for assembly based on parametric data collected at test such as IDD or power/performance measurements. This opportunity is broader than just focusing on die build, and should include the entire post-fab supply chain including board, card, module and system manufacturing.

Key challenges of End-to-End data feed-forward for assembly operations include:

- Cross-company data management
- Build logistics for selecting specific dies/packages
- Robust data availability
- Data security
- Full traceability
- Data format standardization

Adaptive Test for Card/System/Field

Adaptive Test methodologies described in this section for IC-level testing can be equally extended and applied to board and system testing and even field usage. While ICs have traditionally been tested standalone in an almost 'noise-free' ATE environment and/or tested with limited structural tests to see whether they perform to their specifications, the board/system environment can be quite different in terms of noise, timing margin, voltage and functional test trigger conditions that structural tests were unable to produce. Improved board yield and IC DPM can be improved significantly where adaptive test that includes the board/system level performance is able drive enhanced screening both at the IC suppliers test and/or board/system manufacturing test.

The four types of Adaptive Test described (In-situ, Feed-forward, Feed-back and Post-test) can all be extended to include board and system manufacturing.

One of the difficulties in extending the chip-level adaptive test to board/system or even in-field test is to track their test trigger conditions and be able to convert between them. For example, chip-level scan-based logic gate test may not always be applicable for board/system/in-field tests due to the difficulties or impossibilities of controlling the scan chain data, clock pulse, non-stoppable in-field online function executions, etc. Similarly, a functional execution, which can be treated as a functional test, may be hard to convert to a chip-level ATE test because the function execution could involve memory contents, their transactions, logic and I/O data flow, etc. Therefore, tracking the test/failure conditions and the capability to convert between them is the key for adaptive test extension to board/system level.

The color scheme of the table below is:

- White – Manufacturing solutions exist today.
- Yellow – solutions are known – but there are still adoption hurdles
  - Solutions may be widely accepted or solutions may only be company-specific.
- Red – Research & development is needed to develop solutions
  - Or some solutions may only be proprietary and not generally available.

Challenge	Status	Needs
Recipe (Input variables, data treatments, output variables/ responses)	Many good outlier recipes exist. Spatial interpolation over a wafer is becoming popular for sample testing. Opportunity to branch test flow for only fault coverage required by current defect rate	Guidance on best measurements to make
		Guidance on which recipes to apply
		Fault or defect coverage metrics
		Higher level of adaptability where best variance model is automatically discovered instead of chosen beforehand
Decision Rule (Define actions resulting from recipe output)	Actions defined for gross outliers; loosely defined for less extreme events such as downgrade or reconfigure or escalation/ de-escalation of decisions (such as test sampling)	Where to set outlier thresholds
		How to combine the results of multiple outlier definitions (e.g., develop metrics for Quality or Reliability)
		Criteria for rejecting vs. downgrading or reconfiguring
Infrastructure (Foundation to enable execution of recipes and decision rules)	Part traceability enables feed-forward, feed-back but robust environment for data transport, storage and providence is lacking (no commercial solutions currently exist).	Standard data formats amenable to adaptive test
		Move from working with files to working with databases
		Ability to feed data and decisions both forward (with the parts to future stages) and backward (future parts at given stage) in test & assembly manufacturing flow
		Full traceability of adaptive test parameters for each part: limits, content, flows, decision rules, model parameters
		Full part configuration as tested (e.g., redundant cores, partial goods, on-chip tuning, multiple die such as 2.5D/3D in a package)
		Real-time communication among test cell machines and data storage and analysis engines
Evaluation (Execution of test cases to prove viability and benefit)	Receiver-operator curve concept understood by most practitioners but standard methods for experimental definition and ROI interpretation do not exist.	Clear evaluation criteria to build trust in adaptive test methods
		"Gold standard" against which to compare outliers (including all variations of adaptive test flows & settings)
		Good metrics for continuous monitoring of recipe effectiveness.
		Replay capability to evaluate new rules/recipes.
		Quantification of cost of shipping a bad part
Deployment (Implementation and release into production use)	Company-specific implementations currently utilized. But current deployments vary widely in their capabilities and there are still significant implementation hurdles to apply at all Test steps.	Commercial adaptive test platform into which methods can be plugged and recipes specified
		Connections to Manufacturing Execution Systems & Product Data Management systems
		Real-time (unit level) decision making that requires decisions based on off-tester analysis using broad set of data. (and update die result in real-time on the tester)
		Complete visibility across supply chain: fab, test, assembly both internal and external
		Supply chain data integration and processes which automatically detect supply chain issues and implement corrective actions in near real-time

A key emerging requirement is to enable full end-to-end correlation analysis – e.g., from fab data, through product die test, board/card/system test and field operation. The infrastructure to enable this capability broadly is one of the industry's key challenges.

Extending Adaptive Test applications to the board and system level requires extensive data infrastructure, analysis, exchange and security. Companies providing ICs, board design and test need to openly collaborate on a technical and business level to be successful.

Adaptive Test Challenges and Directions

This section highlights the key challenges that the industry must address to fully exploit Adaptive Testing across the supply chain.

Summary

Adaptive Testing has the opportunity to improve product quality & reliability, reduce cost and improve product yield beyond today's capabilities. Almost all companies are starting to use some forms of Adaptive Testing, but there is not a sequential roadmap for implementation and many applications are created in an ad-hoc way.

Adaptive Test methods are evolving over time as new technologies (<10nm, SOI), design methodologies (multichip/stack packaging) and supply chain support models are introduced.

There are a number of challenges that are today limiting the industry's ability to fully exploit Adaptive Testing across the supply chain. These are highlighted in the table in a previous section.

Test and Yield Learning In-Depth Analysis

In the normal sorting function, test provides the essential feedback loop for yield-learning. Product-based diagnostics, product-like test chips and parametric-sensitive test structures all play a key role.

Key Cost of Test Trends

Value derived from diagnostics of actual product hardware is driven by systematic defect mechanisms that are now increasingly complex functions of neighboring shapes, local pattern densities, etc. As a result, some failure mechanisms may be visible only on product. In addition, product-based diagnostics automatically places focus on key yield-limiting failure mechanisms. Volume-based diagnostics are important since individual occurrence of any given systematic defect mechanism may be rare. The pooling of data across many failing die can be important to identify true systematic defect mechanisms.

Product-like test chips can provide some of the same insight for yield-learning, but have the advantage of being available earlier, even when design is on-going. Specifically, rapidly designable, scalable and 100% testable & diagnosable test chips, with and without embedded memory and other key IP blocks, including fast automated design methodologies, are required to accelerate yield ramps and first-time yield success of complex SOCs. Such test chips should play the role of "send-aheads" and be designed on early foundry testsites even while the product design is ongoing. The test chip should be scalable, in that a complete SOC-style (optionally, timing-closed) design is possible with tens or hundreds of standard cells and with a small or large compiled memory and other IP blocks. The test chip should enable both logical and physical layout diversity in order to capture layout topologies found on real product chips. Finally, the test chip must be able to maintain a stable test and diagnosis infrastructure, meaning the same set of ATPG, diagnosis and failure analysis capabilities should be enabled whether the test chip is tiny (<1mm^2) or huge (>100mm^2).

In addition, parametric-related feedback is needed for (1) device and interconnect parameters and (2) design-process interactions. Measurement of device and interconnect parameters have traditionally relied upon test structures, especially scribe-line FETs and interconnect resistance and capacitance monitors. Increasing across-chip variation (intra-die variability) increases the negative impact of scribe-line-to-chip offsets. Moreover, test structures are limited by the number of configurations they can cover. Variations in configurations include both physical variations and electrical variations, such as different gate types and differences in load characteristics. As circuit parametrics are increasingly affected by such configurations, including within-standard-cell and transistor-layout configurations, it becomes necessary to base learning on product test or test of product-like layout configurations. Embedded, distributed monitor circuits such as thermal and VDD sensors, process-monitoring ring oscillators and critical path proxies are now standard on microprocessor-class ICs and can be used to help diagnose parametric fails and understand variability. Understanding variability includes unraveling the structure of variations into spatial and

cross-parameter components (variation in transistor length, Vt, source-drain resistance, etc.) The spatial component includes both die-to-die and within-die components.

Cross-parameter variations, potentially including a spatial component, are important to analog/RF circuits, as well as digital. Methods for understanding/characterizing the manufacturing process and operating environment that are sufficiently sensitive for analog/RF are needed. Moreover, product test is uniquely well-suited to provide feedback on design-process interactions, including those leading to noise-related fails, such as power-grid droop and crosstalk fails.

