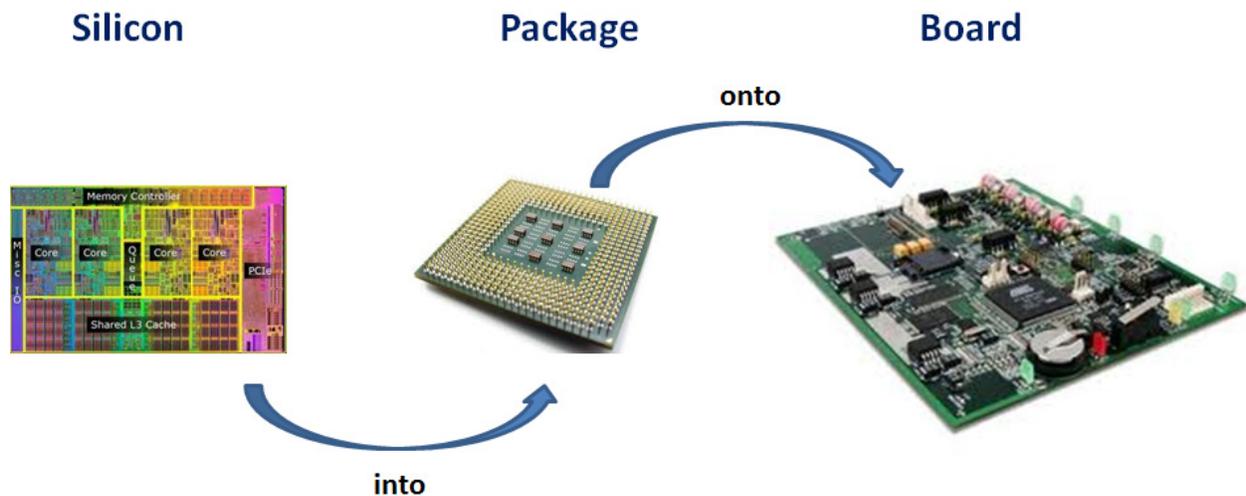


DETECTION AND DIAGNOSIS OF PRINTED CIRCUIT BOARD DEFECTS AND VARIANCES USING ON-CHIP EMBEDDED INSTRUMENTATION



BY ADAM LEY
ALAN SGUIGNA
AL CROUCH

By Adam Ley, Chief Technologist, Non-intrusive Board Test and JTAG



Adam serves customers by ensuring that ASSET’s non-intrusive board test (NBT) methodologies comprise a best-in-class solution to meet the evolving need for improved coverage of board test in the face of ongoing erosion of physical access. Adam is an active participant in IEEE 1149.1, having previously served terms as working group vice chair and as standard technical editor (2001 revision), as well as in nearly all related standards, to include:

1149.4, 1149.5, 1149.6, 1149.7, 1149.8.1, 1500, 1532, 1581, P1149.1.1, iNEMI boundary-scan adoption, PICMG MicroTCA, and SJTAG (system JTAG). Adam’s prior experience prior spanned over a decade at TI where he had roles in application support for TI’s boundary-scan logic products and for test and characterization of new logic families.

Alan Sguigna – Vice President of Sales

Alan Sguigna has more than 20 years of experience in senior-level general management, marketing, engineering, sales, manufacturing, finance and customer service positions. Before joining ASSET, he worked in the telecom industry. He has had profit and loss responsibility for a \$150 million division of Spirent Communications, a supplier of test products and services. Prior to his tenure with Spirent, Mr. Sguigna also served in business development positions with Nortel Networks, overseeing the growth of its voice over Internet protocol (VoIP) products.



Al Crouch



Al Crouch, Chief Technologist – Core Instruments – at ASSET InterTech, is a Senior Member of the IEEE. Al has served as the vice chairman of the IEEE P1687 IJTAG working group that is developing the IJTAG standard and has contributed significantly to the hardware architecture definition. Al is the editor of the P1838 Working Group on 3D test and debug and co-chair of the iNEMI BIST group investigation defining the use of Chip-Embedded Instruments to assist with Board Test.

Table of Contents

INTRODUCTION	5
DEFINITIONS AND BACKGROUND	6
EXAMPLES OF DEFECTS AND VARIANCES	8
BOARD DESIGN DEFECTS	8
BOARD DESIGN VARIANCES	11
BOARD MANUFACTURING DEFECTS	11
<i>PCI Express</i>	11
<i>DDR SDRAM</i>	14
BOARD MANUFACTURING VARIANCES	17
<i>Trace Impedance</i>	18
<i>Trace Surface Roughness</i>	19
CHIP DESIGN DEFECTS	21
CHIP DESIGN VARIANCES	22
<i>Charge-trapping</i>	23
<i>Electro-migration</i>	23
CHIP MANUFACTURING DEFECTS	24
CHIP MANUFACTURING VARIANCES	26
DEFECTS AND VARIANCES COVERAGE	30
PCOLA/SOQ/FAM	30
<i>Component Scoring Guidelines</i>	30
<i>Interconnect Scoring Guidelines</i>	30
<i>Functional Scoring Guidelines</i>	30
STRUCTURAL	31
FUNCTIONAL	31
PERFORMANCE	31
TEST STACK	33
DETECTION TECHNOLOGIES	34
BOUNDARY-SCAN TEST (BST)	34
<i>PCI Express</i>	34
<i>DDR3 and DDR4 SDRAM</i>	38
PROCESSOR-CONTROLLED TEST (PCT)	40

PCT Test Generation 41

PCT Test Coverage..... 41

INTEL® HIGH-SPEED I/O (HSIO) 44

MATRIX OF TESTING CAPABILITIES 47

CONCLUSION 49

LEARN MORE 49

Table of Figures

FIGURE 1: TRANSMIT PAIR OF A PCI EXPRESS LANE..... 12

FIGURE 2: OPEN-CIRCUIT ON PCI EXPRESS TRANSMIT NET 12

FIGURE 3: SHORT TO GROUND ON PCI EXPRESS TRANSMIT NET 13

FIGURE 4: SHORT BETWEEN TWO PCI EXPRESS ADJACENT TRANSMIT NETS 14

FIGURE 5: BLOCK DIAGRAM OF DDR3 DIMM 15

FIGURE 6: DDR3 DIMM DQS STUCK TO GND 16

FIGURE 7: DDR3 DIMM TWO DQ SHORTED TOGETHER 17

FIGURE 8: STRIPLINE CROSS-SECTION, ILLUSTRATING TRACE SURFACE ROUGHNESS 19

FIGURE 9: CIRCUIT BOARD MANUFACTURING PROCESS VARIANCES 21

FIGURE 10: SILICON TRACE ELECTRO-MIGRATION FAULTS 24

FIGURE 11: WAFER MANUFACTURING PROCESS VARIANCE AND DEFECT CLUSTERS 25

FIGURE 12: SILICON, PACKAGES, AND CIRCUIT BOARDS 25

FIGURE 13: FLIP-CHIP CROSS-SECTION ON CIRCUIT BOARD 26

FIGURE 14: CHIP TRACE VARIANCES DUE TO LITHOGRAPHY AND PLANARIZATION PROCESSES 27

FIGURE 15: CHIP MARGINS ACROSS LOTS EXPRESSED AS GAUSSIAN DISTRIBUTIONS 27

FIGURE 16: CHIP EYE MASK AS FUNCTION OF VARIANCES AND EMPIRICAL DATA 28

FIGURE 17: CHIP EYE DIAGRAM IN RELATION TO EYE MASK 29

FIGURE 18: THE TEST STACK: TEST CATEGORIES AND TEST COVERAGE 33

FIGURE 19: PCI EXPRESS LANE BETWEEN PROCESSOR AND PCIe SLOT 35

FIGURE 20: PCIe BOUNDARY SCAN CELL IMPLEMENTATION 36

FIGURE 21: PCIe BOUNDARY SCAN IMPLEMENTATION WITH PASSIVE COPPER LOOPBACK 37

FIGURE 22: PCI EXPRESS TRANSACTION LAYER PACKET (TLP) 45

FIGURE 23: MATRIX OF MEMORY AND SERIAL I/O DEFECTS VERSUS TEST COVERAGE TECHNOLOGY 47

ASSET and ScanWorks are registered trademarks while the ScanWorks logo is a trademark of ASSET InterTech, Inc. All other trade and service marks are the properties of their respective owners.

