INTERNATIONAL TECHNOLOGY ROADMAP FOR SEMICONDUCTORS

2005 EDITION

TEST AND TEST EQUIPMENT

THE ITRS IS DEVISED AND INTENDED FOR TECHNOLOGY ASSESSMENT ONLY AND IS WITHOUT REGARD TO ANY COMMERCIAL CONSIDERATIONS PERTAINING TO INDIVIDUAL PRODUCTS OR EQUIPMENT.

TABLE OF CONTENTS

Introduction and Scope of the 2005 edition 1
Key Drivers, Difficult Challenges, and Future Opportunities 1
Key Drivers 3
Device Trends 3
Increasing Test Process Complexities 4
Continued Economic Scaling of Test 4
Difficult Challenges (in priority order) 5
(1) Test For Yield Learning (summary) 5
(2) Screening For Reliability 6
(3) Increasing Systemic Defects 6
(4) Potential Yield Losses 7
Future Opportunities 8
Test Program Automation (not ATPG) 8
Simulation and Modeling 8
Convergence of Test and System Reliability Solutions 8
Test For Yield Learning Focus Topic 9
Physical Failure Analysis (PFA) 9
Software-Based Diagnosis and Signature Analysis 9
Manufacturing Test Cost Focus Topic 10
Base Cost Trend 13
Channel Cost Trend 13
Power Cost Trend 13
Interface Cost Trend 13
Multi-site Trend 14
Other Cost Trends 17
Important Areas of Concern 17
Test Technology Requirements 17
Introduction 17
System Integration—SOC and SIP Test Challenges and Implications 18
System on a Chip 19
System in a Package 23
Logic 25
High Volume Microprocessor Trends Drivers 25
System Trends Drivers 25
DFT Trends Drivers 26
High Volume Microprocessor Test Requirements 26
Important Areas of Concern 28
High Speed Input/Output Interface 28
Important Areas of Concern 29
Memory 31
Commodity Dram Testing 31
Commodity Flash Testing 32
Embedded Dram and Flash Testing 35
Analog and Mixed-signal 36

Important Areas of Concern 37
Radio Frequency 40
Important Areas of Concern 40
Reliability Technology Requirements 40
Burn-In Requirements..42
Wafer Level Burn-in..42
Test Handler and Prober Technology Requirements 47
Device Interface Technology Requirements..52
Probe Cards..52
Trends Affecting Probe Card Technologies 52
Probe Card Technology Requirements..53
Pitch and Interconnect Deformation..54
Multi-DUT 55
Electrical Performance 55
Thermal Performance..55
Unit Cost and Cost of Ownership..55
Lead-Time 55
Cleaning 56
Test Sockets 60
Electrical Requirements 60
Mechanical Requirements..61

LIST OF FIGURES

Figure 28 Test Cost Drivers ... 11
Figure 30 Channel Cost Range Trends for Product Segments ... 12
Figure 31 Test Cell Cost / Unit vs. Interface Cost Trend ... 13
Figure 32 Organization of Core Semiconductor Technologies from System Integration and Applications Points of View ... 18
Figure 33 The Production Process with WLBI Compared with Package Burn-in ... 43

LIST OF TABLES

Table 22 Summary of Key Test Drivers, Challenges, and Opportunities ... 2
Table 23a Multi-site Wafer Test (Package Test) for Product Segments—Near-term Years ... 15
Table 23b Multi-site Wafer Test (Package Test) for Product Segments—Long-term Years ... 16
Table 24 SOC Model ... 19
Table 25a System on Chip Test Requirements—Near-term Years ... 21
Table 25b System on Chip Test Requirements—Long-term Years ... 22
Table 26a Logic Test Requirements—Near-term Years ... 27
Table 26b Logic Test Requirements—Long-term Years ... 28
Table 27a Commodity DRAM Test Requirements—Near-term Years ... 32
Table 27b Commodity DRAM Test Requirements—Long-term Years ... 32
Table 28a Commodity Flash Memory Test Requirements—Near-term Years ... 33
Table 28b Commodity Flash Memory Test Requirements—Long-term Years ... 34
Table 29a Embedded Memory (DRAM and Flash) Test Requirements—Near-term Years ... 35
Table 29b Embedded Memory (DRAM and Flash) Test Requirements—Long-term Years ... 36
Table 30a Mixed-signal Test Requirements—Near-term Years ... 38
Table 30b Mixed-signal Test Requirements—Long-term Years ... 39
Table 31a Burn-in Requirements—Near-term Years ... 45
Table 31b Burn-in Requirements—Long-term Years ... 46
Table 32a Handler (Memory—Pick and Place) Requirements—Near-term Years ... 47
Table 32b Handler (Memory—Pick and Place) Requirements—Long-term Years ... 48
Table 33a Handler (Logic—Pick and Place) Requirements—Near-term Years ... 49
Table 33b Handler (Logic—Pick and Place) Requirements—Long-term Years ... 49
Table 34a Handler (Network and Communications—Pick and Place)—Near-term Years ... 50
Table 34b Handler (Network and Communications—Pick and Place)—Long-term Years ... 50
Table 35a Prober (Logic MPU—Pick and Place) Requirements—Near-term Years ... 51
Table 35b Prober (Logic MPU—Pick and Place) Requirements—Long-term Years ... 51
Table 36 Test Handler and Prober Difficult Challenges ... 52
Table 37 Probe Card Difficult Challenges—Near-term Years ... 54
Table 38a Wafer Probe Technology Requirements—Near-term Years ... 57
Table 38b Wafer Probe Technology Requirements—Long-term Years ... 59

TEST AND TEST EQUIPMENT

INTRODUCTION AND SCOPE OF THE 2005 EDITION

The 2005 Test Roadmap has undergone substantial evolution since the 2003 edition. Most significant is the reorganization of the Test Technology Requirements section to better reflect the reality of increasing integration of previously disparate chip designs. Looking forward over the ITRS horizon, each device under test (DUT) could be considered a system-on-a-chip (SOC) or system-in-package (SIP) containing one or more of the following “cores”: logic, I/O, memory, analog, RF, etc., each with unique test requirements. The revamped Test Technology Requirements section opens with an overview of these SOC/SIP test challenges including the introduction of a generic SOC model and continues with sections dedicated to the test requirements and challenges for each type of core.

There have also been some notable additions, deletions, and consolidations from previous years. The microprocessor, ASIC and microcontroller test requirements sections have been consolidated into a single logic test requirements section. New sections this year are radio frequency (RF) Testing and Test Sockets. Also, the design for test (DFT) Tester Requirements section has been removed from this year’s edition. This is not to imply any diminished importance for DFT or low-cost test solutions for devices that depend heavily on DFT to meet high volume manufacturing test requirements. On the contrary, DFT advancements and low-cost test solutions will be essential to hold the line on test costs. The dilemma is these test solutions must be more flexible than ever to accommodate the aforementioned integration trends and the emergence of new cores such as RF that are driving new test capabilities. For this reason, the DFT implications and test challenges will be addressed throughout the chapter and on a core technology basis

The previous revision in 2003 noted the beginning of a significant shift from traditional test equipment architectures to “universal slot” architectures with high levels of test instrument encapsulation and modularity. This trend continues to unfold in various forms. Most major test equipment suppliers are looking to leverage design modularity wherever possible, with implementations ranging from modular architectures that span multiple test equipment architectures within a single test equipment supplier to open architectures that aim to enable the ability to mix and match test instruments from multiple suppliers into a single tester hardware and software environment. Given this generic trend, semiconductor manufacturers will have more options and greater flexibility than ever before to tailor their test equipment solutions to their testing needs and this provides even more impetus for reorganizing the test requirements in this year’s roadmap revision around different semiconductor technology cores that can be integrated together in a myriad of different ways as opposed to a handful of market segments.

This document represents significant contributions from a wide cross-section of the industry as noted in the acknowledgements; however, the Test Technology Working Group is always looking for additional participation—please contact the working group chair if you have interest in participating!

KEY DRIVERS, DIFFICULT CHALLENGES, AND FUTURE OPPORTUNITIES

Beginning with this 2005 edition of the ITRS Test Chapter, the working group members re-organized this important section of the test chapter. The key change is to split the previous Difficult Challenges section into Key Drivers and Difficult Challenges, as well as add a Future Opportunities section. This split will distinguish the drivers, that is, primary boundary conditions that define the scope of solutions for upcoming manufacturing test for semiconductor components, from key technical and business challenges. At a high level, these boundary conditions actually represent expectations or even requirements of the test process, while the challenges represent current and upcoming key roadblocks, strategic inflection points, and opportunities for the future

For many years, the mission of semiconductor manufacturing test has been described as “screening defects” and to a lesser extent or within certain business segments “speed binning” or “speed classification.” It is interesting to note that some of the most important test challenges are now actually centered on some of the more subtle historical missions of 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 entire semiconductor business, both in terms of enabling the cadence or timely delivery of future processes and cost effective products, but also in terms of meeting customer expectations for reliability.

Within the Difficult Challenges section, the challenges are listed in order of perceived importance or priority, for example, test for yield learning followed by screening for reliability followed by increasing systemic defects. In contrast, there is no specific intent in the ordering of the Key Drivers, whereas they are all boundary conditions or requirements that the semiconductor test solutions must meet. Table 22 summarizes all of the key test drivers, challenges and opportunities.

Table 22 Summary of Key Test Drivers, Challenges, and Opportunities

KEY DRIVERS (NOT IN ANY PARTICULAR ORDER)	
Device trends	• Increasing device interface bandwidth (both number of signals and signal data rates) • Increasing device integration (SOC, SIP, MCP, 3D packaging) • Integration of emerging and non-digital CMOS technologies (RF, Analog, Optical, MEMs) • Package form factor and electrical / mechanical characteristics • Device characteristics beyond one sided deterministic stimulus/response model
Increasing test process complexity	• Increased device customization and line item complexity during the test process • Increasing "distributed test" to maintain cost scaling • Increased data feedback for tuning manufacturing • Higher order dimensionality of test conditions (e.g., adding multi-power, multi-voltage, multi-freq topologies to single valued T, V, freq)
Continued economic scaling of test	• Physical limits of further test parallelism • Managing (logic) test data volume • Effective limit for speed difference of HVM ATE versus DUT • Acceptable increases for interface hardware and (test) socket costs • Trade-off between the cost of test and the cost of quality
DIFFICULT CHALLENGES (IN ORDER OF PRIORITY)	
Test for yield learning	• Critically essential for fab process and device learning below optical device dimensions
Screening for reliability	• Increasing implementation challenges and efficacies of burn-in, IDDQ, and Vstress • Erratic, non deterministic, and intermittent device behavior
Increasing systemic defects	• Testing for local non-uniformities, not just hard defects • Detecting symptoms and effects of line width variations, finite dopant distributions, systemic process defects
Potential yield losses	• Tester inaccuracies (timing, voltage, current, temperature control, etc) • Overtesting (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
FUTURE OPPORTUNITIES (NOT IN ANY ORDER)	
Test program automation (not ATPG)	Automation of generation entire test programs for ATEs
Simulation and modeling	Simulation and modeling of test interface hardware and instrumentation seamlessly integrated to the device design process
Convergence of test and system reliability solutions	Re-use and fungability of solutions between test (DFT), device, and system reliability (error detection, reporting, correction)

ATE—automatic test equipment ATPG—automatic test pattern generation BIST—built-in self test HVM—high volume manufacturing MCP—multi-chip packaging MEMs—micro-electromechanical systems

KEY DRIVERS

As previously mentioned, the Key Drivers in the Test chapter are considered *boundary conditions* within which the semiconductor test function must operate within and still deliver acceptable component quality and reliability at continued test cost scaling.

DEVICE TRENDS

- *Device Interface Bandwidth*—In the 2001 edition of the ITRS Roadmap, a change in direction was denoted to deliver increasing system performance with rapidly increasing (compared to previous trends) component I/O speeds. In the 2003 and 2005 editions, the Test working group noted that while the speeds are not climbing as rapidly as the 2001 predictions, the penetration of Gb/s interfaces is permeating throughout a wider spectrum of the semiconductor component market segments (memories, CPUs, chipsets, etc.). High speed serial and differential I/O protocols will continue to rise in speed and continue to force innovation in the DFT and HVM test processes to maintain continued cost scaling

- *Increasing Device Integration*—SOC and SIP integration continue to rise in prevalence throughout a number of business segments. Increased device integration forces a re-integration of test solutions to maintain scaling of test costs and product quality. The optimized test solutions for stand-alone RAMs, cores, and other blocks typically do not scale linearly without modification, additional DFT, or new partitioning in the integrated device test solutions. In particular, additional DFT in die or even in-package may be required to provide access to and testing of embedded blocks or cores, or additional distribution or re-integration of the HVM test instrumentation may be required. Techniques for known good die (KGD) that provide high quality dice for multi-die packaging also become very important and an essential part of the test techniques and cost trade-offs.

- *Integration of Emerging and Non-Digital CMOS Technologies*—In the recent past, mixed signal device circuits have been increasingly delivered on die integrated to digital CMOS, posing key challenges for ATE, instrumentation, and test manufacturing flow integration choices. It also poses new challenges and opportunities for DFT innovation, where it had not been previously prevalent (as it had in digital logic and memories). Integration of RF circuits is already starting to make a similar rise, as are more radically different types of semiconductor devices, e.g., MEMs which are already showing up in integrated forms, and optical, which are just around the corner from integration to larger CMOS digital die. As in previous types of integration, the test mission for these devices revolve around access to potentially embedded blocks and integration of radically different test methods into a cost effect manufacturing process.

- *Package Form Factor and Electro-Mechanical and Thermal Characteristics*—The envelope of package form factors is pushing both outward and inward, with higher complexity form factors in MCP and SIP for multi-die on the high end, but also chip scale packaging pushing towards the small end for systems targeted for the smallest platform form factors (e.g., handhelds). In addition, multi-functionality of die packaging with inclusion of heat sinks and spreaders, and perhaps even voltage regulation and power management functionalities, appears to be on the increase. With increasing power a major concern, form factor thermal transfer characteristics and uniformities (thermal gradient coefficient-junction to ambient, junction to case [theta-JA, theta-JC], etc.) become even more critical to the test process. These expansions of the form factor technology envelopes necessitate improvements and delivery of key test sub-systems: handling for packaging test, Ohmic contacting technologies, and test sockets (the later described later in the chapter for the first time).

- *Device Characteristics Beyond One-Sided Stimulus Response Model*—The history of semiconductor testing and ATE architecture and usage is built upon a foundation of deterministic device behavior. In addition, digital CMOS testing has historically been based on simplifying the environmental topologies of V_{cc}, temperature, and frequency, so that worst case (one-sided), one iteration tests guaranteed the performance over a wider and continuous range of environment topologies. When these topologies were only two or even three variables, they were manageable, and fairly easy to characterize, and to optimize for HVM test. However, a number of device characteristics threaten the dimensionality of this one-sided test paradigm, threatening a non-linear effect on non-recurring engineering (NRE) development costs for test as well as potentially limiting continued test cost scaling for HVM. Device features such as extended self-repair and correction, and built-in variable, and multi-power management modes all potentially add tremendous complexities to the device test conditions contour. In addition, from the device design, architecture, and behavior front, their looms the increasing future possibility of non-deterministic device behaviors (e.g., asynchronous logic architectures) where correct device behavior (from end user system standpoint) from die to die under similar

conditions is 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 set of digital logic 1s and 0s delivered on the same synchronization, would not suffice to test such behaviors. While these device architectures and behaviors are not widely prevalent today, they do pose a major paradigm shift, or innovation challenge for the semiconductor test process in the not too distant future.

INCREASING TEST PROCESS COMPLEXITIES

- *Increased Device Customization and Line Item Complexity (during the test process)*—Increasingly, the test process is expanding to include not only the metrologies surrounding the question "is this a good die?" but also actual process steps that modify, differentiate, or customize specific die. Examples of these include memory block (and other) redundancy/repair, laser marks, burning of on die fuses or read-only memories (ROMs), or other programming of product features. In addition, some business segments are seeing substantial increases in product line items derived from the same die base, with differentiation and physical segregation as part of the manufacturing test flow increasing. All of this places increasing demand on the manufacturing test process and expands equipment (e.g., handling, fuse blowing, etc.) and factory automation and integration infrastructures.

- *Increased "Distributed" Test to Maintain Cost Scaling*—Distribution of test content and manufacturing instrumentation and equipment platform requirements enabled by a wider application of DFT techniques will continue to be a widespread industrial test process integration scheme used to optimize test manufacturing capital expenditures, operational test cost, and outgoing product quality. Distribution and differentiation of different types of machines for design verification versus manufacturing test will continue. Typical vectors for this differentiation will center around the speed, accuracy, number and types of channels on verification versus manufacturing tools. Integration of a wider array of digital CMOS circuits, mixed signal and even optical and or mechanical silicon together in future products will continue to make distributed test solutions attractive rather than to increase the platform complexity of ATEs back towards a "do all" platform, at least for manufacturing test. Hand in hand with this will be increases in the complexity (number of steps and interactions) within the manufacturing test process flow itself.

- *Increased Test Data Feedback to Tune Manufacturing*—Test data usage for purposes beyond identifying whether a given die is good or defective has become essential for several reasons and drives a need for expanded, revamped, and better integrated test data systems and infrastructure. In one sense the output of the component test process is data, i.e., the results of the various applied metrologies of the test process across the product dice 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,

- *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 adds 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 for manufacturing test for test costs by the product of the applied test content and the number of iterations of its applications 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 will be reaching non-linear limits in the next one to two 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 be developed that enable further increases in DUT parallelism beyond what is currently envisioned

- *Managing (Digital Logic) Test Data Volume*—Unconstrained, 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 (ATEs) 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, e.g., 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., SOCs, SIPs) 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 ATEs versus HVM manufacturing ATEs 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 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.

- *Acceptable Increases for 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 driving this, such as higher speed (GB/s) and more complex DUT I/O protocols, increased DUT parallelism, higher signal and power pin counts, and increased power delivery and signal channel fidelity requirements. While this might be an alarming trend by itself, this trend needs to be considered in the context of the overall test process cost, and whether or not it enables a continued overall economic scaling of manufacturing test. Keeping these proportional increases of interface hardware costs within an acceptable range may also be dependent on technology drivers and boundary conditions, e.g., the continued extension of the use of FR_4 materials within an acceptable numbers of layers as the primary materials platform for such hardware

DIFFICULT CHALLENGES (IN PRIORITY ORDER)

(1) TEST FOR YIELD LEARNING (SUMMARY)

Test's peripheral role as a feedback loop for understanding underlying defect mechanisms, process marginalities, and as an enabler for rapid fab process yield learning and improvement has traditionally been considered a secondary role to screening hard defects. With the increasing reduction in feature (and defect) sizes well below optical wavelengths, rapidly increasing failure analysis throughput times, reduction in failure analysis efficacy, and approaching practical physical limits to other physical techniques (pica, laser probes), the industry is reaching a strategic inflection point for the semiconductor business where the criticality of DFT and test enabled diagnostics and yield learning becomes paramount. In other words, the yield learning rates of the past per process generations are not sustainable by historical fault isolation and failure analysis methods per se. Rather, they need to be augmented by more universal deployment of enabling on die circuitry (DFT, etc.) across and throughout products, as well as improvements in the on-die instrumentation and diagnostic software tools with respect to fault isolation specificity. Where it may have been sufficient to isolate to failing bit in an array or failing gate in logic in the past, there is a real business need to be able to isolate electrically to at least the failing transistor or interconnect in the future, or suffer the economic consequences of reduced yield improvement

learning rates on new process technologies, perhaps even to the point of lowering the asymptotic maximum achievable yields per die size on future processes. Given the paramount importance of this "mission" for component test, the Test team has not only summarized its relevance here as the top challenge for test, but also included a detailed section on this topic in the early part of the test chapter.

(2) SCREENING FOR RELIABILITY

A lesser-publicized mission for semiconductor test beyond the primary "screening defects" or telling a good part from a bad one, has been to screen out infant mortality reliability 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 reduced product operationally margin (collapsing V_{dd}/V_{cc} 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 latest defects) leakage levels are much higher than under normal application conditions. In addition, reducing product V_{cc} 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 ASICs to SOCs to commodity memories, has been greatly challenged by the very same trend in proportional increases in background leakage, with 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 I_{cc}/I_{dd} combinations and comparisons about results of multiple tests to maintain efficacy against latest 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 reporting a real drop off in effectiveness is expected in the next one to two technology generations for IDDQ. Similarly, voltage stress, or applying V_{cc}/V_{dd} and patterns well above nominal range to accelerate latent defects, also has been losing effectiveness, notably as the differential between V_{stress} and $V_{nominal}$ has continued to get smaller at each generation. In the long run, new techniques for providing this reliability screening function will likely be needed, at least 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.

(3) INCREASING SYSTEMIC DEFECTS

The industry faces new manufacturing-imperfection-related test and yield-learning challenges that result from changing processing technology, changing circuit sensitivities and modeling limitations.

1. Process technology advancements are changing the population of physical defects that affect circuit functionality. For example, smaller or higher aspect-ratio vias are more susceptible to incomplete etch, which may lead to a greater prevalence of resistive vias. Similarly, the change from subtractive aluminum processes to damascene Cu may cause a decrease in particle-related blocked-etch metal shorts in favor of an increase in particle-related blocked-deposition metal opens. In addition, the introduction with Cu interconnects of a liner metal that can remain intact when a line open occurs may increase the occurrence of resistive line opens. The introduction of low-κ dielectrics may lead to an increase in the occurrence of possibly latent, resistive bridges and smaller transistors may increase the importance of degradation mechanisms, such as negative bias temperature instability (NBTI).
2. Increasing design-process interactions may increase the prevalence of systematic defects. Such defects may occur only in certain circuit/layout configurations, for example because of such issues as pattern density, pattern proximity

and imperfections in optical proximity correction (OPC) algorithms. Although they are systematic they may appear random because of their rarity and the complexity of the conditions required for their occurrence.

3. Changing circuit sensitivities are likely to make defects that were benign in the past become killer defects in the future. For example, shorter clock cycles mean defects that cause 10s or 100s of picoseconds of delay are more likely to cause circuit failures. Furthermore power-optimized and/or synthesized designs will result in fewer paths with significant timing margin, which implies that random delay-causing defects will be more likely to cause failure. Similarly increasing noise effects, such as crosstalk and power/ground (GND) bounce, decrease noise and timing margins and again increase circuit susceptibility to defects.
4. Finally, modeling complexity threatens the ability of electronic design automation (EDA)/design to ensure the circuit's functionality under all process conditions. That decreased ability combined with increasing process variability may increasingly result in circuits with parametric failure modes reaching the test floor.

All aspects of the test process, including fault modeling, test generation, test coverage evaluation, DFT solutions, test application and diagnosis, must handle these realistic and changing populations of manufacturing imperfections. New classes of defects may not be detectable using traditional test methods for some products; for example, small delay defects for ASICs. Promising strategies include realistic defect-based fault modeling, out-of-spec testing such as low-V_{dd} or temperature, defect-oriented test generation and diagnostics, such as those based on layout-based identification of defect sensitivities, statistical methods, adaptive test, and techniques to allow continued use of IDDQ and very-low-voltage testing. To ensure these and other techniques are accurately targeted and effective, high-fidelity information about the occurrence and properties of the population of manufacturing imperfections will be needed, therefore methods for understanding and characterizing the defects must be developed. Test-structure-based methods alone may not be sufficient. Methods based on product-level test results must continue to be developed as well.