Top challenges for test-based yield-learning include:

- Better resolution for cell-internal defects. Latest advances in structural testing and scan-based logic/layout-aware diagnosis methods are adequately addressing interconnect and via defects. Statistical approaches built into volume-based diagnostics are able to predict interconnect-related defect modes without an over-dependence on Physical Failure Analysis (PFA). Innovation is required, however, for cell-internal-defect-targeted diagnostics to be able to identify systematic fail modes inside standard cells. Observations derived from production silicon suggest a shift toward a larger percentage of the defect distribution being cell-internal defects, as opposed to interconnect-related. Current best methods for cell-internal defect diagnostics are cell truth-table and gate-exhaustive model-based, with the truth tables established via SPICE simulations of modeled cell-internal parasitics. These methods suffer from aliasing issues and over-reliance on potentially inaccurate modeling of cell-internal defects used in SPICE simulations. In addition, diagnosis resolution needs to be better due to limitations in the PFA process.
- Managing design data for yield learning. A tremendous amount of design data can be brought to bear for yield learning purposes, but it is often not organized effectively for this purpose. In addition, with hierarchical design and DFT flows, the overall management of this data at most companies today is
- ad-hoc, limiting its effective use.
- Inadequacy of LEF/DEF as the basis of layout-aware diagnosis. LEF/DEF suffices for the purposes of layout-aware ATPG but is too early in the design cycle to be used effectively for layout-aware diagnosis. LEF/DEF is less likely to closely resemble the final mask shapes due to complex OPC, boolean and retargeting steps.
- Yield-Learning in a OSAT/Fabless/Foundry environment. Yield-learning capabilities must be cognizant of the environment that has become the dominant model for our industry. If the technology cannot deal with the security and logistical concerns of this environment, it cannot be effective. Factory integration issues must be addressed. Data capture and management capabilities must support increasing reliance on statistical analysis and data mining of volume production data for yield-learning. Secure mechanisms for yield-data flow for distributed design, manufacture and test, including fabless/foundry and 3rd party IP, are needed. Standard test data formats, such as STDF-V4-2007 for scan fail data, and infrastructure to support their transmittal are needed to support automation and sharing of data. Specifically, data exchange standards are needed between the Fabless and the OSAT/Foundries to share system-level test feedback and correlation to wafer-level test and measurement data to (1) improve IC quality and reliability, (2) correlate process variations and parametric variability, and (3) reduce overkill. Distributed design, manufacture and test also creates an emerging role for methodologies and tools to help determine which areas that problems reside, e.g., design house, foundry or OSAT.
- Test for ZERO DPPM/Automotive in advanced node technologies. A change of mindset away from structural test coverage only is required to guarantee functional safety for automotive ICs. Mission-mode in-situ MBIST and LBIST and Design Failure Mode and Effects Analysis [DFMEA] are already part of ISO26262 (an automotive-specific set of standards for designing and testing electronics that focuses on safety critical components), but test architecture research is required to minimize the die-footprint increase due to added circuit redundancy. Moreover, Automotive may require root cause reports to be produced quickly for field failures. This requirement is another driver for rapid diagnosis and root cause analysis.
- Test and data-collection time increases due to longer scan chains. These increases drive a need for focus on LBIST methodologies and scan compression for both test and diagnosis.
- Faster Memory BIST bitmapping.

- Guidance for trading off test and data collection time against improved failure diagnostics.

References and Industry Links

Below is a list of references that provide information about methods and applications of Adaptive Testing.

Part Average Testing (PAT) is described in the AEC-Q001 documents at http://www.aecouncil.com/AECDocuments.html.

K. Huang, N. Kupp, C. Xanthopoulos, J.M. Carulli, Y. Makris, “Low-Cost Analog/RF IC Testing Through Combined Intra- and Inter-Die Correlation Models,” Design & Test, Volume 32, Issue 1, pp. 53-60, 2015.

A. Ahmadi, K. Huang, A. Nahar, B. Orr, M. Pas, J.M. Carulli Jr., Y. Makris, “Yield Prognosis for Fab-to-Fab Product Migration,” IEEE VLSI Test Symposium, Session 3B, p. 3.3, 2015.

S. Zhang, X. Li, R.D. Blanton, J. Machado da Silva, J.M. Carulli, K.M. Butler, “Bayesian Model Fusion: Enabling Test Cost Reduction of Analog/RF Circuits via Wafer-level Spatial Variation Modeling,” International Test Conference Proceedings, p. 19.1, 2014.

U. Guin, K. Huang, D. DiMase, J.M. Carulli Jr., M. Tehranipoor, Y. Makris, "Counterfeit Integrated Circuits: A Rising Threat in the Global Semiconductor Supply Chain," Proceedings of the IEEE, Volume 102, Issue 8, 2014.

Xanthopoulos, K. Huang, A. Poonawala, A. Nahar, B. Orr, J.M. Carulli Jr., Y. Makris, “IC Laser Trimming Speed-Up through Wafer-Level Spatial Correlation Modeling,” International Test Conference Proceedings, p. 19.2, 2014.

Ahmadi, K. Huang, S. Natarajan, J.M. Carulli Jr., Y. Makris, “Spatio-Temporal Wafer-Level Correlation Modeling with Progressive Sampling: A Pathway to HVM Yield Estimation,” International Test Conference Proceedings, p. 18.1, 2014.

Tikkanen, J.; Siatkowski, S.; Sumikawa, N.; Wang, L.-C.; Abadir, M.S., "Yield optimization using advanced statistical correlation methods," Test Conference (ITC), 2014 IEEE International , vol., no., pp.1,10, 20-23 Oct. 2014

K. Huang, N. Kupp, J. Carulli and Y. Makris "Handling discontinuous effects in modeling spatial correlation of wafer-level analog/RF tests", Proc. Design Autom. Test Eur. Conf., pp.553 -558 2013

K. Huang, N. Kupp, J. Carulli and Y. Makris "On combining alternate test with spatial correlation modeling in analog/RF ICs", Proc. IEEE Eur. Test Symp., 2013

S.K. Devarakond, S. Bhattacharya, J. McCoy, A. Nahar, J.M. Carulli, A. Chatterjee, “Predicting Die-to-Die Process Variations from Wafer Test Data: A Feasibility Study,” IEEE Latin American Test Workshop, April 2013.

K. Huang, N. Kupp, J.M. Carulli, Y. Makris, “Process Monitoring through Wafer-level Spatial Variation Decomposition,” International Test Conference Proceedings, p. 5.3, 2013.

C-K Hsu, F. Lin, K-T Cheng, W. Zhang, X. Li, J.M. Carulli Jr., K.M.. Butler, “Test Data Analytics - Exploring Spatial and Test-Item Correlations in Production Test Data,” International Test Conference Proceedings, p. 5.2, 2013.

K. Huang, J.M. Carulli, Y. Makris, “Parametric Counterfeit IC Detection via Support Vector Machines,” International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems, p. 1.1, 2012.

N. Kupp, K. Huang, J. Carulli and Y. Makris "Spatial correlation modeling for probe test cost reduction", Proc. IEEE/ACM Int. Conf. Comput.-Aided Design, pp.23 -29 2012

N. Kupp, K. Huang, J.M. Carulli, Y. Makris, “Spatial Estimation of Wafer Measurement Parameters Using Gaussian Process Models,” International Test Conference Proceedings, p. 5.1, 2012.

K. R. Gotkhindikar, W.R. Daasch, K.M. Butler, J. M. Carulli Jr., A. Nahar, “Die-level Adaptive Test: Real-Time Test Reordering and Elimination,” International Test Conference Proceedings, p. 15.1, 2011.

Sumikawa, N.; Drmanac, D.G.; Wang, Li.-C.; Winemberg, L.; Abadir, M.S., "Forward prediction based on wafer sort data — A case study," Test Conference (ITC), 2011 IEEE International , vol., no., pp.1,10, 20-22 Sept. 2011

Sumikawa, N.; Drmanac, D.; Li-C Wang; Winemberg, L.; Abadir, M.S., "Important test selection for screening potential customer returns," VLSI Design, Automation and Test (VLSI-DAT), 2011 International Symposium on , vol., no., pp.1,4, 25-28 April 2011

Drmanac, D.; Sumikawa, N.; Winemberg, L.; Wang, L.-C.; Abadir, M.S., "Multidimensional parametric test set optimization of wafer probe data for predicting in field failures and setting tighter test limits," Design, Automation & Test in Europe Conference & Exhibition (DATE), 2011 , vol., no., pp.1,6, 14-18 March 2011

Sumikawa, N.; Drmanac, D.; Wang, Li.-C.; Winemberg, L.; Abadir, M.S., "Understanding customer returns from a test perspective," VLSI Test Symposium (VTS), 2011 IEEE 29th , vol., no., pp.2,7, 1-5 May 2011

W. Zhang "Virtual probe: A statistical framework for low-cost silicon characterization of nanoscale integrated circuits", IEEE Trans. Comput.-Aided Design Integr. Circuits Syst., vol. 30, no. 12, pp.1814 -1827 2011

Drmanac, D.; Laisne, M.; Wang, L.C., "Wafer probe test cost reduction of an RF/A device by automatic testset minimization — A case study," Test Conference (ITC), 2011 IEEE International , vol., no., pp.1,10, 20-22 Sept. 2011

S. Reda and S. R. Nassif "Accurate spatial estimation and decomposition techniques for variability characterization", IEEE Trans. Semicond. Manuf., vol. 23, no. 3, pp.345 -357 2010

E. Marinissen, A. Singh, D. Glotter, M. Esposito, J. Carulli, A. Nahar, K. Butler, D. Appello, and C. Portelli, "Adapting to adaptive testing," Design, Automation and Test in Europe Conference & Exhibition (DATE), pp. 556-561, 2010.

