INTERNATIONAL TECHNOLOGY ROADMAP FOR SEMICONDUCTORS

2009 EDITION

TEST AND TEST EQUIPMENT

TABLE OF CONTENTS

LIST OF FIGURES

TEST AND TEST EQUIPMENT

INTRODUCTION AND SCOPE OF THE 2009 EDITION

manufacturing test—reliability and yield learning. It is also important to note that the impact of these challenges affect not only on the manufacturing test process itself, but are essential to the entire semiconductor business, both in teeisof

Table TST1 Summary of Key Test Drivers, Challenges and Opportunities

Key Drivers (not in any particular order)	

KEY DRIVERS

As previously mentioned, the key drivers in the Test chapter are considered *boundary conditions* within which the

not deterministic within the time/vector synchronization standpoint. These behaviors, while correct from an end use customer and system standpoint would break the historical HVM test stimulus-response model, where one

and manufacturing lot populations. The need for better integrated usage of this test output (data) for fab process yield learning, maverick material identification, and feedback within a more distributed manufacturing test are all becoming more critical and even essential applications moving forward, not just nice to haves.

- *Dynamic test flows via "Adaptive Test"*—An emerging method to reduce overall test cost is to break the paradigm of running a constant test flow and instead, instituting a method whereby results of recently tested die or units can be used to determine the probability of need for a particular test. If the "importance" of a test or group of tests is deemed statistically low, then the test(s) can either be temporarily dropped or have other tests substituted. Probability triggers can also be used to increase the tests that are executed.

- *Higher Order Dimensionality of Test Conditions*—Historically, component manufacturing test has used testing at a simple matrix of points or corner conditions of two or three environmental variables (typically Voltage, Temperature, and Frequency) to guarantee the wider multi-dimension contour and interior space of product specifications beyond the contour outlined by the test points. Components in a broad range of markets from battery application platforms to client computing and servers are adding numerous schemes of power management that quickly add exponential complexity to the worst-case test conditions contour. For example, the complexities of component schemes may include multiple or variable power modes (sleep, hibernation, etc) or even multiple and *in situ* responsive variable Vcc and frequency control systems to achieve optimized power-performance in the application. The challenge of determining, characterizing, and optimizing (reducing) the larger set of environmental test points for component manufacturing test is daunting with the additional dimensions of these variables. Keeping pace with this complexity requires additional validation efforts and innovative methods to validate the quality of the more complex environmental test points. This is needed to ensure the lack of holes and predictability of more complex worst case conditions in the product test validation and test development phase. This increased complexity of test environment set points will also challenge the continued economic scaling of manufacturing test cost by the product of the applied test content and the number of iterations of its application at the various set points in the manufacturing flow.

CONTINUED ECONOMIC SCALING OF TEST

- *Physical Limits of Further Parallelism*—Over recent generations, continued increases in parallelism (in number of DUTs tested in parallel at a test insertion), particularly for commodity memories, but also for digital logic, have been a primary means of continuing the economic scaling of test in the context of devices with more transistors, increased functionality, and higher I/O and core speeds. In the current test tool and interface hardware integration paradigms, further increases in DUT test parallelism are reaching non-linear limits and the impact will be further seen in future generations. These are driven by the practical limits of how many electrical channels can be squeezed into the physical space between the parallel DUTs and the test instrumentation while yet maintaining acceptable physical and electrical proximities between the two. These approaching physical limits will require alternate means be used or expanded in order to maintain continued economic scaling of test, or alternately, new paradigms of DUT, handling, contacting, and test instrumentation integration must be developed that enable further increases in DUT parallelism beyond what is currently envisioned.

- *Managing (Digital Logic) Test Data Volume*—Increased digital logic die complexity and content drives proportional increases on the test data volume (number and width of vectors). Unconstrained, this additional test data volume drives increases in test capital and operational costs by requiring additional vector memory depth per digital channel of the test tools (ATE) and by increasing test time per DUT. Currently, a number of logic test vector compression schemes are being developed and applied in a variety of ways on the test databases themselves (for scan based tests) or via compression hardware (DFT) on the product die itself. Moving forward, compression will become more ubiquitous across component business segments, driven by the increasing product complexities and higher levels of product integration (e.g., SoC, SiP) and may ultimately require increases in the rate of compression (i.e. the compression ratio of the test database versus uncompressed).

- *Effective Limit for Speed Difference of HVM ATE versus DUT*—In recent years, a primary means of achieving economic scaling of test has been an increasing gap between application device speed versus manufacturing test instrumentation and applied test. This is one key aspect of distributed test, or test partitioning, and is enabled by a variety of DFT techniques such as I/O loopback and special test modes on the DUT. Maintaining this delta between application speed and design verification ATE versus HVM ATE enables a continued economic scaling for test, but there are limits to how much slower or how much less accuracy (e.g., for signal edge placement) is needed in combination with DFT to assure adequate quality of the DUT in the end user customer application. As device speeds and I/O edge rates increase, the manufacturing test instrumentation will lag but will also need to slowly increase, but probably in larger step function increments than the DUT families themselves, as

opportunities for ATE fleet replacements or expansions offer the most optimal opportunities to keep abreast in cost and quality effectiveness.

- *Managing Interface Hardware and (test) Socket Costs*—The portion of costs based on test and probe interface hardware and test sockets is an increasing proportion of the overall test cost. There are a number of factors

include multiple-adjacent-cell SRAM upsets due to single radiation events and erratic shifts in minimum-operating-voltage, problematic for low power applications.

All aspects of the test process, including defect modeling, test generation, test-coverage evaluation, DFT solutions, test application and diagnosis, must handle these realistic and changing populations of manufacturing/operating imperfections. Promising strategies include out-of-spec testing such as low-VDD or temperature, statistical methods, adaptive test, and realistic defect-based modeling/targeting.

(3) SCREENING FOR RELIABILITY

A lesser-publicized mission for semiconductor test beyond the primary "screening defects" or telling a good unit from a bad one has been to screen out infant mortality of the product population to acceptable levels. Another way of describing it more aligned to the "defect" mission of Test is to call this essential function "screening t > 0 defects" where "0" on the timeline represents the date product moves to the customer from the component provider. Historically, different semiconductor business segments have used a variety of techniques from burn-in to IDDQ, to voltage stressing during the manufacturing process to identify and screen sufficient numbers of the less reliable sub-populations or product dice to meet customer quality expectations. From a similar set of causes, notably increasing background leakage currents, and reducing product operational margin (collapsing Vdd/Vcc with process scaling), all of these techniques are becoming both less effective and more expensive to varying degrees. Where burn-in equipment and techniques had remained essentially unchanged and re-used over many process generations from the early 80s to the mid 90s, beginning in the late 90s burn-in production systems (in product segments that leverage this technique) have been one of the areas of the largest increases in test capital and interface hardware costs. At the burn-in elevated voltage and temperatures needed to accelerate latent defects, leakage levels are much higher than under normal application conditions. In addition, reducing product Vcc and temperature margins limit the range these conditions can be used to accelerate the latent defect populations. IDDQ, which has been widely used in product segments from ASIC to SoC to commodity memories, has been greatly challenged by the very same trend in proportional increases in background leakage, which greatly reduces the signal to noise ratio of "normal" static current levels versus DUTs that contain latent defects. In fact, this has been a challenge at least since the 250 nm DRAM half pitch, where many companies, began using more advanced techniques, for example, IDDQ "delta" and other Boolean static Icc/Idd combinations and comparisons for results of multiple tests to maintain efficacy against latent defects and reduce invalid yield "overkill" cost effects. Commodity memories, which have tended to have somewhat lower intrinsic leakage levels per technology generation, have managed to extend and depend on the various IDDQ techniques a little farther than have other product families, but even they are now forecasting a real drop off in effectiveness fir IDDQ in the next one to two technology generations. Similarly, voltage stress, or applying Vcc/Vdd and patterns well above nominal range to accelerate latent defects, also has been losing effectiveness as the differential between *Vstress* and *Vnominal* has continued to get smaller at each generation. In the long term, new techniques for providing this reliability screening function will likely be needed in some business segments and product types sometime in the next few generations. Some of the newer concepts under exploration are improved Boolean and distribution algorithms among various test results both intra-die test (results from the same die) as well as inter-die (wafer neighborhood analysis, analyses within lots, adaptive test limits, analyses across lots, etc.). Another vector that will likely play a larger role moving forward is correction, whether by self-test and self-correction, which is being implemented on some embedded RAMs today, or by error coding detection and correction (ECC) techniques. The challenge here will not be on memories, but rather how and when similar capabilities or alternate approaches to self-correction might be practically applicable and affordable for logic.

(4) POTENTIAL YIELD LOSSES

Manufacturing yield loss occurs whenever any test or inspection process rejects as faulty a device that would function correctly in the target system. Causes of yield loss include:

- Tester inaccuracies (timing, voltage, current, temperature control, etc.)
- Over-testing (e.g., delay faults on non-functional paths)
- Mechanical damage during the testing process
- Defects occurring in test-only circuitry, e.g., BIST
- Some IDDQ-only failures
- Faulty repairs of normally repairable circuits
- Overly aggressive statistical post-processing

Tester-inaccuracy related yield losses are being mitigated to some extent by the use of alternative test methods to do at-speed functional test. DFT methodologies must mature to provide better coverage of the "collateral" defects currently best identified by at-speed functional test vectors through advanced pattern application methods and novel fault models. Care

must be taken to avoid adding to the over-testing burden in this development path. Further work on appropriate fault models is also required.