(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.)
- Overtesting (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 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 overtesting 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 parametric drift during test require guardbands that can be large enough to affect yield. Increasing DUT parametric variation with advanced lithography is pushing the need for more effective BIST or other alternative test strategies such as on-product parametric measurement circuits.

Similar yield loss issues have recently surfaced with inaccurate launch-capture delay tests. Delay path measurement errors of fifteen picoseconds have been observed in delay path measurements—this is 10% of recently announced internal clock periods, indicating a possibility of either yield loss or test escapes, as there is as yet no known way of adding margin testing to delay path measurement. Delay path measurements also contribute to yield loss if inappropriate delay paths are measured. Tools must be developed to avoid yield loss due to measurement of false paths.

IDDQ failures are known to indicate a variety of interconnection and other defects. If, however, IDDQ-only failures are not severe, experience has shown that only a few percent change with burn-in or otherwise prove to be reliability issues. This can be reckoned as yield loss in non-critical applications.

Statistical post-processing is becoming a significant tool for identifying potential failures. However, since it consists of discarding "good die in bad neighborhoods," for example, it is inherently a potential yield loss mechanism.

In a larger sense, however, the concept of yield loss may be understood to include discarding large numbers of properly fabricated devices or interconnect because a small number of improperly fabricated devices are faulty. Yield loss of this sort is currently mitigated in high-density memories by providing redundant rows and columns; similar strategies have been used successfully at system levels using higher levels of reconfigurability and redundancy. Continued development of reconfigurable circuits for defect avoidance or repair focusing on repair techniques that do not cause collateral damage is needed.

Research and development regarding ways to incorporate on-line testing and repairs has a rich history in systems technology. Migrating the results of these efforts into ICs has the potential of reducing yield loss reckoned in this larger sense. More efforts in pursuit of robust design, online testing, and *in situ* repair technology are required.

FUTURE OPPORTUNITIES

TEST PROGRAM AUTOMATION (NOT ATPG)

While the EDA industry has been providing a range of capabilities around scan DFT and more recently including scan compression and even some DFT synthesis capabilities, for example, for embedded memories, there is an area where the semiconductor suppliers software NRE costs present a productivity and automation opportunity: generation of actual test programs. Across the industry today, the product or test engineering communities spend much of their time creating new customer software code bases, that is, unique and custom individual production test programs. There are a lot of innovative software products for different pieces of this, provided primarily by ATE suppliers along with the ATEs, as well as a lot of additional software tooling down in house by the semiconductor companies themselves. There are even historical or new test program data structure standards (test data format (tdf) or standard test interface language (STIL)) that provide some common structural opportunities. However, the basic fact remains is there are no widely available turn-key software products that produce and validate whole main code test programs and yet most of the 1000s of these in use across the industry today do basically the same kinds of things on similar tools across companies.

SIMULATION AND MODELING

Extending the design phase simulation systems and modeling to cover not only die packaging, but also test interface hardware and perhaps even the ATE instrumentation circuits themselves (both parametric and logical) will enable shorter product and manufacturing test validation cycles. This extention will also enable higher reliability pre-silicon verification of the DUT DFT scheme integrated to the manufacturing test instrumentation. Providing such a seamlessly integrated simulation and modeling environment in the pre-silicon design phase would better enable not only accuracy of product performance, but also help tune the test process, and reduce design turns due to minor discrepancies in DFT and test interface modes and hardware.

CONVERGENCE OF TEST AND SYSTEM RELIABILITY SOLUTIONS

Some of the most widely proliferated DFT techniques (such as scan) were actually originally motivated by the need to provide better system reliability for high-end enterprise computing systems. The same scan hardware was re-usable for manufacturing test screening purposes and has proliferated throughout the industry. In the future, there are a number of physical and circuit phenomena that are challenging lower level intra-die reliability such as alpha particle and cosmic ray 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 FOR YIELD LEARNING FOCUS TOPIC

In addition to the normal sorting function, test provides an essential feedback loop for understanding the failure mechanisms inherent in deep submicron fabrication processes. Test is the main source of feedback regarding killer defects, unacceptable parametric variations, and design-process interactions. Complex failures due to design-process interactions, combined with decreasing design margins that push processing technology to its limit, make feedback from test increasingly important to reaching yield-learning goals. Test must increase its ability to support cost-effective defect isolation, process measurements, and failure root cause determination.

PHYSICAL FAILURE ANALYSIS (PFA)

CMOS technology migration will severely challenge the traditional failure analysis process. The traditional process, comprising the steps of fault localization, deprocessing, and physical characterization/inspection, will increasingly be too slow and difficult for routine analysis. Instead, with development in key areas such as software-based defect localization and signature analysis as described below, destructive failure analysis will move closer to a sampling/verification role. To play this continually important role while keeping pace with technology, improvements and breakthroughs to existing tools/techniques are required—especially better use of tester-based diagnostic and signature analysis tools to direct them. PFA capability gaps are detailed in the following prioritized list.

1. *Circuit probing*—Micro-positioner probing, in-chamber, and atomic-force-microscope-based will be required at each roadmap generation to probe minimum sized transistors (4 probe points) and SRAM cells (5 or more points). This requirement is driven by the need to validate increasingly subtle and/or invisible defects and to characterize individual transistor parameters.
2. *Localization*—There are increasing limitations posed by shrinking dimensions and increasing depth of the metal stack. Failure analysis faces the dilemma that critical defects can be so small that they can only be seen by transmission electron microscopy (TEM) but TEM requires a very thin sample section containing the critical defect that cannot be obtained because the defect cannot be seen or localized with enough precision. As killing defect sizes shrink and the use of inspection tools such as TEM that require significantly better location of the defect increases, it become more critical to be able to identify a failing area to a much smaller volume for analysis.
3. *Computer-aided design (CAD)/EDA tools*—Failure analysis is increasingly reliant on scan methodology to either leads directly to a fault location or to complement another diagnostic tool to improve precision. Significant improvements are required to facilitate the use of both CAD and EDA tools in the failure analysis environment.
4. *Optical Diagnostic tools, particularly for timing information*—The shift in photon energies to longer wavelengths with lower supply voltages will require improvement and/or development in picosecond imaging circuit analysis (PICA)/time resolved photoemission (TRE) detector technology and to solid immersion lens (SIL) optics. Invasive laser-based probes like laser voltage probing (LVP), especially in SOI devices, are also a concern, and the inherent spatial resolution of just under 1 µm is inadequate for distinguishing signals coming from individual transistors. Unless improvements are made in these areas, a radically different technology will be required.
5. *Handling new materials*—New materials pose a number of issues for failure analysis. These include sample preparation issues, focused ion beam (FIB) edits and cross sections, electron beam imaging and interactions with the various failure analysis tools. Each new material requires development across a broad spectrum of failure analysis capabilities, including the determination of invasiveness by each physical analysis tool, especially in regards to the fragility of the new process materials.
6. *FIB editing*—The ability to perform FIB edits of circuits to support prototypes and reduce the number of design revisions is limited by both required beam placement accuracy and aspect ratio of holes. In addition the ability to perform edits on the new materials in upcoming technologies is unproven.

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 increasingly centralized, reducing total market potential and attractiveness especially for smaller suppliers and start-ups. There is a need for increased and more-cohesive consortium inputs to tool manufacturers to help reduce market-related risk.

SOFTWARE-BASED DIAGNOSIS AND SIGNATURE ANALYSIS

As challenges to PFA become more severe, alternatives are needed. A key alternative/supplement to traditional hardware-based fault localization is software-based fault localization. Software-based fault localization methodologies and tools are needed to handle diagnostics for fails detected by all major test methodologies, including scan-based and BIST-based test;

functional; IDDQ and, especially important, AC (delay) test. Localizing faults in embedded or stand-alone memories, a relatively easy task, is no longer sufficient given the lack of dense upper level metal. The tools must handle all realistic physical defects, including resistive bridges, resistive contacts/vias, and opens. Moreover, scan-based diagnostics must comprehend the fact that a relatively large fraction of yield is lost in scan integrity tests. Defects can also affect clock and other infrastructure nets, and these are difficult to diagnose with today's methods. Methods to diagnose problems related to parametric (non-defect) and reliability failures must also be developed. These methods may be based on both test structures and product-level tests. DFT techniques such as BIST must be designed with special consideration to support the necessary data gathering. IDDQ measurement devices need to support the accuracy levels required by diagnostics. Tester response data capture capabilities and data management systems must meet the demands of these methodologies. Specifically, ATE should allow for unlimited collecting of scan data. Diagnostic data collection allowing localization to a single or few failing net candidates should not add significantly to overall test time.

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

1. Production-worthy data collection, trading off resolution against test cost overhead. Concerns may include test data compression and BIST approaches. An absolute minimum requirement is failing core identification. Average test time overhead should be less than 1%.
2. Extensive data gathering on selected engineering or monitor wafers or lots. Granularity must be sufficient to build accurate fault type Paretos and support tool commonality analysis. Throughput time must be short enough to provide timely feedback to the fabrication process on sufficient volume and must support both time-zero and reliability failures. Tools should identify not just failing nets, but failing layers. Such analysis may involve integrating layout information and/or in-line test results into fault localization. Typical test time should be on the order of seconds. During early production, a more significant part of the material may be exposed to such extensive logging.
3. Individual die analysis that identifies defects to a single transistor or section of conductor no longer than 10 µm and identifies the failing layer. Such analysis may involve special-purpose diagnostic-resolution-enhancing ATPG and fail data collection and/or analog re-simulation and may be followed-up by failure analysis. Analysis time may be considerably longer than in the previous two cases.

Factory integration issues must be addressed to support these diagnostic needs. Data capture and management capabilities must support increasing reliance on statistical analysis and data mining of volume production data for yield learning. Mechanisms for yield data flow for distributed design and manufacture, including fabless/foundry and third party IP, must also be developed. For example, data management strategies are required to collect consistent data across multiple products containing the same design cores.

Hardware-based fault localization tools may be used to complement and supplement those above as appropriate. The spatial resolution of these techniques is predominantly fixed at about 0.5 µm by the near infrared light used for imaging and overlay (e.g., timing-resolved emission, emission microscopy, laser probing and thermally-induced voltage alteration (TIVA), etc.). Unless some other method is developed for backside imaging, this constraint must be dealt with by integrating hardware based fault isolation tools with improved computer aided design capability for overlay and signal tracing. CAD navigation must be both spatial and time-based, that is, linked to simulation waveforms.

Finally, signature analysis techniques must be developed to significantly reduce or eliminate the need for physical failure analysis. Statistical methods are needed to select failing die of a particular class to accurately pre-sort and prioritize input to physical failure analysis. In the longer term, methods must be developed to identify root cause based on test information without resorting to physical failure analysis. A key enabling technology is characterization test methods that are capable of distinguishing individual defect types. DFT support, such as used by the demonstrated MUX-based technique to make SRAM bit lines accessible for parametric analysis, may be required. Integration of electrical characterization with layout data and test-structure/in-line test results are also key capabilities.

MANUFACTURING 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 low performance logic test equipment that is available today. However, significant work remains to translate similar improvements to the broad market of analog and RF. 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.

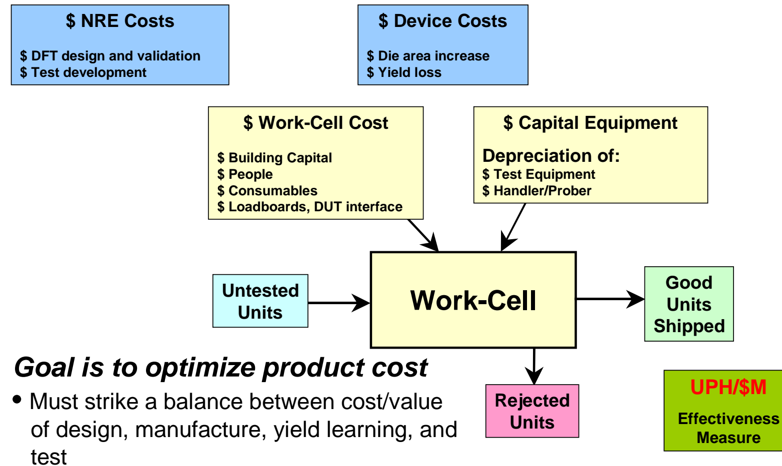


Figure 28 Test Cost Drivers

The cost of semiconductor test to the organization has many drivers. See Figure 28. The importance of these drivers varies substantially from device to device. Test development costs are more important for products with a lower volume. The cost of DFT area depends on whether the product is pad limited or core limited. The acceptable cost of test is very market specific and must be determined by balancing the value of test with its cost. Figure 29 presents the test quality trade-offs for an arbitrary chip.

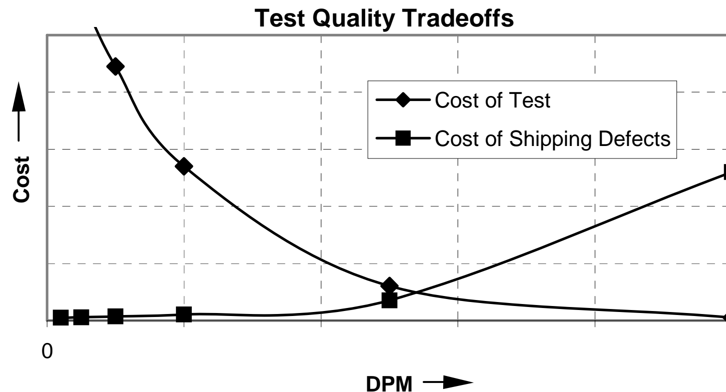


Figure 29 Test Quality Trade-offs: An Arbitrary Example to Illustrate the Trade-offs

Typically, the cost of test increases exponentially with an improvement in DPM. Many semiconductor manufacturers observe an increase in extreme "0 DPM" quality requirements. Despite the indicated trends, many manufacturers are able to achieve the low DPM target while controlling cost. In such cases, a significant part of the cost of test will likely be spent on the tail of the distribution. As technology progresses we expect that defect tolerance techniques will become more pervasive. For certain segments it may become economical to lower the cost of test by accepting increased DPM in tolerant circuits. However, test is not just screening, but a significant value of test is realized in reducing time-to-volume by improving yield learning curves. Test cost should not be minimized independently, but in the context of achieving the lowest overall manufacturing cost over a period of time.

ATE capital cost has traditionally been measured using a simple cost-per-digital pin approach. Although this is a convenient metric, it is misleading because it ignores base system costs associated with equipment infrastructure and central instruments as well as the scaling that occurs with reducing pin-counts and number of sites. Moreover, it is not aligned with the current trend in ATE platforms, where the same base infrastructure can be used for very diverse sets of test channels. The following equation expresses test cell capital cost in terms of the relevant cost drivers of future test technology:

$$C_{CELL} = C_{BASE} + C_{INTERFACE} + C_{POWER-SUPPLIES} + C_{TEST-CHANNELS} + C_{OTHER}$$

In this equation, C_{BASE} equals the base cost of a test system with zero pins/channels (e.g., it includes the cost of the mechanical infrastructure, back-plane, tester operating system software, and central instruments). $C_{INTERFACE}$ includes all costs required for interfacing with the device, e.g., the cost of interface electronics, sockets, and probe-cards (including spare probe-cards). C_{POWER} equals the cost of the power supplies. $C_{TEST\text{-}CHANNELS}$ equals the cost of the instruments (such as digital, analog, RF, memory test instruments). C_{OTHER} includes the remaining costs (e.g., floor-space). Practical considerations may limit the overall performance breadth that can be cost-effectively achieved by a given C_{BASE} infrastructure and should be taken into consideration in the overall test cell planning. For example, a low-end system may have an air-cooled infrastructure, whereas the high-end system will use liquid cooling. Test scenarios are evaluated by dividing the capital cost and performance metrics. For example, an important figure of merit is the *Units per Hour per Cost (UPH/$M)*, that is, the number of shipped devices per hour (*throughput*) over the total cost.

Figures 30 and 31 show trends in channel and interface costs for 2005 to 2020 for devices in each product segment.

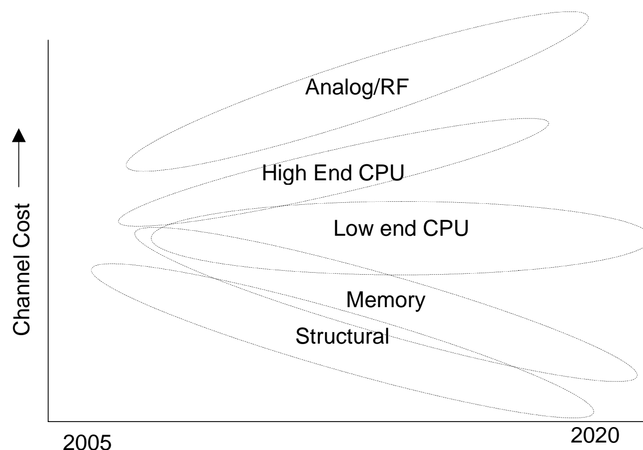


Figure 30 Channel Cost Range Trends for Product Segments

Figure 31, particularly, indicates the rapidly rising interface cost—which must be contained over time to avoid dominating the overall test cell cost. The next sections will describe these cost trends in more detail.

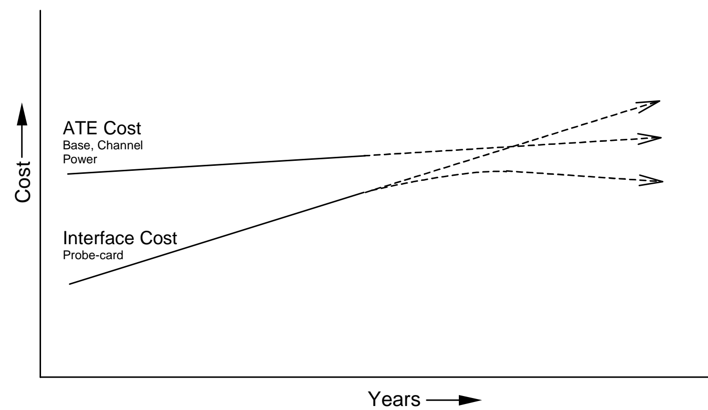


Figure 31 Test Cell Cost / Unit vs. Interface Cost Trend

.

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.

The cost of testing high-speed I/Os is becoming significant. In telecom applications, SONET is going from 2.5 Gbps to 10 Gbps/40 Gbps. In computing applications, serial ATA is going from 1.5 Gbps to 6 Gbps in 2008, and PCIExpress is going from 2.5 Gbps to 5 Gbps around 2006. Moreover, techniques likes simultaneous bi-directional signaling may become significant, adding to the test cost. Looking forward, high-speed I/O DFT techniques and new test methodologies will become a more important part of the solution to control the cost of test.

POWER COST TREND

With increasing degrees of multi-site, the cost of power supplies will increase. Especially if reduced pin-count techniques are deployed, the power cost per site may dominate the channel cost per site. The cost increase in power supplies may be contained by innovations in power supply and power delivery technology. Note that some DFT techniques increase demand on power supplies to achieve shorter test times.

INTERFACE COST TREND

Controlling the interface cost is essential for successful cost scaling using multi-site test: A dominating interface cost that increases exponential with the number of sites may defeat the purpose of increasing the number of sites. The interface cost becomes very challenging with high bandwidth (2 Gbit/s) and/or high multi-sites (128 sites). There is a need to develop consistent cost models that cover the wide range of probe card technologies in the market place. Long probe card lead times cause significant cost problems, especially for the advanced technologies. Probe card lead times will be

reduced by a factor of 2 within this roadmap's horizon. For certain products, it may become economical to skip wafer test or only do a simple low performance test. Looking forward, high-speed I/O DFT techniques will become a more pervasive part of the solution to control interface cost.

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, see Figure 31, (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%.

Shared ATE resources that reduce the channel cost, may cause a low multi-site efficiency (for example, this is apparent in mixed-signal / RF test). Moreover, 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, a 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 test.

Tables 23a and 23b present the expected trend in the number of sites for an arbitrary device in each segment as well as the minimum multi-site efficiency targets. A custom economic model should be deployed to identify the optimized roadmap to scale test cost of a custom device. Note that there are multiple trajectories/approaches that achieve the test cost targets.

Table 23a Multi-site Wafer Test (Package Test) for Product Segments—Near-term Years

Year of Production		2005	2006	2007	2008	2009	2010	2011	2012	2013
DRAM ½ Pitch (nm)		80	70	65	57	50	45	40	36	32
MPU/ASIC ½ Pitch (nm)		90	78	68	59	52	45	40	36	32
High Performance ASIC/MPU										
Wafer test	Number of sites	8	8	8	16	16	16	16	32	32
	Multi-site efficiency [%]	95%	95%	95%	98%	98%	98%	98%	98%	98%
Package test	Number of sites	4	4	4	8	8	8	8	16	16
	Multi-site efficiency [%]	92%	92%	92%	95%	95%	95%	95%	95%	95%
Low Performance Microcontroller										
Wafer test	Number of sites	64	96	128	256	512 / Full wafer	512 / Full wafer	512 / Full wafer	512 / Full wafer	512 / Full wafer
	Multi-site efficiency [%]	99.7%	99.8%	99.9%	99.9%	100%	100%	100%	100%	100%
Package test	Number of sites	128	256	256	512	512	512	512	512	512
	Multi-site efficiency [%]	99.9%	99.9%	99.9%	100%	100%	100%	100%	100%	100%
Mixed-signal										
Wafer test	Number of sites	4	4	4	8	8	16	16	16	16
	Multi-site efficiency [%]	85%	85%	85%	90%	90%	95%	95%	95%	95%
Package test	Number of sites	4	8	8	16	16	16	16	16	64
	Multi-site efficiency [%]	85%	90%	90%	95%	95%	95%	95%	95%	95%
Commodity Memory										
Wafer test	Number of sites	128	256	512	512	1024 / Full wafer	1024 / Full wafer	1024 / Full wafer	1024 / Full wafer	1024 / Full wafer
	Multi-site efficiency [%]	75%	85%	90%	90%	90%	95%	95%	95%	95%
Package test	Number of sites	128	256	256	512	512	1024	1024	1024	1024
	Multi-site efficiency [%]	90%	95%	95%	95%	97%	97%	97%	97%	97%
RF										
Wafer test	Number of sites	2	2	4	4	8	8	16	16	16
	Multi-site efficiency [%]	75%	75%	85%	85%	90%	90%	95%	95%	95%
Package test	Number of sites	4	4	8	16	32	48	64	64	64
	Multi-site efficiency [%]	75%	90%	92%	95%	97%	97%	97%	97%	97%

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

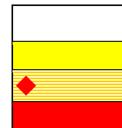
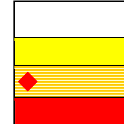


Table 23b Multi-site Wafer Test (Package Test) for Product Segments—Long-term Years

Year of Production	*2014*	*2015*	*2016*	*2017*	*2018*	*2019*	*2020*
DRAM ½ Pitch (nm)	*28*	*25*	*22*	*20*	*18*	*16*	*14*
MPU/ASIC ½ Pitch (nm)	*28*	*25*	*22*	*20*	*18*	*16*	*14*
High Performance ASIC/MPU							
Number of sites (wafer test)	**32**	**64**	**64**	**128**	**256**	**256**	**512 / Wafer scale**
Number of sites (package test)	**16**	**32**	**32**	**64**	**128**	**128**	**128**
Low Performance Microcontroller							
Number of sites (wafer test)	**512**	**768**	**768**	**1024 / Wafer scale**	**1024 / Wafer scale**	**1024 / Wafer scale**	**1024 / Wafer scale**
Number of sites (package test)	**512**	**768**	**768**	**768**	**768**	**768**	**1024**
Mixed-signal							
Number of sites (wafer test)	**16**	**32**	**32**	**64**	**128**	**128**	**256**
Number of sites (package test)	**64**	**128**	**128**	**256**	**256**	**256**	**512**
Commodity Memory							
Number of sites (wafer test)	**1024 / Wafer scale**	**2048 / Wafer scale**	**2048 / Wafer scale**	**2048 / Wafer scale**	**2048 / Wafer scale**	**2048 / Wafer scale**	**2048 / Wafer scale**
Number of sites (package test)	**1024**	**2048**	**2048**	**2048**	**2048**	**2048**	**2048**
RF							
Number of sites (wafer test)	**32**	**32**	**32**	**64**	**128**	**128**	**256**
Number of sites (package test)	**128**	**128**	**128**	**256**	**256**	**256**	**512**

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



OTHER COST TRENDS

The dramatic increase in SIP solutions that integrate memory, logic, and RF have further increased pressure to reduce the cost of test for mixed technology designs as well as improve the wafer test quality of KGD.