Chen, J.; Jing Zeng; Wang, L.-C.; Mateja, M., "Correlating system test fmax with structural test fmax and process monitoring measurements," Design Automation Conference (ASP-DAC), 2010 15th Asia and South Pacific , vol., no., pp.419,424, 18-21 Jan. 2010

Chen, J.; Jing Zeng; Wang, Li.-C.; Mateja, M.; Rearick, J., "Predicting multi-core system Fmax by data-learning methodology," VLSI Design Automation and Test (VLSI-DAT), 2010 International Symposium on , vol., no., pp.220,223, 26-29 April 2010

D. Mannath "Structural approach for built-in tests in RF devices", Proc. Int. Test Conf., 2010

Chen, J.; Wang, Li.-C.; Po-Hsien Chang; Jing Zeng; Yu, S.; Mateja, M., "Data learning techniques and methodology for Fmax prediction," Test Conference, 2009. ITC 2009. International , vol., no., pp.1,10, 1-6 Nov. 2009

Drmanac, D.; Bolin, B.; Wang, L.-C.; Abadir, M.S., "Minimizing outlier delay test cost in the presence of systematic variability," Test Conference, 2009. ITC 2009. International , vol., no., pp.1,10, 1-6 Nov. 2009

K.M. Butler, J.M. Carulli, J. Saxena, A. Nahar, R. Daasch, "Multidimensional Test Escape Rate Modeling," Design & Test of Computers, Volume 26, Issue 5, pp. 74-82, 2009.

A. Nahar, K. Butler, J. Carulli, and C. Weinberger, "Quality improvement and cost reduction using statistical outlier methods," IEEE International Conference on Computer Design, pp. 64-69, 2009.

A. Nahar, K.M. Butler, J.M. Carulli, C. Weinberger, "Statistical Outlier Method Applications", IEEE International Workshop on Defect and Data-Driven Testing, 2009.

Wu, S.H.; Drmanac, D.; Wang, Li.-C., "A Study of Outlier Analysis Techniques for Delay Testing," Test Conference, 2008. ITC 2008. IEEE International , vol., no., pp.1,10, 28-30 Oct. 2008

R. Aitken, "Defect or Variation? Characterizing Standard Cell Behavior at 90 nm and Below," IEEE Transactions on Semiconductor Manufacturing, vol. 21, pp. 46-54, 2008.

H.-G. Stratigopoulos and Y. Makris "Error moderation in low-cost machine-learning-based analog/RF testing", IEEE Trans. Comput.-Aided Design Integr. Circuits Syst., vol. 27, no. 2, pp.339 -351 2008

F. Ferhani, N. Saxena, E. McCluskey, and P. Nigh, "How Many Test Patterns are Useless?" VLSI Test Symposium, pp. 23-28, 2008.

K.M. Butler, J.M. Carulli, J. Saxena, "Modeling Test Escape Rate as a Function of Multiple Coverages," International Test Conference Proceedings, p. 21.2, 2008.

F. Liu "A general framework for spatial correlation modeling in VLSI design", Proc. Design Autom. Conf., pp.817 -822 2007

L. Fang, M. Lemnawar, and Y. Xing. Cost Effective Outliers Screening with Moving Limits and Correlation Testing for Analogue ICs. In Proceedings IEEE International Test Conference (ITC), October 2006. Paper 31.2.

T.J. Anderson, J.M. Carulli, "Modeling & Monitoring of Product DPPM with Multiple Fail Modes," International Reliability Physics Symposium, p. 545, 2006.

K. M. Butler, S. Subramaniam, A. Nahar, J. M. Carulli Jr., T. J. Anderson, W. R. Daasch, "Successful Development and Implementation of Statistical Outlier Techniques on 90nm and 65nm Process Driver Devices," Proc. 2006 Int. Reliability Physics Symp., pp. 552-559, Oct. 2006.

J.M. Carulli, T.J. Anderson, "The impact of multiple failure modes on estimating product field reliability," Design & Test of Computers, Volume 23, Issue 2, p. 118, 2006.

A. Nahar, R. Daasch, and S. Subramaniam, "Burn-in Reduction Using Principle Component Analysis," Proc. 2005 IEEE Int. Test Conf., Paper 7.2, Nov. 2005.

L. Ning, A. Nahar, R. Daasch, K. M. Butler, J. M. Carulli Jr., S. Subramaniam, "Burn-in Reduction using Robust Canonical Correlation Analysis," Proc. SRC TECHCON 2005, Paper 9.1, Oct. 2005.

Z. Yuan and Y. Yang "Combining linear regression models: When and how?", J. Amer. Stat. Assoc., vol. 100, no. 472, pp.1202 -1214 2005

R. Turakhia, B. Benware, R. Madge, T. Shannon, R. Daasch, "Defect Screening Using Independent Component Analysis on IDDQ," Proc 2005 IEEE VLSI Test Symp., pp. 427-432, May 2005.

W. C. Riordan, R. Miller, E. R. St. Pierre, "Reliability Improvement and Burn In Optimization Through the use of Die Level Predictive Modeling," Proc. 2005 IEEE Int. Rel Physics Symp., pp. 435-445, Apr. 2005.

R. Madge, B. Benware, R. Turakhia, R. Daasch, C. Schuermyer, and J. Ruffler, "In search of the optimum test set - adaptive test methods for maximum defect coverage and lowest test cost," Proceedings International Test Conference, pp. 203-212, 2004.

S.S. Sabade, D. M. H. Walker, "CROWNE: Current Ratio Outliers with Neighbor Estimate", DFT 2003, pp 132-139.

C. Schuermyer, B. Benware, K. Cota, R. Madge, R. Daasch, L. Ning, "Screening VDSM Outliers Using Nominal and Subthreshold Supply Voltage IDDQ," Proceedings IEEE Int. Test Conference, 2003, pp. 565-573, Sep 2003.

Automotive Electronics Council, "Guidelines for Part Average Testing (PAT)," AEC Q001, rev. C, 18 July 2003, http://www.aecauncil. comAECDocuments.html.

P. Buxton, P. Tabor "Outlier Detection for DPPM Reduction," Proceedings IEEE Int. Test Conference, 2003, pp. 818-827, Sep-Oct 2003.