Tester inaccuracies in parametric measurements (timing, voltage, current, and temperature control) and DUT parajEMC ETEMC BT

induced soft errors, erratic circuit behaviors, and other increasingly intermittent non-predictable lower level behaviors. In response, there is likely to be further extensions of device functionalities to provide additional detection and correction. Like scan, the development and deployment of lower level intrinsic mitigation schemes designed to provide improved system reliability, such as error correction used on RAMs today, and other capabilities of the future, are likely re-usable for test purposes as well (detecting, correcting defects, even systemic ones, on top of the more infrequent or non-predictable intermittent behaviors they would be originally designed to provide protection against). At the very least there will be integration issues with the test process (for example, for redundancy/repair for memories) but more optimistically they could likely be aligned to provide more effective, more efficient, or reduced overhead to future semiconductor test processes.

TEST AND YIELD LEARNING

In addition to the normal sorting function, test provides the essential feedback loop for understanding the characteristics of manufactured chips. Test must increase its capabilities to support failure root cause determination, cost-effective defect isolation, process measurements and design-process sensitivities.

ELECTRICAL-TEST -BASED DIAGNOSIS

Test-based-learning is needed for both (1) defects and (2) parametrics and variability. Defect-learning is needed for both random and systematic defects (see "Defects and Failure Mechanisms"). Test-structure-based defect learning methods suffer from both traditional area-related limitations and by the number of physical design configurations they are able to cover. As susceptibility to localized failures becomes a function of complexities that include OPC algorithms, nearby shapes and neighborhood densities, it is increasingly attractive to base learning on product-test diagnosis, which inherently reflects the physical design configurations of importance to the product.

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) enhances the negative impact of scribe-line-to-chip offsets. Moreover, test structures are limited by the number of physical and electrical configurations they are able to cover. As circuit parametrics are increasingly affected by such configurations, it becomes attractive to base learning on product test. Monitor circuits such as thermal and VDD sensors and process-monitoring ring oscillators are increasingly being embedded and distributed on-chip and can be used to help diagnose parametric fails and understand variability. Such circuits may increase in importance as modeling capabilities relied upon during design fail to keep pace with technology changes and more parametric problems reach the test floor. 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. Product-test-based parametric yield-learning is a relatively immature area in need of significant progress.

Electrical-test-based diagnosis methods are needed for circuit types including memories, logic and scan chains. Memory techniques, the most mature, are no longer adequate alone. RAMs have been superseded by microprocessors as technology leaders, they fail to utilize all metal layers and they represent only limited layout configurations. Logic diagnosis has been an active area of recent research and development, but still is challenged in terms of accuracy, resolution and large data requirements. Integration of layout information is commonly employed. Successful techniques for integrating in-line inspection results have been developed, but suffer from in-line data availability and ever-smaller, potentially non-visible, killer defects. Volume diagnostics, which combines results from many failing die, is a promising strategy for addressing remaining challenges. Multi-die statistical analyses, such as those employed in volume diagnostics, hold promise also for distinguishing systematic from random failure mechanisms. Identifying systematic failure mechanisms is important for process improvement, DFM guidance (including hot spot identification) and ensuring test quality. Challenges in making the distinction include setting expectations for random failures against which systematic are compared. Expectations based on design analysis, including traditional critical area analysis, suffer from the need to make failure-likelihood assumptions or depend upon defect models that may not be accurate given complicated design-process interactions. Another emerging role of tools is to provide guidance as to whether or not PFA is appropriate for a given failing part of group of parts, especially given the increasing prevalence of non-visible failure mechanisms (non-visible defects or parametric problems) and the related fact that PFA success rate and cost varies by failure mechanism. The yield impact of broken scan chains has made their diagnosis a recent focus, but more widely applicable and accurate methods are needed. Moreover, defects can affect clock and other infrastructural nets, which are difficult to diagnose with today's methods. The tools must handle all realistic physical defects, including resistive bridges, resistive contacts/vias and opens. Methods are also needed to handle diagnostics for fails detected by all major test methodologies, including scan-based and BIST-based test; functional; IDDQ and delay test.

The tools and methodologies should support several levels of software-based diagnosis:

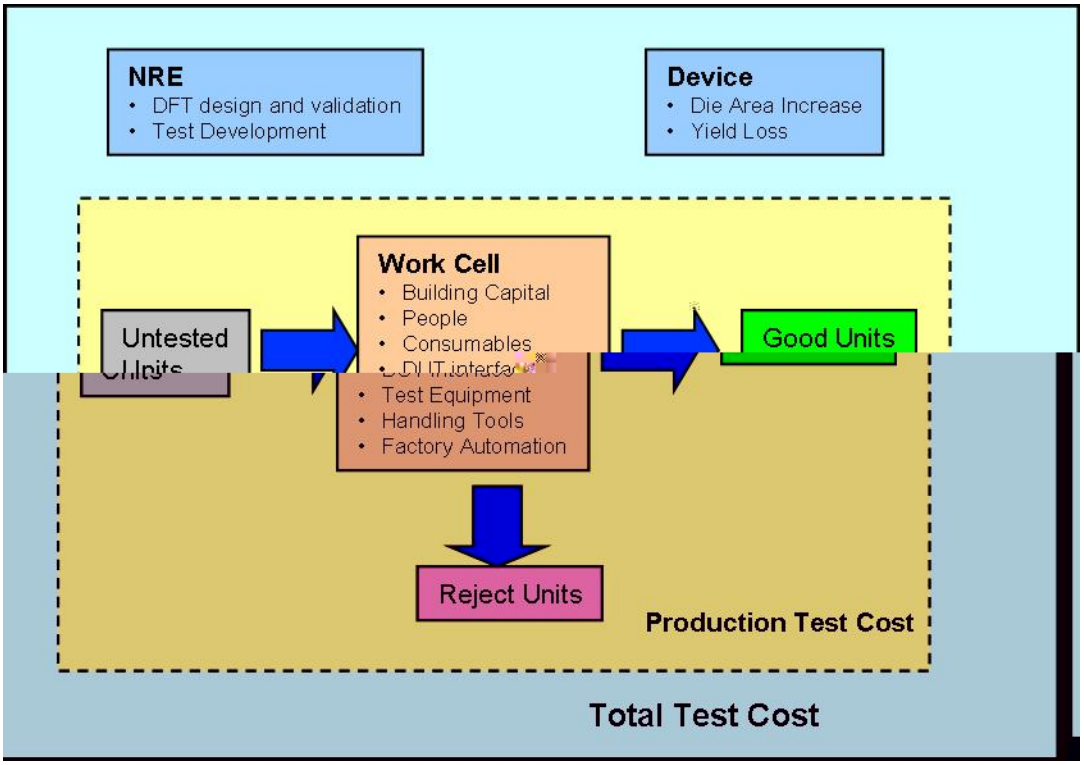
1.

5. **On-chip timing measurement—**Lower supply voltages are causing hot electron based photon emission to exponentially drop in intensity and shift to longer wavelengths. Improvements in time resolved emission (TRE) technology and solid immersion lens (SIL) optics are required to maintain capability. Emerging laser-based probes like laser voltage probing (LVP) are promising. Rapidly increasing active power is also challenging cooling solutions to not compromise optical access to the chip backside or induce vibration. Unless improvements are made in these areas, a radically different technology will be required.
6. **Sample Prep and Fixtures—**A number of developments in packaging have made backside access for timing and fault isolation difficult to impossible. For example, control of stress-relief induced warpage and cracking is required to enable package and die preparation and fixtures for tools like TRE and OBIRCH. Stacked die packages and 3D chips will further challenge sample preparation methods.

The development of new capabilities for failure analysis has become increasingly expensive and high-risk, especially as throughput decreases and Failure Analysis value-proposition goes down. As a result, tool installation is more centralized, reducing total market potential and attractiveness, especially for smaller suppliers and start-ups. The existing model of incidental (or accidental) tool development and R&D investment driven by one or two companies may need to be replaced by a shared consortium whereby cost and risk are spread across a number of stakeholders.

TEST COST FOCUS TOPIC

Significant progress continues in the reduction of manufacturing test cost, however much work remains ahead. Semiconductor test technology continues the trend towards higher channel integration and higher degrees of multi-site testing, enabled by advanced probe card technologies, new handling technologies, and design for testability techniques. Evidence of this trend is clear in the growing number of low performance ATE systems and (BOST) load-board solutions in standalone memory test and logic test. Even though a lot of efforts reduce cost of test, 30% of the respondents already consider the cost of test as one of their biggest concern and 85% expect cost of test to become their biggest concern going forward. Significant work remains to translate similar improvements to the broad market of analog and RF. Current DFT techniques are point solutions in certain areas but not generic solutions. The continuing focus on cost of test will result in a better understanding of cost trade-offs between test methodologies, ATE architectures, and distributed test across



multiple insertions among other considerations, resulting in overall test cost reduction.