Test development time and cost will be reduced further by DFT techniques; test standards (to support test content reuse, test program inter-operability, compression, and manufacturing agility); automatic generation of test patterns (such as structural test approaches), and the programs that use them.

Memory DFT, like BIST and built-in self repair (BISR), will become pervasive: DFT will be essential to control the cost of test in a world where Commodity DRAM bit density quadruples every three years and Flash doubles every two years.

For certain segments, new manufacturing process flows may become economically justified. For example, it is expected that by the year 2006 and beyond, embedded flash and DRAM will be tested and repaired on a memory tester, while the logic blocks will be tested on the logic tester.

IMPORTANT AREAS OF CONCERN

1. Increases in the number of sites place severe demands on ATE architectures and probe card technologies. Research and development must continue to bring to market cost-effective probe technologies directed at product offerings and multi-site test trends. Lead times for advanced probe cards are an issue. RF radiation may become an issue for massive RF parallel test.
2. With increasing adoption of test compression, it is expected that the digital test cost percentage of the overall cost of SOC test will be reduced significantly.
3. The relatively high cost of analog and RF test instruments and the long test times associated with testing these circuits remain key challenges. To enable parallel test, multiple instruments are required with fast execution of DSP test algorithms like fast Fourier transform (FFT) (or other correlation tests). A secondary consideration for mixed-signal multi-site test is the load board circuits required for package test, especially for complex packages. DFT techniques for mixed-signal and RF devices remain development needs. Because of the high costs, mixed-signal resources (and post-processing steps) are frequently shared, significantly reducing multi-site efficiency.
4. 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 scaling.
5. New DFT techniques are required to test pins that are not connected during massive parallel test. Moreover, in setups where reduced pin-count techniques are deployed, power supply cost may dominate channel cost. The power supply cost may need to be contained by innovations in power supply and power delivery technology.
6. Although a low multi-site efficiency is not a problem for low site counts, for massive parallel test the impact can defeat the purpose of multi-site test. To continue multi-site scaling, ATE architectures may require dedicated channels/instruments per site, because shared channels/instruments may limit multi-site efficiency. To allow dedicated channels/instruments in a cost effective way, the channel cost needs to be reduced for certain segments.
7. Many semiconductor manufacturers observe an increase in extreme “0 defects per million (DPM)” quality requirements. Despite the indicated trends, many manufacturers have been able to achieve the low DPM target while controlling cost. Looking forward, continuing controlling the cost with ever-lower DPM requirements is one of the main challenges.

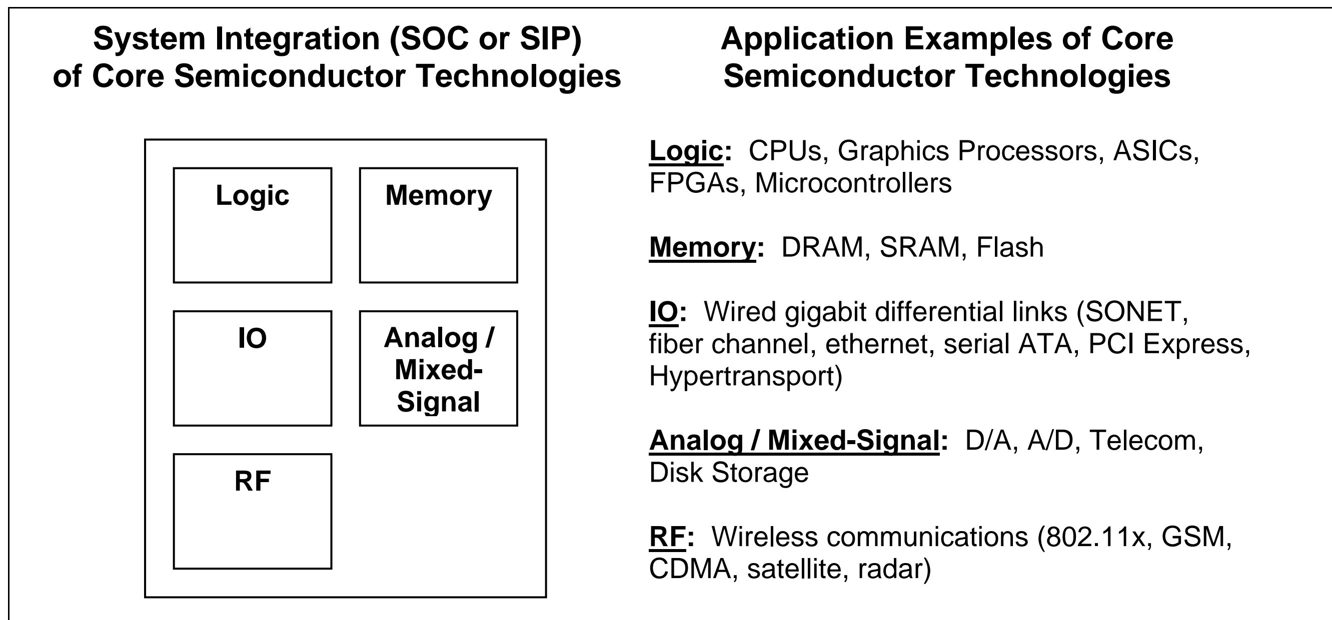
TEST TECHNOLOGY REQUIREMENTS

INTRODUCTION

Over the past 25 years, semiconductor test technology requirements have been driven primarily by relentlessly increasing performance and transistor counts. A fundamental shift is underway driven by the emergence of new market demands (for example, mobility, security, ease of use, ease of system management, low power, etc.). This in turn is fueling the integration of different semiconductor technologies in more ways and in a greater set of applications than ever before. This in itself is a huge challenge to test as it is ultimately the application requirements or specifications that determine test technology requirements, but it would be impossible to capture a comprehensive set of applications and their associated test requirements trends within this chapter. Therefore, the identify core semiconductor technology building blocks have

been identified to provide a framework for describing the test challenges and trends associated with each core technology as well as for describing the test challenges associated with integrating these core technologies together either as a SOC or a SIP.

Each core semiconductor technology has certain applications associated with it and some of these will be used as a basis for extracting long-term trends. In particular, the ITRS publishes key technology attribute trends for CPUs, ASICs, DRAMs, and Flash memory. These will be referenced where appropriate in the core technology sections. Figure 32 shows the core semiconductor technologies addressed in this chapter as well as examples of associated applications. The application mapping is intentionally loose as many of the examples listed may contain multiple core technologies. The core technologies are differentiated mainly by their inherent functional differences and thus their different test requirements. Two emerging core technologies that are not included in this revision, but may be in the future are MEMS and optical.



A/D—analog to digital CDMA—code division multiple access CPUs—central processing unit D/A—digital to analog GSM—global standard for mobile

Figure 32 Organization of Core Semiconductor Technologies from System Integration and Applications Points of View

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

An SOC design consists of multiple individual design blocks, or cores, using different technologies (logic, memories, 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. A fundamental challenge of SOC test is the need to combine test requirements from multiple sources with differing testability methods. Opportunities exist to define standards for core test to conform to a hierarchical methodology. When an IP core is obtained from a third party provider, its predefined test solution must be adopted. Many EDA tools already leverage a standard format for logic cores; and this format must be extended to other core types, such as analog cores.

In the SOC test roadmap, it is desirable that the significant indices of DFT, such as DFT area investment, test pattern length, or test data volume, be shown not only qualitatively, but also quantitatively. Obviously, quantification requires a model. In Table 24, an SOC model based on the following assumptions is targeted:

1. The chip size is kept stable at 64 mm^2 in each technology generation.
2. The number of transistors increases twice every three years, according to ITRS technology trend.
3. The memory part grows faster than the logic part in SOC (Logic part 15% / year, total 26% / year.). The transistor count of each part is shown in Table 24.
4. Memory bit count = transistor count in memory cell array/ 6 transistors.
5. Memory cell array area size is 70% of memory device area according to the system driver chapter.
6. Number of memories = memory bit count/32 Kbits.
7. Word count in memory = 32 K words on average.
8. Number of Flip-Flops = transistor count in logic part * 50%/40 transistors.

Table 24 SOC Model

			130 nm	90 nm	65 nm
Transistor Count (Million)	High Frequency Logic Part	Logic	4.7	7.1	10.9
		Memory	8.6	19.5	42.3
	Low Frequency Logic Part	Logic	6.8	10.3	15.7
		Memory	19.6	42.5	89.9
	Total		39.7	79.4	158.9

For logic cores, sophisticated DFT methods such as *random pattern logic BIST* or *compressed deterministic pattern test* are required to reduce large amount of test data. Each method should be adopted considering the pros and cons regarding DFT area investment, design rule restrictions, and associated ATE cost. DFT area mainly consists of test controllers and test points. It is kept constant over the years by using a hierarchical approach. Regarding ATE cost, random pattern logic BIST is available with a small number of ATE pins assigned and, therefore, it will be especially effective for wafer level burn-in. As physical synthesis techniques, which consider physical design during logic synthesis from a register transfer level (RTL) design to a gate level design, and RTL DFT solutions become widely used, DFT area investment estimation at RTL design is required.

To keep the test quality level of embedded cores against deep sub-micron defects, such as resistive opens or small delay defects, additional delay test is required. Therefore, the number of test patterns increases inevitably and imperatively, and test application time may be as much as thirty times larger than today's in 2010. This means various techniques for the significant reduction of test time, such as test pattern compaction, scan architecture improvement, and scan shift speed acceleration, are strongly needed.

The conventional delay fault models represented by the transition fault model are not accurate enough to predict the defect level of SOC. Though the transition delay fault model achieves comparatively high fault coverage, it tends to activate short paths. Therefore, the delay test based on the transition fault model is not accurate enough to detect small

size delay defects. A high accuracy delay fault model that can detect small delay defects is required. Furthermore, as crosstalk will be a crucial issue in the timing design, a viable crosstalk fault model is also required.

As process technology advances, the ratio of memory device area in SOC rapidly increases. This increases embedded memory devices, memory bit count, and consequently memory BIST logic gates. On the other hand, as the density of memories and their operating frequencies grows, memory BIST technology needs innovations on the functionality and the performance as follows:

1. To cover new types of defects in memory devices that appear in a nanometer process, testing algorithms in BIST controllers should evolve from a generic fixed one to either a selective combination of generic algorithms and test conditions or a dedicated optimal algorithm for a given memory design and defect set. Furthermore, a highly programmable BIST that enables flexible composition of the tests should be developed.
2. Practical embedded repair technologies, such as built-in redundancy allocation (BIRA) and BISR technologies should be developed. BIRA analyzes the BIST results and allocate redundancy elements for yield improvement. BISR performs the actual reconfiguration (hard-repairing) on-chip.
3. At-speed testing and repairing of memory cores in high-frequency parts of SOC should be developed.

These requirements result in more sophisticated test and repair quality, so they will cause growth in logic infrastructure. Although memory repair with redundancy elements can be performed using the conventional direct access method and memory tester, it is very difficult to apply this on complex SOCs with increasing number of embedded memory instances. Therefore, it is required to develop a BIRA/BISR technology that uses at-speed BIST, and performs on-chip repair analysis.

Based on Tables 25a and 25b, the embedded test and repair logic size will be up to 35 K gates per million bits of embedded memory in 2005. This contains BIST, BIRA, and BISR logic, but does not include the repair programming devices such as optical or electrical fuses. The ratio of embedded test and repair logic size to the number of memory bits must not increase over the next decades. This ratio is not an easily achievable requirement considering the factors discussed above. In particular, when the memory redundancy architecture becomes more complex, it will be difficult to implement the repair analysis with a limited amount of logic. Therefore, breakthrough in embedded test and repair architecture is required.

Cost effective test of analog cores is a critical issue. Analog BIST techniques provide a potential solution but are not mature enough for general use. Failure analysis techniques for analog cores must be developed.

Structured use of IP core wrappers and test access mechanisms must be adopted for individual cores within a SOC. These should be developed carefully to enable functional, at-speed, and interconnect testing. In order to ensure ease of SOC test integration, a standard wrapper architecture and information model, such as IEEE 1500[1], needs to be utilized for each IP core. Furthermore, the information methods should be standardized with the interface language for interoperability of EDA tools. One such effort is the core test language (CTL) of IEEE P1450.6.[2] The high complexity of SOC design creates design and test throughput and test quality challenges. EDA tools must be developed to aid management of this complexity.

[1] *1500-2005 IEEE Standard Testability Method for Embedded Core-based Integrated Circuits.*
[2] *P1450.6, Draft Standard for Standard Test Interface Language (STIL) for Digital Test Vector Data—Core Test Language (CTL).*

Table 25a System on Chip Test Requirements—Near-term Years

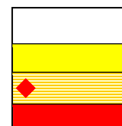
Year of Production	*2005*	*2006*	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*	*2013*
DRAM ½ Pitch (nm) (contacted)	*80*	*70*	*65*	*57*	*50*	*45*	*40*	*36*	*32*
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	*90*	*78*	*68*	*59*	*52*	*45*	*40*	*36*	*32*
MPU Physical Gate Length (nm)	*32*	*28*	*25*	*22*	*20*	*18*	*16*	*14*	*13*
Embedded Cores									
Standardization of core test data [1]	Standard format on EDA/ATE	Standard format on EDA/ATE	Standard format on EDA/ATE	Extension to analog cores	Extension to analog cores	Extension to Analog Cores	Extension to Analog Cores	Extension to Analog Cores	Extension to Analog Cores
Embedded Cores: Logic									
Test logic insertion at RTL design	Partially	Partially	Fully	Fully	Fully	Fully	Fully	Fully	Fully
Testability analysis and overhead estimation at RTL design	Ad hoc	Fully	Fully	Fully	Fully	Fully	Fully	Fully	Fully
BISR for logic cores	Minimal	Some	Some	Some	Some	Some	Some	Logic BISR	Logic BISR
Embedded Cores: Logic – Random Pattern Logic BIST									
Area investment beyond scan (%) [2]	3.1	3.1	3.1	3.1	3.1	3.1	3.1	3.1	3.1
Embedded Cores: Logic – Compressed Deterministic Pattern Test									
Area investment beyond scan (%) [3]	1.1	1.2	1.3	1.4	1.5	1.6	1.6	1.7	1.8
Test pattern length (number of captures) [4]	1.0	1.3	1.6	2.0	2.5	3.2	4	5	6
Test pattern length compression ratio [5]	1.0	1.4	2.0	3.1	5.2	9.2	18	23	30
Test data volume compression ratio [6]	100	120	150	210	300	460	770	860	980
Embedded Cores: Memory									
Area investment of SRAM BIST (Kgates/Mbits)	35	35	35	35	35	35	35	35	35
Repairing mechanism of embedded SRAM cells [7]	BISR	BISR	BISR for Row & Col R/D	BISR for Row & Col R/D	BISR for Row & Col R/D	BISR for Row & Col R/D	BISR for Row & Col R/D	BISR for More Sophisticated R/D	BISR for More Sophisticated R/D
Embedded Cores: Analog									
DFT (BIST, BOST) for analog cores	Limited use (PLL, ADC, etc.)	Limited use (PLL, ADC, etc.)	Limited use (PLL, ADC, etc.)	Limited use (PLL, ADC, etc.)	Partial	Partial	Partial	Partial	Partial
Design for failure analysis of analog cores	Ad hoc	Ad hoc	Ad hoc	Ad hoc	Ad hoc	Ad hoc	Ad hoc	Ad hoc	Ad hoc
Core Access									
Use of standard interface on IP core access	Partially	Partially	Partially	Partially	Partially	Partially	Partially	Partially	Partially
Analog-mixed signal core access	Direct Access	Direct Access	Direct Access	Direct Access	Analog wrapper [8]	Analog wrapper [8]	Analog wrapper [8]	Standard analog wrapper [8]	Standard analog wrapper [8]
SOC Level Testing									
Test strategy for IP core-based design [9]	Partially automated	Partially automated	Partially automated	Fully automated	Fully automated	Fully automated	Fully automated	Fully automated	Fully automated
DFT selection for cores	DFT selection for cores	DFT selection for cores	DFT selection for cores	Selection for cores/fully automated EDA tool	Selection for cores/fully automated EDA tool	Selection for cores/fully automated EDA tool	Selection for cores/fully automated EDA tool	Selection for cores/fully automated EDA tool	Selection for cores/fully automated EDA tool
DFT at higher level design [10]	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Fault Model for SOC									
Fault model for SOC level fault coverage [11]	Standard fault models	Standard fault models	Standard fault models	New standard fault model, its coverage [12]	New standard fault model, its coverage [12]	New standard fault model, its coverage [12]	New standard fault model, its coverage [12]	New standard fault model, its coverage [12]	New standard fault model, its coverage [12]
Delay fault model with high accuracy	Partially	Partially	Fully usable	Fully usable	Fully usable	Fully usable	Fully usable	Fully usable	Fully usable
X-talk fault model	No	No	No	Partially	Fully usable	Fully usable	Fully usable	Fully usable	Fully usable
Manufacturing									
Diagnosis interface/data [13]	Standard format and methods on IP core	Standard format and methods on IP core	Standard format and methods on IP core	Automated SOC diagnosis	Automated SOC diagnosis	Automated SOC diagnosis	Automated SOC diagnosis	Automated SOC diagnosis	Automated SOC diagnosis

BOST—built off-chip self test

Table 25b System on Chip Test Requirements—Long-term Years

Year of Production	2014	2015	2016	2017	2018	2019	2020
DRAM ½ Pitch (nm) (contacted)	28	25	22	20	18	16	14
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	28	25	22	20	18	16	14
MPU Physical Gate Length (nm)	11	10	9	8	7	6	6
Embedded Cores							
Standardization of core test data [1]	Extension to Analog Cores	Extension to Analog Cores	Extension to Analog Cores	Extension to Analog Cores	Extension to Analog Cores	Extension to Analog Cores	Extension to Analog Cores
Embedded Cores: Logic							
Test logic insertion at RTL design	Fully	Fully	Fully	Fully	Fully	Fully	Fully
Testability analysis and overhead estimation at RTL design	Fully	Fully	Fully	Fully	Fully	Fully	Fully
BISR for logic cores	Logic BISR	Logic BISR	Logic BISR	Logic BISR	Logic BISR	Logic BISR	Logic BISR
Embedded Cores: Logic – Random Pattern Logic BIST							
Area investment beyond scan (%) [2]	3.1	3.1	3.1	3.1	3.1	3.1	3.1
Embedded Cores: Logic – Compressed Deterministic Pattern Test							
Area investment beyond scan (%) [3]	1.9	2	2.1	2.1	2.1	2.1	2.1
Test pattern length (number of captures) [4]	8	10	13	16	20	25	32
Test pattern length compression ratio [5]	40	55	68	83	100	130	160
Test data volume compression ratio [6]	1,140	1,360	1,450	1,560	1,680	1,810	1,960
Embedded Cores: Memory							
Area investment of SRAM BIST (Kgates/Mbits)	35	35	35	35	35	35	35
Repairing mechanism of embedded SRAM cells [7]	BISR for More Sophisticated R/D	BISR for More Sophisticated R/D	BISR for More Sophisticated R/D	BISR for More Sophisticated R/D	BISR for More Sophisticated R/D	BISR for More Sophisticated R/D	BISR for More Sophisticated R/D
Embedded Cores: Analog							
DFT (BIST, BOST) for analog cores	Partial	Partial	Partial	Partial	Full use	Full use	Full use
Design for failure analysis of analog cores	Partially structural	Structural	Structural	Structural	Structural	Structural	Structural
Core Access							
Use of standard interface on IP core access	Partially	Fully	Fully	Fully	Fully	Fully	Fully
Analog-mixed signal core access	Standard analog wrapper [8]	Standard analog wrapper [8]	Standard analog wrapper [8]	Standard analog wrapper [8]	Standard analog wrapper [8]	Standard analog wrapper [8]	Standard analog wrapper [8]
SOC Level Testing							
Test strategy for IP core-based design [9]	Fully automated	Fully automated	Fully automated	Fully automated	Fully automated	Fully automated	Fully automated
DFT selection for cores	Selection for cores/fully automated EDA tool	Selection for cores/fully automated EDA tool	Selection for cores/fully automated EDA tool	Selection for cores/fully automated EDA tool	Selection for cores/fully automated EDA tool	Selection for cores/fully automated EDA tool	Selection for cores/fully automated EDA tool
DFT at higher level design [10]	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fault Model for SOC							
Fault model for SOC level fault coverage [11]	New standard fault model, its coverage [12]	New standard fault model, its coverage [12]	New standard fault model, its coverage [12]	New standard fault model, its coverage [12]	New standard fault model, its coverage [12]	New standard fault model, its coverage [12]	New standard fault model, its coverage [12]
Delay fault model with high accuracy	Fully usable	Fully usable	Fully usable	Fully usable	Fully usable	Fully usable	Fully usable
X-talk fault model	Fully usable	Fully usable	Fully usable	Fully usable	Fully usable	Fully usable	Fully usable
Manufacturing							
Diagnosis interface/data [13]	Automated SOC diagnosis	Automated SOC diagnosis	Automated SOC diagnosis	Automated SOC diagnosis	Automated SOC diagnosis	Automated SOC diagnosis	Automated SOC diagnosis

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



Definitions for Tables 25a and b:
[1] The standardization of test data format needs to reduce turn-around-time of test program development.
[2] Area investment of random pattern logic BIST consists of BIST controller and test points.
[3] Area investment of compressed deterministic pattern test logic consists of controller and test points.
[4] Required number of test pattern length (number of captures), which is normalized based on the number on 2005.
[5] Test pattern compression ratios are values that are necessary to suppress the total test pattern length (sum of those for stuck-at, transition, path delay, and X-talk tests) in the table within the required test pattern length.
[6] Ratio of compressed test data volume in the tester memory against conventional scan test data volume with same fault coverage.
[7] Hard repair that uses optical or electrical fuse devices for the programming.
[8] Extended wrapper structure to access to embedded analog-MS cores, not chip-level analog boundary-scan.
[9] The strategy contains test control integration, test scheduling for low power consumption, test time, and test pin reduction.
[10] Behavioral level, HW/SW co-design, high-level synthesis with testability analysis.
[11] The standardization of fault model and fault coverage needs to popularize IP Cores.
[12] A method to obtain overall test quality measure of SOC considering all embedded deices; logic, memory and analog.
[13] The standardization of diagnosis data format and interface needs to reduce turn-around-time of failure analysis.