T.S. Barnett and A.D. Singh. Relating Yield Models to Burn-In Fall-Out in Time. In Proceedings IEEE International Test Conference (ITC), pages 77-84, October 2003.
S.S. Sabade, D. M. Walker, "Evaluation of Effectiveness of Median of Absolute Deviations Outlier Rejection-based IDDQ Testing Burn-in Reduction" Proceedings 2002 VLSI Test Symp., pp. 81-86, Apr-May 2002.
D. I. Bergman, H. Engler, "Improved IDDQ Testing with Empirical Linear Prediction", Proc. 2002 IEEE Int. Test Conf, page 954, Nov. 2002.
W.R. Daasch, J. McNames, R. Madge, and K. Cota. Neighborhood Selection for IDDQ Outlier Screening at Wafer Sort. IEEE Design & Test of Computers, 19(5):74-81, September-October 2002.
W. R. Daasch, J. McNames, R. Madge and K. Cota "outlier screening at wafer sort", IEEE Design Test Comput., vol. 19, no. 5, pp.74 -81 2002
P. N. Variyam, S. Cherubal and A. Chatterjee "Prediction of analog performance parameters using fast transient testing", IEEE Trans. Comput.-Aided Design Integr. Circuits Syst., vol. 21, no. 3, pp.349 -361 2002
T.S. Barnett, A.D. Singh, M. Grady, and K. Purdy. Redundancy Implications for Early-Life Reliability: Experimental Verification of an Integrated Yield-Reliability Model. In Proceedings IEEE International Test Conference (ITC), pages 693-699, October 2002.
J. Saxena, K. M. Butler, J. Gatt, Raghuraman R., S. P. Kumar, S. Basu, D. J. Campbell, J. Berech, "Scan-Based Transition Fault Testing: Implementation Challenges," Proc. 2002 IEEE Int. Test Conf., pp. 1120-1129, Oct. 2002.
R. Madge et al. Screening MinVDD Outliers Using Feed-Forward Voltage Testing. In Proceedings IEEE International Test Conference (ITC), pages 673-682, October 2002.
R. Madge, M. Rehani, K. Cota, and W.R. Daasch. Statistical Post-Processing at Wafer Sort - An Alternative to Burn-In and a Manufacturable Solution to Test Limit Setting for Sub-Micron Technologies. In Proceedings IEEE VLSI Test Symposium (VTS), pages 69-74, May 2002.
T.S. Barnett, A.D. Singh, M. Grady, and K. Purdy. Yield-Reliability Modeling: Experimental Verification and Application to Burn-In Reduction. In Proceedings IEEE VLSI Test Symposium (VTS), pages 75-80, May 2002.
S.S. Sabade, D.M.H. Walker, "Improved wafer-level spatial analysis for IDDQ limit setting, Proc. 2001 IEEE Int. Test Conf., pp 82-91, Nov. 2001.
R. Daasch, K. Cota, J. McNames, R. Madge, "Neighbor Selection for Variance Reduction in IDDQ and Other Parametric Data" Proceedings IEEE Int. Test Conference, 2001, pp. 92-100, Oct 2001.
S. Benner and O. Boroffice, "Optimal production test times through adaptive test programming", Proceedings International Test Conference, pp. 908-915, 2001.
R. B. Miller, W. C. Riordan, "Unit Level Predicted Yield: A Method of Identifying High Defect Density Die at Wafer Sort," Proc 2001 IEEE Int. Test Conf., pp. 1118-1127, Oct.-Nov. 2001.
K. Butler and J. Saxena, "An empirical study on the effects of test type ordering on overall test efficiency," Proceedings International Test Conference, pp. 408-416, 2000.
E. Riess, "Automatic Monitoring of Electrical Parameters in the Semiconductor Industry Based on ROC," IEEE Trans. On Sys., Man, and Cybernetics-Part A: Sys. And Humans, vol. 30, No. 6, pp. 853-857, Nov. 2000.
C. Thibeault, "Diagnosis Method Using AIDDQ Probabilistic Signatures: Theory and Results", Journal of Electronic Testing: Theory and Applications 2000, Vol 4., pp 339-353.
A. Kinra, H. Balachandran, R. Thomas, J.M. Carulli, “Logic mapping on a microprocessor,” International Test Conference Proceedings, p. 701, 2000.
W. R. Daasch, J. McNames, D. Bockelman, K.Cota, R. Madge, "Variance Reduction Using Wafer Patterns in IDDQ Data" Proceedings IEEE Int. Test Conference, 2000, pp. 189-198, Oct 2000.
J.M. Carulli Jr., D.C. Wrobbel, A. Mehta, K.E. Krause Jr., B.E. Campbell, F.A. Valente, “Frequency Distribution Modeling for High-Speed Microprocessors Using On-chip Ring-oscillators,” SPIE Proceedings, Inline Methods & Monitors for Process & Yield Improvement, p. 146, 1999.
A. E. Gattiker, W. Maly, "Current Signatures: Application," Proc. 1998 IEEE Int. Test Conf, pp. 1168-1177, Oct. 1998.
B. Stine, D. Boning, and J. Chung, "Analysis and decomposition of spatial variation in integrated circuit processes and devices," IEEE Transactions on Semiconductor Manufacturing, vol. 10, pp. 24-41, 1997.
A. Singh, P. Nigh, and C. M. Krishna, "Screening for known good die (KGD) based on defect clustering: An experimental study," in Proceedings International Test Conference, 1997, pp. 362-371.
A.D. Singh, P. Nigh, and C.M. Krishna. Screening for Known Good Die based on Defect Clustering: An Experimental Study. In Proceedings IEEE International Test Conference (ITC), pages 324-331, November 1997.
T. W. Williams, R. H. Dennard, R. Kapur, M. R. Mercer, W. Maly, "IDDQ Test: Sensitivity Analysis of Scaling," Proc. 1996 IEEE Int. Test Conf, pp. 786-792, Oct. 1996.
A. Singh and C. Krishna, "On the effect of defect clustering on test transparency and IC test optimization," IEEE Transactions on Computers, vol. 45, no. 6, pp. 753-757, Jun. 1996.
A.D. Singh and C.M. Krishna. On Optimizing VLSI Testing for Product Quality Using Die Yield Prediction. IEEE Transactions on Computer-Aided Design, 12(5):693-709, May 1993.

J. M. Soden, C. F. Hawkins, R. K. Gulati, and W. Mao, "IDDQ Testing: A Review," J. Electronic Testing:Theory and Applic. (JETTA), vol.3, no.4, pp. 291-303, Dec. 1992.
C. Crapuchettes, "Testing CMOS IDD on Large Devices," Proc. 1987 IEEE Int. Test Conf., pp. 310-315, Sep. 1987.

Industry Links
CAST www.semi.org/en/collaborative-alliance-semiconductor-test
DR Yield dryield.com
Optimal+ www.optimalplus.com
PDF Solutions www.pdf.com
Mentor Graphics (previously Galaxy) www.mentor.com/products/silicon-yield/quantix/
Qualtera www.Qualtera.com

Section 10: 2.5D & 3D Device Testing

This section will address six key test challenges, based on the evolution of 2.5D/3D, from complex die stacks through SiP. These test challenges include: test flows, cost and resources; test access; testing heterogeneous die individually and in a single stack/package; debug and diagnosis of failing stacks/die; DfX (Design for Test, Yield, Cost); and power. It is important to note that 2.5D/3D is not yet a mature and mainstream technology, and, because of that, it is difficult at this time to make any predictions regarding 2.5D/3D test flows. 2.5D and 3D technologies are characteristic of a system, and because of that, should be tested like a system: testing the complete package at an application level and diagnosing failures at the die and interconnect level.

Memory die stacks (Wide I/O, High Bandwidth Memory and Hybrid Memory Cube) were precursors to 2.5D and 3D. Both of these technologies have provided insights to requirements and challenges associated with 3D and 2.5D test. The best that can be gleaned from these technologies at this time is that reliance on BIST and boundary-scan based technologies, and use of fault tolerance with simple configurations, tends to produce relatively high yields at the stack level. As these adjacent technologies become more mature and as more 2.5D/3D-TSV applications emerge, more and better data will enable better predictions and decision making, with respect to 2.5D/3D-TSV test processes.

Executive Summary

2.5D and 3D/TSV are the next evolution beyond SiP/SoC. There have been significant advances from both academics (research) and industry (standards and working models/test chips) to identify and resolve challenges to testing 3D/TSV devices. In the medium to long term, as TSV-based die stacking becomes more prevalent and more complex/exotic die stacks appear, test challenges will also become more difficult. It is certain that new and additional Design-For-Test features will be needed to mitigate increased tester resource and time requirements, as well as increased test complexity, due to large numbers of different die in the same package. This section will discuss challenges for testing, including cost (dollars, resources and yields); design for test; test access; debug and diagnosis; and heterogeneous device testing.

Difficult Challenges for Test: 5 years

While the current state of 2.5D/3D seems to be maturing, new enabling and supporting technologies will require advances in test access, capabilities and costs. These emerging technologies will provide significant challenges for testing 2.5D and 3D technologies, in addition to supportive technologies. The challenges below represent potential impacts to test, including increased costs, test times, and reduced yields and reliability.

2.5D Test Challenges (Short Term)

- **Known Good Die Test**: While logic blocks in the die can be partitioned and effectively tested, testing interactions between the logic blocks requires an application-based test.
- **System level test/diagnosis/repair**: Emulating a system-level test environment can be considerably costly and time consuming. Complex integration can produce new and challenging defects for test.
- **Access to the individual die in the 2.5D assembly**: Access may include multiple protocols: IEEE-1149 based (including IEEE 1687.1 and IEEE P1838 proposed standards), photonics, various component and system-level protocols, including I2C and SPI.
- **Interposer testing**: Point to Point testing can be accomplished primarily by probing (see the probing section). Multipoint interposer testing requires significantly more probing (more time and higher costs) and requires embedded logic to coordinate point-to-multipoint connections. Academic research from Duke University[1] has proposed solutions to point-to-multipoint connections on the interposer.
- **High speed interconnects (photonics)/signal integrity**: while testing photonics interconnects is mature, the cost of test is significant, primarily due to equipment and knowledge.
- **Discrete components**: more specific to high speed interconnects through the interposer. Point to point, high speed interconnects may compromise probe-based connectivity test.
- **Impact of emerging technologies with respect to test**: New technologies can impose new defects on the 2.5D. A table presented by Li Li at ECTC describes the relationship to the new technologies (Die thinning, TSV, C2C connection, BS-RDL, micro-bumps, Large volume of copper, increased power density and large thermal gradients, removal of IO structure) and the failure/defect mechanisms. See Table 1 below.

- **Innovations in Wafer Level Packaging/Wafer Level Fan Out**: similar to the statements above: advances in Wafer Level Packaging open the door for new defects and new requirements for test. Aspects of wafer level packaging impinge on interposer technologies.

3D Test Challenges (Short Term)

- **Known Good Die test**: Known good die testing becomes more important in a stack integration, where a defective logic die could compromise the entire die stack. System/application level testing at the die level has significant challenges. Significant Design-for-Test and Built-in Self-Test features are required to complement and simplify system/application level test. Cost of test ($$ and time) are significant.
- **System level test/diagnosis/repair**: System and application level test of a 3D die stack may require significant test/tester resources in order to emulate the component/test environment. These resources may consume significant time and money.
- **Diagnosis of system level failures** on the stack will have significant challenges. On a positive note, a diagnostic die may be added to the stack without significant effort. Redundant die may help as well.
- **TSV/interconnect testing**: TSV testing can be done pre-bond, microvoid and pinhole defects. Mid-bond testing becomes quite challenging. Some probing could help. A boundary-scan stack architecture may be a possibility. An IEEE standard is in the works (IEEE P1838 – this will be covered later as part of "technical issues"). Application and/or system test most likely will not be feasible or practical with part of the stack still missing. Test results may not be informative.
- **Probing 3D die stacks**: This area seems to be maturing. Companies like Tezzaron have probing figured out, since microbumps also seem to be maturing.
- **Stack repair**: While stack repair is now possible, cost for stack repair and re-test would be prohibitive. Redundant die may simplify stack repair.
- **High speed signaling/signal integrity/interconnects**: Signal integrity measurements from die to die are challenging. Design for Signal Integrity Test, at the edge of the die, may be a solution for SI measurement. Costs could be significant at the die level and higher at the stack level. There may be impact to stack level power.
- **KGD/KGS**: Known Good Die is critical from a supply chain perspective. Known Good Stack also requires a stack level system/application test.
- **Wafer Level Packaging** – uncertain if there will be an impact to test.