Figure TST1 Test Cost Drivers

The cost of semiconductor test to the organization has many drivers as shown in Figure TST1. To better understand

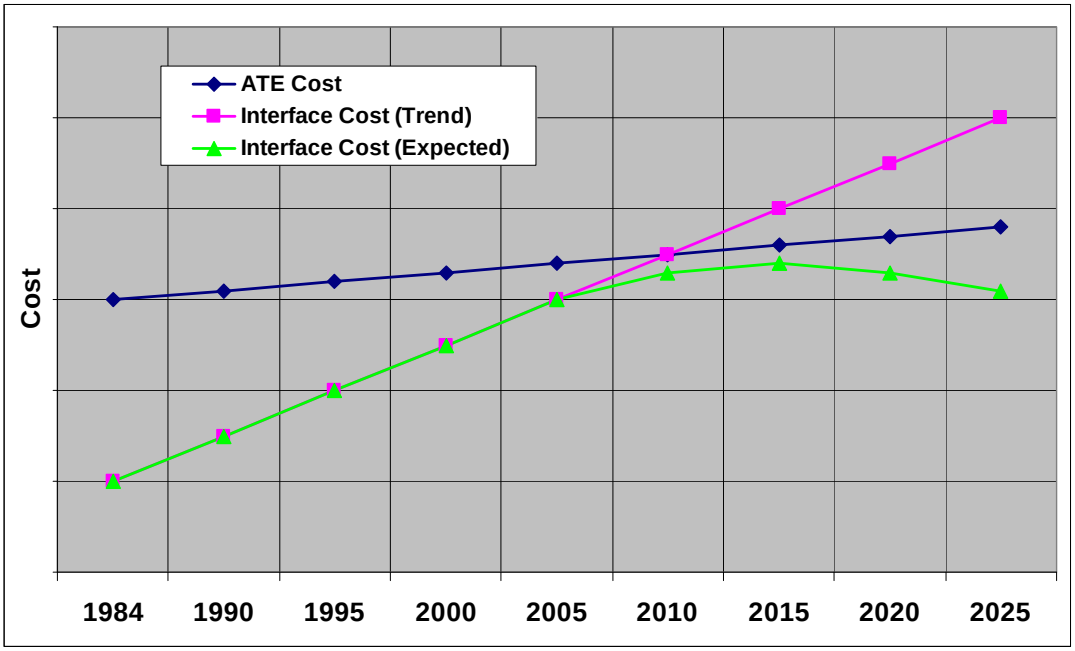


Figure TST2 Test Cell Cost / Unit versus Interface Cost Trend

A web survey was conducted to determine the top test cost reduction techniques that are currently used and those that may be deployed in the future. The results were:

CURRENTLY DEPLOYED COST REDUCTION TECHNIQUES

1. Multi-site & reduced pin-count
2. Structural Test & Scan
3. Compression/BIST/DFT
4. Yield Learning
5. Concurrent Test

COST REDUCTION TECHNIQUES THAT MAY BE DEPLOYED IN THE FUTURE

1. Wafer-level at-speed testing
2. Advanced embedded instruments
3. Adaptive Test
4. New contacting technologies
5. Build-in Fault Tolerance

BASE COST TREND

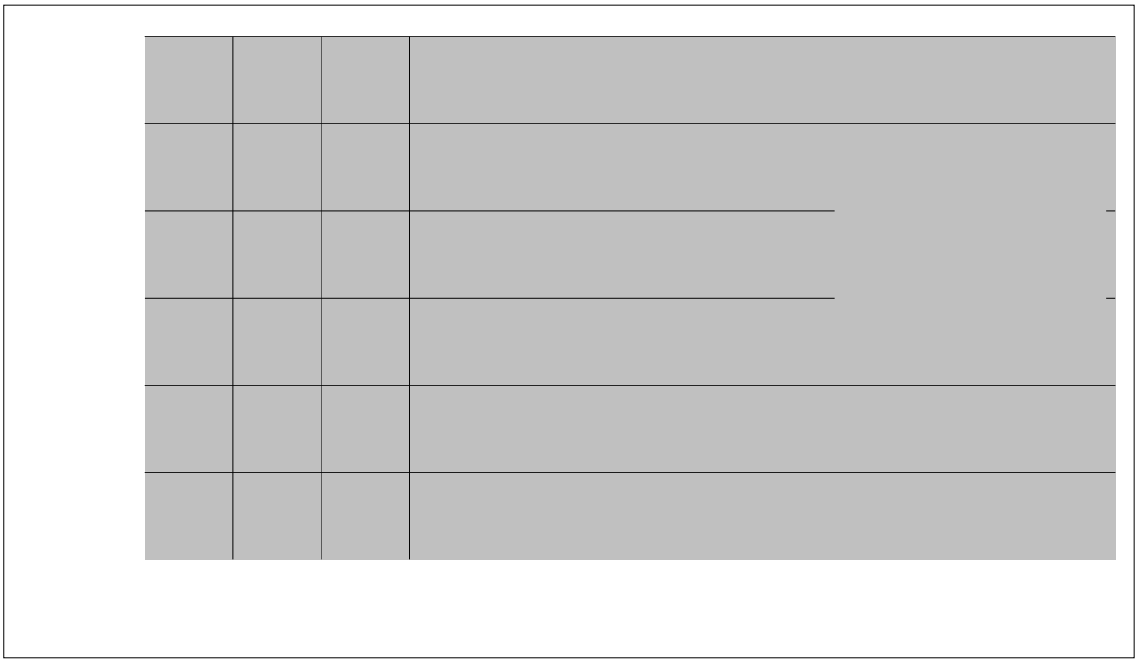
The total base is expected to decrease slightly over time. Platform strategies will extend the lifetime of the base infrastructure. Moreover, cost may move from the base infrastructure to the instruments. Multi-site test increases throughput and distributes the base cost across multiple dies, thereby reducing the base cost per site (and making the base cost less of a concern). For successful cost scaling using multi-site test it is important that the ATE infrastructure allows dedicated resources because shared resources may limit throughput. The trend of massive parallel test in memory will continue. Moreover, new probe card technologies and handler technologies, will enable massive parallel test in other segments (for both wafer and package test).

CHANNEL COST TREND

Continuing reduction in channel cost is essential for successful cost scaling using multi-site test: A dominating channel cost per site reduces the advantage of distributing the base-cost among many sites, whereas sharing expensive channels across multiple sites limits throughput. The channel cost is expected to decrease through continued integration within the tester electronics, and also by increased DFT adoption that reduces the ATE pin's performance requirements. Additionally, reduced pin-count test strategies utilizing small test ports can reduce the channel cost per site.

The relatively high cost of analog and RF test instruments, and the long test times associated with testing of these circuits, remain key challenges. DFT methodologies for analog and mixed-signal test are required.

16 Test and Test Equipment



- High-speed serial interfaces are penetrating ASIC and SoC markets. Jitter testing results in high test-times and equipment capital costs. As the number of interfaces increase, the cost problem will increase linearly. New test methods need to be developed to manage the cost scalinf prof2(t)10uet

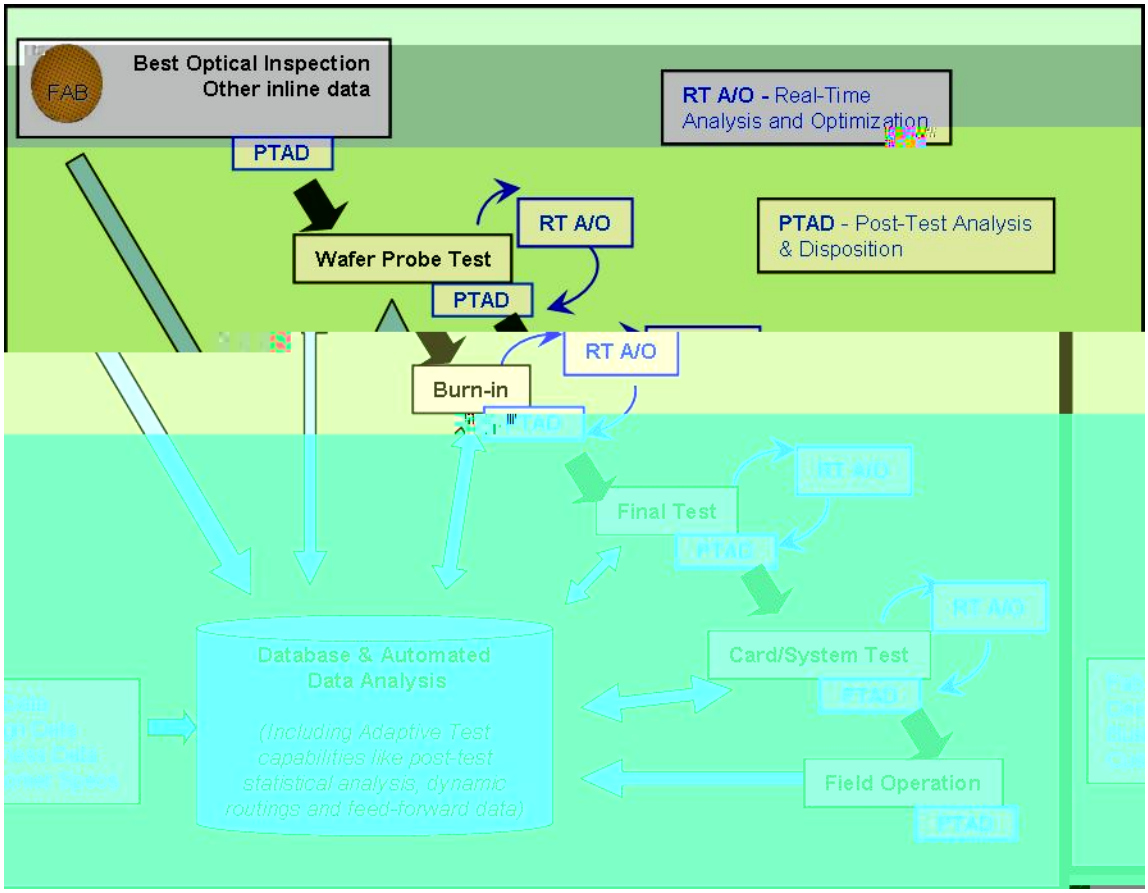


Figure TST4 High-level diagram of Adaptive Test Flow

PROCESS VARIABILITY – ITS IMPACT ON TESTING

The effects of process variation remain the overarching theme of the ITRS. For example, the ITRS Design chapter uses variation/variability and uncertainty over one hundred times. If design could control for all variation and uncertainty, then test would leverage these controls and largely ignore variation at test. Similarly, if systems could accommodate the variability of its components, then conventional test guard-banding would deliver useable parts. These levels of control and accommodation are unknown and hence test must identify and screen for variation by designing new test methods capable of constantly adjusting test conditions, test content and test flows.