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" 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, especially when final assembly may include die thinning and stacking (which can damage/change KGD die). 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 voltage stress, etc. 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 fully test the assembled module. The multiple test insertions may result in test escapes. 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 times 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 as a 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 Post Processing

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 impossible using traditional on-tester limits. The advantages of reliability screening at a test insertion instead of burn-in are savings in time, fixtures, equipment, and handling. The KGD implications are that screens can be performed at the wafer level with standard probes and testers, so every device can be considered fully conditioned in compliance with data sheet specifications and shipped quality and reliability targets for that process regardless of the final package in which the device is to be shipped. Using off-tester statistical methods the test measurements (for example, I_{dd}, V_{ddmin}, F_{max}) of each die are recorded instead of being binned. These measurements can be recorded for different test conditions, pre and post stress testing, and at different temperatures. Pass or fail criteria are determined based on statistical analysis of the

measurements recorded using off-tester post processing algorithms. Outliers to the statistical distribution are graded based on their statistical likelihood of being system failures or early life failing devices, and the inkless wafer maps are modified accordingly. The challenge for testing using statistical methods is to meet an acceptable trade-off between the potential failing population and the intrinsic yield loss.

SUBSEQUENT PROCESSING AFFECTS THE QUALITY OF THE DIE

The processing that occurs during assembly can damage some technologies. Wafer thinning is one example: when DRAM wafers are thinned, a shift in the refresh characteristics has been observed. A die from a wafer that was fully tested at wafer level may fail the exact same test after being thinned and assembled into a SIP or MCP. The thermal processing steps that are present in the assembly process can also lead to a change in the refresh characteristics of individual bits. This phenomenon, known as variable retention time (VRT), is impossible to screen prior to the assembly process.

A key challenge is to re-establish the quality levels achieved by the die supplier. This can be accomplished through additional post assembly testing, invoking redundant elements in the individual failing die within the multi chip package or using components that are specifically designed for multi chip applications.

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” is chosen as the primary reference because the most trend data is available for it. 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.

Tables 26a and 26b below include two distinct sections. They are (1) High Volume Microprocessor Trends Drivers and (2) High Volume Microprocessor Test Requirements. The former represents fundamental underlying roadmap trends and the latter projected test requirements after adjusting for system and DFT trends. The system and DFT trends drivers that are most likely to have a significant impact to logic test requirements will be discussed as will a summary of important areas of concern at the end of this section.

HIGH VOLUME MICROPROCESSOR TRENDS DRIVERS

The trends in the first part of Tables 26a and 26b 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 the table 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 Tables 26a and 26b are the nominal device V_{dd} 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-core architectures may necessitate more sophisticated independent power delivery to each core in order to fully optimize power consumption. Eventually it

is likely that this will need to be done on-chip. Overall, this trend would simplify the problem of delivering power to devices under test, but it also could create many new test issues since precise voltage control and current measurement have always been an important aspect of testing high power devices.

Another important system trend is the continuous increase of chip-to-chip data bandwidth required over time. This translates into increasing chip I/O data rates and / or an increasing numbers of I/Os. In order to reliably achieve speeds much greater than one giga-transfer per second (GT/s), it is necessary to incorporate high-speed serial signal techniques such as differential signaling or embedding the clock together with the data. For example, this is already occurring with PCI Express and Serial ATA interfaces and will proliferate to all I/O ports (e.g., front side bus) for microprocessors over the ITRS timeframe. Refer to the High Speed Input/Output Interface section for more detailed discussion on the test requirements and challenges for testing these interfaces.

DFT TRENDS DRIVERS

In order for test cost not to increase proportionally with chip scale trends, continuing improvements to DFT coverage and effectiveness will be crucial. The general trend toward multiple reusable cores offers many exciting DFT possibilities. Could the cores be tested in parallel? Could the cores test each other? Could one or more cores be redundant so that "bad" cores could be disabled by test or even in end-use? Also, could there be opportunity for general purpose test cores—sort of an on-chip ATE? It is very difficult to predict exactly how this will evolve and how this will impact manufacturing test requirements. However, there are some clear trends:

1. Structural, self test and test data compression techniques will continue and will be important for containing test data volume increases as well as constraining the device I/O interface during test.
2. DFT will continue to be essential for localizing failures to accelerate yield learning.
3. 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 alleviate the need for external algorithmic pattern generator capability and, similarly, I/O DFT features are increasingly being employed to alleviate the need for sophisticated I/O testing capabilities in high volume manufacturing test.
4. DFT will increasingly be required to ensure device deterministic behavior or accommodate device non-deterministic behavior. Power management features, I/O communication protocols, and device self repair mechanisms are a few examples of desirable non-deterministic behaviors.

HIGH VOLUME MICROPROCESSOR TEST REQUIREMENTS

The second parts of Tables 26a and 26b enumerate the high volume microprocessor test requirements over the ITRS horizon. I/O data rates are bounded on the low end by the need to provide slow speed access for structural or DFT-based testing and on the high end by the native speed of the interface. There will be considerable overlap of I/O types, including on the same device, through at least the near term horizon. Power consumption has been effectively capped at a maximum of 300 W for client microprocessor applications and a maximum of 400 W for server microprocessor applications through the end of the roadmap. Similarly, equipment vector memory requirements are expected to grow modestly over time.

Table 26a Logic Test Requirements—Near-term Years

Year of Production	*2005*	*2006*	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*	*2013*
DRAM ½ Pitch (nm) (contacted)	*80*	*70*	*65*	*57*	*50*	*45*	*40*	*36*	*32*
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	*90*	*78*	*68*	*59*	*52*	*45*	*40*	*36*	*32*
MPU Physical Gate Length (nm)	*32*	*28*	*25*	*22*	*20*	*18*	*16*	*14*	*13*
Part 1: High Volume Microprocessor Trends Drivers									
Functions per chip at production (million transistors [Mtransistors])	**193**	**193**	**386**	**386**	**386**	**773**	**773**	**773**	**1,546**
Chip size at production (mm^2)	**111**	**88**	**140**	**111**	**88**	**140**	**111**	**88**	**140**
Number of processor cores	**2**	**2**	**4**	**4**	**4**	**8**	**8**	**8**	**16**
Nominal V_{dd} Range (V)	**0.9–1.1**	**0.9–1.1**	**0.8–1.1**	**0.8–1.0**	**0.8–1.0**	**0.7–1.0**	**0.7–1.0**	**0.7–0.9**	**0.6–0.9**
Chip-to-board (off-chip) speed (high-performance, for peripheral buses) (MHz)	**3125**	**3906**	**4883**	**6103**	**7629**	**9536**	**TBD**	**14901**	**18626**
Part 2: High Volume Microprocessor Test Requirements									
I/O data rate (GT/s)	**0.1 - 3**	**0.1 - 6**	**0.1 - 6**	**0.1 - 6**	**0.2 - 12**	**0.2 - 12**	**0.2 - 12**	**0.2 - 15**	**0.2 - 15**
I/O types	**Slow speed scan / DFT, source synchronous, clock forwarding, clock embedded**				**Slow speed scan / DFT, clock forwarded, clock embedded**				
Total device maximum power consumption at test (W) — Client	**200**	**200**	**200**	**200**	**300**	**300**	**300**	**300**	**300**
Total device maximum power consumption at test (W) — Server	**200**	**250**	**300**	**300**	**300**	**300**	**300**	**300**	**300**
Number of power supplies per site	**1–4**	**1–4**	**1–6**	**1–6**	**1–6**	**1–4**	**1–4**	**1–3**	**1–3**
Power supplies voltage range (V)	**0.7–2.0**	**0.7–2.0**	**0.6–2.0**	**0.6–2.0**	**0.6–2.0**	**0.6–12**	**0.6–12**	**0.6–12**	**0.6–12**
Scan vector memory (Mbit per pin)	**64-256**	**64-256**	**64-256**	**64-256**	**64-512**	**64-512**	**64-512**	**64-512**	**64-1024**
Functional vector memory (M-vectors per pin)	**16-128**	**16-128**	**16-128**	**16-128**	**16-256**	**16-256**	**16-256**	**16-256**	**16-512**

Values for the year 2011 will be determined in the 2006 ITRS Update.

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

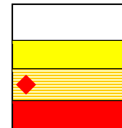
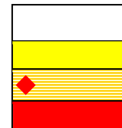


Table 26b Logic Test Requirements—Long-term Years

Year of Production	*2014*	*2015*	*2016*	*2017*	*2018*	*2019*	*2020*
DRAM ½ Pitch (nm) (contacted)	*28*	*25*	*22*	*20*	*18*	*16*	*14*
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	*28*	*25*	*22*	*20*	*18*	*16*	*14*
MPU Physical Gate Length (nm)	*11*	*10*	*9*	*8*	*7*	*6*	*6*
Part 1: High Volume Microprocessor Trends Drivers							
Functions per chip at production (million transistors [Mtransistors])	**1,546**	**1,546**	**3,092**	**3,092**	**3,092**	**6,184**	**6,184**
Chip size at production (mm^2)	**111**	**88**	**140**	**111**	**88**	**140**	**111**
Number of processor cores	**16**	**16**	**32**	**32**	**32**	**64**	**64**
Nominal Vdd Range (V)	**0.6–0.9**	**0.6–0.8**	**0.5–0.8**	**0.5–0.7**	**0.5–0.7**	**0.4–0.7**	**0.4–0.6**
Chip-to-board (off-chip) speed (high-performance, for peripheral buses) (MHz)	**TBD**	**29103**	**36379**	**TBD**	**56843**	**TBD**	**TBD**
Part 2: High Volume Microprocessor Test Requirements							
I/O data rate (GT/s)	**0.2-20**	**0.2-20**	**0.2-20**	**0.2-40**	**0.2-40**	**0.2-40**	**0.2-40**
I/O types	**Slow speed scan / DFT, advanced clock embedded, optical**						
Total device maximum power consumption at test (W) — Client	**300**	**300**	**300**	**300**	**300**	**300**	**300**
Total device maximum power consumption at test (W) — Server	**400**	**400**	**400**	**400**	**400**	**400**	**400**
Number of power supplies per site	**1–3**	**1–3**	**1–3**	**1–3**	**1–3**	**1–3**	**1–3**
Power supplies voltage range (V)	**0.6–12**	**0.6–12**	**0.6–48**	**0.6–48**	**0.6–48**	**0.6–48**	**0.6–48**
Scan vector memory (Mbit per pin)	**64–1024**	**64–1024**	**64–1024**	**64–1024**	**64–1024**	**64–1024**	**64–1024**
Functional vector memory (M-vectors per pin)	**16–512**	**16–512**	**16–512**	**16–1024**	**16–1024**	**16–1024**	**16–1024**

Values for the years 2014, 2017, 2019, and 2020 will be determined in the 2006 ITRS Update.

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



IMPORTANT AREAS OF CONCERN

1. *Power Consumption and Thermal Management*—While it is assumed that device power consumption will be significantly constrained through the roadmap horizon, new innovations in power delivery and thermal control at test may be required to stay ahead of system design optimization.
2. *High Speed I/O Interface*—Test equipment solutions are needed for device characterization and DFT / test equipment solutions that scale economically are needed for high volume manufacturing test.
3. *Mult-core Trends*—create both test challenges and opportunities. Multi-core designs may exacerbate the previous two areas of concern and add test complexity, but also may offer exciting DFT opportunities such as redundancy that could significantly aid test.

HIGH SPEED INPUT/OUTPUT INTERFACE

High frequency I/O technology continues to expand beyond the original transceiver applications for telecomm. Serial I/O interfaces are being widely adopted into back plane applications, short and long-haul communications, and chip-to-chip links for computing applications. Many new industrial standards emerged recently, including new introductions from consortiums and evolution from existing standards. For example, in telecom applications, SONET is going from 2.5 Gbps to 10 Gbps/40 Gbps, fiber channel is going from 1.0625 Gbps/2.125 Gbps and 4.25 Gbps to 8.5 Gbps in 2006, and ethernet is going from 100 Mb/s to gigabit ethernet and 10 gigabit ethernet. In computing applications, serial ATA is ramping quickly from current 1.5 Gbps and 3 Gbps, and targeting 6 Gbps in 2007, and PCI Express is poised to increase from today's 2.5 Gbps to 5 Gbps in 2006. Even source-synchronous bus, which is already approaching its practical limit at 2~3 Gbps, for example 2.8 hypertransport in late 2005, is now reinventing itself with new clocking schemes to reach beyond 5–8 Gbps around 2006. The convergence of classic parallel bus and classic serial bus is seen in the clocking scheme concept, and differential signaling is replacing single-ended signaling. All these high-speed differential links presents many challenges for production test on ATE.

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 provided pin-card-level integrated solutions up to 3.2 Gbps, and several are in the process to introduce 6~6.4 Gbps solutions around 2006. At the present time, testing in the 3.2 Gbps and above remains in the arena of expensive, stand-alone pattern-generators and bit-error-rate detectors. In some hybrid solutions, such lab instruments are bundled with ATE to address the embedded 3.2~12 Gbps serial links in SOC or ASICs. The excessive test time and cost makes this approach impossible for high volume production, but this hybrid approach always provides a leading edge performance much sooner to the market. That becomes a valuable enabler for new product introduction and debug. Another drawback for this hybrid approach has become more obvious as the data-rate goes beyond 3 Gbps. That is the bandwidth restrictions for the long cable required to connect these external instruments. Integrated instruments in the ATE test head are physically much closer to the DUT for signal delivery, but require much high integration level than an external instrument. In the near future high-frequency pin-electronics front-end for these instruments will be moved close to the DUT to deliver/preserve the bandwidth, while the rest part of the instrument being kept out of the test head. This compromise will effectively alleviate the integrated vs. hybrid dilemma. At present time, test solutions for < 3.2 Gbps range from simple internal digital loopback, high-end ATE pin cards, to test modules and golden device on test board. Each of this approach has their pros and cons, but every company is adapting its design to a certain solution. The trade-off is on silicon area used for DFT versus test cost. The tolerance to defect rate of different products is also a major determining factor for the different test choices.

In order to keep the overall system cost down, low cost printed circuit board (PCB) lamination materials such as FR_4 will still be the material of choice for most of the telecomm backplane and computing applications. However, the preference of FR_4 forms a bottleneck in spectral bandwidth. Several techniques are being developed to extend the transmission data-rate under this constraint, such as pre-emphasis/de-emphasis, transmit-side equalization/receiver-side adaptive equalization, and multi-level encoding. All these techniques, if deployed, would further complicate the test requirement significantly for high-speed serial links.

Without proper test methodology and equipment available, many IC manufacturers are forced to use various innovative but limited testing techniques (such as loopback, golden device, and adding components on the DUT board), which normally compromise fault coverage to a certain extend. This is a potential risk for relatively new I/O technologies that are often intentionally designed to push the envelop of the processing technology.

In the near term there is an urgent need for ATE manufacturers to design multi-port, gigabit rate instruments and integrate them into test systems, including control software, in order to keep up with the pace the rapid progression in speed and port count. On the other hand, viable solutions need to be cost effective. In the long run, the existing DFT features need to be extended beyond the current functional pseudo random bit sequence (PRBS) BIST approach to provide more performance related parametric coverage. It is projected that on-chip instrument and built-in design verification techniques will evolve and coexist with off-chip test equipment. An economically ideal distribution of on-chip and off-chip test coverage is yet to be determined. The goal is to minimize manufacturing test cost and efficiently test high port count devices.

IMPORTANT AREAS OF CONCERN

1. *Data Rate Increase*—In the computing industry, the current source synchronous bus is approaching 2.8 Gbps this year. Between 2005 and 2008, the classical source synchronous bus will gradually be replaced by embedded clock, clock forwarding, and simultaneous bi-directional architectures. The new architecture will break through the 2-3 Gbps limit of the classic source synchronous architecture, and bring the maximum computing I/O speed up to 1.5 Gbps to 8.5 Gbps. This trend seems to coincide with the data rate requirements in the telecom backplane applications of 2.5 Gbps to 8.5 Gbps. Because most telecom related backplanes would be limited to the use of low-cost FR-4 as PCB lamination material, the data rate will stay below 10 Gb/s at least until 2012. On the other hand, telecom long haul and short haul transceivers will continue to lead serial link technology to higher data rates. The current 10 Gbps long haul applications are moving to CMOS for lower cost and higher integration. Long-haul 40 Gbps does not see a market demand until 2012. Short-haul applications will reach 10 Gbps before 2006, with no massive demand for 40 Gbps in the next few years.
2. *Port Count Increase*—Low voltage CMOS technologies and low output voltage swings enable massive integration into large ASICs and SOCs. Currently in 2005, up to 100 pairs of 1~3 Gbps backplane style serializer and deserializer (SerDes) are being produced by many IC makers. This port-count will exceed 200 pairs in 2006, while the source synchronous bus will exceed 240 ports around 2006, but for a large percent of applications the port count will be limited to 40 or less. With such a high port count, the traditional rack-and-stack approach with lab instruments becomes impractical. A multi-port ATE solution is required to handle so many serial ports on a single device. Major

ATE suppliers have introduced some products to address this problem, but a lot of work is still needed in cost reduction and functionality enhancements. As mentioned above, the common practice for handling very high port count is still DFT techniques and design margins. The trade off point between test and design margin varies from application to application.

3. *Cost Factor*—Traditionally, most multi-gigabit transceivers were designed as high-performance, high-priced, and high-margin devices with a low level of integration and relatively low production volume. With the introduction of low cost, low power CMOS macro cells, gigabit transceivers have become valued additions to many high-volume and low priced (even commodity) devices. In addition to high port count, a cost-efficient ATE solution that can test all serial ports concurrently is essential for production. The constant trade-off between performance and integration level results in the separation of SerDes devices into two categories: high-performance-level serial transceivers, and high-integration-level gigahertz link macro-cells. The test method for each type should be 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 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 stimulus and capture before a product becomes mature. With accelerating technology improvements, the life cycles for most products are become much shorter, therefore it may become increasingly difficult to verify and optimize DFT circuitry.
4. *Jitter Decomposition Measurement*—The jitter generated by the transmitter is the key parameter to guarantee transmitter quality. Currently, jitter measurement capability on ATE is in its infancy, there is no instrument available that simultaneously satisfies the noise floor, analog bandwidth, and test time requirements for high performance interfaces. 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 bit-error-rate for the serial link (typically 10^{-12}). 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. For example, a digital signal of 3.2 Gbps has a frequency spectrum up to 8~10 GHz. Therefore, associated jitter measurement instruments must provide a very high analog bandwidth to avoid adding erroneous data-dependent jitter to the measurement. Most existing instruments for jitter measurements take more than 20 seconds to capture jitter from a high-speed data stream. Although it is faster to measure jitter based on the device clock, many of these designs do not provide direct access to the internal clock. In that case, jitter can only be measured from the data stream. Although many jitter decomposition techniques have been introduced in the last two years by various companies, few are reliable beyond 2.5~3.2 Gbps range, and it is difficult correlate from one technique to the next.
5. *Jitter Tolerance Test*—Jitter tolerance measures the level of jitter on the input signal that the receiver can tolerate before communication quality, in terms of bit error rate (BER), is degraded. This is a key specification for receiver (Rx) noise immunity. To conduct a jitter tolerance test, jitter must be deliberately injected into the data stream in a controlled fashion. Depending on the architectures of the clocking schemes used, different type of jitter is needed to stress the receiver. Some applications need sinusoidal jitter injection, some require a combination of DJ/RJ/PJ, and others demand jitter injection in terms of peak-to-peak or rms jitter within a certain numbers of cycles. Integrated instruments that can inject all these kinds of jitter do not exist in neither the ATE world nor the lab equipment world, although they are starting to appear on test equipment suppliers' roadmaps. The problem lies in the gap between the standard definition and instrument capabilities that are widely available. As of today, some indirect measurement techniques are more practical for cost effective manufacturing test, until low-cost and integrated instruments becomes available to test jitter to the specs as they are defined. Some standards still require testing with very long patterns that a memory based pattern generation and checker are not practical.
6. *Test Fixture Bandwidth*—The test fixture used to interface the device to the instruments/ATE include a printed circuit board, cable, connector/pogo pin, etc. With frequencies growing and port counts increasing, the ability to deliver high frequency signals to instruments without much loss and distortion becomes a monumental task in test engineering. Once the signal goes beyond 10 Gb/s, the fixture bandwidth requirement reaches 20 GHz. Although, it is possible to implement this for only a few lines, it is difficult to interface several hundred such lines when the instruments are placed several feet away from the DUT. Moving the pin-electronics front-end into the ATE test head will alleviate this problem. The socket and wafer probe are also serious bandwidth bottlenecks for multi-gigahertz testing, where additional R&D is needed. Although there have been significant improvements in socket bandwidth recently, the new high bandwidth socket solutions are challenged for their limited mechanic specs. Non-reliable insertions and inconsistent contact are still the common problems causing test result variation and yield hit. Therefore, a DFT approach to measure the jitter on chip is preferred. This is yet to be developed in the multi-gigabit domain. There is

also another area of active development—providing tools to simulate the whole signal path including socket, PCB, cable and connector.

7. *Synchronization*—Most receivers for serial communications use clock and data recovery circuits (CDR) to extract the clock from a data stream. The phase of the recovered data is not necessarily fixed from part to part, or even from one reset to the next. Highly flexible timing and clocking schemes are required to accommodate this latency variation by conducting phase alignment and frame alignment. In this area, there has been some progress made recently on ATE.
8. *Parametric DFT versus Logic DFT*—To date only basic functional DFT circuits have been implemented for SerDes modules. On-chip BIST mainly consists of a built-in pseudo-random-bit-stream (PRBS) generator and a BER checker. These, however, only provide functional coverage, without parametric test capabilities (such as input and output jitter and voltage levels). However, there has been very convincing progress on the parametric DFT using the increasingly complicated pre-compensation and equalization. BIST circuits for jitter and level tests are still in the research stage.
9. *Advanced Signal Shaping And Encoding*—The adoption of more sophisticated analog techniques such as pre-emphasis, equalization, multilevel encoding and phase encoding could potentially lead the test requirement to a somewhat traditional analog test solution, but much higher in frequency. Although pre-emphasis and equalization are now becoming more popular, it is still not clear when multi-level coding will replace the predominant binary coding in this area.
10. *Receiver Input Sensitivity*—Because of signal integrity issues at GHz, the input sensitivity is not only an issue for the signal source that generates it, but also how it can be calibrated close to the DUT input. For some 10 GbE and some backplane applications, the 15~20 mV spec is challenging for high volume production test on ATE. The DC parametric test requirement is making it harder for ATE instrument manufacturers to achieve the challenging low voltage swing accuracy requirement. This problem is limited to high-end applications.

MEMORY

Memory density will continue to grow over the roadmap period according to Moore's Law. Memory, both DRAM and NAND flash, will continue to be used as a process development vehicle for new digital technologies, and microprocessors and FPGA devices are also being used. Refer to the DRAM, Flash, and Embedded memory Tables 27, 28, and 29, respectively.

COMMODITY DRAM TESTING

In the long term, DRAM bit density will quadruple every three years, but there will be some deviations to the trend as technology issues are resolved. Increasing memory size will drive higher device test time and decreasing manufacturing cell throughput unless offset by DFT. Failure detection, analysis and repair is necessary for commodity DRAMs. To enhance test productivity, new test-oriented architectures will be required. Multi-bit testing, BIST, and built-in self repair will be essential to maintain the production throughput and yield.