Difficult Challenges for Test: 10, 15, 25 years

Over the next 10 to 25 years, it is expected that requirements for speed and power will be significantly higher than the current state. Significant advancement to the integration of massive, high performance, low power die will be required for the future versions of application specific integrated devices. Test will be significantly challenged.

2.5D Test Challenges (Long term)

- **Known Good Die Test**: From ITRS predictions, the number of flops per KGD are significantly higher (exponential) over time. Speed, power and thermal scale with logic. Device cooling during test will become critical. Timing may be an issue through the stack. Test time may be significant.
- **System level test/diagnosis/repair**: System test requires higher speeds. Longer test times due to significantly more logic in the die. IEEE 1149 protocols will be obsolete or antiquated at best. Optical protocols using IEEE P1838 test access protocols will replace current test protocols.
- **Access to the 2.5D assembly**: Photonic protocols, more in line with P1838. Protocol objective is access to interposer of multiple-die stacks. Multiple-die stacks create a significant, interconnected logic pool.
- **Interposer testing**: Testing becomes logic based. Probing becomes more challenging, from Point-to-Point to Large Multipoint.
- **Cost of test** (equipment/resources/time): Cost of test increases as high-speed interconnects pass through the optical interposer. Signal integrity adds cost and time.
- **Testing MEMS- and Sensor-based die**: technology is somewhat analog. Testing is not conventional and requires an active motion, light, sound source. Significant setup and test times while income/prices are low and volumes are extremely high.

3D Test Challenges (Long term)

- **Known Good Die(s):** similar to the 2.5D section, logic per die is increasing over time. Speed, power and thermal are scaling with increasing logic/flops. Heterogeneous die become more and more exotic, making test more challenging, and determining when the die becomes "known good".
- **System level test/diagnosis/repair**: Test equipment cost increases significantly with respect to dollars and time. High speed signaling increases test costs and the possibility for errors/failures. Potential impact on the supply chain.
- **Carbon Nanotubes** replace TSV, CNT test replaces TSV Test: Carbon nanotubes help to sustain logic, speed, power, thermal characteristics. CNT is still academic; however, more prognostication exists for short- to mid-term replacement of TSVs with Carbon Nanotubes. External interfaces convert protocols to CNT interface.
- **Mega Stack testing**: Mega Stacks address the need for high performance data processing. Stacking multiple die becomes feasible as advances in die to die bonding minimize the stack size, while adding more die to the stack. Future Mega stacks may potentially have the same structural integrity as initial die stacks. Multiple redundant die may be necessary for more probable defective die in the stack, or mis-connections between die in the stack.
- **Probing die and stacks**: Advanced Probing technology will be needed as TSVs and micro-bumps become significantly smaller.
- **Stack repair**: Primary repair will come from redundant die. Dis-assembly of the die stack becomes significantly challenging and could potentially destroy the pre-bonded die.
- **Testing MEMS- and Sensor-based die**: technology is somewhat analog. Testing is not conventional and requires an active motion, light, sound source. Significant setup and test times while income/prices are low and volumes are extremely high.

Discussion of Key 3D Test Technical Issues

Future of 3D Integration (Test and DfT focus, Academic Perspective)

The semiconductor industry has been able to meet the demand for high-performance integrated circuits (ICs) with added functionality by relentlessly scaling device sizes. However, it is becoming increasingly difficult to sustain device scaling in an economically viable manner.

A promising way to achieve high-performance ICs with more functionality and reduced die footprint is through 3D integration. Today's 3D integration process is primarily based on die/wafer stacking, as it does not require substantial changes to the existing fabrication flow. In this process, separately manufactured dies/wafers are integrated onto the same package, and through-silicon-vias (TSVs) are used to connect dies to each other. Considerable research efforts have therefore been directed toward the development of TSV-based 3D stacking technology, and products based on this technology have been successfully introduced into the market, e.g., the AMD Fiji chip. However, the keep-out-zone (KOZ) required for TSVs and limitations on the die alignment precision impose limits on the device integration density that can be achieved using TSV-based 3D stacking. A minimum KOZ of 3 μm is required for ICs fabricated at the 20 nm technology node [Kannan et al. 2015], and the die alignment precision is currently limited to 0.5μm.

Monolithic three-dimensional (M3D) integration is receiving considerable interest as a technology for the future, as it has the potential to achieve higher device density compared to TSV-based 3D stacking. In this technology, transistors are processed layer by layer on the same wafer. Sequential integration of transistor layers enables high-density vertical interconnects, known as the interlayer vias (ILVs). Typically, the size and pitch of an ILV is one to two orders of magnitude smaller than those of a TSV [Batude et al. 2012]. To realize such high-density vertical interconnects, the interlayer dielectric (ILD) thickness is being aggressively scaled [Batude et al. 2012; Lee and Lim 2013], and such scaling has been shown to lead to electrostatic coupling between device layers. This is a challenge for test researchers.

Researchers have recently analyzed electrostatic coupling between device layers in M3D ICs and quantified its impact on circuit timing [Koneru 2017]. Device simulations have been carried out to understand the impact of coupling on the threshold voltage of a top-layer transistor for both transistor- and gate-level integration. To realize a new silicon layer over the bottom layer without damaging the underlying interconnects and degrading the properties

of the bottom-layer transistors, several layer-transfer techniques are being explored [Batude et al. 2015; Ishihara et al. 2012].

Low-temperature wafer bonding is a key processing step in these techniques. The condition of the bonding surfaces plays a crucial role in achieving a defect-free bond. Oxide layers are the prime candidates for bonding surfaces due to the presence of hydroxyl (OH) groups that lead to high bond strengths. These oxide layers also act as the ILD. Therefore, defects that arise during the wafer-bonding step can impact the top-layer transistors, as well as the ILVs. It is important to understand and analyze wafer-bonding defects, and develop methods to test for these defects. Research is needed to study the impact of bond defects on the threshold voltage of a top-layer transistor and on the ILVs. In addition, advances in test access and debug/diagnosis for M3D will also be of growing interest as this technology advances.

It has been shown thus far that the impact of coupling and wafer-bonding defects on the threshold voltage of a top-layer transistor is significant, and cannot be ignored, when thickness of the ILD is less than 100 nm. In such scenarios, the paths through the top layer in a gate-level-integrated M3D IC can change depending on the size of the defect and the voltage on the metal lines in the bottom layer. The presence of defects at the bond interface can lead to a change in resistance of an ILV and in some cases lead to an open in the ILV or a short between two ILVs. A resistive open in an ILV or a resistive short between two ILVs can have a significant impact on the path delays. Due to these challenges, existing test-generation methods for small-delay defects are of limited effectiveness when the ILD is less than 100nm.

There is also a paper which includes a discussion regarding supply-chain capability requirements for test and reliability: see [Alfano].

References

[Kannan 2015] S. Kannan, R. Agarwal, A. Bousquet, G. Aluri, and H. S. Chang. 2015. Device performance analysis on 20nm technology thin wafers in a 3D package. In Proceedings of the 2015 IEEE International Reliability Physics Symposium. 4C.4.1–4C.4.5. DOI:http://dx.doi.org/10.1109/IRPS.2015.7112735

[Batude 2012] P. Batude, T. Ernst, J. Arcamone, G. Arndt, P. Coudrain, and P. E. Gaillardon. 2012. 3-D sequential integration: A key enabling technology for heterogeneous co-integration of new function with CMOS. IEEE Journal on Emerging and Selected Topics in Circuits and Systems 2, 4, 714–722. DOI:http://dx.doi.org/ 10.1109/JETCAS.2012.2223593

[Lee and Lim 2013] Y.-J. Lee and S. K. Lim. 2013. Ultrahigh density logic designs using monolithic 3-D integration. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems 32, 12, 1892–1905.

[Koneru 2017] A. Koneru, S. Kannan and K. Chakrabarty, "Impact of electrostatic coupling and wafer-bonding defects on delay testing of monolithic 3D integrated circuits", ACM Journal on Emerging Technologies in Computing Systems, vol. 13, no. 4, Article 54, July 2017.

Alfano, Black, Rearick, Siegel, Su, and Din “Unleashing Fury: A New Paradigm in 3D Design and Testing”

Section 11: Key Drivers and Test Costs

Minimizing costs are a key part of the semiconductor manufacturing process. Test is no exception, although steady improvements in efficiencies over the last 15 years have lowered the typical cost of test as a percentage of IC revenue to less than 2-3%. The primary drivers of increased efficiency have been reductions in capital costs per resource and test times, coupled with increases in parallelism and Built-In Self-Test (BIST) capability. Most SOC device are tested 2 to 16 at a time, and memory devices can have more than 1,000 devices tested at once. Measured as the cost to use capital equipment for test (in terms of cost per hour per device), these decreases in test cost will continue at a relatively consistent rate per year. The figure below shows the historical rate of capital investment in test, interface (consumables) and handling equipment.