Table TST3 "Stages" of Adaptive Test

Stages of Adaptive Test	*Implementation*

development, integrated and peripheral instrumentation, data quality assurance, data communication and delivery, and flexible test flows and manufacturing floors, so that every die is screened appropriately and with the least cost.

A functional abstraction for Adaptive Test allows others to build new and expand upon existing Adaptive Test applications. By necessity, abstraction removes significant detail to highlight the essential characteristics of the system being described. Answers to two related questions form the key to understanding Adaptive Test; "What components are reviewed and approved by the in-house group that controls changes to the test program?" and "What components in the test flow are data-driven by the observed DUT response?" The abstraction isolates two functional elements of Test: DUT response modeling and limit setting. A test program executes combinations of the two to disposition the part, e.g., final binning or continues testing.

ADAPTIVE TEST BUILDING BLOCKS

Adaptive Test depends on a set of infrastructure building blocks to enable dynamic optimization of test content and/or flow. The key elements are:

- Data (various sources, may be real-time, historical or feed-forward, and structured to feed decision algorithms)
- Data Structure & Control Path (deliver targeted data to the analysis engine and decisions to the action point)
- Decision Algorithms (models to automatically adjust test limit, flow, or content based on input data.)
- Analysis Engine (Executes model/algorithm on the data structure to return control signals based on specific product parameters. The engine may be separate or embedded in test program or station controller software.)
- Station controller/test program hooks to communicate with or run the analysis engine and take action.

Adaptive Test will drive additional infrastructure and business processes **that must enable the following:**

- Full traceability (details of how Adaptive Test was applied for each product/part tested)
- Standard models and algorithms (Parts Average Testing, Linear models, Machine learning, sampling, …)
- Recipe management (model form and parameters, target products/lots)
- Analysis functions (to quantify trade-offs in quality, reliability, test time, yield loss, risk assessment)

IMPLICATIONS OF ADAPTIVE TEST FOR TEST DATA AUTOMATION

- Test data storage for large volume of data coming from the complete manufacturing flow
- Real-time retrieval of historic test data (die level feed forward, real-time lot statistics, feed-back)
- Real-time data integrity validation
- File formats and data representation standards for response and traceability
- Algorithm representations

IMPLICATIONS OF ADAPTIVE TEST FOR ATE

- Define/create a (standard) interface at the ATE for a data path transfer of the real-time test inputs/results
- Create a (standard) real-time interface at the ATE for controlling the test flow and hardware configuration for an adapted test flow.
- Define the interface for adjusting test program content real-time (between units or lots).
- Automatic adjustments to hardware configuration for optimized test flow

Table TST4 Implications of Adaptive Test

Challenge	*Required Direction*

Figure TST5 Organization of Cores for System Integration and Applications

In the recent past, these core semiconductor technologies and applications have demanded distinctly different test solutions, each having specific test equipment and interface tooling markets. Increasing integration is blurring these boundaries. It has also raised the stakes for DFT as successful integration is determined not just by "can it be done?" but also "can the integrated whole be tested economically?" The remainder of the Test Technology Requirements section will address the test challenges associated with increasing integration followed by the test requirements of each constituent core technology.

SYSTEM INTEGRATION—SOC AND SIP TEST CHALLENGES AND IMPLICATIONS

While possibly equivalent in theory, SoC and SiP are very different from each other in terms of which technologies tend to be more easily integrated in package versus on chip and each has very different test implications. Recent advancements in assembly and packaging technologies coupled with the difficulty of optimizing the same wafer fabrication process for different core semiconductor technologies have provided a lot of momentum for SiP, causing some to forecast that SiP will be dominant. It may be that wafer fabrication process improvements and design/DFT needs could push SoC to the fore front or there could be hybrids of the two. One thing is clear: integration is a trend that will continue. The only questions are how fast and in what forms. The next two sections will discuss the test challenges and implications associated with SoC and SiP respectively.

SYSTEM ON A CHIP

A SoC design consists of multiple individual design blocks, or cores, using different technologies (logic, memory, analog, RF, etc). This assortment requires a diversity of solutions to test the specific technologies corresponding to these embedded cores. Increasingly, SoC design is relying on a database of pre-existing IP cores that encapsulate the design itself, its embedded test solution, and the interface to other cores. SoC test implies a highly structured DFT infrastructure to observe and control individual core test solutions. SoC test must include the appropriate combination of these solutions associated with individual cores, core test access, and full-chip testing. One fundamental challenge of SoC test is the need to combine test requirements from multiple sources with differing testability methods. Another challenge of SoC test is to highly optimization both quality and cost. The overall quality and cost of SoC should be evaluated and be adjusted to an acceptable level for customers. Some hierarchical approaches or parallel approaches will be required.

The quantitative trends and requirements in *Table TST5* were estimated according to the trends that are shown in the SoC-PE driver model of the ITRS Systems Driver Chapter. The trends include the number of transistors, flip-flops, and

memory bits. Although the estimation has been done assuming a specific SoC model, the requirements should also be

Figure TST7 Impact of Repeated Use of Same Cores

REQUIREMENTS FOR EMBEDDED MEMORY CORES

SYSTEM IN A PACKAGE

In contrast to SoC, SiP offers the option of testing components prior to integration. This is important since integrating one bad component could negate several good components in the SiP, severely limiting SiP yield. In addition, this component testing must typically be done at wafer probe test since integration occurs at assembly and packaging. A key challenge then is identifying good die prior to integration. The term "known good die," or KGD, was coined during the mid-1990s to designate bare die that could be relied upon to exhibit the same quality and reliability as the equivalent single chip packaged device.

In most instances, testing and screening the device in a single chip package format achieves the outgoing quality and reliability figures for IC products shipping today. Wafer probe test is not generally suitable for performance sorting, reliability screening, or effective parallel contacting, so it is generally more efficient to do these tests at the package level using test and burn-in sockets, burn-in chambers, and load boards. Consequently, KGD processing implies that die will be up-binned at probe or with a subsequent insertion of die level tests and screens to meet acceptable quality and reliability targets. The key short term challenges are to determine the quality and reliability targets required in different market segments, develop cost effective tests and reliability screens that can be applied at the wafer or die level, and to develop quality and reliability methods that provide high confidence regarding quality and reliability levels achieved. Longer-term challenges will be to move to a complete self-test strategy with error detection and correction available in the end application.

SiP TESTING AND EQUIPMENT CHALLENGES

SiP products can present many unique challenges to backend manufacturing flows because SiP products often contain die from more than one supplier. This can create problems in the areas of:

- development of a package test strategy to realize both cost and DPM goals
- production flows to accommodate the necessary reliability screening methods (burn-in, voltage stress, etc) of diverse product/process technologies
- failure analysis methodologies for fault localization in order to resolve quality problems and systematic yield issues

SiP test at the package level closely resembles the test problems of complex SoC products, that is, a variety of IP, each with specialized test requirements, which must be consolidated into a single consistent test flow. In the case of SoC, because everything is on one chip and designed together, various block test strategies can be consolidated via the use of test shell wrappers, test control blocks, etc. using strategies such as defined in the IEEE 1500 specifications. In the case of SiP, die suppliers may be reluctant to provide information needed to access special test modes (sometimes considered confidential, especially for commodity memory products) and the individual die may not have the necessary test infrastructure overhead to implement test strategies commonly used for SoC.

Even in the case of SiPs that use only KGD, a certain amount of testing is necessary after final assembly to ensure that the die have been assembled properly. When final assembly may include die thinning and stacking, which can damage/change KGD die, additional testing may be necessary. For the case of fault localization, the ability to narrow the failure to a specific die, and further to a small region of that die, may require full understanding of the detailed test strategies for that die, even if not necessary in normal production..

In the case of reliability screens, some die may require burn-in while others may require only voltage stress. Stress conditions for one die may be inconsistent (or even detrimental) to other die in the same package. Resolution is more difficult since the different die in a SiP product often have totally different processes. One solution is to avoid reliability screens after final packaging but this can increase overall costs (for example, wafer level burn-in is typically more costly than package level burn-in).

When heterogeneous die are assembled into a multi-chip package, several test insertions on different platforms may be required to test the assembled module fully. The multiple test insertions may result in test escapes or yield fallout due to mechanical damage. New testing equipment will be required to accommodate contacting the top side of the package for package stacking. For wafer stacking technologies, better redundancy/repair technologies are needed so that the final stack can be "fixed" to achieve yield/cost targets. Design and production of electronic systems that can detect failed components and invoke redundant elements while in service is a key challenge for SiP reliability.

WAFER TESTING AND EQUIPMENT CHALLENGES/CONCERNS

The probe card technologies in common use today are less than ideal as a "final test" environment. Since much of the performance based speed critical, RF, delay and analog testing is presently performed at package level, a critical challenge for KGD processing is the development of cost-effective, production worthy, reliable and accurate methods of rapidly identifying devices that are defective or will fail early in an application before those devices are transferred to the next level assembly.

Test time for certain technologies, such as display drivers or state of the art DRAM is exceedingly large. Because of the limitations in the wafer probing process, the test throughput is much less than packaged components. The challenges for fully testing DRAM die in a cost effective manner at the wafer level include development of technology that can probe multiple die on a wafer without overlapping previously probed die or stepping off the wafer, and to avoid wasting test time and power on all previously rejected and obviously non-functional die.

WAFER TEST FOR RF DEVICES

A key challenge for applying KGD processes to RF die is development of high performance, fine pitch probe cards. Because of the small size of RF die, the pad pitch is very small. As an example, the pad pitch in some products can go below 75 μm, which is the limit of the actual probe technology today.

In order to obtain good signal integrity during RF probing, a configuration of GND-Signal-GND for RF signals is required. A key challenge for KGD processing of RF devices is to ensure that the GND-Signal-GND configuration is designed into the die to maintain the RF path at controlled impedance, given proper probe card design and RF probing techniques.