INTRODUCTION

Printed circuit board design and test has come a long way in the last decade. On older, simpler designs, structural defects, such as short- or open-circuits, were detected by bed-of-nails type testers. Any such defects that went undetected at the bed-of-nails would usually manifest itself by causing some or all of the design to be non-functional.

But now, in an era of shrinking chip and board geometries, higher speeds, and greater densities, legacy test technologies no longer provide the access and defect coverage they once did. And further, variances that were once within acceptable tolerances are now manifesting themselves in intermittent behavior and performance degradations. The effect of defects and variances is now cumulative and it is difficult to isolate the root cause of a marginal design with legacy technologies. When such designs hang or crash intermittently, or degrade in performance for no apparent reason, customer satisfaction plummets. The cost of this to the manufacturer's brand can be enormous.

Fortunately, a new class of test and detection technologies based upon on-chip embedded instrumentation has emerged to provide root cause insight into design marginalities brought about by the cumulative effects of defects and variances.

This paper examines the nature of chip and board defects and variances and explores the technologies needed to detect and diagnose them.

DEFINITIONS AND BACKGROUND

To be specific, and for the purposes of this paper, we define defects and variances as below.

Defect: an *unexpected* fault which arises from a design or manufacturing process. Examples include board design layout errors, missing bond wires, and short-circuits.

Variance: an *expected* deviation from the norm. Examples include variable process/ voltage/ temperature (PVT) related effects, power distribution network (PDN), simultaneous switching noise (SSN), and interconnect trace surface roughness, among many others.

Defects and variances apply to both semiconductors (**chips**) and printed circuit boards (**boards**). Also, defects and variances can manifest themselves within both the **design** and **manufacturing** processes.

Collectively, chips and boards make up **systems**. Within any given system, the effect of defects and variances is cumulative. Since variances are expected deviations, and to a great extent cannot be controlled, the quantity of defects within any given system will ultimately determine system performance and robustness. Whether the system fails catastrophically, fails intermittently, experiences degraded performance, or otherwise operates normally, is determined by its **operating margin**. In other words, if a system is designed well, with plenty of operating margin, it may operate even in the presence of a significant level of defects.

Alternatively, a bad system design can fail even in the presence of everyday variances and few, if any, defects. The extremes of normal operating conditions, such as temperature, voltage, vibration, humidity, pollution, capacitor aging, gate current leakage over time, etc. will cause marginal designs to fail.

In early hardware designs, defects almost universally led to catastrophic system failure. For example, if there were an open-circuit on a low-speed single-ended clock net, that signal would simply not propagate, causing the clock to fail and that portion of the board to be non-functional. In most cases, this defect would have a highly visible manifestation, and would be easily detected and diagnosed by a conventional functional test.

But, with today's differential signaling on high-speed I/O, and error detection and correction on memories, such defects are far more difficult to detect. Built-in redundancy at the physical layer makes such defects look like variances, which may be within the operating margin envelope of the system as a whole. However, it is still important to detect these, as they will contribute to an additive degradation of overall system performance; and under normal operating conditions, may contribute to system intermittent faults.

EXAMPLES OF DEFECTS AND VARIANCES

Board design defects

High-speed I/O design is a complex topic, and there are many references available on the subject. Examples include [Advanced Signal Integrity for High-Speed Digital Designs](http://www.amazon.com/Advanced-Integrity-High-Speed-Digital-Designs/dp/0470192356)¹ by Stephen H. Hall and Howard L. Heck, and [High Speed Digital Design: A Handbook of Black Magic](http://www.amazon.com/High-Speed-Digital-Design-Handbook/dp/0133957241/ref=pd_sim_b_2)² by Howard Johnson and Martin Graham. As high-speed buses such as PCIe, HDMI, and XLAUI exceeding 5 GT/s are now becoming commonplace on everything from smartphones to notebooks and routers, a board designer's job is becoming more and more complicated.

Here's a partial list of what designers need to worry about when designing PCBs with higher speed buses:

1. Intra-pair total length matching should be ≤ 5 mils.
2. For length running skews > 25 mils, compensation should be made within 600 mils.
3. Serpentine "bumps" can be added to the shorter member of a pair to reduce skew.
4. Where bumps are needed, the original inter-pair spacing should be preserved.
5. Avoid routing nets near voltage regulator-induced noise, or limit the noise induced by fast switching VR nodes.
6. Avoid routing over a split power plane.
7. Avoid 90-degree routing which can result in accumulated common mode noise effects.
8. Fan out narrow traces (which may be necessary to escape a tight pin field area) within 100 mils to reduce impedance discontinuity and trace loss.
9. The "dog bone" (segment between the component pad and an inner layer transition) should be less than 30 mils height and 5 mils width.
10. The maximum routed length under a pin field is 1.5".
11. Don't route over pin fields that have a high magnitude of transient currents, like power delivery pin fields.

¹ <http://www.amazon.com/Advanced-Integrity-High-Speed-Digital-Designs/dp/0470192356>

² http://www.amazon.com/High-Speed-Digital-Design-Handbook/dp/0133957241/ref=pd_sim_b_2

12. If you have to length-match within a pin field, place the serpentine bumps near ground vias, or similar via types (i.e. TX via over TX trace, RX via over RX trace).
13. Avoid layer transitions wherever possible.
14. If you have to do a layer transition, reduce solution space by 5”.
15. Reduce discontinuities caused by layer transitions and via stubs along the signal path.
16. Place vias symmetrically to avoid differential to common mode conversion.
17. Between a via pair, the pitch must be between 25 and 50 mils. The gap between any two different via pairs must be greater than 50 mils.
18. Each high-speed signal via should have a Vss via within a 50 mils gap.
19. Be sure to strip via pads on un-routed layers.
20. Backdrill to remove the via stub of pressfit connectors for boards thicker than 73 mil.
21. Minimize capacitance due to lead length protrusions from THMT components.
22. The same package size of capacitor should be used for each signal in a differential pair.
23. Pad sizes for capacitors are to be the minimum allowed per DFM to minimize parasitic effects.
24. Avoid placing capacitors next to devices that generate heat, such as power FETs (some capacitors perform at less than 50% of their nominal value when exposed to heat).
25. Use of through-hole connectors is preferred (press fit connectors induce vertical coupling cross-talk effects).
26. Some critical pins should never have any trace routing connected whatsoever, lest they act as antennae, transmitting induced crosstalk into adjacent traces.

The list goes on and on, and constitutes hundreds of “rules” for serial I/O design, that in some (but not all) cases, can be bent, but not broken. Further, the realities of board design and layout is that not all rules can be simultaneously enforced. Attempting to do so collides with the demands of bringing a design to market in a timely manner.

Examples of blatant SerDes design error would be, for example, breaking Rule #11, “Don’t route over pin fields that have a high magnitude of transient currents, like power delivery pin fields”,

or routing high-speed signals across a DIMM connection. In such cases, the design is unlikely to function at all, or will exhibit intermittent failures or performance degradation.

In terms of memory, one of the biggest design challenges today revolves around maintaining signal integrity in the presence of power and ground rail fluctuations due to simultaneously switching signals, especially on the new DDR4 memory. DDR4 is a big step from DDR3, much bigger than DDR3 was over DDR2. Speed is going from 2133 Mb/s at the top end to 3200 Mb/s. V_{dd} goes from 1.5V to 1.2V. The Unit Interval (UI) shrinks from 469ps to 313ps. Channel interconnect skew and jitter easily consume 50% of the 2133 Mbps timing budget. These, combined with other factors, including the effects of DQS jitter, edge roll-off, impedance discontinuities, pin-to-pin capacitance variations, crosstalk and inter-symbol interference (ISI), make designs with DDR4 far more challenging.