Considerable test parallelism from the automatic tester equipment will be required over the roadmap. The table reference to the number of devices simultaneously tested refers to the packaged devices tested at-speed. Commodity DRAMs will lag the leading edge specialty DRAMs in I/O bit rates. In the realm above 2 GBPS, there are DUT interface signal integrity issues with the test socket, probe card, and handling. Exceeding 128 devices per test head will be a challenge. I/O bit rates will be at least two times greater than the required clock and tester MHz will be up to eight times greater on performance DRAM. High I/O bit rates create a challenge on how to properly define tester accuracy in the future.

The primary fault models for DRAMs will continue to be cell stuck-at, multi-cell coupling, decoder open, and data retention faults. For 100 nm feature size and below, in-line defect detection will be necessary for product development. With inline defect monitoring, processing of defective wafers could be avoided and test time for wafer sort and package-level test could still be maintained.

Table 27a Commodity DRAM Test Requirements—Near-term Years

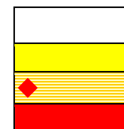
Year of Production	2005	2006	2007	2008	2009	2010	2011	2012	2013
DRAM ½ Pitch (nm) (contacted)	80	70	65	57	50	45	40	36	32
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	90	78	68	59	52	45	40	36	32
DRAM Capacity (Gbits)									
R&D	8	8	16	16	16	32	32	32	64
Mass production	1	2	2	2	4	4	4	8	8
DRAM data rate (Gbs)	0.67	0.80	1.00	1.20	1.20	1.33	1.33	1.50	1.50
Performance DRAM data rate (Gbs)	1.8	2.25	2.5	3	3.5	5	5	6	6
DRAM bit width/device (mass production)	16	16	16	16	16	16	16	16	16
Device CLK rate (GHz)	0.3	0.3	0.4	0.5	0.5	0.6	0.7	0.7	0.8
Overall timing accuracy (ps)	40	40	32	32	25	25	25	Cannot be determined	

CLK—clock signal

Table 27b Commodity DRAM Test Requirements—Long-term Years

Year of Production	2014	2015	2016	2017	2018	2019	2020
DRAM ½ Pitch (nm) (contacted)	28	25	22	20	18	16	14
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	28	25	22	20	18	16	14
DRAM Capacity (Gbits)							
R&D	64	64	128	128	128	256	256
Mass production	8	16	16	16	32	32	32
DRAM data rate (Gbs)	1.8	1.8	2.0	2.0	2.25	2.25	2.5
Performance DRAM data rate (Gbs)	8	8	10	10	12	12	14
DRAM bit width/device (mass production)	16	16	16	16	16	16	16
Device CLK rate (GHz)	0.9	0.9	1.0	1.1	1.1	1.2	1.3
Overall timing accuracy (ps)	Cannot be determined						

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



COMMODITY FLASH TESTING

Flash will double in density ever year in the short term and slow to a doubling every 1.5 years. NAND density has generally been 4× the NOR density at any given technology generation and the forecast shows that trend continuing. NAND bus width has continued to be primarily 8 bit with some products at 16 bit.

NAND and NOR generally do not have the same test solution due to architectural, usage, and bad block handling differences. Bus differences between NAND and NOR have blurred over time and the software process improvement (SPI) and other serial bus definitions are targeted to embedded applications. Further proliferation of bus types is expected due to the customization of flash for applications. Bus width is predominantly 8-bit and 16-bit on densities > 4 Mb. The 32 bit bus width NAND has been dropped from the roadmap for 2005.

Flash is commonly used in battery-powered embedded applications and supply voltage requirements of Flash have been dropping slowly over time, however, the need for internal voltages that are 3–8 times the external supply requirements is expected to continue 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 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.

Table 28a Commodity Flash Memory Test Requirements—Near-term Years

Year of Production	*2005*	*2006*	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*	*2013*
DRAM ½ Pitch (nm) (contacted)	*80*	*70*	*65*	*57*	*50*	*45*	*40*	*36*	*32*
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	*90*	*78*	*68*	*59*	*52*	*45*	*40*	*36*	*32*
NAND Device Characteristics									
Capacity (Gbits)									
R&D	**32**	**32**	**64**	**64**	**128**	**128**	**128**	**256**	**256**
Mass production	**8**	**8**	**16**	**16**	**32**	**32**	**32**	**64**	**64**
Data width (bits)	**16**	**16**	**16**	**16**	**16**	**16**	**16**	**16**	**16**
Power Supplies									
Power supply voltage range	**1.5–5.5**	**1.5–5.5**	**1.5–5.5**	**1.5–3.5**	**1.5–3.5**	**1.5–3.5**	**1.5–3.5**	**1.5–3.5**	**1.5–3.5**
Power supplies per device	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**2**
Maximum current (MA)	**35**	**35**	**35**	**35**	**35**	**35**	**35**	**35**	**35**
Pattern Generator									
Tester channels per device	**24**	**24**	**24**	**24**	**24**	**24**	**24**	**24**	**24**
Timing									
Maximum I/O data rate (Mbs)	**40**	**40**	**50**	**50**	**50**	**55**	**55**	**55**	**60**
Accuracy OTA (ns)	**1.2**	**1.2**	**1**	**1**	**1**	**0.9**	**0.9**	**0.9**	**0.8**
NOR Device Characteristics									
Capacity (Gbits)									
R&D	**4**	**4**	**8**	**8**	**16**	**16**	**16**	**32**	**32**
Mass production	**1**	**1**	**2**	**2**	**4**	**4**	**4**	**8**	**8**
Data width (bits)	**16**	**16**	**16**	**32**	**32**	**32**	**32**	**32**	**32**
Power supply voltage range	**1.0–5.5**	**1.0–5.5**	**1.0–5.5**	**1.0–5.5**	**0.9–3.3**	**0.9–3.3**	**0.9–3.3**	**0.9–3.3**	**0.9–3.3**
Power supplies per device***	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**2**
Maximum current (MA)	**150**	**150**	**150**	**150**	**150**	**150**	**150**	**150**	**150**
Tester channels per test site	**64**	**72**	**72**	**72**	**72**	**72**	**72**	**72**	**72**
Maximum I/O data rate (Mbs)	**166**	**200**	**200**	**200**	**266**	**266**	**266**	**333**	**333**
Accuracy OTA (ns)	**0.3**	**0.2**	**0.2**	**0.2**	**0.18**	**0.18**	**0.18**	**0.15**	**0.15**

OTA—overall ATE timing accuracy

**** Discrete only*

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

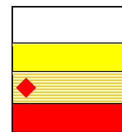
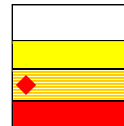


Table 28b Commodity Flash Memory Test Requirements—Long-term Years

Year of Production	*2014*	*2015*	*2016*	*2017*	*2018*	*2019*	*2020*
DRAM ½ Pitch (nm) (contacted)	*28*	*25*	*22*	*20*	*18*	*16*	*14*
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	*28*	*25*	*22*	*20*	*18*	*16*	*14*
NAND Device Characteristics							
Capacity (Gbits)							
R&D	**256**	**512**	**512**	**512**	**1024**	**1024**	**1024**
Mass production	**64**	**128**	**128**	**128**	**256**	**256**	**256**
Data width (bits)	**16**	**16**	**16**	**16**	**16**	**16**	**16**
Power Supplies							
Power supply voltage range	**1.5–3.5**	**1.5–3.5**	**1.5–3.5**	**1.5–3.5**	**1.5–3.5**	**1.5–3.5**	**1.5–3.5**
Power supplies per device	**2**	**2**	**2**	**2**	**2**	**2**	**2**
Maximum current (MA)	**35**	**35**	**35**	**35**	**35**	**35**	**35**
Pattern Generator							
Tester channels per device	**24**	**24**	**24**	**24**	**24**	**24**	**24**
Timing							
Maximum I/O data rate (Mbs)	**60**	**60**	**60**	**70**	**70**	**70**	**70**
Accuracy OTA (ns)	**0.8**	**0.8**	**0.8**	**0.7**	**0.7**	**0.7**	**0.7**
NOR Device Characteristics							
Capacity (Gbits)							
R&D	**32**	**64**	**64**	**64**	**128**	**128**	**128**
Mass production	**8**	**16**	**16**	**16**	**32**	**32**	**32**
Data width (bits)	**32**	**32**	**32**	**32**	**32**	**32**	**32**
Power supply voltage range	**0.9–3.3**	**0.9–3.3**	**0.9–3.3**	**0.9–3.3**	**0.9–3.3**	**0.9–3.3**	**0.9–3.3**
Power supplies per device***	**2**	**2**	**2**	**2**	**2**	**2**	**2**
Maximum current (MA)	**150**	**150**	**150**	**150**	**150**	**150**	**150**
Tester channels per test site	**72**	**72**	**72**	**72**	**72**	**72**	**72**
Maximum I/O data rate (Mbs)	**333**	**400**	**400**	**400**	**400**	**400**	**400**
Accuracy OTA (ns)	**0.15**	**0.12**	**0.12**	**0.12**	**0.12**	**0.12**	**0.12**

***** *Discrete only*

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



EMBEDDED DRAM AND FLASH TESTING

Embedded DRAM bits will double every two years in the near term and slow to 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, inline defect detection will be necessary for product development. With inline defect monitoring, processing of defective wafers will be avoided and test time for wafer sort and package level test will be maintained.

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 from 2003 and onwards.

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 background will be essential for embedded DRAM.

Considerable parallelism in test will be required to maintain test throughput in the face of rising memory densities. It is expected that by the year 2006 and onwards, test will become 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 DRAMs could be tested and repaired on the memory tester, while the logic blocks could still be tested on the logic tester. Embedded SRAM test requirements are captured in the High Performance Microprocessor section of the chapter.

Table 29a Embedded Memory (DRAM and Flash) Test Requirements—Near-term Years

Year of Production	*2005*	*2006*	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*	*2013*
DRAM ½ Pitch (nm) (contacted)	*80*	*70*	*65*	*57*	*50*	*45*	*40*	*36*	*32*
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	*90*	*78*	*68*	*59*	*52*	*45*	*40*	*36*	*32*
Embedded DRAM									
Embedded DRAM size (Mbits) *									
R&D	**256**	**512**	**512**	**512**	**1024**	**1024**	**1024**	**2048**	**2048**
Mass production	**128**	**256**	**256**	**256**	**512**	**512**	**512**	**1024**	**1024**
Failure concerns	**Particle defects; array noise, data retention**					**Particle defects, array noise, sense-amp imbalance**			
Wafer level test	**Single and double insertion**					**In-line defect detection**			
Usage of on-chip test	**100% BIST/BISR**					**100% BIST/100% BISR**			
Embedded Flash									
Embedded flash size (Mbits) *									
R&D	**64**	**128**	**128**	**128**	**256**	**256**	**256**	**512**	**512**
Mass production	**32**	**64**	**64**	**64**	**128**	**128**	**128**	**256**	**256**
Embedded mixed memory size (Mbits) *									
Flash	**32**	**64**	**64**	**64**	**128**	**128**	**128**	**256**	**256**
DRAM	**32**	**64**	**64**	**64**	**128**	**128**	**128**	**256**	**256**
Failure concerns	**Oxide defects; ONO scaling; over-erase**					**Sense-amp imbalance**			
Wafer level test	**Single and double insertion**					**In-line defect detection**			
Usage of on-chip test	**BIST/BIST/DAT**					**BIST/BISR**			

DAT—direct access DFT

Number of bits in mass production is approximately 50% of number of bits in R&D

** Solution space is both on-chip and stacked die*

Manufacturable solutions exist, and are being optimized

Manufacturable solutions are known

Interim solutions are known

Manufacturable solutions are NOT known

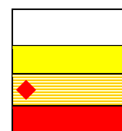
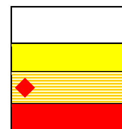


Table 29b Embedded Memory (DRAM and Flash) Test Requirements—Long-term Years

Year of Production	*2014*	*2015*	*2016*	*2017*	*2018*	*2019*	*2020*
DRAM ½ Pitch (nm) (contacted)	*28*	*25*	*22*	*20*	*18*	*16*	*14*
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	*28*	*25*	*22*	*20*	*18*	*16*	*14*
Embedded DRAM							
Embedded DRAM size (Mbits)							
R&D	**2048**	**4096**	**4096**	**4096**	**8184**	**8192**	**8192**
Mass production	**1024**	**2048**	**2048**	**2048**	**4092**	**4096**	**4096**
Failure concerns	**Particle defects, array noise, sense-amp imbalance**						
Wafer level test	**In-line defect detection**						
Usage of on-chip test	**100% BIST/100% BISR**						
Embedded Flash							
Embedded flash size (Mbits)							
R&D	**512**	**1024**	**1024**	**1024**	**2046**	**2048**	**2048**
Mass production	**256**	**512**	**512**	**512**	**1023**	**1024**	**1024**
Embedded mixed memory size (Mbits)							
Flash	**256**	**512**	**512**	**512**	**1023**	**1024**	**1024**
DRAM	**256**	**512**	**512**	**512**	**1023**	**1024**	**1024**
Failure concerns	**Sense-amp imbalance**						
Wafer level test	**In-line defect detection**						
Usage of on-chip test	**BIST/BISR**						

* *Solution space is both on-chip and stacked die*

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



ANALOG AND MIXED-SIGNAL

The trend toward more system functionality on a single die (SOC) or in a single package (SIP) will increasingly blur the lines between traditional digital, analog, RF/microwave and mixed-signal devices. This trend will drive test equipment toward a single platform solution that can test any application. Consequently, the ATE must be modular and expandable across the entire spectrum from digital-only to the full integration of high performance DC/analog/RF/microwave instruments. In previous years the Test chapter section for Mixed-signal included RF. Due to increasing presence and complexity, the RF roadmap has been moved to a separate section. The digital, high speed serial and RF requirements for mixed-signal chips are equivalent to the requirements contained in those respective sections.

The analog/mixed-signal test issues and test technology limiters are higher instrument bandwidth, higher arbitrary waveform generator sampling rates; lower noise floors; and seamless integration of digital and all other instruments, optimized for greater throughput and ease of use into the ATE environment.

The Mixed-Signal Test Requirements table focuses on test instruments rather than specific chip applications. Current analog/mixed-signal testing methodologies require performance-based measurements (that is, using external, outside-the-chip, instruments); therefore, instrumentation requirements reflect the increasing chip performance predicted by the process technology roadmaps. The analog waveform generation and capture requirements are set in two classes: low frequency—basic/minimum requirement 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.

The complexity and breadth of applications demands specialized instruments optimized for a particular chip application. The bandwidth, resolution and noise-floor requirements for these optimized instruments will be in a range bounded by the low frequency and high-end requirements. Often, more than one complex analog function is integrated into a single chip. Instruments for ATE are chasing ever-increasing test requirements, especially for arbitrary waveform generators. The result of this trend is an increased number of instruments in a given test system, which increases cost and creates

significant configuration-management issues for equipment that must be shared across multiple products. This trend of increasing instrument numbers, complexity, and performance is expected to continue, but while 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 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.

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 load boards 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.
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.
3. Currently, most of the analog/mixed-signal portion of an ATE test program is manually generated. Automatic test pattern generators are widely used for generating digital vector patterns. Automated test program generators have recently become available to generate an entire digital-only test program. These tools need to be extended into analog/mixed-signal by supporting such tasks as digital signal generation and capture, integration of these signals into the digital patterns, complete instrument setups and circuit simulation of the chip's analog output along with the load board and test instruments.

Table 30a Mixed-signal Test Requirements—Near-term Years

Year of Production	*2005*	*2006*	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*	*2013*	*2014*
DRAM ½ Pitch (nm) (contacted)	*80*	*70*	*65*	*57*	*50*	*45*	*40*	*36*	*32*	*28*
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	*90*	*78*	*68*	*59*	*52*	*45*	*40*	*36*	*32*	*90*
MPU Physical Gate Length (nm)	*32*	*28*	*25*	*22*	*20*	*18*	*16*	*14*	*13*	*11*
Low Frequency Waveform										
BW (MHz)	**50**	**50**	**50**	**75**	**75**	**75**	**100**	**100**	**100**	**100**
Sample rate (MS/s)	**Moving from Nyquist sample rates to over/under sampling sources/digitizers**									
Resolution (bits)	**DSP computation to 24 bits – however effective number of bits will be limited by noise fllor**									
Noise floor (dB/RT Hz)	**-155**	**-155**	**-155**	**-160**	**-160**	**-160**	**-165**	**-165**	**-165**	**-165**
Very High Frequency Waveform Source										
Level V (pk–pk)	**4**	**4**	**4**	**4**	**Driven by communication physical layer standards – likely to be lower**					
Accuracy (±)	**0.50%**	**0.50%**	**0.50%**	**0.50%**	**Likely to remain the same**					
BW (MHz)	**1500**	**1500**	**1500**	**1800**	**2050**	**2250**	**2700**	**3000**	**3000**	**3000**
Sample rate (MS/s)	**6000**	**6000**	**6000**	**7200**	**8200**	**9400**	**10,800**	**12000**	**12000**	**12000**
Resolution (bits) AWG/Sine†	**8/10**	**8/10**	**8/10**	**8/10**	**8/10**	**8/10**	**8/10**	**10/12**	**10/12**	**10/12**
Noise floor (dB/RT Hz)	**-135**	**-135**	**-135**	**-140**	**-140**	**-140**	**-140**	**-145**	**-145**	**-145**
Very High Frequency Waveform Digitizer										
Level V (pk–pk)	**4**	**4**	**4**	**4**	**Driven by communication physical layer standards – likely to be lower**					
Accuracy (±)	**0.50%**	**0.50%**	**0.50%**	**0.50%**	**Likely to remain the same**					
BW (MHz) (undersampled)	**6400**	**8000**	**9200**	**10800**	**10800**	**12500**	**12500**	**15000**	**15000**	**15000**
Sample rate (MS/s)	**Direct conversion <400 MS/s**							**Direct conversion <600 MS/s**		
Resolution (bits)	**Minimum 12 bits – noise floor is more important**							**Minimum 14 bits – noise floor is more important**		
Noise floor (dB/RT Hz)	**-145**	**-145**	**-145**	**-145**	**-145**	**-145**	**-145**	**-150**	**-150**	**-150**
Time Measurement										
Jitter measurement (ps RMS)	**Will be driven by high-speed serial communication ports**									
Frequency measurement (MHz)	**Will be driven by high-performance ASIC clock rates**									
Single shot time capability (ps)	**Will be driven by high-speed serial communication ports**									
RF/Microwave instrumentation	**Same as RF test requirements – see this new section**									
Special Digital Capabilities										
D/A and A/D digital data rate (MB/s)	**Same as high performance ASIC "off-chip data rate"**									
Sample clock jitter (ps RMS)	**<0.2**	**<0.2**	**<0.2**	**<0.1**	**<0.1**	**<0.1**	**<0.1**	**<0.1**	**<0.1**	**<0.1**

AWG—array waveguide gratings BW—bandwidth pk–pk—peak to peak

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

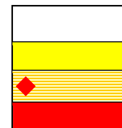
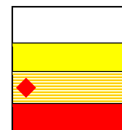


Table 30b Mixed-signal Test Requirements—Long-term Years

Year of Production	*2014*	*2015*	*2016*	*2017*	*2018*	*2019*	*2020*
DRAM ½ Pitch (nm) (contacted)	*28*	*25*	*22*	*20*	*18*	*16*	*14*
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	*28*	*25*	*22*	*20*	*18*	*16*	*14*
MPU Physical Gate Length (nm)	*11*	*10*	*9*	*8*	*7*	*6*	*6*
Low Frequency Waveform							
BW (MHz)	**100**	**100**	**100**	**100**	**100**	**100**	**100**
Sample rate (MS/s)	**Moving from Nyquist sample rates to over/under sampling sources/digitizers**						
Resolution (bits)	**DSP computation to 24 bits – however effective number of bits will be limited by noise fllor**						
Noise floor (dB/RT Hz)	**-165**	**-165**	**-165**	**-165**	**-165**	**-165**	**-165**
Very High Frequency Waveform Source							
Level V (pk–pk)	**Driven by communication physical layer standards – likely to be lower**						
Accuracy (±)	**Likely to remain the same**						
BW (MHz)	**3000**	**3000**	**3750**	**3750**	**3750**	**3750**	**3750**
Sample rate (MS/s)	**12000**	**12000**	**15000**	**15000**	**15000**	**15000**	**15000**
Resolution (bits) AWG/Sine†	**10/12**	**10/12**	**10/12**	**10/12**	**10/12**	**10/13**	**10/14**
Noise floor (dB/RT Hz)	**-145**	**-145**	**-145**	**-145**	**-145**	**-144**	**-143**
Very High Frequency Waveform Digitizer							
Level V (pk–pk)	**Driven by communication physical layer standards – likely to be lower**						
Accuracy (±)	**Likely to remain the same**						
BW (MHz) (undersampled)	**15000**	**15000**	**15000**	**15000**	**15000**	**15000**	**15000**
Sample rate (MS/s)	**Direct conversion <600 MS/s**						
Resolution (bits)	**Minimum 14 bits – noise floor is more important**						
Noise floor (dB/RT Hz)	**-150**	**-150**	**-150**	**-150**	**-150**	**-150**	**-150**
Time Measurement							
Jitter measurement (ps RMS)	**Will be driven by high-speed serial communication ports**						
Frequency measurement (MHz)	**Will be driven by high-performance ASIC clock rates**						
Single shot time capability (ps)	**Will be driven by hihg-speed serial communication ports**						
RF/Microwave instrumentation	**Same as RF test requirements – see this new section**						
Special Digital Capabilities							
D/A and A/D digital data rate (MB/s)	**Same as high performance ASIC "off-chip data rate"**						
Sample clock jitter (ps RMS)	**<0.1**	**<0.1**	**<0.1**	**<0.1**	**<0.1**	**<0.1**	**<0.1**

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



Definitions for Mixed-signal Test Requirements Table 30:
Low Frequency Source and Digitizer—This is the basic, minimum, instrument set of any mixed-signal tester. Telecommunications, advanced audio and wireless baseband will drive these specifications. Differential inputs/outputs are required.
Very High Frequency Waveform Source—Disk storage applications will drive sample rate and bandwidth. Local area network devices will drive sample rate, resolution and amplitude accuracy. Differential outputs are required.
Very High Frequency Waveform Digitizer—Undersampled (down conversion, track-and-hold, etc) bandwidth is shown. The sample rates and bit resolutions are for a direct conversion digitizer, which is usually preceded by the undersampler. Storage and network devices will drive digitizer specifications. Differential inputs are required.
Special Digital Capabilities—For converter testing, the ability to source a digital signal to a D/A and capture a digital signal from an A/D.

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 6–45 GHz band is focused on medium distance communications (Bluetooth moving to ultra wide band (UWB)). The high section: 45-77 GHz is dedicated just to short distance radar applications, particularly automotive.

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 e-sort 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.

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 itself 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. Further improvements may not be needed for some time.
4. The visibility of trends in the high RF ranges (> 6 GHz) is not clear at the moment.

RELIABILITY TECHNOLOGY REQUIREMENTS

Reliability solutions are an optimization of 1) reliability defect density (RDD), 2) learning, reliability screens and test methods (RS&TM) applications, and 3) design for reliability (DFR). The goal of the reliability solution optimization is to provide the best value for the reliability dollar spent, where value is defined as the ratio of customer satisfaction to customer cost.