RELIABILITY SCREENING AT THE WAFER OR DIE LEVEL

Voltage and temperature over time are the known stresses for accelerating silicon latent defects to failure. These are more readily applied at package level than at wafer or die level. Applying these stresses prior to packaging the die is a key challenge for KGD.

Development of a cost-effective full-wafer contact technology with the process capability required for manufacturing is a key challenge for the industry. Contact process capability is a function of not only the contactor technology performance but also the burn-in stress requirements for a given product.

STATISTICAL PROCESSING OF TEST DATA

Techniques using statistical data analysis to identify subtle and latent defects are gaining favor in the industry, especially for device types with low shipping volumes, part number profusion and short product lifetimes that make burn-in an untenable option and for products where intrinsic process variation makes separating good die from defective die

LOGIC

The focus of this section is the testing of CMOS digital logic portions of highly complex digital logic devices such as microprocessors and more generally the testing of logic cores that could stand-alone or be integrated into more complex devices. Of primary concern will be the trends of key logic test attributes assuming the continuation of fundamental roadmap trends. The "high volume microprocessor" and "Consumer SoC" devices are chosen as the primary reference because the most trend data is available for them. Specific test requirements for embedded memory (such as cache), I/O, mixed-signal, or RF are addressed in their respective sections and must also be comprehended when considering complex logic devices that contain these technologies.

HIGH VOLUME MICROPROCESSOR TRENDS DRIVERS

The trends in the first part of the tables are extracted from other parts of the ITRS, and are reproduced here to form the foundation of the key assumptions used to forecast future logic testing requirements. The first two line items in *Table TST6* show trends of functions per chip (number of transistors) and chip size at production. Chip size is held relatively constant aside from incremental yearly reductions within a process generation. The next line item reflects a trend toward multiple core designs to address, in part, what has become the primary microprocessor scaling constraint - power consumption. The ITRS currently assumes a doubling of cores with each process generation. The last two line items in this part of the logic tables are the nominal device Vdd range and off-chip data rate trends.

SYSTEM TRENDS DRIVERS

System trends drivers are very important to consider when projecting future test requirements. For example, one of the most critical system constraints is power consumption. The proliferation of mobile applications, lagging battery technology improvements, system power dissipation issues, and increased energy costs are all contributing to a practical cap on device power consumption. The era of device power increasing unconstrained with increasing performance is over. This does not necessarily mean that performance will be similarly capped, but this is one of the main challenges to be overcome if Moore's Law is to continue. Innovations in transistors, process technology, design architecture, and system technologies could all have a major impact.

One system technology innovation that could impact test would be the integration of voltage regulation on-chip/package. Increasing chip power and increasing number of cores make this ever more likely for at least two reasons. The first reason is that eventually the package limits power consumption by constraining the number of power/ground pins and the maximum current per pin. These constraints can be greatly eased with on-chip regulation since you can then deliver power to the chip at a significantly higher voltage. The second reason is that multi-co of 8(can-1.15 o c)8(o.0837 T8746.886 0 gn arc)

- DFT is required to minimize the complexity of testing embedded technologies such as memories and I/O. For example, memory BIST engines are routinely integrated into the device to

used in high speed interfaces, the adoption was far from simple. For example, most of the Gigahertz interfaces are built in SOC type of devices with mainstream CMOS processes. The large scale integration not only presents a challenge in designing reliable high performance interface IP targeted for a very noisy SoC environment, but also on testing of such kind of high performance interfaces. The other challenge derived from high integration level is the trend for lower Gbps-per-mW, which is becoming a major competitiveness measure in this area. This power reduction requirement challenged the design margins. This impacts the test area by an intricate balance of "guarantee by design" and "production test coverage". As an integrated IP block, Gigahertz interface testing is now tied to the whole SoC chip level testing. For most of these applications, it means high frequency/performance testing is tied together with large amount of CMOS digital pins/logics. Test cost tolerance, high frequency instrument availability, and signal integrity restriction on test hardware design are the new areas for development in the last few years.

As in any technology roadmap analysis, sustainability of the exponential growth trend depends on changes in technology requirements and foundation technology development. As forecasted in 2007, the telecomm, datacom, and storage industries reached a temporary gigahertz plateau and it will require a serious undertake in technology to go beyond the somewhat established 13Gbps range. As of 2009, shrinking process technology is starting to support growth on data rates beyond 13Gbps; such as 14Gbps FC and 16Gbps PCIe. Some next generation focused common electrical interface (CEI) of OIF are expected to reach up to 25Gbps. Unlike the long haul telecomm applications, some of the Gigahertz interface (especially the chip-to-chip interfaces) have chosen to grow in port count (like a bus) until the higher speed technology becomes more cost efficient for use (for example, CEI/OIF, and 40G/100G Ethernet (IEEE 802.3ba standard)). Some applications have opted to change the coding scheme which easies the required generational speed increase, so the expected 3 year doubling requirement is somewhat reduced (for example, PCIe, and FC). Future data rate progression will be slower than that seen in recent history but still requires doubling every few years.

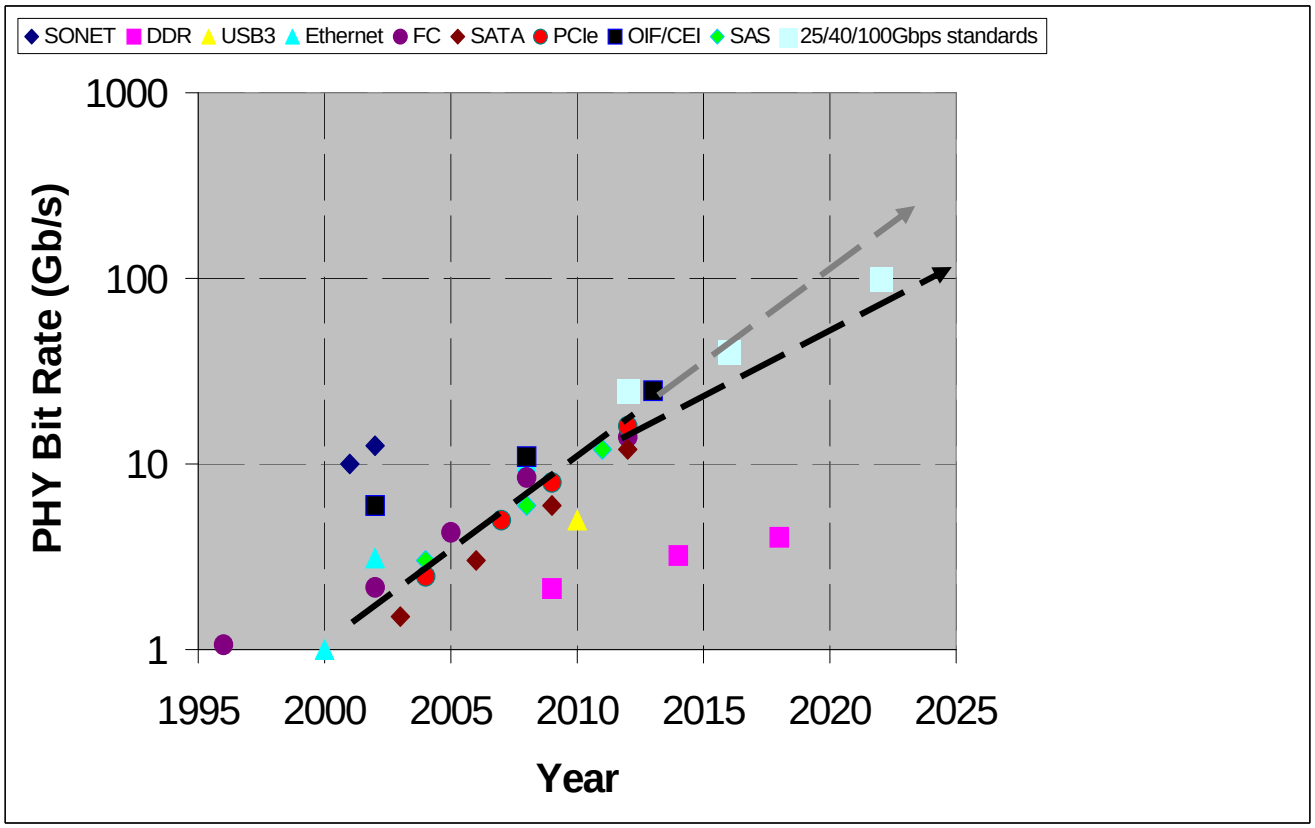


Figure TST9 High Speed Interface Trend

In the past two years, the test and measurement industry has made significant progress in providing high-speed serial link test solutions. Several ATE suppliers provide pin-card-level integrated solutions beyond 6Gbps, and others already have 12Gbps solutions. Throughput for some ATE solutions has greatly improved from the previous generation. More flexible clocking schemes eliminated a lot of overhead for frequency and phase synchronization. Jitter measurement and jitter tolerance tests are no longer an after-thought in these instruments, even though true jitter compliance tests are still challenged by the test cost constraints of many applications. ATE instruments are also beginning to support alignment

pattern type of synchronization and the typical PRBS based patterns used for BIST. From the basic at-speed testing instrument stand point, it is safe to say that ATE test solutions for 6Gbps have been established. Of course, there will always be applications that push beyond the ATE speed limit (such as for high-end PHY) or port count limit (such as for network switching devices).