Most importantly, stability of the power distribution network (PDN) plays a key role in signal integrity and operating margins of designs with DDR4. The maximum ripple of the PDN is specified as +/- 60mV for DDR4 as opposed to +/- 75mV for DDR3. Simultaneous switching noise (SSN) will have a major effect – in the worst case, for example, all 64 bits of a data bus transition simultaneously, with large instantaneous changes in current across the PDNs causing fluctuations in voltage levels that impact the timing margins of the transitioning signals. These simultaneously switching outputs (SSO) affect memory and other serial I/O data integrity on the board.

Nowadays, some board designers use power-aware SI simulation tools to provide some level of assurance that things will work properly. This involves the modeling of the copper shapes that comprise the power and ground planes, as well as the vias that run through them, along with the coupling to the signal traces. These vias essentially act as radial transmission lines that excite the parallel plate plane structures, perturb the power supplied to the chips, and couple noise back onto the signals as well.

In addition, decoupling capacitors must also be modeled and incorporated into the simulation, as does the voltage regulator module (VRM).

The higher speeds, lower voltages, and smaller unit intervals of DDR4 memory make the overall design rules for newer designs far more complex. And with the advent of next-generation memory technologies, including Hybrid Memory Cube (HMC)³ and High Bandwidth Memory (HBM)⁴, board design and test is expected to become more complex.

Board design variances

Given the less-than-infinite time available to create a perfect design which follows all the above known rules, some can be bent rather than broken and still yield a good, high-margin design. However, in reality, not all rules can be precisely followed, resulting in some variances. For example, Rule #7, “Avoid 90-degree routing which can result in accumulated common mode noise effects”, is sometimes impossible to obey, based upon layout constraints. In such cases, you might be able to get away with it numerous times, but the effect is cumulative, contributing to the overall marginality of the system. Eventually, increased crosstalk, inter-symbol interference, jitter, bit errors, and other irregularities will result in reducing or even closing the margin on any given bus.

Board manufacturing defects

The most common board manufacturing defects consist of short-circuits (shorts), open-circuits (opens), and stuck-at faults (which, for the purpose of this discussion, will be included generically within the shorts classification).

Let’s examine these defects and their effects on high-speed serial I/O and memory, by illustrating specific instances as they apply to PCI Express and DDR SDRAM.

PCI Express

Shorts and open circuits on high-speed SerDes buses, such as PCI Express, may have subtle and difficult-to-diagnose effects on system performance. In other words, their presence may go

³ Hybrid Memory Cube Consortium, www.hybridmemorycube.org.

⁴ JEDEC HBM Task Force, <http://www.jedec.org/category/technology-focus-area/3d-ics-0>.

undetected until customers start complaining and warranty returns increase. The reason for this is explained below.

As an example, PCIe is made up of one or more lanes of four wires each, two transmit and two receive. A transmit pair of wires can be illustrated as follows:

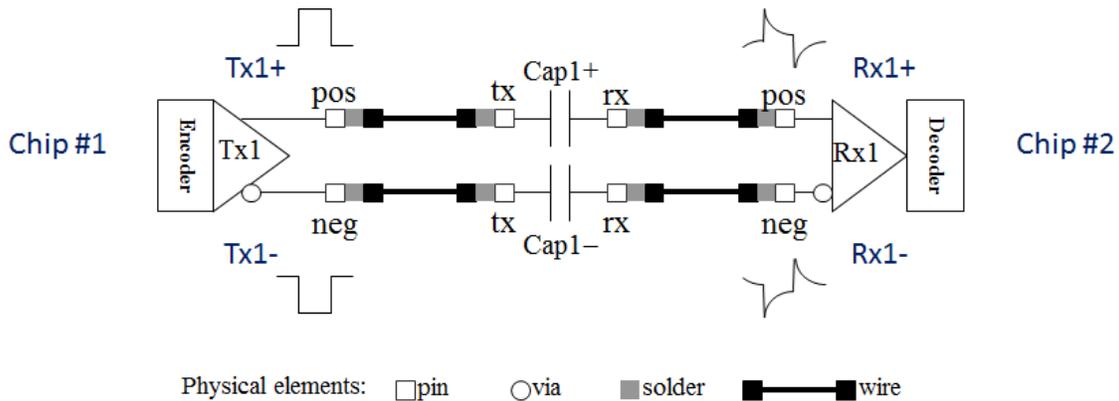


Figure 1: Transmit pair of a PCI Express lane

Now let's review a couple of failure scenarios, and see what happens.

Open-circuit: Missing Capacitor

In this example, a problem during a system assembly caused a capacitor to be left off, or somehow the capacitor was detached or disabled during its lifespan in the field. This open circuit on one net will, however, not necessarily prevent any signal from making its way to the Rx1- net at the receiver, as can be seen below:

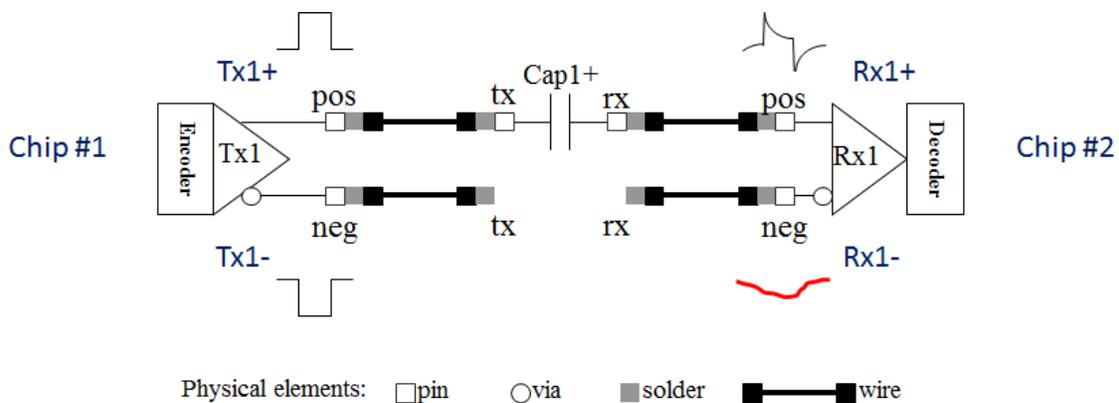


Figure 2: Open-circuit on PCI Express transmit net

Receivers work by reconstructing the differential signaling on the + and – legs of a pair. And in this case there may be sufficient coupling present for that lane to train and operate, albeit at a reduced level of performance. It will be more susceptible to crosstalk, PDN noise, jitter, and ISI. So it will operate with a higher bit error rate (BER). If it crosses the appropriate thresholds, the results will be PHY layer re-inits, datalink layer retransmissions, and ultimately lane drop-outs (either soft (intermittent) or hard).

Short-circuit: Tx1- TO GND

In this example, a short exists between one of the transmit nets and ground. Similar to the example above, this will impair the propagation of the signal to its corresponding receiver.

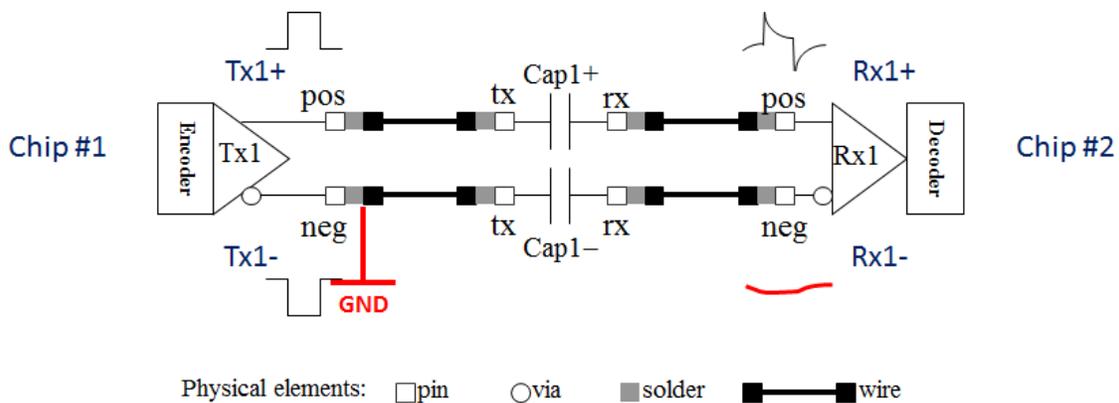


Figure 3: Short to Ground on PCI Express transmit net

But again, the receiver operates by considering the *difference* in the signals received, and it may be able to reconstruct the data stream. Whether the link drops out or continues to operate is of course dependent on a large number of factors and the bit error rate ultimately determines this.