It is notable that, in 2015 for the first time, the cost of consumable material has become the leading capital expenditure relative to ATE-based test. This has to do with the increased cost of interface material (primarily influenced by probe cards and relative items) and the decreasing depreciation period for materials utilized for the production of devices used in the mobile device space where devices have a shorter life span. In this case, material is typically discarded not because it has ceased to function, but rather because the devices it is used to test are replaced by newer versions for end devices like mobile phones.

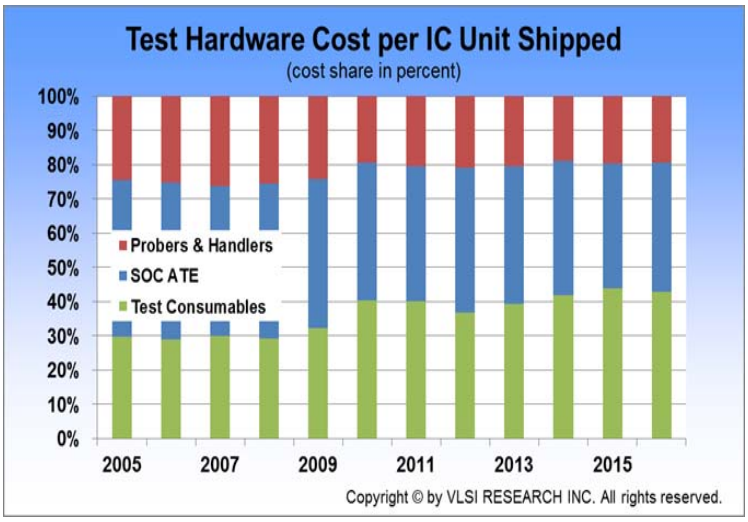


Figure 1: Test Costs as a percentage of device ASP (Used by permission of VLSI Research)

Key Cost of Test Trends

Looking forward, there are several trends which will counterbalance equipment efficiency and serve to cause cost increases:

- Increases in transistor count that outstrip compression technology will increase the amount of external data which must be supplied to the Device Under Test (DUT). Coupled with scan shift rates that are limited by power and thermal concerns, the overall effect will be longer test times. This will be addressed primarily with increased parallelism.
- Device configuration and one-time programming during test is causing more time to be spent during test to perform initial device calibrations or to reconfigure devices based on defects or electrical performance. As silicon geometries shrink and defect densities drive circuit redundancy, repair functions will also add to test costs.
- The eventual drive to multi-die packages will add a requirement for more System Level ("mission mode") testing owing to lack of access to individual die. Without significant Design For Test (DFT) improvements, this type of testing can take much longer than conventional structural test. This will also drive more exhaustive test processes at wafer probe in order to improve the yield of multi-die packages.
- Site count increases at probe test are not able to increase owing to the attendant increase in the cost of consumable material (discussed above) and the limitations of Touch-Down Efficiency (TDE). This is discussed in more detail below.

- An increasing reluctance on the part of IC manufacturers to dedicate silicon area and power to circuitry used exclusively for test.
- The continuing increase of silicon content in automotive applications, especially for safety systems, which drives additional test insertions for fault coverage and temperature-related test.

Even though continuous improvement in equipment efficiency will be offset by new device test requirements, the overall cost of test will continue to decrease. The major contributors to that cost are described below.

Cost of Test as a Part of Overall Manufacturing Cost

While the cost to own and operate test equipment has been reducing, other semiconductor manufacturing costs have been significantly increasing with new silicon technology. Specifically, fab costs for leading-edge processes have increased to about 70-80% of the overall cost of producing a large-scale SOC device. It now costs far more to fab a device than to test it, and that trend will accelerate as new fabrication technologies are deployed.

The figure below represents third-party analysis of the capital and service costs of equipment used in device fabrication, packaging and test.

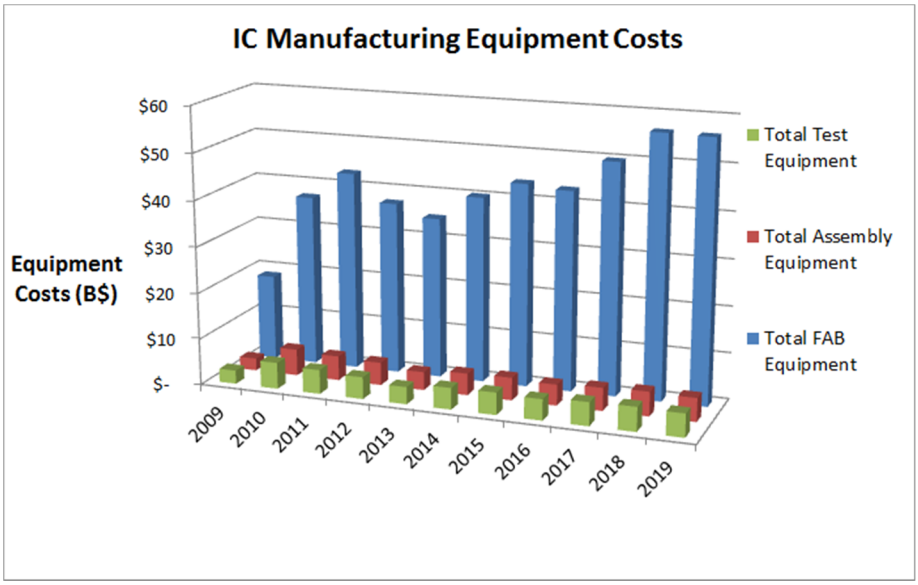


Figure 2: Relative cost of Fab, Packaging and Test Equipment

While it is helpful to focus on the cost of test itself, the overall contribution to a manufacturer's profitability from lower test costs will be very small since test is a small part of the device cost overall. The highest avoidable costs in test are devices that are good but are rejected at test for some reason.

Consider the following, simplified example.

- A device costs $1.00 to manufacture, including Fabrication, packaging, etc.
- Test constitutes 5% of that cost, or $0.05

Reducing the cost of test by 10%, will reduce overall costs by $0.05 X 10% = $0.005 per device

Improving yield by 1% reduces overall cost by $1.00 * 1% = $0.01 per device

While the 10% Cost of Test reduction is good, the yield improvement is better.

The figure below shows the effect on cost of test of traditional cost reduction techniques:

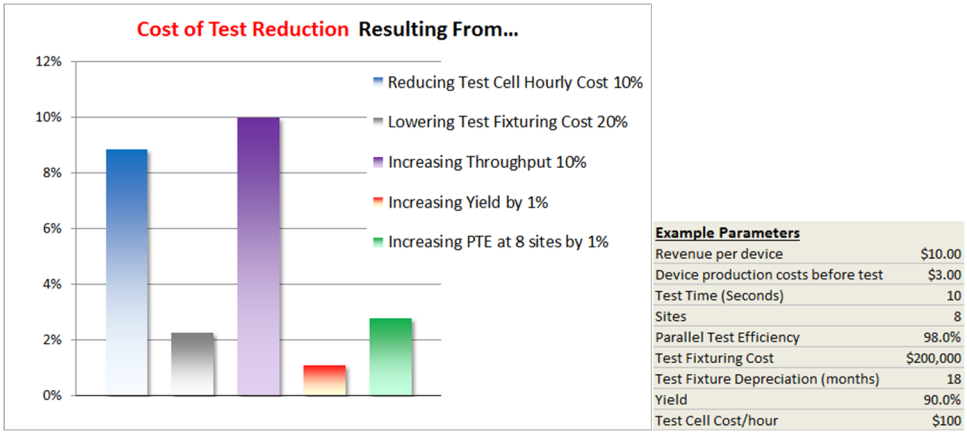


Figure 3: Cost of Test Reduction realized by traditional cost reduction techniques

If one considers the effect on total manufacturing costs, including the cost to scrap devices that are actually good, the cost savings due to improved yield becomes far more significant.