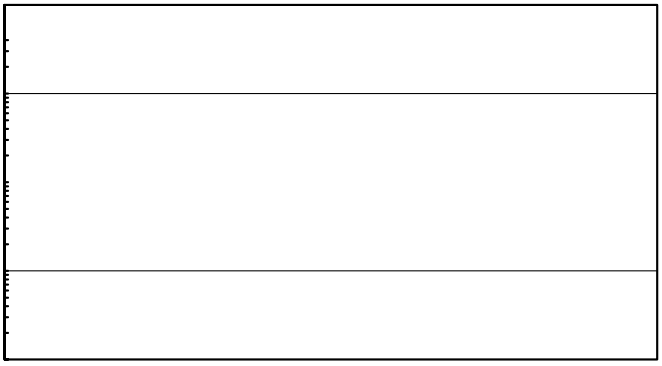
At the present time, high-speed production test solutions range from internal digital loopback, DFT assist digital loopback, external wired loopback, external active loopback, high end ATE pin cards, test modules, golden device on test board, to add-on external instrumentation. Each of these approaches has their own pros and cons, which each company must consider. The trade off is of course on silicon area used for DFT vs. implementation costs vs. test equipment and interface cost. The tolerance to defect rate of different products is also a major determining factor for the different test choices.

Another challenge to test instruments is the wider adoption of on-chip equalizations. In order to keep the overall system cost down, low cost PCB lamination materials such as FR4 will still be the material of choice for most of the telecomm backplane and computing applications. However, the preference for FR4 forms a bottleneck in spectral bandwidth. Several techniques are being developed to extend the transmission data-rate under this constraint, such as transmit-side equalization (single/multi-tap pre-emphasis/de-emphasis) and receiver-side adaptive equalization (analog equalization and digital feedback equalization). In order to test these types of features, test instruments need to go beyond generating and

each type are selected with cost in mind. The economics of high-performance, long haul, communication related products typically allow a more traditional, instrument based test approach or use of a hybrid tester as discussed earlier. Although reliable DFT features or other low cost test techniques are the ultimate solution for large port count SerDes, there is still a strong desire that the tester can provide at-speed characterization or early production. With accelerating technology improvements, the life cycles for most products are becoming shorter, therefore it may become increasingly difficult to verify and optimize DFT circuitry and yield learning.

4. **Jitter Decomposition and Measurement—**The jitter generated by the transmitter is the key parameter to guarantee transmitter quality. Many serial link standards adopt the concept of separating jitter into deterministic jitter (DJ) and random jitter (RJ). The traditional concept of histogram based peak-to-peak jitter has been replaced by the concept of total jitter (TJ), which is associated with a certain BER for the serial link (typically 10^{-12} or smaller). As the data rate keeps increasing to ~ 10 Gbps and higher, jitter amplification, (a new phenomenon during which high-frequency jitter caused by pulse width shrinkage or duty-cycle distortion (DCD) gets amplified by the lossy channel (e.g., FR4 based)), becomes severe and needs to be bounded. In addition to the conventional DJ, RJ, and TJ components, Pulse width jitter (PWJ) and DCD have become the new jitter components to test at 10 Gbps and higher. Furthermore, in some high-volume and high performance I/Os such as PCIe 3.0, a new criterion of whether the jitter is correlated or uncorrelated (to the data pattern) is established to account for the compensation effects of equalization to the correlated jitter. As such DJ and RJ are further separated to correlated DJ (including data–dependent jitter (DDJ)) and uncorrelated DJ and RJ. Correlated DJ (e.g., DDJ), uncorrelated DJ and RJ, along with PWJ and DCD need to be tested to ensure the interoperability and utilization of the jitter margin provided by equalization circuits. There are also other trends to measure jitter in terms of cycle-to-cycle jitter, peak-to-peak jitter or RMS jitter within a certain numbers of cycles, which could be more meaningful for clocking schemes only allowing very short-term jitter accumulation. Jitter measurement also imposed a very stringent signal integrity requirement. On the other hand, with the source synchronous bus entering the Gigahertz range, it introduces another challenge in jitter measurement – the uncorrelated jitter between the clock path and multiple data paths. This is certainly beyond the traditional serial PHY jitter definition.

   To accurately measure the DJ and RJ, the jitter floor instrument or tester needs to be significantly below the DJ and RJ of the DUT to avoid the errors in the measured DJ and RJ. The instrument or tester for jitter measurement had also adopted the DJ and RJ merits for its performance. Most of the high-speed standards specify a ~ 0.3 UI TJ at BER =1e-12, with 0.15 UI from DJ and 0.15 UI from RJ. Following this jitter budget allocation distribution, and assuming that a 10% (a relatively loose accuracy target) TJ accuracy for an instrument or tester, then its DJ and RJ accuracy limits can be estimated via dual Dirac model (i.e., DJ distribution is a dual Dirac, and RJ distribution is a Gaussian), and the TJ, DJ, and RJ accuracy limits as a function of data rate are shown in Figure TST10. At 10 Gbps, the TJ, DJ, and RJ accuracy requirements are 3ps, 1.5ps, and 0.11ps respectively according to Figure TST10. These jitter accuracy targets can be met by most leading edge laboratory instruments. If large TJ error is allowed, then DJ and RJ accuracy precision will be relaxed proportionally. It is worth mentioning that this figure intends to give a math and physics based TJ, DJ, and RJ accuracy guideline for a jitter measurement instrument or tester, rather than a specification or hard requirement.



8. **Calibration firmware—**Because of the power and area requirements for GHz I/O design, the analog portion of

remain a constant. I/O voltage decreases are pushing the operation 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 is required. 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 DRAM bits will not match the density growth rate of commodity DRAM and NAND. The major concern for a merged logic-DRAM design in a dual-gate process will be array noise and sense-amp imbalance. For the 100 nm DRAM half pitch and below, DFT for inline defect detection will be necessary for product development.

Embedded Flash memory bits will grow exponentially in the near term and then double every two years in the later years of the roadmap. More devices will include both DRAM and Flash memory. Oxide reliability, sense-amp imbalance, and oxide-nitride-oxide (ONO) scaling will be the major concerns in flash memories in the future.

To enhance test productivity, new test-oriented architectures will be required. Built-in self-test and built-in self-repair will be essential to test embedded DRAM and embedded Flash memories and to maintain production throughput and yield. 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 could be tested and repaired on the memory tester, while the logic blocks are tested on the logic tester. Embedded SRAM test requirements are captured in the High Performance Microprocessor section of the chapter.

ANALOG AND MIXED-SIGNAL

The economic benefit of monolithic integration (SoC) and single package integration (SiP) is well established. This integration has combined digital, analog, power management, mixed signal, and RF/microwave circuitry 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.

Another important trend impacting mixed signal and analog testing is the compelling economics of multi-site testing, also called parallel test. To support parallel test, many more instrument channels of each interface type are required to keep test cell throughput high.

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 non-deterministic and must be 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.

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.

This section focuses on analog/mixed-signal test requirements. The Mixed-Signal Test Requirements table focuses on test instruments rather than specific chip applications. The test instrumentation must often cover more than one device market segment to provide sufficient utilization in a single tester configuration, so the requirements for multiple segments are aggregated into a few instrument categories. The analog waveform generation and capture requirements are set in two classes: low frequency—basic/minimum requirements for a mixed-signal ATE and very high frequency high-end requirements. Where appropriate, the mixed-signal instrument requirements are linked to other sections and tables in the roadmap.

There are two important trends. The first is to deliver adequate quality of test. Most analog/mixed-signal testing is done functionally. 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.

The second key trend is to enable the economics of parallel test through instrumentation density and parallel test efficiency, a measure of the overhead in testing multiple parts. The level of parallelism shown in *Table TST2* "Multi-site Test for Product Segments" indicates 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 are lagging. No proven alternative to performance-based analog testing exists 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 or elimination of the need for external instrumentation.

Table TST9 Mixed-signal Test Requirements

IMPORTANT AREAS OF CONCERN

1. Time-to-market and time-to-revenue issues are driving test to be fully ready at first silicon. The analog/mixed-signal test environment seriously complicates the test fixtures and test methodologies. Noise, crosstalk on signal traces, added circuitry, load board design complexity, and ATE hardware/software issues 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.
2. 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, will need 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.
3. Improvements in analog/mixed-signal DFT and BIST are needed to support item 1 and 2 above.

RADIO FREQUENCY

Four main RF frequency areas are distinguished. The low frequency range up to 3 GHz is dominated today by long distance cell phone communications technology. With wireless client communications protocol (WiMax) this will transfer over time to the higher frequency bands. The 3-6 GHz range is used by satellite TV and 802.11a networking. The 6–45 GHz band is focused on medium distance communications (Bluetooth moving to ultra wide band (UWB)). The high section, 45-94 GHz, is used primarily for short distance radar applications, particularly automotive, but 60GHz will also be used for congested area (short range) WLAN and satellite to satellite communication. CMOS devices are expected to be able to support 60GHz.

The most important movement is the increase in frequency: using higher frequency bands in line with the 802.11 and 802.16 communication standards. For the high frequency ranges (> 12 GHz) classical test techniques (non-modulated) are expected to fulfill the requirements.

An important requirement for test is full synchronization between the power/digital/AC baseband part of the tester and the RF instruments. Error vector magnitude measurements are a prerequisite. An important trend is the move of RF into SoC and SiP solutions. This requires not only test solutions for the RF parameters, but also one in combination with high-end digital and mixed signal requirements. For SiP, the wafer probe capability becomes important. To cope with the economics of RF becoming a real commodity, multi-site will also be mandatory and will increase in this application area. The tooling (load boards, sockets, probe cards) is also critical to ensure signal integrity to and from the DUT. Looking at these challenges, the need for specific design-for-test for RF and finding lower cost alternatives to functionally testing RF devices is expected to increase heavily in the near term.