Short-circuit: Tx1- to Tx2+

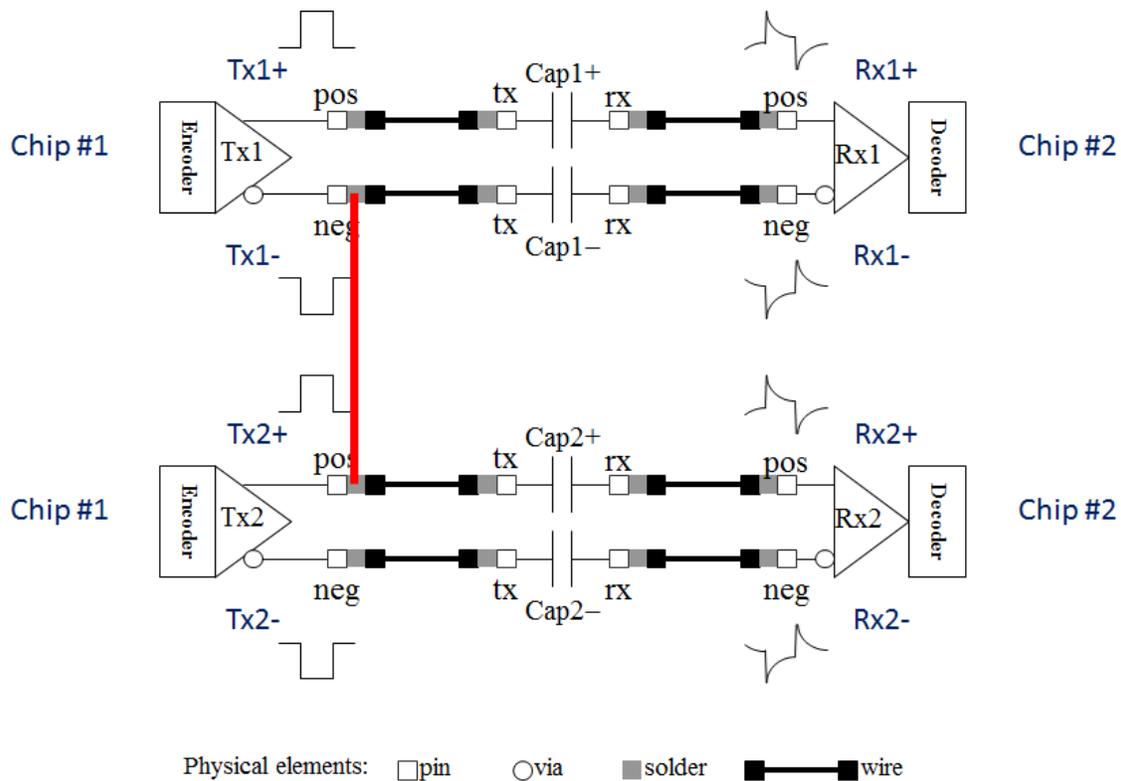


Figure 4: Short between two PCI Express adjacent transmit nets

In this example, the negative leg of a transmit lane (Tx1-) is shorted to the positive leg of an adjacent transmit lane (Tx2+). But, the receiver operates by rejecting common mode noise, and energy from Tx2+ is already coupled to both Tx1+ and Tx1-, so some of the additional coupling energy will still be rejected. So, again, the bus may continue to operate, but its performance will be impaired. The degree of impairment due to the enhanced BER will determine if the bus simply undergoes numerous packet retransmissions and PHY layer re-initializations, impacting throughput; or if it reduces in link width and/or speed, and whether such reduction is intermittent or permanent.

DDR SDRAM

DDR SDRAM, although a parallel bus as opposed to serial, exhibits similar behavior.

A simplified high-level block diagram of the pinout of a standard 240-pin DDR3 DIMM is given below as an example:

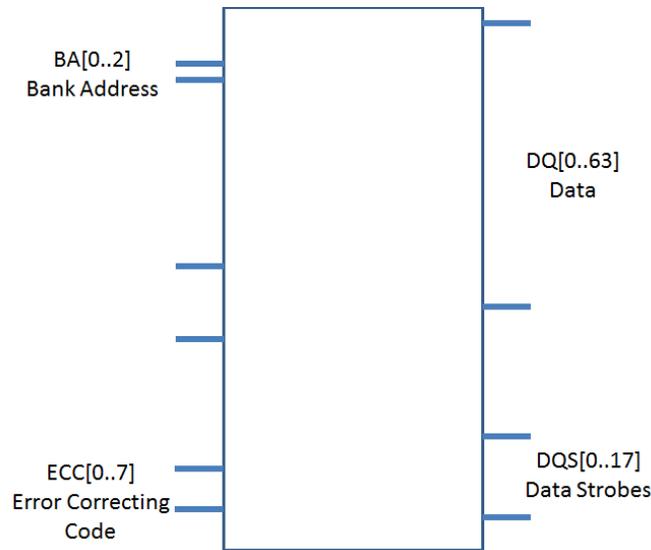


Figure 5: Block diagram of DDR3 DIMM

DDR3 memory on high-end systems differs from serial I/O in many fundamental ways. It is a parallel bus, as opposed to serial. Error detection and correction is via ECC (Error Correcting Code) memory at the “physical layer”, as opposed to high-speed serial buses which relegate such tasks to the data link layer (using cyclic redundancy checks – CRCs) and upper portions of the protocol stack. And unlike serial buses which usually use embedded clocking, separate strobe differential, but not AC-coupled, pairs (the DQS signals) are assigned per nibble of data (the DQ signals). And the strobes are not continually running “clocks” as they are in serial I/O; but rather they turn on as needed, and act as source-synchronous clocks.

To understand the behavior of a DIMM when there is a defect present, it is important to understand the process whereby the memory is first initialized (also known as “trained”). A BIOS or boot loader will run the minimal amount of code when a system is first booted in order to ensure that the memory is basically functional. So in general the BIOS will sync up DQ and DQS to optimize the system at the center timing and voltage point, and then it will do a basic test of the DQ at location 0 within each rank. So, if there are gross defects, the BIOS will either (a) disable the channel, or (b) hang the boot process with a “memory failure” post code.

If the system quietly disables the channel, this may pose a problem to conventional functional memory testers because they may not be aware of the issue. And hanging the boot process is also

an issue because, as we all know, when the screen (normal terminal output) goes dark, it takes a level of expertise to figure out what has gone wrong.

There are a number of defects which may act as variances, as follows:

A Short-Circuit on DQS0+

Since DQS are differential pair, they are immune to common mode noise, and the receivers operate by considering the difference in amplitude between the positive and negative nets of the pair. There may be enough residual signal even with one leg stuck at GND, for example, to ensure that the memory timing requirements are met – most of the time, so the boot loader will allow the memory to train. A higher number of bit errors will occur, however, most of which will be invisibly detected and corrected by the ECC, but which will impact memory performance.

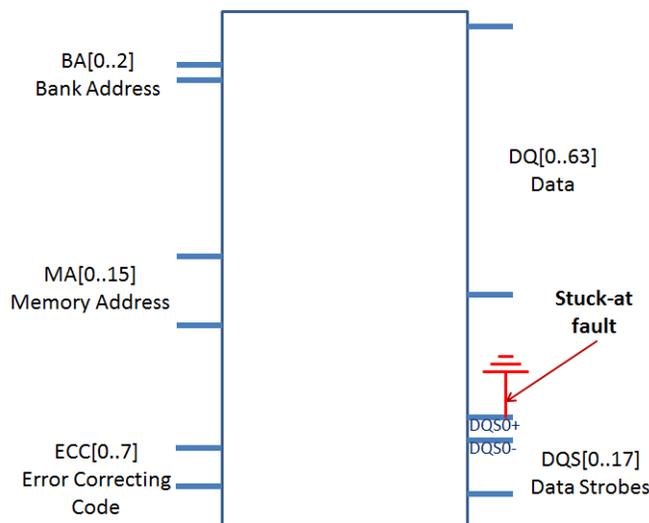


Figure 6: DDR3 DIMM DQS stuck to GND

Two DQ Shorted Together

At first glance it might seem that this kind of defect would be easily detectable. But, examining this further, consider that the boot loader memory training process may not be reliably pattern-sensitive. In a perfect world, the net signal of a shorted '1' and a '0' is 0.5, but the internal voltage biases of the receivers may in fact cause them to miss the defect, and read back what was written. So the memory may or may not train; and if it does train correctly, the system will then subsequently fail under load.