In reliability circles, customer satisfaction is measured by the field failure rate or failures in time (FITs). The cost of reliability has two components: manufacturing operations costs and yield. As such, these two components of the reliability cost equation are the primary challenges facing every reliability solution provider. In turn, manufacturing operations costs are also driven by two fundamental components—burn in duration and equipment sophistication. The industry is still searching for a means to accelerate latent defects outside of the traditional elevated voltage and temperature methods. It follows that much progress has been made in detection techniques, but acceleration remains all about applying elevated voltage and temperature.

Applying voltages and temperatures in excess of the application specs most often demand providing solutions to leakage induced power (electrical power delivery and thermal/heat dissipation). The component of reliability cost reduction associated with yield is severely biased towards elimination of "overkill"/"false rejects," which in many ways are tied to derivatives of the power solution. However, the primary source of false rejects stems back to the stress methodology, through the modeling assumptions, and ultimately finds its root in what one believes about "escapes" from the manufacturing stress process. The argument proceeds like this: the majority of market applications are most concerned with the early life component of the failure rate. Most latent defects that "escape" acceleration will fail early in the product life. In order to best guarantee a part received stimulus—and therefore did not "escape" stress—is simply to

measure the outputs during stress. Defining terms, measuring outputs is called *in situ stress*, while measuring no outputs is *dynamic stress*. Obviously the "escapes" component is less for *in situ*, and hence the early life failure rate is lower. As anticipated however, this lower failure rate does not come without cost. *In situ* stress requires functionality at stress conditions, which in turn shrinks the performance distribution. Completing the original thread, measuring outputs during stress also introduces a component of yield loss. Due to process variation, some portion of the distribution does not have sufficient margin to function at stress voltages or temperatures, however these same parts operate fine at application conditions. Although these parts contain no reliability defects, in situ stress will fail these perfectly functional parts—hence "over-kill." These same parts with "marginal margin" are the target of the advances in detection techniques mentioned earlier. Achieving reliability requires trade-offs. In most instances performance and yield hang in the balance.

Reliability defect density learning rate is the most cost effective means of achieving the reliability demands of the marketplace. In itself, it is the by-product of the fundamental core practice in achieving profitability in microelectronics = yield learning rate. Defect learning is addressed in Defect Modeling and Physical Defects section—and although historical data has overwhelmingly supported the premise that the component of defects that are "reliability unique" has been small—recent advances in technology may be changing the picture. The section on defect learning will always be directly applicable to RDD learning, however the high voltages and temperatures of defect acceleration are causing us to peer over the edge of device physics and materials science. Stress conditions are no longer dictated by "technology nominal" specs but by system application conditions. Technology's recent inability to meet marketplace performance demands at reasonable power has forced systems designers to increase system application conditions (voltage and temperature) to compensate. Shifts in array V_{min} operating range, NBTI-driven performance margin, and gate oxide integrity (time dependent dielectric breakdown (TDDB)) as a result of the application of stress conditions still remain largely unexplained. As such, they dictate compensatory actions and/or reliability failure rate modifications. Even the standard thinking of metal electromigration for C_4 and BEOL wiring requires careful scrutiny when confronted with the 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 in the forms of error detection/correction, 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 wearout were to be avoided. The adage in the past was "one must be able to accelerate defects while avoiding the onset of wearout". However this is becoming increasingly more difficult in the face of stretched system applications conditions; sub-100 nm oxides; NBTI; marginal margin (that is, array V_{min}); hundreds of amps and Watts, miles of copper wire, and jillions 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.

Non-traditional approaches to burn-in include system level burn-in, wafer level burn-in, and strip/array burn-in. On high reliability applications, system level burn-in complements or replaces traditional device level burn-in. Wafer level burn-in technology continues to be developed, but has not been able yet to make significant inroads against traditional packaged level burn-in. The challenge here is to eliminate socketed burn-in and find ways to perform simultaneous multiple wafer level burn-in using scan/logic and memory BIST (MemBIST). Strip/array burn-in is becoming more important as more packages are receiving massively parallel test in either strip or array format.

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 the application of high enough temperature over enough 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-controller or SOC through DFT functions such as scan or BIST.

Key challenges are to quantify how to measure the effectiveness of these options and to develop standards that define 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 environment for wafers 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 environment as a massively parallel testing opportunity.

The need for WLBI is increasing. Infant mortality rate is getting worse due to transistor scaling effects and new processing technology / materials for devices. Decreasing operating voltages and margins for devices are reducing the ability to use voltage acceleration / voltage stress testing to guarantee reliability. KGD is becoming 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, detection and removal of defective devices prior to the packaging process eliminates packaging scrap costs based on intrinsic device defects.

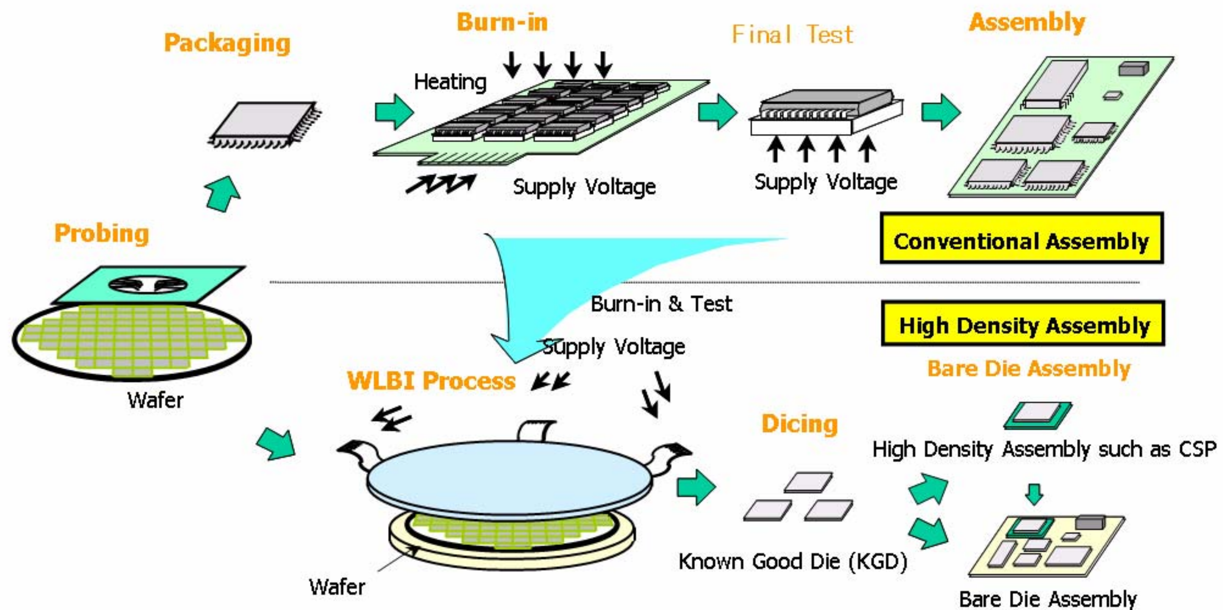


Figure 33 The Production Process with WLBI Compared with Package Burn-in

PROBING TECHNOLOGY FOR WLBI

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 efficiently concentrate pressure 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.

Micro pogo-pin contactor consists of a CTE matched probe housing and pogo-pins with moving plungers at both sides. The pogo-pins stand vertically and have enough compliance (maximum of 0.5 mm) and independent travel to accommodate height variations between adjacent contacts. The probe pitch is 0.14 mm to 1.27 mm, with inductance of 0.48 nH. The target materials on the device are solder bump, Al, Al-Cu, and under-bump metallurgy (UBM). Pin location arrangement has no limit. It is possible to clean the plunger tip and to change each pogo individually.

Other contactors such as spring-functioned material on wafer pads as a wafer level package technology are expected for whole wafer contactor usage. For contactor roadmaps, DRAM is selected as the target application due to its large predominance in general memory burn-in. DFT is considered for system LSI.

OTHER WLBI TECHNOLOGY CONSIDERATIONS

The current consumption of a wafer increases by sub-threshold leakage from shorter transistor channel lengths and an increased number of transistors per unit area. The high temperature of burn-in also increases sub-threshold leakage. Therefore, the burn-in equipment must be capable of supplying over 1000 A of current per wafer. Also, to appropriately manage current, wafer temperature control/uniformity becomes necessary. Finally, the burn-in equipment must be able to accommodate different quality distributions across each wafer.

BIST is capable of decreasing number of pins under test per device, but die shrinks and tighter pad pitches more than offset this advantage by increasing the total number of die and pads per wafer. The increased number of pins being tested also increases the force required to contact the wafer. In order to enable the use of WLBI through DFT functions such as

scan, BIST, and JTAG[3], the number of tested pins per device and total cost per device must be decreased and performance of the WLBI technology must be improved.

The probing technology for WLBI is described above. However, probing technology faces several challenges in order to meet the technology trend for the future. When the probe pitch is studied with TPS contactor technology for the devices with LOC and peripheral, a pitch around 70 µm seems feasible. For a pitch less than 70 µm, MEMS technology by use of photolithography is an option. This technology, however, does not yet have a solution for 300 mm wafers. While probing technology for tighter pitches is required, the intelligent use of DFT during pad layout may provide some relief by bypassing every other pad in order to effectively double the probe pitch as compared to pad pitch. Application to high pin count and low force probing due to low-κ materials will also be required. This will help drive new probing technology.

[3] *an IEEE standard 1149 boundary scan*

Table 31a Burn-in Requirements—Near-term Years

Year of Production	*2005*	*2006*	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*	*2013*
DRAM ½ Pitch (nm) (contacted)	*80*	*70*	*65*	*57*	*50*	*45*	*40*	*36*	*32*
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	*90*	*78*	*68*	*59*	*52*	*45*	*40*	*36*	*32*
MPU Physical Gate Length (nm)	*32*	*28*	*25*	*22*	*20*	*18*	*16*	*14*	*13*
Clock input frequency (MHz)	400	400	400	400	400	400	400	400	400
Off-chip data frequency (MHz)	75	75	75	75	75	75	75	75	75
Power dissipation (W per DUT)	600	600	600	600	600	600	600	600	600
Power Supply Voltage Range (V)									
High-performance ASIC / microprocessor / graphics processor	0.5–2.5	0.5–2.5	0.5–2.5	0.5–2.5	0.5–2.5	0.5–2.5	0.5-2.5	0.5–2.5	0.5–2.5
Low-end microcontroller	0.7–10.0	0.7–10.0	0.7–10.0	0.7–10.0	0.7–10.0	0.5–10	0.5–10	0.5–10	0.5–10
Mixed-signal	0.5–500	0.5–500	0.5–500	0.5–500	0.5–500	0.5–500	0.5–500	0.5–500	0.5–500
Maximum Number of Signal I/O									
High-performance ASIC	384	384	384	384	384	384	384	384	384
High-performance microprocessor / graphics processor / mixed-signal	128	128	128	128	128	128	128	128	128
Commodity memory	72	72	72	72	72	72	72	72	72
Maximum Current (A)									
High-performance microprocessor	400	450	450	450	450	450	450	450	450
High-performance graphics processor	40	80	100	150	200	200	200	200	200
Mixed-signal	20	20	20	20	20	30	30	30	30
Burn-in Socket									
Pin count	2500	3000	3000	3000	3000	3000	3000	3000	3000
Pitch (mm)	0.4	0.4	0.3	0.3	0.3	0.2	0.2	0.2	0.2
Power consumption (A/Pin)	2.0	3.0	3.0	4.0	4.0	5.0	5.0	5.0	5.0
Wafer Level Burn-In									
Maximum burn-in temperature (ºC)	150±3	150±3	175±3	175±3	175±3	175±3	175±3	175±3	175±3
Pad Layout – Linear									
Minimum pad pitch (µm)	80	80	65	65	65	65	65	65	65
Minimum pad size (µm)	60	60	50	50	50	50	50	50	50
Maximum number of probes	70k	70k	70k	70k	70k	70k	70k	70k	70k
Pad Layout – Periphery, Area Array									
Minimum pad pitch (µm) *1	100	100	100	80	80	80	80	80	80
Minimum pad size (µm)	45	40	40	35	35	35	35	30	30
Maximum number of probes	125k	125k	150k	150k	150k	150k	150k	150k	150k
Power consumption (W/DUT – Low-end microcontroller, DFT/BIST SOC *2)	10	10	10	10	10	20	20	20	20
Vector memory depth (M vectors – DFT/BIST SOC *2)	16	32	32	64	64	64	64	64	64

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

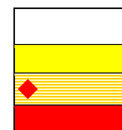
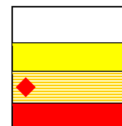


Table 31b Burn-in Requirements—Long-term Years

Year of Production	*2014*	*2015*	*2016*	*2017*	*2018*	*2019*	*2020*
DRAM ½ Pitch (nm) (contacted)	*28*	*25*	*22*	*20*	*18*	*16*	*14*
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	*28*	*25*	*22*	*20*	*18*	*16*	*14*
MPU Physical Gate Length (nm)	*11*	*10*	*9*	*8*	*7*	*6*	*6*
Clock input frequency (MHz)	**400**	**400**	**400**	**400**	**400**	**400**	**400**
Off-chip data frequency (MHz)	**75**	**75**	**75**	**75**	**75**	**75**	**75**
Power dissipation (W per DUT)	**600**	**600**	**600**	**600**	**600**	**600**	**600**
Power Supply Voltage Range (V)							
High-performance ASIC / microprocessor / graphics processor	**0.5–2.5**	**0.5–2.5**	**0.5–2.5**	**0.5–2.5**	**0.4–2.5**	**0.4–2.5**	**0.4–2.5**
Low-end microcontroller	**0.5–10**	**0.5–10**	**0.5–10**	**0.5–10**	**0.5–10**	**0.5–10**	**0.5–10**
Mixed-signal	**0.5–500**	**0.5–1000**	**0.5–1000**	**0.5–1000**	**0.5–1000**	**0.5–1000**	**0.5–1000**
Maximum Number of Signal I/O							
High-performance ASIC	**384**	**384**	**384**	**384**	**384**	**384**	**384**
High-performance microprocessor / graphics processor / mixed-signal	**128**	**128**	**128**	**128**	**128**	**128**	**128**
Commodity memory	**72**	**72**	**72**	**72**	**72**	**72**	**72**
Maximum Current (A)							
High-performance microprocessor	**450**	**450**	**450**	**450**	**450**	**450**	**450**
High-performance graphics processor	**200**	**200**	**200**	**200**	**200**	**200**	**200**
Mixed-signal	**30**	**30**	**30**	**30**	**30**	**30**	**30**
Burn-in Socket							
Pin count	**3000**	**3000**	**3000**	**3000**	**3000**	**3000**	**3000**
Pitch (mm)	**0.2**	**0.1**	**0.1**	**0.1**	**0.08**	**0.08**	**0.08**
Power consumption (A/Pin)	**5.0**	**5.0**	**6.0**	**6.0**	**6.0**	**6.0**	**6.0**
Wafer Level Burn-In							
Maximum burn-in temperature (ºC)	**175±3**	**175±3**	**175±3**	**175±3**	**175±3**	**175±3**	**175±3**
Pad Layout - Linear							
Minimum pad pitch (µm)	**65**	**50**	**50**	**50**	**50**	**50**	**50**
Minimum pad size (µm)	**50**	**40**	**40**	**40**	**40**	**40**	**40**
Maximum number of probes	**70k**	**140k**	**140k**	**140k**	**140k**	**140k**	**140k**
Pad Layout – Periphery, Area Array							
Minimum pad pitch (µm) *1	**80**	**60**	**60**	**60**	**60**	**60**	**60**
Minimum pad size (µm)	**30**	**25**	**25**	**25**	**25**	**25**	**25**
Maximum number of probes	**150k**	**300k**	**300k**	**300k**	**300k**	**300k**	**300k**
Power consumption (W/DUT – low-end microcontroller, DFT/BIST SOC *2)	**20**	**20**	**20**	**20**	**20**	**20**	**20**
Vector memory depth (M vectors – DFT/BIST SOC *2)	**64**	**128**	**128**	**128**	**256**	**256**	**256**

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



TEST HANDLER AND PROBER TECHNOLOGY REQUIREMENTS

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.

Increased parallelism at wafer probe drives a greater span of probes across the wafer surface ultimately resulting in full wafer test across a 300 mm wafer. The increased probe count is driving interface complexity to route signals. Prober and probe card architecture will need to evolve to simplify the interface.

Reducing the cost of sort and 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 short test times), faster handler speed or process improvements such as asynchronous test or continuous lot processing.

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, handlers will need the capability to very accurately pick and place, small, fragile parts, yet apply similar or increasing insertion force without inducing damage.

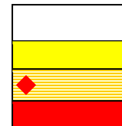
Temperature ranges are expanding to meet more stringent end use conditions and there is a need for better control of the junction temperature at the start of test. Power dissipation overall appears to be increasing, but multi-core technology is offering relief in some areas.

It is unlikely that there will be one handler that is all things to all users: integration of all of the technology to meet thermal, throughput, placement accuracy and special handling needs while being cost effective in competitive environment is a significant challenge.

Table 32a Handler (Memory—Pick and Place) Requirements—Near-term Years

Year of Production	*2005*	*2006*	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*	*2013*
DRAM ½ Pitch (nm) (contacted)	*80*	*70*	*65*	*57*	*50*	*45*	*40*	*36*	*32*
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	*90*	*78*	*68*	*59*	*52*	*45*	*40*	*36*	*32*
MPU Physical Gate Length (nm)	*32*	*28*	*25*	*22*	*20*	*18*	*16*	*14*	*13*
Parallel testing	**64–128**	**128–256**	**128–256**	**128–256**	**128-256**	**128–256**	**128–256**	**128–512**	**128–512**
Index time (S)	**3–5**	**3–5**	**2–5**	**2–5**	**2–4**	**2–4**	**2–4**	**2–4**	**2–4**
Throughput (devices per hour)	**8–10K**	**8–10K**	**8–10K**	**8–12K**	**8–12K**	**12–20K**	**12–20K**	**12–20K**	**12–20K**
Sorting	**5–9**	**5–9**	**5–9**	**5–9**	**5–9**	**5–9**	**5–9**	**5–9**	**5–9**
Temperature set point range (°C)	**-55 to 155**	**-55 to 175**	**-55 to 175**	**-55 to 175**	**-55 to 175**	**-55 to 175**	**-55 to 175**	**-55 to 175**	**-55 to 175**
Temperature accuracy (°C)	**±2**	**±2**	**±2**	**±2**	**±1.5**	**±1.5**	**±1.5**	**±1.5**	**±1.5**
Number of pins/device	**40–250**	**40–250**		**6–400**	**40–400**	**6–400**	**40–400**	**6–400**	**6–400**
Pin pitch (mm)	**0.5–1.0**	**0.4–1.0**	**0.4–1.0**	**0.3–1.0**	**0.3–1.0**	**0.3–1.0**	**0.3–1.0**	**0.3–1.0**	**0.3–1.0**
Ball edge to package edge clearance (mm)	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Minimum package thickness (mm)	**0.3–1.8**	**0.3–1.8**	**0.3–1.8**	**0.2–1.8**	**0.2–1.8**	**0.2–1.8 (kitless)**	**0.2–1.8 (kitless)**	**0.2–1.8 (kitless)**	**0.2–1.8 (kitless)**
Conversion time (minutes)	**40**	**1 (kitless)**	**1 (kitless)**	**1 (kitless)**	**1 (kitless)**	**1 (kitless) GEM–HSEM**	**1 (kitless) GEM–HSEM**	**1 (kitless) GEM–HSEM**	**1 (kitless) GEM–HSEM**

GEM—generic equipment model

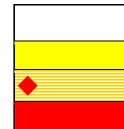


Manufacturable solutions exist, and are being optimized
Manufacturable solutions are known
Interim solutions are known
Manufacturable solutions are NOT known

Table 32b Handler (Memory—Pick and Place) Requirements—Long-term Years

Year of Production	*2014*	*2015*	*2016*	*2017*	*2018*	*2019*	*2020*
DRAM ½ Pitch (nm) (contacted)	*28*	*25*	*22*	*20*	*18*	*16*	*14*
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	*28*	*25*	*22*	*20*	*18*	*16*	*14*
MPU Physical Gate Length (nm)	*11*	*10*	*9*	*8*	*7*	*6*	*6*
Parallel testing	128 - 512	128–512	128–512	128 - 512	128–512	128 - 512	128–512
Index time (S)	2–4	2–4	2–4	2–4	2–4	2–4	2–4
Throughput (devices per hour)	12–20K	12–20K	12–20K	12–20K	12–20K	12–20K	12–20K
Sorting	5–9	5–9	5–9	5–9	5–9	5–9	5–9
Temperature set point range (°C)	-55 to 175	-55 to 175	-55 to 175	-55 to 175	-55 to 175	-55 to 175	-55 to 175
Temperature accuracy (°C)	±1.5	±1.5	±1.5	±1.5	±1.5	±1.5	±1.5
Number of pins/device	6–400	6–400	6–400	6–400	6–400	6–400	6–400
Pin pitch (mm)	0.3–1.0	0.3–1.0	0.3–1.0	0.3–1.0	0.3–1.0	0.3–1.0	0.3–1.0
Ball edge to package edge clearance (mm)	0	0	0	0	0	0	0
Minimum package thickness (mm)	0.2–1.8 (kitless)	0.2–1.8 (kitless)	0.2–1.8 (kitless)	0.2–1.8 (kitless)	0.2–1.8 (kitless)	0.2–1.8 (kitless)	0.2–1.8 (kitless)
Conversion time (minutes)	1 (kitless) GEM–HSEM	1 (kitless) GEM–HSEM	1 (kitless) GEM–HSEM	1 (kitless) GEM–HSEM	1 (kitless) GEM–HSEM	1 (kitless) GEM–HSEM	1 (kitless) GEM–HSEM

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



Notes for Tables 32a and b:

Index time was done from test end signal reception from tester to the test start signal transmission of handler.

Units per hour (UPH) calculated with zero-second test time and no lot-size generated interruptions.

Sort is number of stackable JEDEC tray sleeves used for output of devices.

± assumes a normal distribution centered at the temperature with 3 standard deviations equal to the ± number.

Allowable temperature rise due to a step power pulse of the corresponding power density.

Asynchronous capability is defined as the capability of the handler to input, socket and output devices independently with multiple test sites-no gang socketing.

Uninterrupted tray flow requires the handler operation to not be halted when loading/unloading trays.

Auto-Retest requires units to be retested automatically without the need for operator intervention. This is different from a simple reprobe in that the part must be socketed on a different change kit head (if possible).

Electromigration interference (EMI) event field is a measurement of electric emissions due to electrostatic discharge (ESD) events during normal handler operation.