Table TST10 RF Test Requirements

IMPORTANT AREAS OF CONCERN

1. The increase in performance and frequency in combination with the economics of test will put a strong focus on novel design-for-test and alternative test techniques to be developed in the coming years.
2. RF will much more frequently be embedded into products via SoC or SiP techniques. Combination of RF tests with (high-end) digital and mixed signal will be more common. RF test on wafer level will increase. Next to the test system, there will also be emphasis on the tooling (load boards, sockets, and probe cards) to cope with signal integrity.

3. Source and measurement accuracy for phase noise and signal detectors are adequate today but must improve in the near term. Phase noise at 100 KHz needs to improve from today's -120 dBc/Hz to at least -130 within the next year or two.
4.

radical currents and powers conjured up by stress conditions. The industry's ride on the "Performance Juggernaut" isn't over quite yet.

DFR also has three key components: 1) technology design, 2) chip design (logical and physical), and 3) system design. In each of the three, the DFR work must strive for *defect tolerance*. In the case of technology design, leakage induced power mitigation maintains an edge in importance over defect tolerance. Regarding chip design and DFR, power mitigation and fault tolerance are at par in design priority. Redundant element analysis and power dissipation analysis burn considerable design engineering horsepower. At the system level, defect tolerance exists in the forms of error detection/correction and redundant elements.

In the arena of *reliability screens and test methods*, the literature is rich with techniques and methodologies with champions and supporting/compelling/biased data. Debates vary, depending upon the technology generation, chip/circuit type, design style, performance target, reliability requirements, and defect type. As long as excessive voltage and temperature retain the throne of defect acceleration, RS&TM will challenge the best and brightest minds in power delivery and thermal solutions. One must be able to accelerate defects while avoiding destroying the device—which is a change in precedence. In years past, stress conditions or actions that invoked or even hinted upon wear-out were to be avoided. The adage in the past was "one must be able to accelerate defects while avoiding the onset of wear-out." However this is becoming increasingly more difficult in the face of stretched system applications conditions; sub-10 nm oxides; NBTI; marginal margin (that is, array Vmin); hundreds of amps and Watts, miles of copper wire, and billions of interconnects.

RS&TM are best categorized by separating them into wafer applications and package (or module) applications and then further segregation into "*detection*" and "*acceleration*" techniques. This tiered structure will help to dilute the perennial argument between test and reliability regarding whether a field return is a test escape or an early life reliability failure.

Regardless of operational process step (wafer or package), acceleration techniques invariably must deal with potent power implications simply because acceleration requires temperature and/or voltage far in excess of application conditions—and leakage varies exponentially with both. The same is not true for detection techniques. In many instances, detection techniques employ conditions that reduce leakage (that is, VLV (very low voltage) or VLT (very low temperature), and in instances where detection requires application conditions that exacerbate leakage, those conditions typically do not approach the level of acceleration conditions.

BURN-IN REQUIREMENTS

Technical challenges for the burn-in process are driven by increasing device pin count, decreasing package pitch, increasing device functionality and operating frequencies, dramatically increasing leakage current, and eroding voltage/thermal acceleration. Several alternate techniques such as IDDQ, high voltage stress, and wafer mapping are being used to try to improve device reliability, since many reliability failure modes are proving to be resistant to burn-in.

Burn-in system technology must continue to drive down costs, in particular for high power devices. The minimum device core voltage continues to decrease. Scan requires very deep vectors for large memories, while high power requires individual device thermal and power management. The burn-in process (system/driver/burn-in board/socket) will be challenged to meet speeds of the newest technology devices without some form of internally generated clock. Devices without DFT are requiring increasing I/O. The growing need for KGD continues to drive efforts for wafer level burn-in, KGD carriers, or additional stress during probe.

Device power and signal requirements are driving burn-in boards toward higher board layer counts, smaller traces, less space for routing, more complex processes and materials, higher test costs, and board reliability issues. Tight pitch on future devices will require new cost-effective, innovative interfaces between the burn-in sockets and the burn-in boards.

Burn-in sockets are undergoing major design challenges as they must accommodate increasing contact count, decreasing pitch, higher currents, and higher frequencies. At the same time, sockets are a key component of an overall thermal solution designed to prevent high power devices from self-destructing. A major challenge for socket manufacturers is to maintain low costs and short lead times while providing the technology to meet these new demands. Horizontally actuated contact design will be displaced below 0.5 mm pitch ball grid array (BGA) by vertically actuated contacts as pin count increases and existing socket materials fall short of increased mechanical stress requirements. New designs and new materials will be required for higher current carrying capabilities. Socket design will need to accommodate looser packaging specs in areas such as warpage and package dimensions, while coping with increased package size, thinner/more fragile packages, and reduced/non-standard/mixed pitches. Contact design will need to provide greater strength without a loss of electrical/mechanical performance.

Approaches to burn-in include traditional unit level burn-in, system level burn-in, wafer level burn-in, and strip/array burn-in (Figure TST11). On high oapplicati ons, system level burn-in complements or oplaces traditional dvic level burn-in. Wafer level burn-in technology continues to be dvelopd, but has not been able yet to make significant inroads against traditional packaged level burn-in. The challenge here is to ate socketted burn-in and find ways to perform simultaneous multiple wafer level burn-in using scan/logic and memory BIST (MBIST). Strip/array burn-in is becoming more important as more packages are receiving massively parallel test in either strip or array format.

Table TST11 Burn-In Test Requirmnts

WAFER LEVEL BURN-IN

There is no standard definition of what constitutes wafer level burn-in (WLBI). Some vendors use the term "burn-in" to refer to the application of a simple DC stress that applies opposite voltage potential to the internal nodes of a DRAM. Some say that WLBI requires full wafer contact and th a pplication of high gh temperature over gh time to activate thermal defects, while also applying voltage stress with the device operating in "normal" mode. Some vendors enable the use of WLBI for low-end micro-controllers or SoC through DFT functions such as scan or BIST.

Key challengs are to quantify how to measure the effectiveness of these options and to dvelop standards that dfine WLBI and the methods that are used to confirm the effectiveness of this wafer level treatment. The challenge for DRAM in particular, as a device well suited for WLBI, is to provide a burn-in vironment for wafe rs that provides the same functionality, is as effective as package-level burn-in, and yet is no more costly. The concept is to leverage the time spent in burn-in by using the burn-in vironment as a massively parallel testing opportunity.

The need for WLBI is increasing. The infant mortality rate is getting worse due to transistor scaling effects and new processing technology / materials for dvices. Decreasing operating voltages and margins for dvices are rducing the ability to use voltag acceleration / voltag stress testing to guarantee roming a more significant need by the customers due to requirements for chip scale packaging and multi-chip modules. Decreased cycle time and the need for faster feedback of yield / defect information to the wafer fab can be assisted by moving burn-in earlier in the overall semiconductor process. Finally, dtection and rmova l of dctive dvices prior to the packaging process eliminates packaging scrap costs based on intrinsic dvice ds.

Figur TST11 The Production Process with WLBI Compared with Package Burn-in

PROBING TECHNOLOGY FOR WAFER LEVEL BURN-IN

Contactors for whole wafer contact include TPS probe and micro pogo-pin contactor. TPS probe consists of a substrate board, membrane with bumps, and PCR sheet, where the PCR sheet between two components absorbs the uneven height of bumps to achieve uniform and stable contact. Significant features of this system include the ability to concentrate pressurently at each bump top and the ability to achieve over 20,000 bumps contact with Al pads by control of

bump material and surface condition. Materials with coefficient of thermal expansion (CTE) similar to Si (such as glass and ceramics) are used for the substrate board to prevent CTE mismatch.

Table TST12 Test Handler and Prober Difficult Challenges

	Temperature control and temperature rise control due to high power densities during test

Coordination or test coverage is an ongoing challenge with heterogeneous designs. Solutions may involve the addition of test structures to assemblies or instantiated dies.

PROBE CARDS

Wafer probe technologies face complex electrical and mechanical challenges driven by product specifications, test implementation requirements, test productivity goals, and reduced test cost demands. Across the device spectrum, these challenges include: higher average power demands, higher frequency response (bandwidth), rising pin counts across tighter pitches and smaller pads/bumps, increasing switching currents (di/dt), alternative pad/bump metallurgies and increasing test parallelism. Research and development of new or improved probe technologies is required to meet these challenges to ensure that the basic probing requirement of ensuring reliable, sound, and cost-effective electrical contact to the device(s) under test is achieved. Recent developments in contactless probing technology may help address future density, speed and 3D requirements for probe.

Improvements in passive probe cards are at appear to be approaching mechanical and electrical limits for increased functionality. Intelligent Probe Cards are potentially capable of solving problems of both parallelism and speed. Advances in materials MEMS and heterogeneous integration suggest that intelligent probe cards can be made economically with parallelism and performance to match DUT technologies. There are some extant commercial offerings of intelligent probe cards for use in 3D packaging test.

The tables contained in this section derive trends based on product families similar to the layout of the Test Technology Requirements section above.

TRENDS AFFECTING PROBE CARD TECHNOLOGIES

Along with addressing the key challenges listed below, research and development is urgently required to bring to the market cost-effective probe technologies directed at trends in product offerings and the testing environment.

The continuing volume growth of bumped devices, often with I/O in area arrays, points to the escalating demand for "vertical" style probe card technologies, with a rising need in multi-DUT configurations as well. Multi-row wirebond also supports this vertical style need and is particularly challenging due to tighter pitches.

Some microprocessor products and high end ASIC devices are driving power levels to 500 Watts and 1000 Watts with associated current/probe and thermal issues. Current/needle is also an issue for cantilever and MEMS technology as wire bond devices move into higher technology silicon.

Manufacturing test of devices has moved to parallel test. For some product groups (e.g., memory), wafer probe technologies are available that handle parallel testing of 512 and more devices. Probe technologies capable of full wafer contacting are in use already for 200 mm and300 mm wafers. Increasing the contacts/DUT for these massively parallel probes is the next challenge.