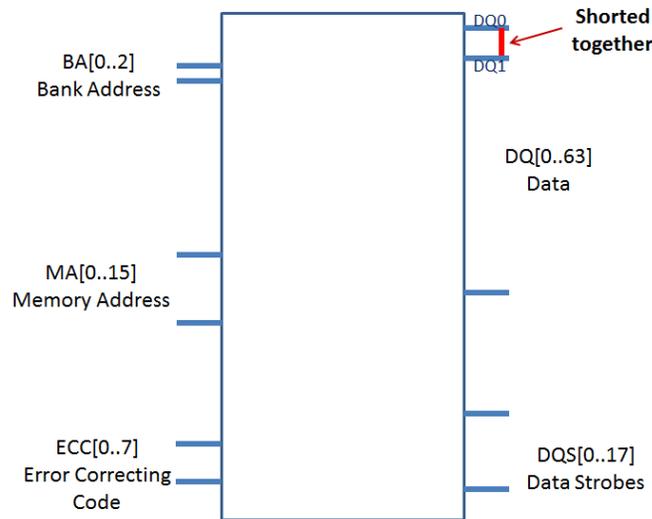


Figure 7: DDR3 DIMM two DQ shorted together

Board manufacturing variances

Despite following all of the design rules and recommended practices, variances and flaws introduced during the production of a PCB can modify the performance of transmission circuits over an extended production run. Any one variance may not cause the PCB to fall outside the acceptable tolerance range, but the cumulative effects of multiple variances may downgrade the performance of the system to an unacceptable level.

Some of the variances expected as part of the board manufacturing process are:

- Trace impedance
- Trace surface finish
- Breakout
- Pinhole, nicks, cuts
- Under etching
- Over etching
- Mouse bites
- Spurs
- Spurious copper
- Incompletely plated vias
- Annual ring breakout
- Voids

- “Head-in-pillow”
- Plating thickness
- Delamination

The first two will be examined in detail as typical examples.

Trace Impedance

The ratio of the conductor width to the distance from the power planes plays an important role in controlling impedance. The dielectric constant of the material that separates the trace from the power planes is also an important factor. These dimensions all work together to affect the inductance and capacitance of the trace and, consequently, its overall impedance.

Variation in etching time and the amount of etching material applied will affect the width of the trace. The typical tolerance on trace width is +/-10 per cent. For this example, the width (W) can range from 0.0045-in. to 0.0055-in.

The processes used for creating the core FR-4 and for pressing the pre-impregnated (pre-preg) layers of FR-4 are also inexact. For the purposes of this example, it is assumed that the core thickness is 0.005-in. (+/-0.001-in.) and that the finished pre-preg thickness is 0.008-in. (+/-0.002-in.).

The thickness of a trace is determined by the weight specification. Half-ounce copper foil is 0.000675-in. thick and it is rarely specified with a tolerance. The thickness is relatively easy to control. For this example, thickness is considered constant. The dielectric constant is 4.3 with a tolerance of +/- 0.1, or a range of 4.2 to 4.4.

Impedance is derived from a formula as described in [How to avoid poor SerDes performance caused by circuit board manufacturing variances](#)⁵. Using the example trace described above and assuming all specifications are exact, the impedance would be 47.77 ohms or close to the ideal, but the cumulative effects of worst-case extremes can result in an impedance range of 63.39

⁵ <http://www.asset-intertech.com/Products/High-Speed-I-O-Validation/HSIO-Software/Manufacturing-Variance-e-Book>

ohms (thick dielectrics, narrow trace and low dielectric constant) to 36.28 ohms (thin dielectrics, wide trace and high dielectric constant).

Of course, the distribution of this impedance range between 36.28 ohms and 63.39 ohms is Gaussian (that is, it follows a bell curve), but an impedance range this wide is enough to cause significant variation in signal reflections which result in losses and distortion at the input of the receiving device. The reflection coefficient for the example Stripline, assuming source and load impedances of 50 ohms and considering no variance from the PCB specifications, is 0.002281, while the reflection coefficients for the two worst case extremes are -0.09413 and 0.15889.

In other words, based on PCB variances that are within commonly accepted tolerances, as much as 15 percent of the energy in the signal may never enter the receiver. This is enough to have very noticeable effects on signaling margin on this trace.

Trace Surface Roughness

This example assumes that the cross section of a Stripline trace is rectangular. In practice, it is not. Process differences between suppliers, or even between different batches of PCBs from one supplier, can cause significant variations in the shape and smoothness of a trace's cross sectional perimeter.



Figure 8: Stripline cross-section, illustrating trace surface roughness

Some of the process variations that can affect the shape of a trace are over etch/under etch, substrate effects, imaging quality, oxide treatments and micro-etch oxide alternatives, and foil treatments. Assuming that the PCB manufacturer has assured that the nominal trace width is within tolerance, some of these effects can be minor. However, trace surface roughness is proving to be a significant factor in high-speed transmission lines and this variance may not be detected by common PCB tests.

Surface roughness has both good and bad effects. The metal foil's surface is intentionally made rough so that the metal will have a stronger bond to the dielectric materials during core layer production and during the pressing of core and pre-preg layers. Better bonds ensure against delamination of the layers, which can cause fatal (and possibly latent) failures on transmission lines. But increasing surface roughness also increases insertion loss.

The skin effect causes current density at higher frequencies to concentrate near the outer edge of a conductor. As frequency increases, the skin depth decreases. At 5GHz, the frequency used in this example, the skin depth is only 0.92 microns and even less for the harmonics. Surface roughness is measured in microns. Roughness typically measures from 1 micron rms (root-mean-square) to 8 microns rms. When surface roughness exceeds the skin depth, the effects of roughness on characteristic impedance and resistance (per unit length) is more pronounced. As current travels down the length of the trace on the outer shell of the metal, the rough outer surface affects the distance the current travels, which increases the resistance. The rough path of the current also forces many changes in the direction of current flow, which increases the inductance of the trace. This is analogous to driving a car over a mountain range vs. driving it on a flat and straight road. This phenomenon is accentuated at the higher harmonics of the signal. As a result, the shape of the signal will be affected.

If a PCB manufacturer changes the foil type or changes suppliers, or if the processes for ensuring good adhesion to pre-preg layers is altered, surface roughness can change from batch to batch. Due to the non-uniform nature of surface roughness, the effects are difficult to simulate and measure.

A visual representation of other manufacturing process variances and defects is as below

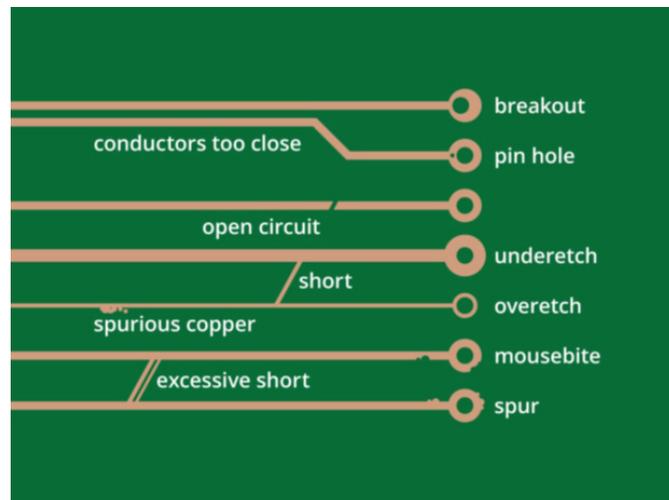


Figure 9: Circuit board manufacturing process variances

As may be evident, some of the defects and variances are localized to possibly one board, others may be isolated to one production line, while still others are related to a given production facility.