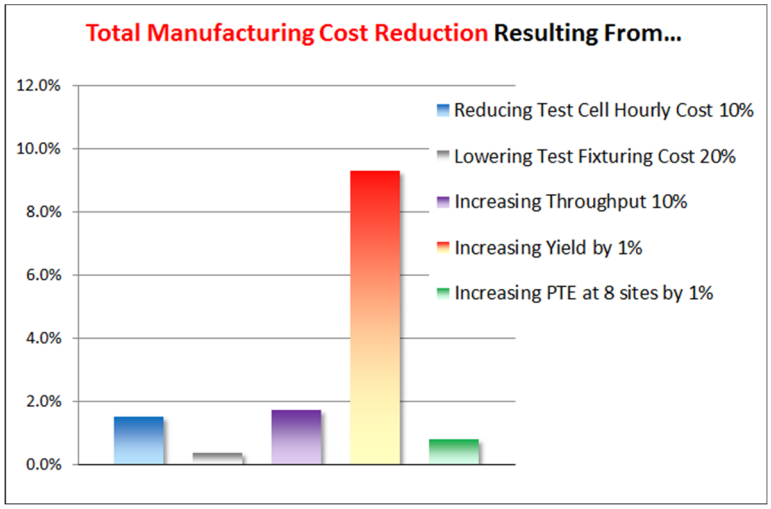


Figure 4: Total Cost of Manufacturing Reduction realized by traditional cost reduction techniques

The risk of yield loss is increasing over time for several reasons:

- Trends such as the reduction of power supply voltages and more complex RF modulation standards will drive higher accuracy requirements for test equipment. Test equipment accuracy is typically added as a "guardband" in testing, reducing the range of acceptable measurements. If measured DC and AC values become smaller and there is no improvement in test accuracy, this guardband will cause more marginal (but good) devices to be scrapped.
- As noted earlier, many devices, especially for mobile applications, require some sort of calibration or trim during the test process to improve DC and AC accuracy. This dramatically increases both the number of measurements made and the accuracy required of the test equipment. The requirements increase the chance of discarding devices that would otherwise have been good.
- Faster production ramps and short IC product life cycles will reduce the amount of time available to optimize measurements for the majority of devices produced.

The remainder of this section will examine Costs associated with owning and operating test equipment. It must be stressed that reducing these costs must be done in the context of the overall cost to produce devices and balance reduction in test costs with potential reductions in product yield.

Test Cost Models and Cost Improvement Techniques

The cost of semiconductor test has many drivers, which is further complicated for multi-die SiP precuts as shown in Figure 5.

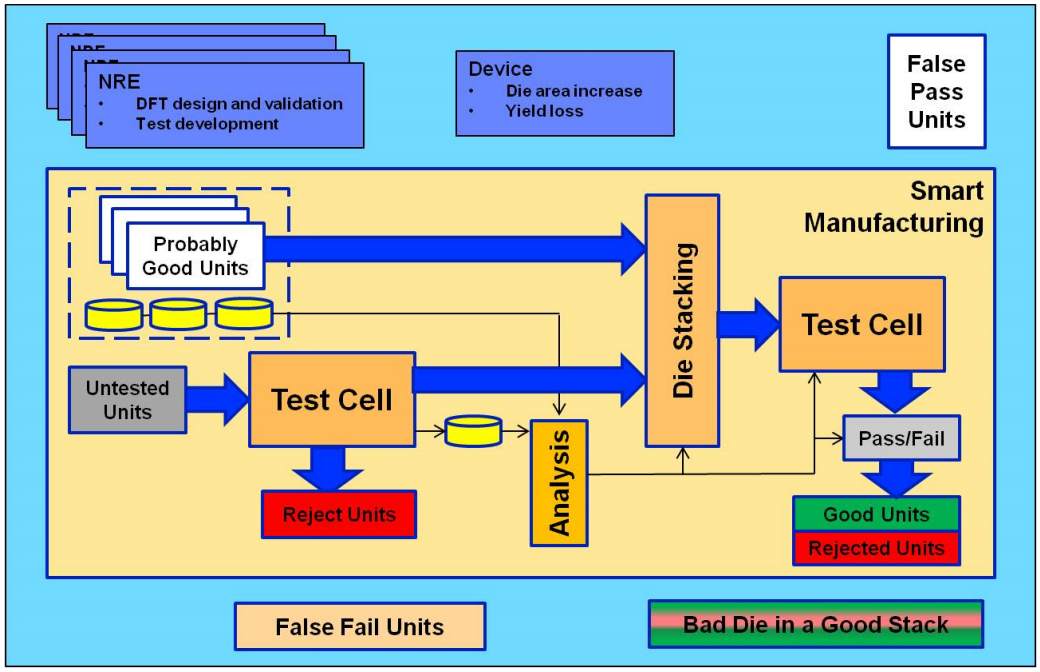


Figure 5: Multi-die Flow

Current Top Cost Drivers

The traditional drivers of Test Costs typically include (In rough order of impact to Cost)

- Device Yield
- Test Time, site count and Parallel Test Efficiency (PTE)
- Overall Equipment Utilization
- ATE Capital & Interface Expenditures
- Facility/Labor costs
- Cost of Test Program Development
- Cost of die space used for Test-only functions

Future Cost Drivers

- Increased test time due to larger scan patterns
- Increased testing at wafer to produce Known Good Die (KGD)
- Addition of system-level testing
- Increased cost of handling equipment to support high site count or singulated die
- Increasing use of device calibration/trimming at test or device repair with redundant components

Currently Deployed Cost Reduction Techniques

- Multi-site & reduced pin-count
- Structural Test and Scan
- Compression/BIST/DFT and BOST
- Yield Learning & Adaptive Test
- Concurrent Test
- Wafer-level at-speed testing

Cost Reduction Techniques that may be Deployed in the Future

- Advanced embedded instruments
- New contacting technologies
- In-system level testing to detect latent defects and potentially repair
- Built-in fault-tolerance

Multi-site Trend

As discussed in the previous sections, the most important way to reduce cost of test is increasing the number of sites. The effectiveness of increasing the number of sites is limited by (1) a high interface cost, (2) a high channel and/or power cost, and (3) a low multi-site efficiency M:

$$M = 1 - \frac{(T_N - T_1)}{(N-1)T_1}$$

where N is the number of devices tested in parallel (N>1), T_1 is the test-time for testing one device, and T_N is the test time for testing N devices in parallel. For example, a device with a test time T_1 of 10 seconds tested using N=32 sites in T_N =16 seconds has a multi-site efficiency of 98.06%. Hence, for each additional device tested in parallel there is an overhead of (1-M) = 1.94%.

Typical site counts for various device types are shown in the ITRS "Site Count Table 2017". We also looked at the roadmap plans from 2013 and compared them to 2017 ("Site Count Comparison 2013 to 2017"). This clearly shows how increased device complexity as well as device interface complexity and costs have constrained efforts to expand site counts as quickly as desired.

Minimizing costs are a key part of the semiconductor manufacturing process. Test is no exception, although steady improvements in efficiencies over the last 15 years have lowered the typical cost of test as a percentage of IC revenue to less than 2-3%. The primary drivers of increased efficiency have been reductions in capital costs per resource and test times, coupled with increases in parallelism and Built-In Self-Test (BIST) capability. Most SOC devices are tested 2 to 16 at a time, and memory devices can have more than 1,000 devices tested at once. Measured as the cost to use capital equipment for test (in terms of cost per hour per device), these decreases in test cost will continue at a relatively consistent rate per year. The figure below shows the historical rate of capital investment in test, interface (consumables) and handling equipment.

It is notable that, in 2016, for the first time, the cost of consumable material had become the leading capital expenditure relative to ATE-based test. This has to do with the increased cost of interface material (primarily influenced by probe cards and relative items) and the decreasing depreciation period for materials utilized for the production of devices used in the mobile device space where devices have a shorter life span. In this case, material is typically discarded not because it has ceased to function, but rather because the devices it is used to test are replaced by newer versions.