Innovation in test is required for effective use of new interconnect technologies such as TSV or proximity communications.

Table TST15 Probing Difficult Challenges

- *Probe technologies to support peripheral fine pitch probe of 23 µm peripheral staggered pad probes at effective pitches of 20/40, and fine pitch (45 µm) for dual row, non-staggered probing on all four die sides.*

This section explores the challenges of probe technologies including those that are independent of the devices being probed. These include the resulting behavior of the probe when/after contacting the wafer, the design of the probe card to realize the productivity benefits of probing multiple die at the same time and the environment that the probe card is expected to operate within.

PITCH AND INTERCONNECT DEFORMATION

I/O density requirements are driving pad/bump sizes to ever-smaller sizes. It is well known that on the leading edge, wirebond pad pitches are under 30 µm (with resulting pad sizes naturally less than that). It is a formidable challenge for traditional probe technologies to scale down continually since with this scaling comes a parallel scaling-down of the permissible probe mark.

The use of cantilever probe cards for probing wirebond technologies, though still today's leading solution, is seen to be reaching practical limits in pitch and scrub within the nearer term horizon. Thus the newer emerging technologies, many using "semiconductor-like" processes (e.g., MEM and membrane structures) offer solutions for reduced pitch scrub requirements.

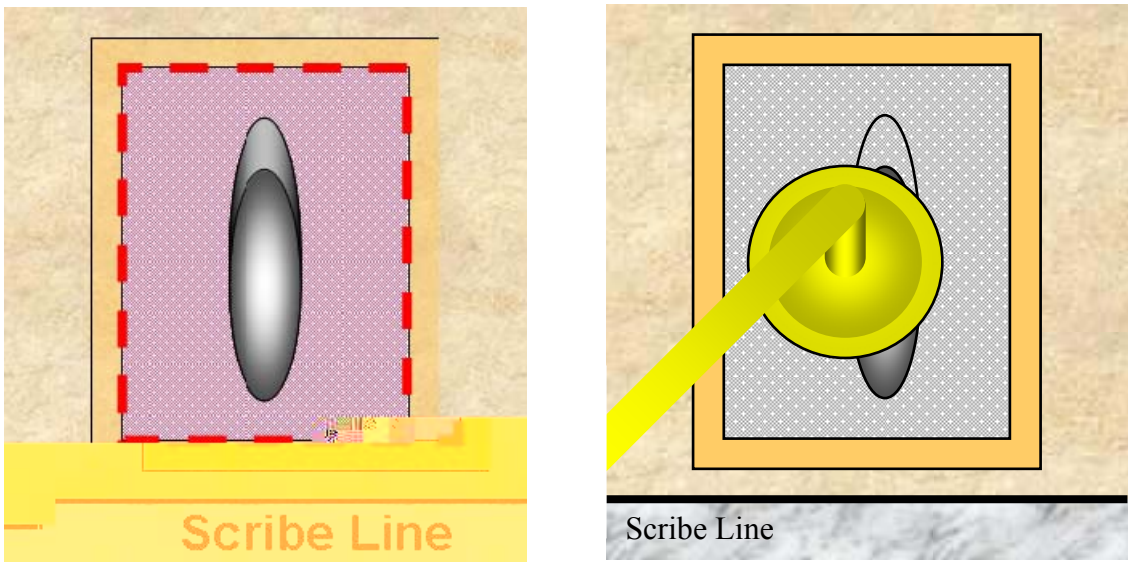


Figure TST12 Probing and Wirebond Contacting a Bond Pad

Area array solder bumps are seeing growing application and driving the commensurate need/demand for vertical probing technologies. As the pitch/bump dimensions get smaller, curre

There appears to be growth in the current carrying capability of individual probes contacts. At the same time the aggregate total current across the DUT is expected to rise w

Table TST17 contains the test socket technology requirements. The requirements have been divided into contacting NAND, DRAM, and SoC devices that are contained in TSOP

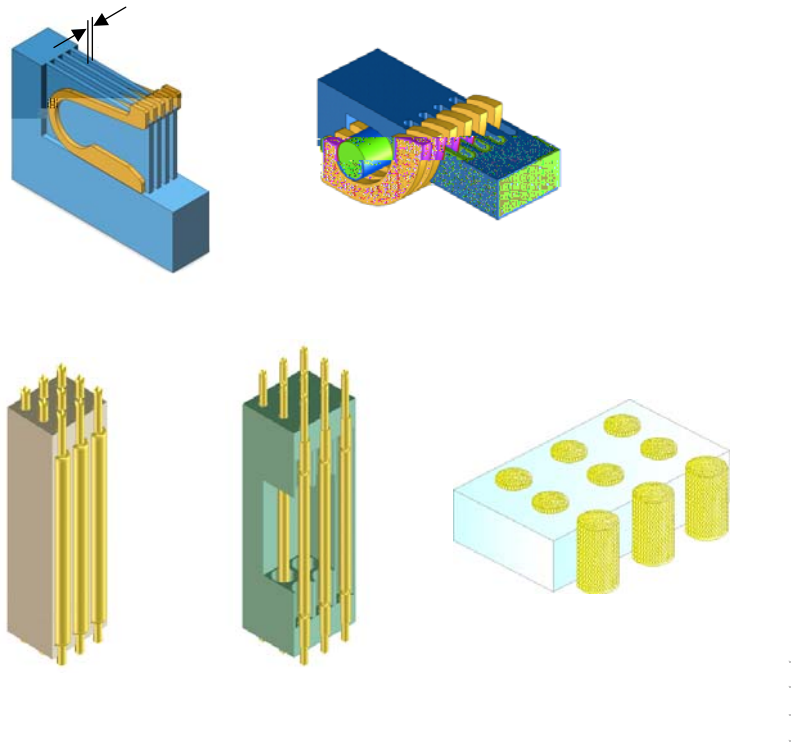


Figure TST13 Contactor Types

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 on over the roadmap are expected to exceed 20 GHz, which will greatly challenge impedance matching and the 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 insure 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 raised mechanical requirements.

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 the 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-dir.), 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.

Table TST17 Test Socket Technology Requirements

SPECIALTY DEVICES

The test roadmap is not all inclusive so it does not contain test requirements for all devices. Many of the test requirements for some omitted devices fall within the bounds specified for devices within this roadmap. Other devices stretch the bounds specified in this chapter and need to be mentioned for completeness. Devices included in the specialty device section are high volume devices that are generally contained in and driven by the requirements of mobile communication and computing. The intent of this section is to document the challenges of specialty devices. For 2009, LCD display drivers, imaging devices and MEMS accelerometers are featured in this section.

LCD display drivers are unique because of die form factor, which can have a 10:1 aspect ratio (or greater), and by the high number of very narrow pads requiring contact for test. In 2009, LCD display drivers with pads that are 8 µm ×150µm and pitches of 20 µm are planned to be in production. The 8 µm pad width in use in 2009 is less than half of the 25 µm 2022 pad width specified for commodity devices and the pad aspect ratio of 15:1 also exceeds the worst case pad aspect ratio for commodity devices of 1.3:1. Massive parallelism is not needed for these devices due to short test time, but a repeatable economic contacting solution is needed for the probe environment.

Imaging devices are required for every digital camera and have become a standard feature in cell phones and mobile computing devices. Digital cameras with 12 mega pixels of resolution are typical consumer devices in 2009 and will increase in resolution in the future. Automotive industry applications and surveillance need fast frame rates of 60 fps in 2009. Each pixel of the image sensor has a micro lens (Figure TST14) to increase the light intensity and improve the signal to noise ratio and must be tested for consistency, sharpness, and contrast across its detection spectrum. Current production solutions are generally proprietary and composed of a pupil projection system used to project images to the sensor under test over a range of various angles. Back-side illumination BSI is another approach to increase incident light intensity and fill factor of sensor cell of each pixel. Detected image data is processed to ensure picture quality. Image sensors convert the analog information of the picture into digital output, so the test result is non-deterministic. Image sensors coupled with image processing digital logic in a single chip pose significant test challenges when assembled into module form, since typical test access is very limited for module testing, and when embedded in a system such as a cell phone, simple functional tests may be all that is possible.

MEMS accelerometers are increasingly being incorporated into personal electronic devices and automotive application. Some smart phones and personal digital assistants contain accelerometers for user interface control. The cost of accelerometer devices dropped over 10% in 2009, but reduction of test cost is challenged by the difficulty of achieving high parallelism package test and difficulty in enabling probe test data to gain yield improvement.

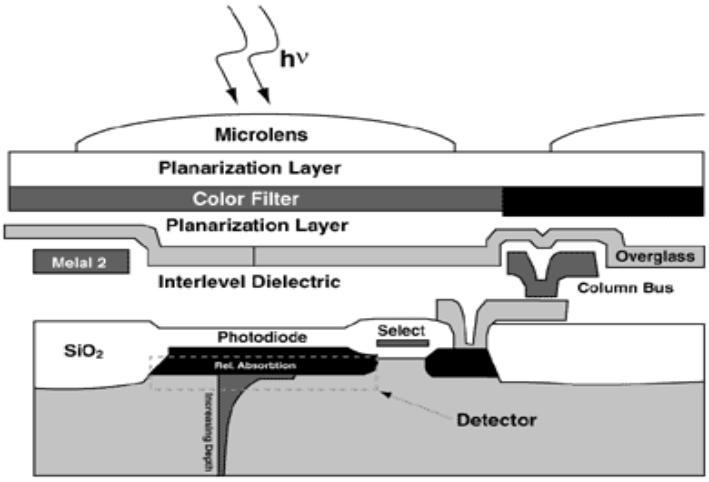


Image sensor structure cross section

Figure TST14 Image Sensor Cell