Chip design defects

As with board design, chip design is complex and subject to many “rules”, far too numerous to list here. Requirements of timing, area, power, signal integrity, routing, clocking, and others are prescribed within many design tools, but just as in board design, not all rules are known or documented, nor can they be followed.

When the rules are broken, results can be severe. A documented example of this is Intel’s \$700 million Cougar Point SATA bug, as described in the [AnandTech interview](http://www.anandtech.com/show/4143/the-source-of-intels-cougar-point-sata-bug)⁶ with Intel’s Steve Smith.

In early 2011, Intel discovered a design issue on their Cougar Point chipset, and took an approximately \$700 million charge against earnings to repair and replace affected parts and systems. This was not the largest product recall in history, but it was significant.

⁶ <http://www.anandtech.com/show/4143/the-source-of-intels-cougar-point-sata-bug>

In the interview, the root cause of the SATA problem was traced back to a transistor in the 3Gbps PLL clocking tree. This transistor was biased with too high of a voltage, which could result in a failure of the SATA ports 2 through 5 over time. In fact, the problem could be coaxed out by running the part at elevated temperatures and voltage. Intel discovered this problem itself with thermal chamber testing. The differential AC-coupled SATA physical layer uses embedded strobes derived from the PLL to clock the 8b/10b encoding, so leakage and drift in the PLL logic ultimately led to clocking marginalities and an increasing number of re-transmits over time and the associated performance hit. Ultimately, the ports would fail in months, years or possibly never.

Some design defects are inevitable, and may or may not result in a product recall, if not discovered as part of the design or system test process.

Chip design variances

Chips are in some ways mirror images of boards, but on a highly reduced scale. So, many of the variances experienced in the PCB world, such as varying impedances, trace surface roughness, pinholes, and others, are reproduced in the silicon world. Let's examine a couple which are specific to nano-geometries: charge trapping and electro-migration.

As metal oxide semiconductor field effect transistors (MOSFETs) scale to ever-smaller geometries, speed and transistor density increase, and active power per transition decreases – all desirable characteristics in today's world. But, the natural aging process becomes accelerated as this scaling continues. Let's define aging in this context as being degradation in the signal integrity (SI) within a chip, which in turn leads to more bit errors and reduced performance over time. Of course, aging will affect all attributes of a chip's performance, but SI is of particular interest due to the significant impact to system overhead at higher levels of the stack when errors are introduced at the SerDes PHY layer.

Charge-trapping

Some examples of charge trapping include RTN, BTI and HCI. An excellent article⁷ in *EDN* describes the effects of random telegraph noise (RTN) and bias temperature instability (BTI). RTN occurs when a hole or an electron is captured in an oxide trap and the captured charge is emitted from the trap. As this charge capture and charge emission continues, the drain current (I_d) fluctuates, which causes the threshold voltage (V_{th}) to shift. RTN gets worse with increasing temperature.

BTI is another example of charge-trapping which decreases I_d and shifts V_{th} . BTI in particular has a permanent component, which almost never is recovered from.

Hot carrier injection (HCI) is a similar phenomenon where an electron or a “hole” gains sufficient kinetic energy to overcome a potential barrier necessary to break an interface state. This can result in damage to the encasing dielectric material if the hot carrier disrupts its atomic structure. The presence of such mobile carriers in the oxides triggers numerous physical damage processes that can drastically change the device characteristics over prolonged periods.

Charge trapping degrades the chip performance over time, until ultimately the thresholds collapse.

Electro-migration

Another source of reliability issues is electro-migration within a chip’s interconnects. Electro-migration is the transport of material caused by the gradual movement of the ions in a conductor due to the momentum transfer between conducting electrons and diffusing metal atoms.

Although electro-migration damage ultimately results in failure of the affected IC, the first symptoms are intermittent glitches, which are almost impossible to diagnose. Differential pair used for I/O (within a chip and on a board) are in fact somewhat self-healing, in that open

⁷ <http://www.edn.com/design/test-and-measurement/4419732/1/Solve-MOSFET-characteristic-variation-and-reliability-degradation-issues>

circuits may yet have sufficient coupling to allow successful data transmission, albeit at a higher error rate.

An electron microscope view of interconnect breakdown due to electro-migration is below:

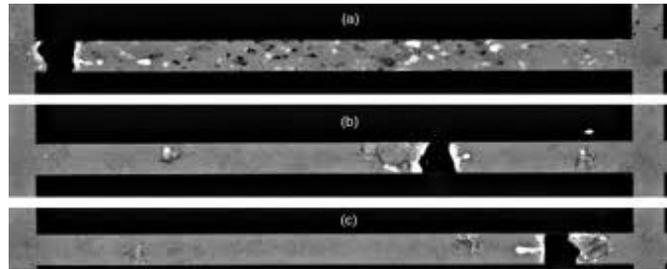


Figure 10: Silicon trace electro-migration faults

Certainly, the semiconductor industry intensively researches these reliability issues, and many mitigating technologies are in place to extend the life of ICs. As usual, it is a race between shrinking geometries, form factors and process nodes, offset by new technology innovation.

Chip manufacturing defects

The impact of manufacturing defects has become much more pronounced at and below the 90nm process node. Prior to 90nm, wafer yield profiles were mostly impacted by imperfections in the silicon and dirt – hence clean room science. Defects tended to be in the range of $2/\text{cm}^2$ (two defects per square centimeter) and tended to cluster, so taking basic precautions tended to produce acceptable yields. And keeping die small also statistically improved yields. This is illustrated below:

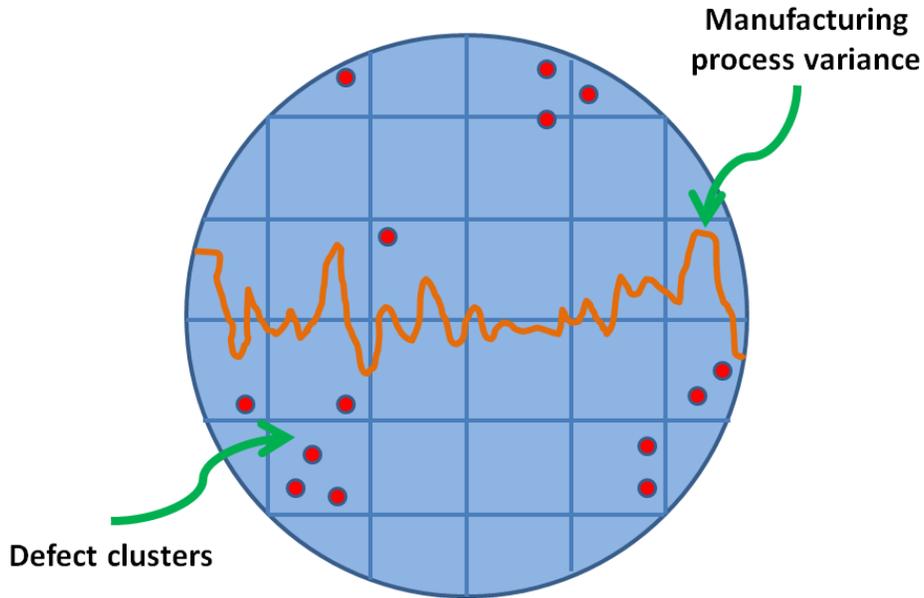


Figure 11: Wafer manufacturing process variance and defect clusters

The above figure shows the correlation between defects and the overall manufacturing process variance across the wafer.

But at process nodes below 90nm, the chip yield profiles are almost entirely dependent upon manufacturing processes. Wafer processing now includes processes such as phase-shift mask lithography, chemical-mechanical planarization and other complex steps, which are far more damaging to the routing layers. And as chips are simply miniature versions of boards, the same sort of opens, shorts and stuck-at faults that plague board production also affect chip manufacturing.