Table 33a Handler (Logic—Pick and Place) Requirements—Near-term Years

Year of Production	*2005*	*2006*	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*	*2013*
DRAM ½ Pitch (nm) (contacted)	*80*	*70*	*65*	*57*	*50*	*45*	*40*	*36*	*32*
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	*90*	*78*	*68*	*59*	*52*	*45*	*40*	*36*	*32*
MPU Physical Gate Length (nm)	*32*	*28*	*25*	*22*	*20*	*18*	*16*	*14*	*13*
Parallel testing < 10 sec test time	**8**	**16**	**16**	**16**	**16**	**32**	**16**	**32**	**64**
Parallel testing > 10 sec test time	**16**	**32**	**32**	**32**	**32**	**64**	**64**	**128**	**128**
Index time (S)	**0.3–0.4**	**0.3–0.4**	**0.25–0.3**	**0.25–0.3**	**0.25**	**0.25**	**0.25**	**0.25**	**0.25**
Throughput (devices per hour)	**8–12K**	**8–12K**	**9–14K**	**9–14K**	**12–20K**	**12–20K**	**12–20K**	**12–20K**	**12–20K**
Sorting	**3–6**	**3–6**	**3–6**	**3–6**	**3–6**	**3–6**	**3–6**	**3–6**	**3–6**
Temperature set point range (ºC)	**-55 to 155**	**-55 to 175**	**-55 to 175**	**-55 to 175**	**-55 to 175**	**-55 to 175**	**-55 to 175**	**-55 to 175**	**-55 to 175**
Tj accuracy at start of test (ºC)	**± 0.5**	**± 2**	**± 2**	**± 2**	**± 2**	**± 2**	**± 2**	**± 2**	**± 2**
Total thermal load (Watts) – MPU	**125**	**125**	**150**	**150**	**175**	**200**	**200**	**200**	**200**
Thermal Watt density (Watts/cm^2) – MPU	**130**	**130**	**175**	**175**	**200**	**225**	**225**	**225**	**225**
Maximum socket load per unit (kg)	**24**	**27**	**50**	**50**	**35**	**60**	**35**	**60**	**60**
Asynchronous capability	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Number of pins or lands/device	**750**	**750**	**800**	**800**	**850**	**850**	**850**	**850**	**850**
Pin/land pitch (mm)	**1.1**	**1.1**	**0.3**	**0.3**	**0.8**	**0.3**	**0.3**	**0.3**	**0.2**
Conversion time (minutes)	**30**	**30**	**15**	**15**	**15**	**5**	**5**	**5**	**5**
Uninterrupted tray loading/auto-2A	**No**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Reliability (hours)	**400**	**600**	**800**	**1000**	**1000**	**1000**	**1000**	**1200**	**1200**

Table 33b Handler (Logic—Pick and Place) Requirements—Long-term Years

Year of Production	*2014*	*2015*	*2016*	*2017*	*2018*	*2019*	*2020*
DRAM ½ Pitch (nm) (contacted)	*28*	*25*	*22*	*20*	*18*	*16*	*14*
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	*28*	*25*	*22*	*20*	*18*	*16*	*14*
MPU Physical Gate Length (nm)	*11*	*10*	*9*	*8*	*7*	*6*	*6*
Parallel testing < 10 sec test time	**64**	**64**	**64**	**64**	**64**	**64**	**64**
Parallel testing > 10 sec test time	**128**	**128**	**128**	**128**	**256**	**256**	**256**
Index time (S)	**0.25**	**0.25**	**0.25**	**0.25**	**0.25**	**0.25**	**0.25**
Throughput (devices per hour)	**12–20K**	**20–28K**	**20–28K**	**20–28K**	**20–28K**	**20–28K**	**20–28K**
Sorting	**3–6**	**3–6**	**3–6**	**3–6**	**3–6**	**3-6**	**3–6**
Temperature set point range (ºC)	**-55 to 175**	**-55 to 175**	**-55 to 175**	**-55 to 175**	**-55 to 175**	**-55 to 175**	**-55 to 175**
Tj accuracy at start of test (ºC)	**± 2**	**± 2**	**± 2**	**± 2**	**± 2**	**± 2**	**± 2**
Total thermal load (Watts) – MPU	**200**	**250**	**250**	**250**	**300**	**300**	**300**
Thermal Watt density (Watts/cm^2) – MPU	**225**	**250**	**250**	**250**	**250**	**250**	**250**
Maximum socket load per unit (kg)	**60**	**65**	**65**	**65**	**75**	**75**	**75**
Asynchronous capability	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Number of pins or lands/device	**850**	**900**	**900**	**900**	**1000**	**1000**	**1000**
Pin/land pitch (mm)	**0.3**	**0.2**	**0.2**	**0.3**	**0.2**	**0.2**	**0.2**
Conversion time (minutes)	**5**	**5**	**5**	**5**	**5**	**5**	**5**
Uninterrupted tray loading/auto-2A	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Reliability (hours)	**1000**	**1200**	**1400**	**1400**	**1400**	**1400**	**1400**

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

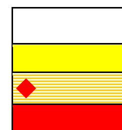


Table 34a Handler (Network and Communications—Pick and Place)—Near-term Years

Year of Production	*2005*	*2006*	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*	*2013*
DRAM ½ Pitch (nm) (contacted)	*80*	*70*	*65*	*57*	*50*	*45*	*40*	*36*	*32*
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	*90*	*78*	*68*	*59*	*52*	*45*	*40*	*36*	*32*
MPU Physical Gate Length (nm)	*32*	*28*	*25*	*22*	*20*	*18*	*16*	*14*	*13*
Parallel testing	**4**	**8**	**8**	**8**	**8**	**8**	**16**	**16**	**16**
Index time (S)	**0.3–0.4**	**0.3–0.4**	**0.3–0.4**	**0.3–0.4**	**0.25–0.3**	**0.25–0.3**	**0.25**	**0.25**	**0.25**
Throughput (devices per hour)	**4–6K**	**8–12K**	**8–12K**	**8–12K**	**9–14K**	**9–14K**	**12–20K**	**12–20K**	**12–20K**
Sorting	**3–6**	**3–6**	**3–6**	**3–6**	**3–6**	**3–6**	**3–6**	**3–6**	**3–6**
Set point range (ºC)	**-45 to +150**	**-45 to +150**	**-45 to +150**	**-45 to +150**	**-45 to +150**	**-45 to +150**	**-45 to +150**	**-45 to +150**	**-45 to +150**
Temperature accuracy (ºC)	**-0.5**	**-0.5**	**-0.5**	**-0.5**	**-0.5**	**-0.5**	**-0.5**	**-0.5**	**-0.5**
Allowable device temperature rise (ºC)	**20**	**20**	**20**	**20**	**20**	**20**	**20**	**20**	**20**
Maximum socket load per unit (kg)	**16**	**20**	**24**	**27**	**30**	**30**	**35**	**35**	**35**
Asynchronous capability	**No**	**No**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Number of pins or lands/device	**700**	**700**	**750**	**750**	**800**	**800**	**850**	**850**	**850**
Pin/land pitch (mm)	**1.2**	**1.2**	**1.1**	**1.1**	**1**	**1**	**0.8**	**0.6**	**0.6**
Conversion time (minutes)	**30**	**30**	**30**	**30**	**15**	**15**	**15**	**5**	**5**
Uninterrupted tray loading/auto-2A	**No**	**No**	**No**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Reliability (hours)	**80**	**100**	**100**	**168**	**168**	**500**	**500**	**1000**	**1000**

Table 34b Handler (Network and Communications—Pick and Place)—Long-term Years

Year of Production	*2014*	*2015*	*2016*	*2017*	*2018*	*2019*	*2020*
DRAM ½ Pitch (nm) (contacted)	*28*	*25*	*22*	*20*	*18*	*16*	*14*
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	*28*	*25*	*22*	*20*	*18*	*16*	*14*
MPU Physical Gate Length (nm)	*11*	*10*	*9*	*8*	*7*	*6*	*6*
Parallel testing	**16**	**16**	**16**	**32**	**32**	**32**	**32**
Index time (S)	**0.25**	**0.25**	**0.25**	**0.25**	**0.25**	**0.25**	**0.25**
Throughput (devices per hour)	**12–20K**	**12–20K**	**12–20K**	**20–28K**	**20–28K**	**20–28K**	**20–28K**
Sorting	**3–6**	**3–6**	**3–6**	**3–6**	**3–6**	**3–6**	**3–6**
Set point range (ºC)	**-45 to +150**	**-45 to +150**	**-45 to +150**	**-45 to +150**	**-45 to +150**	**-45 to +150**	**-45 to +150**
Temperature accuracy (ºC)	**-0.5**	**-0.5**	**-0.5**	**-0.5**	**-0.5**	**-0.5**	**-0.5**
Allowable device temperature rise (ºC)	**20**	**20**	**20**	**20**	**20**	**20**	**20**
Maximum socket load per unit (kg)	**35**	**35**	**35**	**35**	**35**	**35**	**35**
Asynchronous capability	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Number of pins or lands/device	**850**	**850**	**850**	**900**	**900**	**900**	**1000**
Pin/land pitch (mm)	**0.6**	**0.6**	**0.6**	**0.4**	**0.4**	**0.4**	**0.4**
Conversion time (minutes)	**5**	**5**	**5**	**5**	**5**	**5**	**5**
Uninterrupted tray loading/auto-2A	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Reliability (hours)	**1000**	**1000**	**1000**	**1000**	**1000**	**1000**	**1000**

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

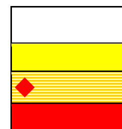


Table 35a Prober (Logic MPU—Pick and Place) Requirements—Near-term Years

Year of Production		*2005*	*2006*	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*	*2013*
DRAM ½ Pitch (nm)		*80*	*70*	*65*	*57*	*50*	*45*	*40*	*36*	*32*
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)		*90*	*78*	*68*	*59*	*52*	*45*	*40*	*36*	*32*
MPU Printed Gate Length (nm)		*45*	*40*	*35*	*32*	*28*	*25*	*22*	*20*	*18*
MPU Physical Gate Length (nm)		*32*	*28*	*25*	*22*	*20*	*18*	*16*	*14*	*13*
Wafer diameter (mm)		**300**	**300**	**300**	**300**	**300**	**300**	**300**	**300**	**300**
Pad pitch	Peripheral (mm)	**40–100**	**40–100**	**40–100**	**30–80**	**30–80**	**30–60**	**30–60**	**30–60**	**30–60**
	Bump (mm)	**30**	**30**	**30**	**30**	**30**	**20**	**20**	**20**	**20**
Wafer thickness (mm)		**80–775**	**80–775**	**80–775**	**80–775**	**80–775**	**50–1000**	**50–1000**	**50–1000**	**50–1000**
Maximum I/O pads		**3000**	**3000**	**3000**	**4000**	**4000**	**5300**	**5300**	**5300**	**5300**
Chuck positioning accuracy	X & Y (µm)	**4**	**4**	**4**	**2**	**2**	**2**	**2**	**2**	**2**
	Z (µm)	**2**	**2**	**1**	**1**	**1**	**0.5**	**0.5**	**0.5**	**0.5**
Probe-to-pad alignment (µm)		**4.5**	**4.5**	**4.5**	**4.5**	**4.5**	**3.5**	**3.5**	**3.5**	**3.5**
Maximum chuck force (kg)		**50**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
Parallel testing		**16**	**32**	**32**	**32**	**32**	**64**	**64**	**64**	**128**
Set point range (ºC)		**-30 to +85**	**-30 to +85**	**-30 to +85**	**-30 to +85**	**-30 to +85**	**-45 to +125**	**-45 to +125**	**-45 to +125**	**-45 to +125**
Total power (Watts)		**150**	**150**	**200**	**200**	**250**	**250**	**250**	**250**	**250**
Power density (Watt/cm^2)		**90**	**90**	**90**	**90**	**120**	**120**	**120**	**120**	**120**

Table 35b Prober (Logic MPU—Pick and Place) Requirements—Long-term Years

Year of Production		*2014*	*2015*	*2016*	*2017*	*2018*	*2019*	*2020*
DRAM ½ Pitch (nm)		*28*	*25*	*22*	*20*	*18*	*16*	*14*
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)		*28*	*25*	*22*	*20*	*18*	*16*	*14*
MPU Printed Gate Length (nm)		*16*	*14*	*13*	*11*	*10*		
MPU Physical Gate Length (nm)		*11*	*10*	*9*	*8*	*7*	*6*	*6*
Wafer diameter (mm)		**450**	**450**	**450**	**450**	**450**	**450**	**450**
Pad pitch	Peripheral (mm)	**30–60**	**30–60**	**30–60**	**30–60**	**30–60**	**30–60**	**3–60**
	Bump (mm)	**20**	**20**	**20**	**20**	**20**	**20**	**20**
Wafer thickness (mm)		**50–1000**	**50–1000**	**50–1000**	**50–1000**	**50–1000**	**50–1000**	**50–1000**
Maximum I/O pads		**5300**	**5300**	**5300**	**5300**	**5300**	**5300**	**5300**
Chuck positioning accuracy	X & Y (µm)	**2**	**2**	**2**	**2**	**2**	**2**	**2**
	Z (µm)	**0.5**	**0.5**	**0.5**	**0.5**	**0.5**	**0.5**	**0.5**
Probe-to-pad alignment (µm)		**3.5**	**3.5**	**3.5**	**3.5**	**3.5**	**3.5**	**3.5**
Maximum chuck force (kg)		**100**	**100**	**100**	**100**	**100**	**100**	**100**
Parallel testing		**128**	**128**	**128**	**128**	**128**	**128**	**128**
Set point range (ºC)		**-45 to +125**	**-45 to +125**	**-45 to +125**	**-45 to +125**	**-45 to +125**	**-45 to +125**	**-45 to +125**
Total power (Watts)		**250**	**300**	**300**	**250**	**300**	**300**	**300**
Power density (Watt/cm^2)		**120**	**120**	**120**	**120**	**120**	**120**	**120**

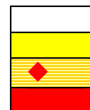
Manufacturable solutions exist, and are being optimized
Manufacturable solutions are known
Interim solutions are known

Manufacturable solutions are NOT known

Table 36 Test Handler and Prober Difficult Challenges

	CHALLENGE	ISSUE / GOAL
Memory handler	Package form factors	Variety of sizes, thicknesses, and ball pitches requires kitless handlers with thin-die handling capability
	Ball-to-package edge gap	As this decreases from 0.6 mm to 0 mm, new handling and socketing methods must be introduced
	Massive parallelism	Parallelism at x128 and up to x256 on roadmap, thermal, and alignment challenges
Logic handler	Thermal control	Improved temperature control and temperature rise control due to high power densities during test
	Operations improvements	Continuous lot processing (lot cascading), auto-retest, asynchronous device socketing, low-conversion times
	ESD	Products more sensitive to ESD, while on-die protection circuitry increases cost.
	Packaging technology	Lower stress socketing, low-cost change kits, higher I/O count, heat lids change thermal characteristics
	Through put and multi-site	Increase in multi-site handling capability for short test time devices (1–7 seconds), thus increase in throughput
Netcom handler	New packaging technologies	Known good die solutions (KGD), stacked die packaging, thin die packaging
	Temperature control	Wide range tri-temperature soak requirements (-45ºC to 150ºC) increases system complexity
	Operations improvements	Continuous lot processing (lot cascading), auto-retest, low conversion times, asynchronous operation
	EMI/RF up to 40 GHz	Shielding issues associated with high frequency testing (>10 GHz)
Logic prober	Thermal contact resistance between wafer and chuck	The high thermal resistance and variation in contact resistance across chuck are required to improve temperature control and reduce temperature rise of device under test
	Heat dissipation at elevated temperature	Heat dissipation of >100 Watts at > 85ºC is a configuration gap in the prober industry
	Probe card optical standardization	With advancement in probe card technology a new optical alignment methodology must be developed

DEVICE INTERFACE TECHNOLOGY REQUIREMENTS

As device I/O bandwidth and power demands increase there is a corresponding requirement for high performance power and signal delivery during electrical test. These requirements drive challenges for the assemblies used to interface the test equipment to the device under test. The highest performance interfaces require complete power and signal path modeling from the source instrument to the die, requiring accurate simulation models of the test instrument, path, probe or socket, and die. Shrinking die and package geometries as well as manufacturing productivity further complicate these interfaces with decreasing pitch, increasing pin count and multi-DUT requirements.

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.

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/Os 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 ASICs 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 MEM's technology as wire bond devices move into higher technology silicon.

Increasingly, manufacturing test of devices is moving to parallel test. For some product groups (e.g., memory), current wafer probe technologies handle parallel testing of 32 to 256 and more devices. Probe technologies capable of full wafer contacting are in use already for 200 mm wafers and the ability to get to 300 mm full contact is imminent. Increasing the contacts/DUT for these massively parallel probes is the next challenge.

Parallel probing requirements are also increasing for some high pin count products, e.g., ASICs. To achieve high parallelism wafer probing, the number of probed I/Os per die will need to be restricted using DFT techniques.

Wafer probe electrical models that integrate models of other elements in the path from tester to DUT will be required of probe card suppliers. These models will be needed to conduct simulations of increasingly complex ATE to DUT interface networks to optimize performance at the DUT.

As new or advanced probe technologies are entering the marketplace, issues of single-sourcing, order to delivery time, probe lifetime, application support, and reparability are important and essential considerations in the selection of a probe card for use in volume production. Standard layout designs are envisioned that will provide compatibility and its associated cost and lead-time benefits.

PROBE CARD TECHNOLOGY REQUIREMENTS

Many probe card technology types are available in the marketplace, each with suitability (technical and/or test operations driven) for probing certain device types and limitations that prevent more widespread use. There is no single probe technology capable of addressing the requirements across the entire device spectrum.

Table 37 Probe Card Difficult Challenges—Near-term Years

CHALLENGE	ISSUE / GOAL
Geometry	Probe technologies to support peripheral fine pitch probe of 25 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. Fine pitch vertical probe technologies to support 130 µm pitch area array solder bump and 50 µm pitch staggered pad devices. Multi-site pad probing technologies with corner pitch capability below 125 µm. Reduction of pad damage at probe commensurate with pad size reductions (or better). Alternative probe technology for 75 µm on 150 µm pitch dense array (vertical probe; bumped device). Increasing probe array planarity requirements in combination with increasing array size.
Parallel test	Need a probe technology to handle the complexity of SOC devices while probing more than one device. Current probe technologies have I/O limitations for bumped device probes.
Probing at temperature	Reduce effects on probes for non-ambient testing -40 to 150°C; especially for fine-pitch devices. For effects on Handlers and Probers, see that section.
Product	Probe technologies to direct probe on copper bond pads including various oxidation considerations. Probe technologies for probing over active circuitry (including flip-chip).
Probe force	Reduce per pin force required for good contact resistance to lower total load for high pin count and multi DUT probe applications. Evaluation and reduction of probe force requirements to eliminate die damage, including interlayer dielectric damage with low-κ dielectrics.
Probe cleaning	Development of high temperature (85 C–15°C) *in situ* cleaning mediums/methods, particularly for fine pitch, multi-DUT and non-traditional probes. Reduction of cleaning requirements while maintaining electrical performance to increase lifetime.
Cost and delivery	Fine pitch or high pin count probe cards are too expensive and take too long to build. Time and cost to repair fine pitch or high pin count probe cards is very high. The time between chip design completion (“tape-out”) and the availability of wafers to be probed is less than the time required to design and build a probe card in almost every probe technology except traditional cantilever. Space transformer lead times are too long, thus causing some vertical probe technologies to have lengthy lead-times.
Probe metrology	Tools are required that support fine pitch probe characterization and pad damage measurements. Metrology correlation is needed—repair versus on-floor usage.
High power devices	Probe technologies will need to incorporate thermal management features capable of handling device power dissipations approaching 1000 Watts and the higher currents (>= 1.5 amp) flowing through individual probe points.
Contact resistance	Probe technologies that achieve contact resistance <.5 Ohms initially and throughout use are needed.
High frequency probing	Traditional probe technologies do not have the necessary electrical bandwidth for higher frequency devices. At the top end are RF devices, requiring up to 40 GHz.

This section explores the challenges of probe technologies including those that are independent 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 continually scale down 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., MEMs and membrane) may offer solutions for reduced pitch scrub requirements.

Area array solder bumps are seeing growing application and the commensurate need/demand for vertical probing technologies. As the pitch/bump dimensions get smaller, current vertical technologies, typically an array of guided wires, may also see their practical limit, thus requiring development of newer technologies.

MULTI-DUT

Productivity gains are often realized when testing (probing) more than one device in parallel. Memory testing has been a leader in this area, with leading edge approaching 500 devices in parallel. As Table 38 indicates virtually all memory testing is done in multi-DUT fashion. The move to multiple DUT testing within other product categories is already underway and is accelerating: with the use of DFT and "smart test" techniques, 16, 32, and even 64 DUTs is realizable and up to 4 DUT for high end microprocessors.

Multi-DUT probing requirements drive the need for more and more probe contacts across an ever-growing area. Today some new contact/probe technologies claim full wafer contact capability for 300 mm wafers. Ultimately increasing the contacts/DUT to hundreds will be required.

ELECTRICAL PERFORMANCE

Wafer probe technology—the probe card—provides electrical contact between the device(s) under test on a wafer and the test system electronics. The probe card must faithfully transmit/deliver device under test power and signals from/to the test system.

Within this ITRS document information can be found concerning device operating voltages and AC Characteristics. Additionally, within this Test and Test Equipment chapter tester performance information is provided on a wide range of electrical characteristics that may be helpful in understanding requirements for wafer probing.

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 with growing circuit densities and pin counts. Of note is that there are some selected applications that are seeing the need for higher and growing current carrying capability, approaching 1.5 amp and more. Of note is that peak values for transient currents are growing as well.

Contact resistance is always a closely watched probe technology parameter. It is influenced by many factors such as pad/bump metallurgy, contamination from pads/bumps, multi-DUT "off-stepping," contact force, scrub, cleaning, etc. The values shown in the requirements table reflect contact resistance under 'normal' usage conditions over the practical lifetime of the probe. Initial and after cleaning requirements are often considerably lower, typically in the 200 milli-Ohm range or lower. There is a growing requirement for lower contact resistance values for longer periods (numbers of touchdowns) before cleaning.

THERMAL PERFORMANCE

Though stable through the roadmap horizon, the thermal environment for the probe is demanding. With low end chuck set-point requirements well below the freezing point and the upper end past the boiling point, the total range is wide - placing difficult demands on selecting materials that handle the extremes, but possibly more notably to deal with temperature co-efficient of expansion issues and high current demands.

Additionally, handling the heat produced by very high transient current heating effects and/or by high power products may drive the need for active thermal management within probers as well as an improved wafer to chuck thermal interface.

UNIT COST AND COST OF OWNERSHIP

Probe card unit cost and cost of ownership (CoO) trends are not currently covered in this roadmap document. Though individual member companies may have their own approaches to unit cost and cost of ownership measurements and goals, there is a need to develop consistent models that can be used industry wide and cover the wide range of probe card technologies that are in the marketplace.

LEAD-TIME

Driven by the accelerating pace of new design introductions and "shrinks," lead-time requirements for initial orders and re-orders are trending rapidly downward. During this roadmap's horizon, lead-times are reduced by ~50%. The growing percentages of wafers that are tested in multi-DUT fashion, with more complicated probe assemblies, magnify the task of achieving the desired lead-times. Strategies, and perhaps technologies, that enable realizing lead-time reduction are needed.

CLEANING

Generally, online cleaning frequency for cantilever type probes rises slightly through the roadmap horizon, however increasing probe usage (touchdowns) before taking offline for cleaning is being seen for many of the product families. The goal is better utilization of the test systems and the probe.

For vertical probes, the rapidly growing number of touchdowns before online cleaning reflects a desire to reduce the vertical technologies' online cleaning frequency to more closely match/better cantilever technologies. Similar to cantilever technologies, the touchdowns before offline cleaning is increasing but across all product categories.

Notably, in some instances there is a move to eliminate online cleaning for memory products in the outer years of this roadmap's horizon. This is likely reflective of the design and/or complexity of probes with pin counts approaching full wafer contact.