Figure 6: Importance of Multi-Site Efficiency in Massive Parallel Test

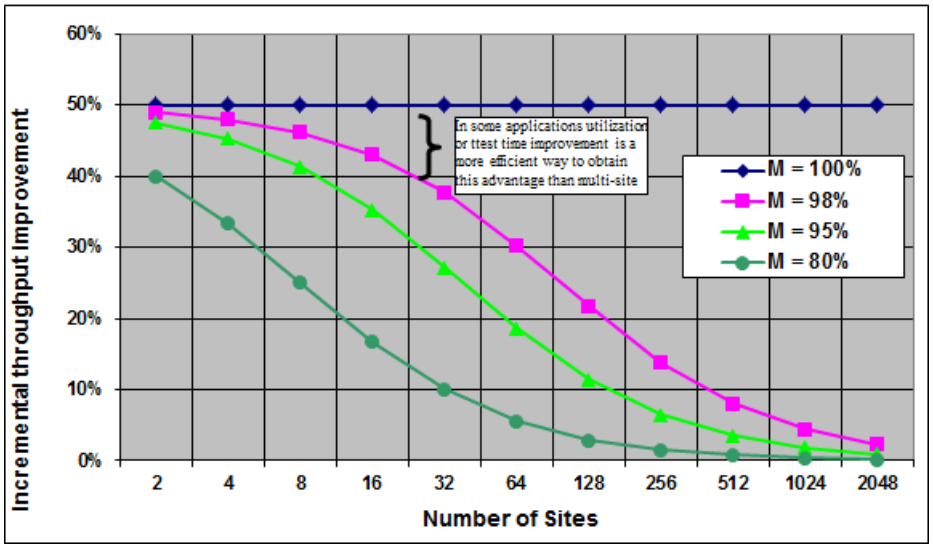


Table 1: Multi-site Test for Product Segments

		2018	2019	2020	2021	2026	2030	Drivers
High Performance MPU, ASIC (1)								
Wafer test	Number of sites	8	8	8	8	8	8	
Package test	Number of sites	16	16	16	16	16		MPU
SoC (2)								
Wafer test	Number of sites	8	8	8	16	32	32	SoC
Package test	Number of sites	16	16	16	32	64		SoC
Low Performance - MCU, MPU, ASIC (3)								
Wafer test	Number of sites	64	64	128	128	256	256	MCU
Package test	Number of sites	16	16	32	32	64	64	MCU
Mixed-signal, & Communications								
Wafer test	Number of sites	16	16	16	16	32	32	Mixed
Packaged Test	Number of sites	16	16	16	16	32	32	Mixed
DRAM Memory								
Wafer test [note 4]	Number of sites	3000	3000	3000	3000	3000	3000	DRAM
Packaged Test	Number of sites	1024	2048	2048	2048	2048		DRAM
At Speed DRAM Memory								
Wafer Test Parallelism	Number of sites	128	128	128	128	128	128	DRAM
3D Stacked Memory (Wide I/O, HBM, HMC)								
Wafer test	Number of sites	3000	3000	3000	3000	3000	3000	DRAM
Packaged Test	Number of sites	1024	2048	2048	2048	2048		DRAM
Commodity Flash Memory (NAND)								
Wafer test	Number of sites	2048	2048	2048	2048	2048		NAND
Packaged Test	Number of sites	2048	2048	2048	2048	2048		NAND
Stack test [note 6]	Number of sites	4	4	4	4	4	4	NAND
LCD Driver								
Wafer test (Small panel) (5)	Number of sites	6	6	8	8	8	8	LCD
Wafer test (Large panel) (5)	Number of sites	12	12	16	16	16	16	LCD
RF								
Wafer & Packaged test [7]	Number of sites	**32**	**32**	**32**	**32**	64	64	RF
CIS								
Wafer test	Number of sites	64	96	96	128	256	512	CIS
MEMS - Inertial Sensor (Consumer)								
Wafer test	Number of sites	64	64	64	128	512	1024	MEMS
Final test	Number of sites	98	128	256	256	512	1024	MEMS
MEMS - Inertial Sensor (Automotive & Industrial)								
Wafer test	Number of sites	4	4	8	8	16	32	MEMS
Final test	Number of sites	8	8	8	8	16	32	MEMS
MEMS - Microphone								
Wafer test	Number of sites	16	16	16	32	64	128	MEMS
Final test	Number of sites	49	144	144	144	256	512	MEMS

Manufacturable solutions exist, and are being optimized

Manufacturable solutions are known

Interim solutions are known

Manufacturable solutions are NOT known

Notes for Table 1:

1. Assumes I/O count of 250 for MPU and 1000 for ASIC
2. Assumes I/O count of 300
3. Assumes I/O count of 100
4. Wafer test uses Reduced Pin Interface
5. Assumes define Small panel as hand-held display application with one LCD device per each set and Large panel as TV display application with multiplex LCD devices per set
6. Engineering Testing
7. Maximum according to # active RF ports/device

As one continues to increase the number of sites, a low multi-site efficiency has a larger impact on the cost of test. For example, 98% efficiency is adequate for testing two and four sites. However, much higher efficiency is needed for testing 32 sites. At 98% efficiency, going from testing a single site to testing four sites will increase a 10s test time to 10.8s. However, going from testing a single site to testing 32 sites will increase a 10s test time to 16.4s, that is, significantly reducing the potential advantage of multi-site as shown in Figure 6. There are more efficient ways to reduce overall cost of test than going to the next setup with more sites in certain cases. Especially for high-mix, low-volume applications, there are many tester utilization challenges. In these setups, frequently, lower degrees of multi-site is preferable because test time improvement of techniques to improve utilization have a higher impact on the overall cost of test.

Touch-Down Efficiency (TDE) is defined as the number of wafer touch-downs required to test all devices on a wafer, relative to the theoretical minimum. TDE is influenced for the most part by the die size (and therefore the number of die per wafer) and the pattern used to probe. For example, if a device is tested 10 sites at a time, and there are 1,000 die per wafer, then ideally a probe card would have to touch down 100 times in order to tester the wafer and be 100% efficient. If, due to the mismatch between the round shape of the wafer and the linear or rectangular pattern of the probe card, the probe card must touch down 110 times to test the 1,000 devices, then the TDE is closer to 90%. This is shown in the figures below.

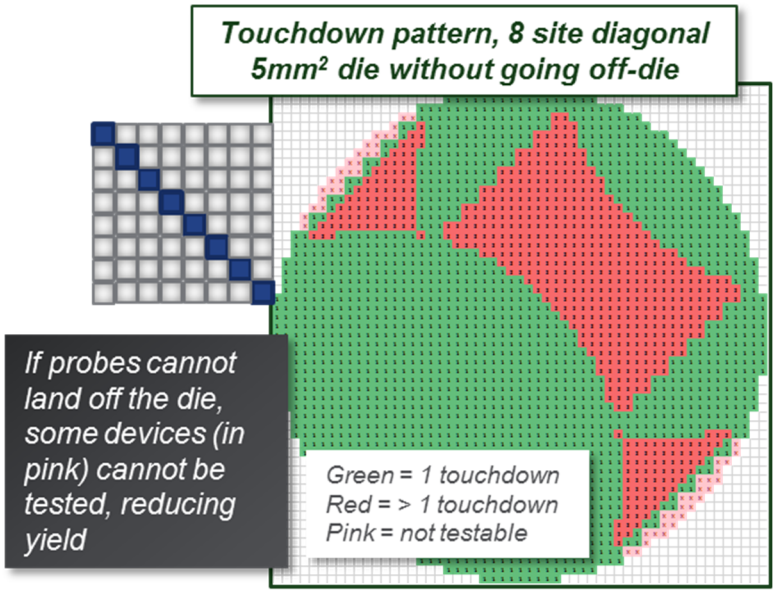


Figure 7: Probe Pattern of 5mm2 die using 8-site probe pattern

As die size of complex device increases, the TDE will continue to degrade as shown below in Figure 8. This degradation of efficiency will negate any advantages of increased site count and will eventually increase Cost of Test as shown in the example below. In this case, there are gaps in the probe pattern to allow for the inclusion of electrical components on the probe card required for the proper operation of the Device Under Test.

TDE inefficiencies will primarily be address by the development of singulated die testing technology. There is significant work underway to allow die to be reassembled in silicon panels that have a rectangular shape as opposed to the round shape of the original silicon wafer. The deployment of this technology will re-start the increase in site count at probe that is currently stalled due to interface costs and TDE limitations.

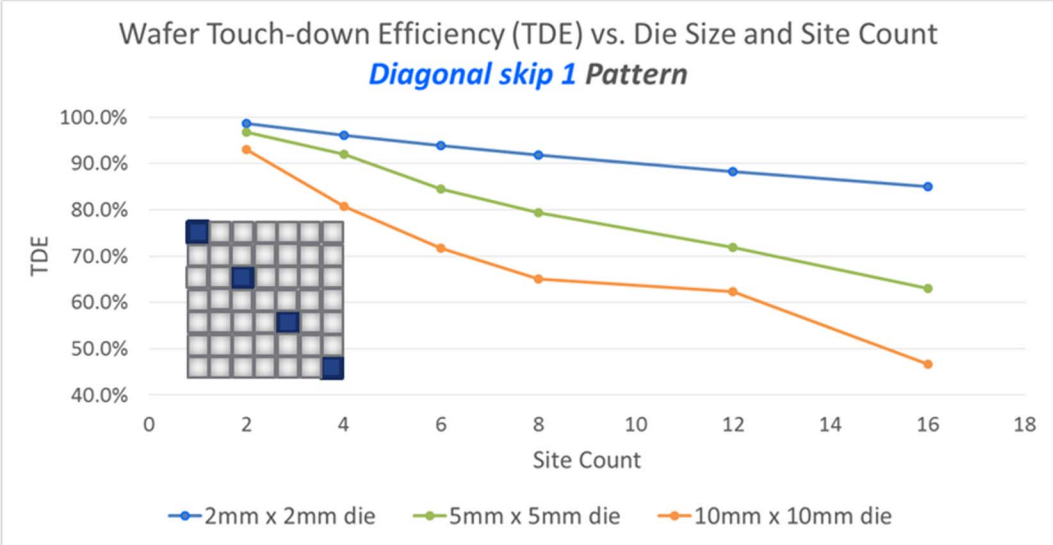


Figure 8: Touch-Down Efficiency as function of die size using a 4-site probe pattern

Summary

Major conclusions are:

- Cost of test has been declining for some time, but the rate of reduction has slowed down.
- Major reason for the slower rate of cost reduction are:
  - Packaging trends that drive more test at the wafer probe insertion where site counts are lower.
  - Increased cost of consumable material, which now dominates tester capital cost in terms of test cell costs.
  - Desire for higher yield, which has a much larger impact on overall device production costs than test costs alone.
  - Desire for higher device quality, especially for automotive applications, which necessitates more test.
- Potential solutions to decrease test costs are:
  - New probing technology which allows test of singulated die.
  - New PCB and Interposer technology to lower the cost and complexity of consumable material.
  - Factory automation.
  - Cost reduction of system-level testing.

Edited by Paul Wesling