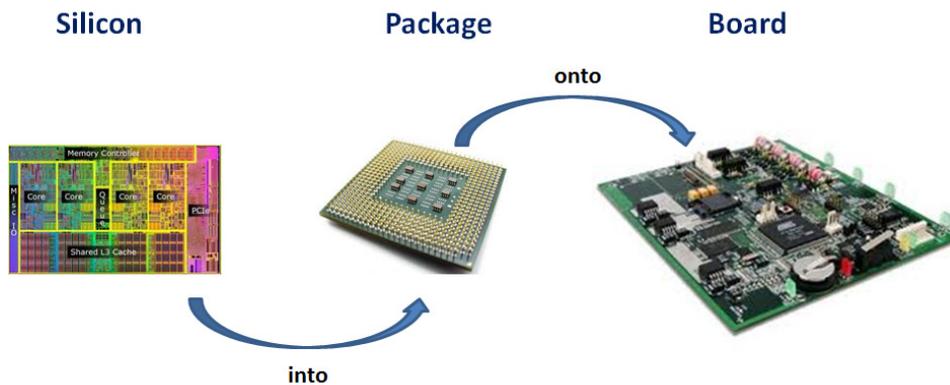


Figure 12: Silicon, packages, and circuit boards

Looking at a cross-section of a flip-chip bonded to its package which in turn is soldered to a board, it can be seen that there are plenty of opportunities for structural defects:

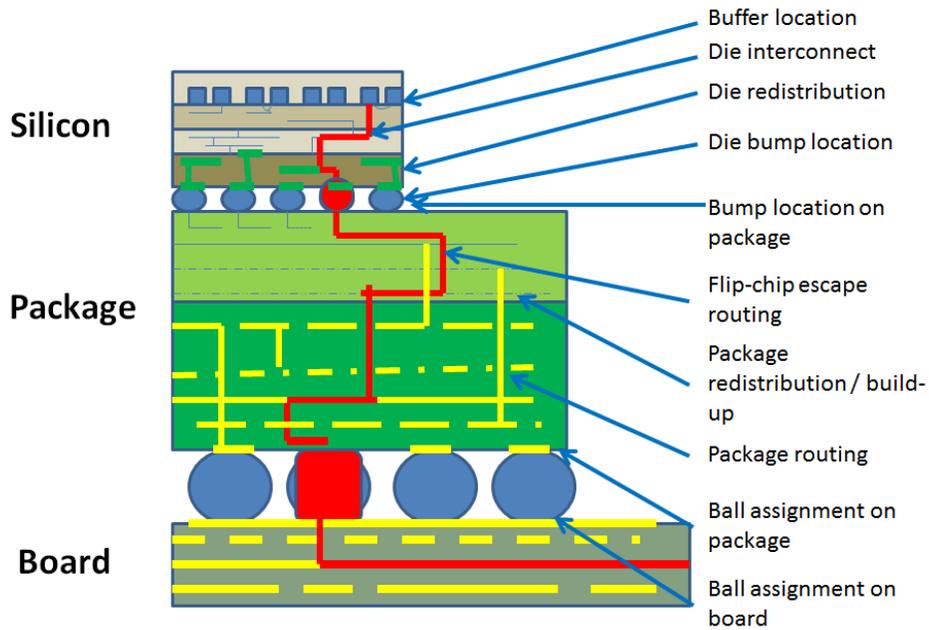


Figure 13: Flip-chip cross-section on circuit board

As we shall see in detail shortly, some of these faults are catastrophic; that is, the chip is simply “bad” and will be non-functional. Other defects will allow the chip to operate, but with degraded performance, or intermittent failures.

Chip manufacturing variances

In addition to outright structural faults introduced into the chip assembly process, variances are derived from the same lithography and planarization processes. Just looking at the traces, this can be seen below:

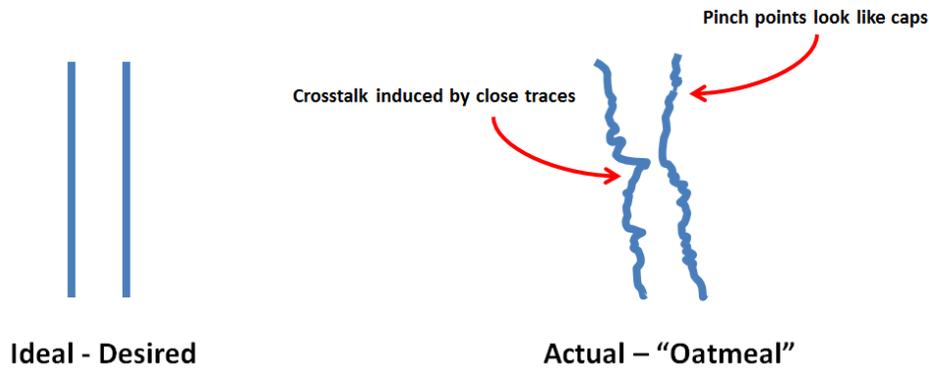


Figure 14: Chip trace variances due to lithography and planarization processes

The cumulative effect of defects and variances within devices, due to either design or manufacturing processes, introduces marginality within the chip. The impact on the chip’s performance can be represented as a set of Gaussian distributions given for different lots.

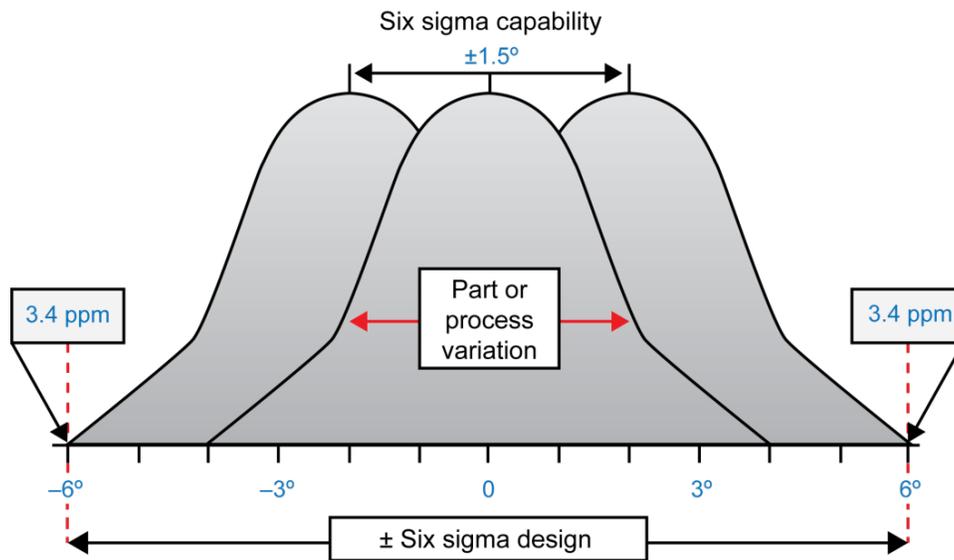


Figure 15: Chip margins across lots expressed as Gaussian distributions

Chip manufacturers must test rigorously for performance across hundreds or even thousands of devices, knowing that these must perform adequately when integrated into boards and systems. Because, fundamentally, semiconductor suppliers cannot generate income unless their customers are able to ship product, they need to ensure that their devices can withstand operating conditions that might be encountered within any given environment. In other words, chip marginality cannot exceed a given defined “guard band”.