Table 38a Wafer Probe Technology Requirements—Near-term Years

Year of Production	*2005*		*2006*		*2007*		*2008*		*2009*		*2010*		*2011*		*2012*		*2013*	
DRAM ½ Pitch (nm) (contacted)	*80*		*70*		*65*		*57*		*50*		*45*		*40*		*36*		*32*	
I/O Pad Size (µm)	**X**	**Y**	**X**	**Y**	**X**	**Y**	**X**	**Y**	**X**	**Y**	**X**	**Y**	**X**	**Y**	**X**	**Y**	**X**	**Y**
Wirebond	**35**	**60**	**30**	**55**	**30**	**55**	**25**	**45**	**25**	**45**	**25**	**45**	**25**	**45**	**20**	**35**	**20**	**35**
Bump	**75**	**75**	**75**	**75**	**60**	**60**	**60**	**60**	**50**	**50**	**50**	**50**	**50**	**50**	**50**	**50**	**50**	**50**
Scrub (% of I/O)	**AREA**	**DEPTH**	**AREA**	**DEPTH**	**AREA**	**DEPTH**	**AREA**	**DEPTH**	**AREA**	**DEPTH**	**AREA**	**DEPTH**	**AREA**	**DEPTH**	**AREA**	**DEPTH**	**AREA**	**DEPTH**
Wirebond	**25**	**50**	**25**	**50**	**25**	**50**	**20**	**40**	**20**	**40**	**20**	**40**	**20**	**40**	**20**	**40**	**20**	**40**
Bump	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**
Multi-DUT Volume (% of Total Product Type Wafers Probed)																		
Memory (DRAM)	**99.9**		**99.9**		**99.9**		**99.9**		**99.9**		**99.9**		**99.9**		**99.9**		**99.9**	
ASIC	**50**		**60**		**75**		**75**		**75**		**75**		**75**		**75**		**75**	
Microprocessor	**50**		**60**		**75**		**75**		**75**		**75**		**75**		**75**		**85**	
RF	**45**		**50**		**50**		**60**		**60**		**60**		**60**		**60**		**60**	
Mixed-signal	**60**		**60**		**75**		**75**		**80**		**80**		**80**		**80**		**80**	
Size of Probed Area (mm^2)																		
Memory (DRAM)	**50% of wafer**		**50% of wafer**		**100% of wafer**		**100% of wafer**		**100% of wafer**		**100% of wafer**		**100% of wafer**		**100% of wafer**		**100% of wafer**	
ASIC	**2050**		**2050**		**2050**		**2400**		**2400**		**2400**		**2400**		**2400**		**2400**	
Microprocessor	**2050**		**2050**		**2050**		**2400**		**2400**		**2400**		**2400**		**2400**		**2400**	
RF	**900**		**900**		**1225**		**1225**		**1225**		**1225**		**1225**		**1225**		**1225**	
Mixed-signal	**1413**		**1413**		**1600**		**1600**		**1600**		**1600**		**1600**		**1600**		**1600**	
Number of Probe Points /Touchdown	**Signal**	**Total**	**Signal**	**Total**	**Signal**	**Total**	**Signal**	**Total**	**Signal**	**Total**	**Signal**	**Total**	**Signal**	**Total**	**Signal**	**Total**	**Signal**	**Total**
Memory (DRAM)	**14500**	**18700**	**14500**	**18700**	**17000**	**20000**	**17000**	**20000**	**17000**	**20000**	**17000**	**20000**	**17000**	**20000**	**20000**	**25000**	**20000**	**25000**
ASIC	**1050**	**4000**	**1050**	**5000**	**1050**	**5000**	**1200**	**6000**	**1500**	**7500**	**1500**	**7500**	**1500**	**7500**	**3000**	**9000**	**3000**	**9000**
Microprocessor	**1024**	**10000**	**1024**	**15000**	**1024**	**20000**	**1024**	**20000**	**1024**	**20000**	**1024**	**20000**	**1024**	**20000**	**2000**	**30000**	**2000**	**30000**
RF	**250**	**450**	**250**	**450**	**350**	**630**	**350**	**630**	**350**	**630**	**350**	**630**	**350**	**630**	**350**	**630**	**350**	**630**
Mixed-signal	**450**	**600**	**450**	**600**	**510**	**680**	**510**	**680**	**510**	**680**	**510**	**680**	**510**	**680**	**510**	**680**	**510**	**680**
Maximum Current (mA)	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**
Memory (DRAM)	**100**	**<10**	**125**	**<10**	**125**	**<10**	**125**	**<10**	**125**	**<10**	**125**	**<10**	**125**	**<10**	**125**	**<10**	**125**	**<10**
ASIC	**200**	**<10**	**300**	**<10**	**400**	**<10**	**500**	**<10**	**500**	**<10**	**500**	**<10**	**500**	**<10**	**1000**	**<10**	**1000**	**<10**
Microprocessor	**800**	**<10**	**1000**	**<10**	**1200**	**<10**	**1200**	**<10**	**1500**	**<10**	**1500**	**<10**	**1500**	**<10**	**1500**	**<10**	**1500**	**<10**
RF	**200**	**<10**	**225**	**<10**	**225**	**<10**	**225**	**<10**	**225**	**<10**	**225**	**<10**	**225**	**<10**	**225**	**<10**	**225**	**<10**
Mixed-signal	**250**	**<10**	**275**	**<10**	**275**	**<10**	**275**	**<10**	**275**	**<10**	**275**	**<10**	**275**	**<10**	**275**	**<10**	**275**	**<10**
Maximum Resistance (Ohm)	**Contact**	**Series**	**Contact**	**Series**	**Contact**	**Series**	**Contact**	**Series**	**Contact**	**Series**	**Contact**	**Series**	**Contact**	**Series**	**Contact**	**Series**	**Contact**	**Series**
Memory (DRAM)	**<0.5**	**<3**	**<0.5**	**<3**	**<0.5**	**<3**	**<0.5**	**<3**	**<0.5**	**<3**	**<0.5**	**<3**	**<0.5**	**<3**	**<0.5**	**<3**	**<0.5**	**<3**
ASIC	**<0.5**	**<3**	**<0.5**	**<3**	**<0.5**	**<3**	**<0.5**	**<3**	**<0.5**	**<3**	**<0.5**	**<3**	**<0.5**	**<3**	**<0.5**	**<3**	**<0.5**	**<3**
Microprocessor	**<0.5**	**<2**	**<0.5**	**<2**	**<0.5**	**<2**	**<0.5**	**<2**	**<0.5**	**<2**	**<0.5**	**<2**	**<0.5**	**<2**	**<0.5**	**<2**	**<0.5**	**<2**
RF	**<0.5**	**<1.5**	**<0.5**	**<1.5**	**<0.5**	**<1.5**	**<0.5**	**<1.5**	**<0.5**	**<1.5**	**<0.5**	**<1.5**	**<0.5**	**<1.5**	**<0.5**	**<1.5**	**<0.5**	**<1.5**
Mixed-signal	**<0.5**	**<1.5**	**<0.5**	**<1.5**	**<0.5**	**<1.5**	**<0.5**	**<1.5**	**<0.5**	**<1.5**	**<0.5**	**<1.5**	**<0.5**	**<1.5**	**<0.5**	**<1.5**	**<0.5**	**<1.5**
Chuck Set-point (^oC)	**Min.**	**Max.**	**Min.**	**Max.**	**Min.**	**Max.**	**Min.**	**Max.**	**Min.**	**Max.**	**Min.**	**Max.**	**Min.**	**Max.**	**Min.**	**Max.**	**Min.**	**Max.**
Memory (DRAM)	**-40**	**140**	**-40**	**150**	**-40**	**150**	**-50**	**180**	**-50**	**180**	**-50**	**180**	**-50**	**180**	**-50**	**180**	**-50**	**180**
ASIC	**25**	**110**	**0**	**140**	**0**	**140**	**-10**	**140**	**-10**	**140**	**-10**	**140**	**-10**	**14**	**-10**	**14**	**-10**	**14**
Microprocessor	**-30**	**135**	**-30**	**135**	**-30**	**135**	**-30**	**135**	**-30**	**135**	**-30**	**135**	**-30**	**135**	**-30**	**135**	**-30**	**135**
RF	**5**	**120**	**5**	**120**	**5**	**120**	**5**	**120**	**5**	**120**	**5**	**120**	**5**	**120**	**5**	**120**	**5**	**120**
Mixed-signal	**25**	**125**	**25**	**125**	**25**	**125**	**25**	**125**	**25**	**125**	**25**	**125**	**25**	**125**	**25**	**125**	**25**	**125**

Year of Production	*2005*		*2006*		*2007*		*2008*		*2009*		*2010*		*2011*		*2012*		*2013*	
DRAM ½ Pitch (nm) (contacted)	*80*		*70*		*65*		*57*		*50*		*45*		*40*		*36*		*32*	
Soak Time (minutes)																		
Memory (DRAM)	10		8		8		8		7		7		7		7		7	
ASIC	8		7		7		7		6		6		6		6		6	
Microprocessor	10		10		10		9		9		9		9		9		9	
RF	10		10		9		9		9		9		9		9		9	
Mixed-signal	10		10		9		9		9		9		9		9		9	
Order Lead-time—Single DUT (weeks)	1st Order	Re-Order	1st Order	Re-Order	1st Order	Re-Order	1st Order	Re-Order	1st Order	Re-Order	1st Order	Re-Order	1st Order	Re-Order	1st Order	Re-Order	1st Order	Re-Order
Memory (DRAM)	6	3	5.5	3	5	3	4	2	4	2	4	2	4	2	4	2	4	2
ASIC	2.5	1.5	2.5	1.5	2.5	1.5	2	1	2	1	2	1	2	1	2	1	2	1
Microprocessor	2.5	1.5	2.5	1.5	2.5	1.5	2	1	2	1	2	1	2	1	2	1	2	1
RF	4	2	3.5	1.5	3.5	1.5	3	1	3	1	3	1	3	1	3	1	3	1
Mixed-signal	3	2	2.5	1.5	2.5	1.5	2	1	2	1	2	1	2	1	2	1	2	1
Order Lead-time—Multi-DUT (weeks)	1st Order	Re-Order	1st Order	Re-Order	1st Order	Re-Order	1st Order	Re-Order	1st Order	Re-Order	1st Order	Re-Order	1st Order	Re-Order	1st Order	Re-Order	1st Order	Re-Order
Memory DRAM)	7	4	6	3	5	3	4.5	2.5	4.5	2.5	4.5	2.5	4.5	2.5	4.5	2.5	4.5	2.5
ASIC	5	2	4	1.5	3.5	1.5	3	1	3	1	3	1	3	1	3	1	3	1
Microprocessor	4	2	4	1.5	3.5	1.5	3	1	3	1	3	1	3	1	3	1	3	1
RF	5	3	4.5	2	4	1.5	4	1	4	1	4	1	4	1	4	1	4	1
Mixed-signal	4	2	3.5	2	3	1.5	3	1	3	1	3	1	3	1	3	1	3	1
Touchdowns Before Clean (Cantilever)	Online	Offline	Online	Offline	Online	Offline	Online	Offline	Online	Offline	Online	Offline	Online	Offline	Online	Offline	Online	Offline
Memory (DRAM)	400	20,000	400	20,000	450	20,000	450	20,000	450	20,000	450	20,000	450	20,000	450	20,000	450	20,000
ASIC	3,250	60,000	3,500	60,000	3,500	60,000	3,500	60,000	3,500	60,000	3,500	60,000	3,500	60,000	3,500	60,000	3,500	60,000
Microprocessor	1,250	50,000	1,500	50,000	1,500	50,000	1,500	50,000	1,500	50,000	1,500	50,000	1,500	50,000	1,500	50,000	1,500	50,000
RF	1,000	100,000	1,000	100,000	1,000	100,000	1,000	100,000	1,000	100,000	1,000	100,000	1,000	100,000	1,000	100,000	1,000	100,000
Mixed-signal	2,000	200,000	2,000	200,000	2,000	300,000	2,000	300,000	2,000	300,000	2,000	300,000	2,000	300,000	2,000	300,000	2,000	300,000
Touchdowns Before Clean (Vertical)	Online	Offline	Online	Offline	Online	Offline	Online	Offline	Online	Offline	Online	Offline	Online	Offline	Online	Offline	Online	Offline
Memory (DRAM)	1,500	20,000	2,000	25,000	2,000	25,000	2,000	25,000	2,000	25,000	2,000	25,000	2,500	27,500	2,500	27,500	2,500	27,500
ASIC	1,500	17,500	2,000	20,000	2,000	20,000	2,000	20,000	2,000	20,000	2,000	20,00	2,500	22,500	2,500	22,500	2,500	22,500
Microprocessor	1,500	80,000	2,000	100,000	2,000	100,000	2,000	100,000	2,000	100,000	2,000	100,000	2,500	100,000	2,500	100,000	2,500	100,000
RF	100	20,000	100	25,000	100	25,000	125	25,000	125	25,000	125	25,000	125	27,500	125	27,500	125	27,500
Mixed-signal	1,500	85,000	2,000	87,500	2,000	87,500	2,000	87,500	2,000	87,500	2,000	87,500	2,500	90,000	2,500	90,000	2,500	90,000

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

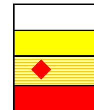
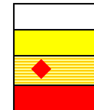


Table 38b Wafer Probe Technology Requirements—Long-term Years

Year of Production	*2014*		*2015*		*2016*		*2017*		*2018*		*2019*		*2020*	
DRAM ½ Pitch (nm) (contacted)	*28*		*25*		*22*		*20*		*18*		*16*		*14*	
I/O Pad Size (µm)	**X**	**Y**	**X**	**Y**	**X**	**Y**	**X**	**Y**	**X**	**Y**	**X**	**Y**	**X**	**Y**
Wirebond	20	35	15	25	15	25	15	25	15	25	15	25	15	25
Bump	50	50	50	50	50	50	50	50	50	50	50	50	50	50
Scrub (% of I/O)	**Offline**	**DEPTH**	**AREA**	**DEPTH**	**AREA**	**DEPTH**	**AREA**	**DEPTH**	**AREA**	**DEPTH**	**AREA**	**DEPTH**	**AREA**	**DEPTH**
Wirebond	20	40	20	40	20	40	20	40	20	40	20	40	20	40
Bump	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Multi-DUT Volume (% of Total Product Type Wafers Probed)														
Memory (DRAM)	99.9		99.9		99.9		99.9		99.9		99.9		99.9	
ASIC	75		75		75		75		75		75		75	
Microprocessor	85		85		85		85		85		85		85	
RF	60		80		60		60		60		60		60	
Mixed-signal	80		80		80		80		80		80		80	
Size of Probed Area (mm^2)														
Memory (DRAM)	100% of wafer		100% of wafer		100% of wafer		100% of wafer		100% of wafer		100% of wafer		100% of wafer	
ASIC	2400		2400		2400		2400		2400		2400		2400	
Microprocessor	2400		2400		2400		2400		2400		2400		2400	
RF	1225		1225		1225		1225		1225		1225		1225	
Mixed-signal	1600		1600		1600		1600		1600		1600		1600	
Number of Probe Points /Touchdown	**Signal**	**Total**	**Signal**	**Total**	**Signal**	**Total**	**Signal**	**Total**	**Signal**	**Total**	**Signal**	**Total**	**Signal**	**Total**
Memory (DRAM)	20000	25000	20000	25000	20000	25000	20000	25000	20000	25000	20000	25000	20000	25000
ASIC	3000	9000	3000	9000	3000	9000	3000	9000	3000	9000	3000	9000	3000	9000
Microprocessor	2000	30000	2000	30000	2000	30000	2000	30000	30000	6000	30000	6000	30000	6000
RF	350	630	350	630	350	630	350	630	350	630	350	630	350	630
Mixed-signal	510	680	510	680	510	680	510	680	510	680	510	680	510	680
Maximum Current (mA)	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**	**Probe Tip**	**DC Leakage**
Memory (DRAM)	125	<10	125	<10	125	<10	125	<10	125	<10	125	<10	125	<10
ASIC	1000	<10	1000	<10	1000	<10	1000	<10	1000	<10	1000	<10	1000	<10
Microprocessor	1500	<10	1500	<10	1500	<10	1500	<10	1500	<10	1500	<10	1500	<10
RF	225	<10	225	<10	225	<10	225	<10	225	<10	225	<10	225	<10
Mixed-signal	275	<10	275	<10	275	<10	275	<10	275	<10	275	<10	275	<10
Maximum Resistance (Ohm)	**Contact**	**Series**	**Contact**	**Series**	**Contact**	**Series**	**Contact**	**Series**	**Contact**	**Series**	**Contact**	**Series**	**Contact**	**Series**
Memory (DRAM)	<0.5	<3	<0.5	<3	<0.5	<3	<0.5	<3	<0.5	<3	<0.5	<3	<0.5	<3
ASIC	<0.5	<3	<0.5	<3	<0.5	<3	<0.5	<3	<0.5	<3	<0.5	<3	<0.5	<3
Microprocessor	<0.5	<2	<0.5	<2	<0.5	<2	<0.5	<2	<0.5	<2	<0.5	<2	<0.5	<2
RF	<0.5	<1.5	<0.5	<1.5	<0.5	<1.5	<0.5	<1.5	<0.5	<1.5	<0.5	<1.5	<0.5	<1.5
Mixed-signal	<0.5	<1.5	<0.5	<1.5	<0.5	<1.5	<0.5	<1.5	<0.5	<1.5	<0.5	<1.5	<0.5	<1.5
Chuck Set-point (°C)	**Min.**	**Max.**	**Min.**	**Max.**	**Min.**	**Max.**	**Min.**	**Max.**	**Min.**	**Max.**	**Min.**	**Max.**	**Min.**	**Max.**
Memory (DRAM)	-50	180	-50	180	-50	180	-50	180	-50	180	-50	180	-50	180
ASIC	-10	14	-10	14	-10	14	-10	14	-10	14	-10	14	-10	14
Microprocessor	-30	135	-30	135	-30	135	-30	135	-30	135	-30	135	-30	135
RF	5	120	5	120	5	120	5	120	5	120	5	120	5	120
Mixed-signal	25	125	25	125	25	125	25	125	25	125	25	125	25	125
Soak Time (minutes)														
Memory (DRAM)	7		7		7		7		7		7		7	
ASIC	6		6		6		6		6		6		6	
Microprocessor	9		9		9		9		9		9		9	
RF	9		9		9		9		9		9		9	
Mixed-signal	9		9		9		9		9		9		9	
Order Lead-time—Single DUT (weeks)	**1st Order**	**Re-Order**	**1st Order**	**Re-Order**	**1st Order**	**Re-Order**	**1st Order**	**Re-Order**	**1st Order**	**Re-Order**	**1st Order**	**Re-Order**	**1st Order**	**Re-Order**
Memory (DRAM)	4	2	4	2	4	2	4	2	4	2	4	2	4	2

Year of Production	*2014*		*2015*		*2016*		*2017*		*2018*		*2019*		*2020*	
DRAM ½ Pitch (nm) (contacted)	*28*		*25*		*22*		*20*		*18*		*16*		*14*	
ASIC	2	1	2	1	2	1	2	1	2	1	2	1	2	1
Microprocessor	2	1	2	1	2	1	2	1	2	1	2	1	2	1
RF	3	1	3	1	3	1	3	1	3	1	3	1	3	1
Mixed-signal	2	1	2	1	2	1	2	1	2	1	2	1	2	1
Order Lead-time—Multi-DUT (weeks)	1st Order	Re-Order	1st Order	Re-Order	1st Order	Re-Order	1st Order	Re-Order	1st Order	Re-Order	1st Order	Re-Order	1st Order	Re-Order
Memory DRAM)	4.5	2.5	4.5	2.5	4.5	2.5	4.5	2.5	4.5	2.5	4.5	2.5	4.5	2.5
ASIC	3	1	3	1	3	1	3	1	3	1	3	1	3	1
Microprocessor	3	1	3	1	3	1	3	1	3	1	3	1	3	1
RF	4	1	4	1	4	1	4	1	4	1	4	1	4	1
Mixed-signal	3	1	3	1	3	1	3	1	3	1	3	1	3	1
Touchdowns Before Clean (Cantilever)	Online	Offline	Online	Offline	Online	Offline	Online	Offline	Online	Offline	Online	Offline	Online	Offline
Memory (DRAM)	450	20,000	450	20,000	450	20,000	450	20,000	450	20,000	450	20,000	450	20,000
ASIC	3,500	60,000	3,500	60,000	3,500	60,000	3,500	60,000	3,500	60,000	3,500	60,000	3,500	60,000
Microprocessor	1,500	50,000	1,500	50,000	1,500	50,000	1,500	50,000	1,500	50,000	1,500	50,000	1,500	50,000
RF	1,000	100,000	1,000	100,000	1,000	100,000	1,000	100,000	1,000	100,000	1,000	100,000	1,000	100,000
Mixed-signal	2,000	300,000	2,000	300,000	2,000	300,000	2,000	300,000	2,000	300,000	2,000	300,000	2,000	300,000
Touchdowns Before Clean (Vertical)	Online	Offline	Online	Offline	Online	Offline	Online	Offline	Online	Offline	Online	Offline	Online	Offline
Memory (DRAM)	2,500	27,500	2,500	27,500	2,500	27,500	2,500	27,500	2,500	27,500	2,500	27,500	2,500	27,500
ASIC	2,500	22,500	2,500	22,500	2,500	22,500	2,500	22,500	2,500	22,500	2,500	22,500	2,500	22,500
Microprocessor	2,500	100,000	2,500	100,000	2,500	100,000	2,500	100,000	2,500	100,000	2,500	100,000	2,500	100,000
RF	125	27,500	125	27,500	125	27,500	125	27,500	125	27,500	125	27,500	125	27,500
Mixed-signal	2,500	90,000	2,500	90,000	2,500	90,000	2,500	90,000	2,500	90,000	2,500	90,000	2,500	90,000

Manufacturable solutions exist, and are being optimized
Manufacturable solutions are known
Interim solutions are known

Manufacturable solutions are NOT known

TEST SOCKETS

The socket devices are responsible for transferring all the electrical properties between DUT and PCB/tester through mechanical contact mechanism in order to determine the electrical characteristics of DUT. As the semi-conductor designing and manufacturing capabilities progress in the recent years, the testing process keeps raising the electrical and mechanical requirements. 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 requirements but also by the mechanical. The multi-physics problems have made socket designs progressively challenging along the higher requirements.

ELECTRICAL REQUIREMENTS

The 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. In particular, the process of record (POR) data rate is in the neighborhood of ~7 GHz. Even higher data rate requirements on new chip technologies are expected (e.g., ~20 GHz), which will greatly challenge impedance matching and the potential signal loss. On the other hand, as package size, solder resist opening, and pitches become smaller as pin counts get higher, the smaller pins are required to fit within tighter mechanical constraints will greatly aggravate contact resistance and signal integrity issues. One of the critical parameters to stabilize electrical contact and insure of low contact resistance is the contact force per pin, which generally ranges from 20 ~ 30 grams. As pitches go down, it is unavoidable to deploy smaller/slender pins, which may not be able to sustain high 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 even 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 on the order of a few microns. Consequently, 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 POR contact ranges between 0.3 mm and 0.5 mm. However, as required pin sizes get smaller, it may not be feasible anymore to maintain the same stroke as the POR and thus the compression issue may become the bottleneck of electrical contact performance.

The pin reliability, such as pin tip wear-out, has also experienced great challenges because the tight geometric constraints prevent adding redundant strength to the pins, while the testing environment becomes more difficult by such things as higher temperatures, higher currents, smaller pin tip contacts, etc.