The guard band is defined by validating the operating margins of hundreds of samples. Typically, voltage (eye height) and timing (eye width) are varied by test IP within the silicon to gauge the size and shape of the chip's eye. An example of some of the factors taken into consideration in defining the minimum overall eye height, which is part of the guard band or eye mask, is shown below:

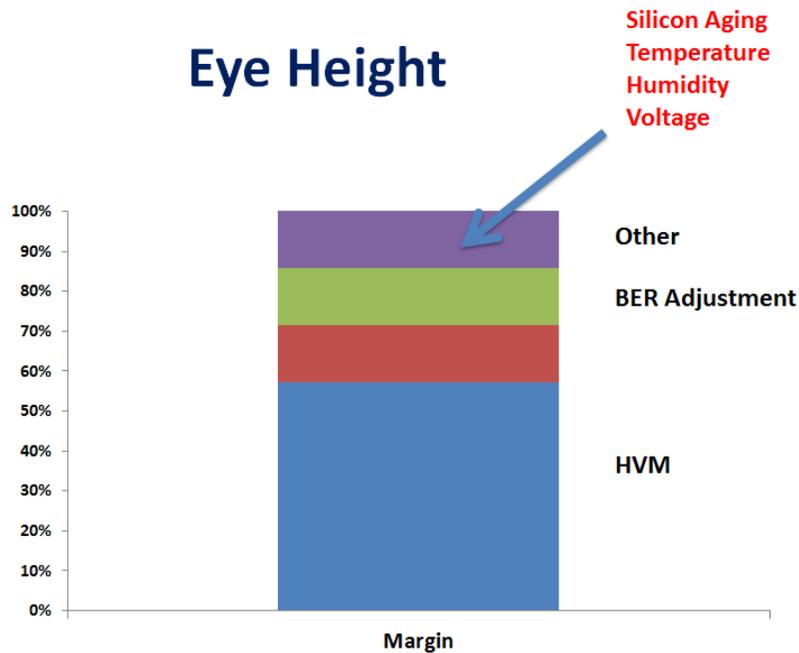


Figure 16: Chip eye mask as function of variances and empirical data

The size of the eye mask is defined statistically by:

- HVM: High Volume Manufacturing, which takes into account expected chip manufacturing process variations
- Small Sample Size: a buffer given by the fact that it is possible to only discretely and rigorously test a small number (in the hundreds, at best) of parts within a production run
- BER Adjustment: typically, for buses that run at Gbps speeds, a bit error rate (BER) measurement is made at each margining point to detect correctable or uncorrectable faults. Give that, for example, the acceptable bit error rate threshold of Intel QuickPath Interconnect is 1×10^{14} , an extremely long period of time (in the hours or days) would be needed at each margining point (defined voltage and timing step) to detect such errors. The BER Adjustment extrapolates the error rate.

- Other: Adjustments are also made for silicon aging, humidity, pollution, process/voltage/temperature (PVT), and other factors.

Semiconductor suppliers' internal tools gather data across a statistically large enough sample size to define the size and shape of the eye mask, to be used in margining tests done by their customers, as in the following:

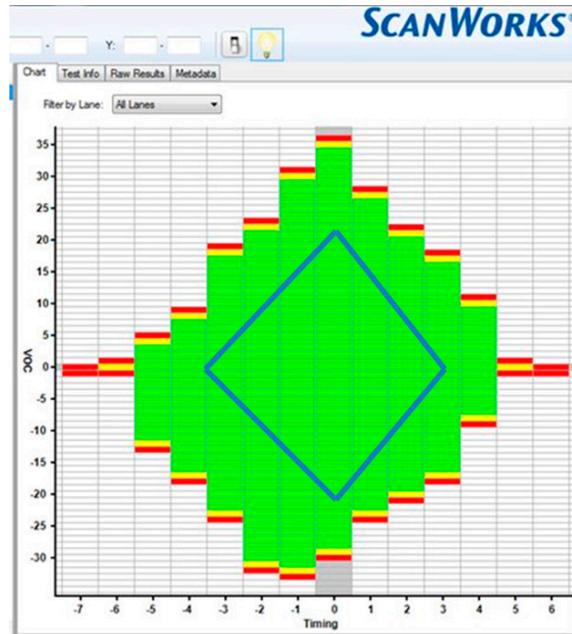


Figure 17: Chip eye diagram in relation to eye mask

DEFECTS AND VARIANCES COVERAGE

PCOLA/SOQ/FAM

Given the above treatment of defects and variances, it is important to quantify their overall impact on a board design. Engineers must have a repeatable, deterministic means of quantifying test coverage on these defects and variances. This was attempted within the International Electronics Manufacturing Initiative (iNEMI) PCOLA/SOQ/FAM framework⁸ as described below:

Component Scoring Guidelines

P	Presence	Does the test determine the presence of the part?
C	Correctness	Does the test determine that the part is correct?
O	Orientation	Is the part oriented properly or of the correct polarity?
L	Live	Is the part electrically functional for basic activity?
A	Alignment	Can the test identify lateral displacement or minor rotation?

Interconnect Scoring Guidelines

S	Shorts	On a given interconnect, can shorts within a shorting radius be detected?
O	Opens	If there is an open on the pin/trace will there be a test failure?
Q	Quality	Encompasses excess solder, insufficient solder, poor wetting, voids, etc.

Functional Scoring Guidelines

F	Feature	Can presence or absence of a feature be detected?
A	At-speed	Can the pin/interface/feature be tested at min/mid/max speeds?
M	Measurement	Can a measurement be taken that confirms performance to a given bit error rate (BER) threshold, packet loss, write latency, etc.

It should be pointed out that the A or Alignment attribute is usually only obtainable via inspection technologies, as opposed to electrical test technologies, and thus shall not be covered in this paper. Further, the important Q attribute is subsumed under the M attribute, since such

⁸

http://thor.inemi.org/webdownload/projects/Board_and_Systems_Mfg_Test/Functional_Test/Functional_Test_Cov_SOW_V1-0.pdf

defects are usually only detectable by test technologies which take measurements such as Bit Error Rate computations, for example.

For the sake of further classification, the individual PCOLA/SOQ/FAM attributes are placed into three broad categories of test coverage:

Structural

The P, C, O, L, S and O attributes belong to a broad category designated as Structural. What this relates to is the correct or incorrect assembly of the board: are the parts in place? Do they have the right polarity? Are they soldered down properly? And so on. Structural testing is usually the first test step performed on a board within the production environment, after the inspection step. Technologies which provide structural test coverage include flying probe, in-circuit test, boundary-scan test, and others.

Functional

F and A (within FAM) designate attributes which identify whether the system operates correctly within normal operating parameters, such as nominal voltage, time, temperature, and so on. Often, in manufacturing production environments, this is a go/no-go test which is the basis upon which the decision to ship product is made. All features that would normally be used by the consumer are verified or at least those which are easily tested for within the production test time window.

Functional tests are typically a combination of off-the-shelf at-speed test tools combined with whatever specific applications are necessary to functionally verify the operation of the features unique to that board or system.

Performance

Performance-based testing is the within the M or Measurement domain; that is, it is necessary to take measurements to determine if a system is performing at the level expected. For example, it may be desirable to test whether a PCI Express Gen3 lane is running at full line rate with a bit error rate (BER) of less than 1 in 10^{12} . This could be done by running at nominal time and

voltage for significant periods of time (hours, days) or by using a performance extrapolation to margin the design based upon an eye mask prescribed by the silicon supplier.

TEST STACK

Given the above definitions and classifications, it's possible to visually correlate the PCOLA/SOQ/FAM methodology to structural, functional and performance factors, and to map each of these to a set of non-intrusive board test tools which can discover defects and variances. This can be presented visually below in a test stack

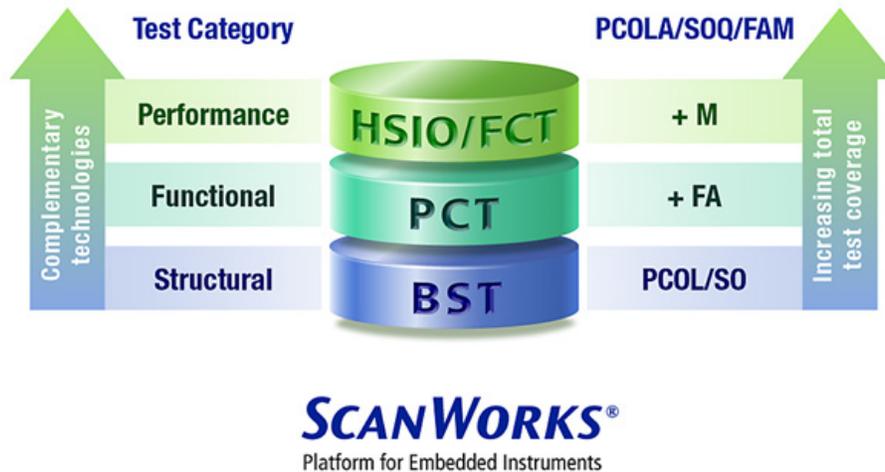


Figure 18: The Test Stack: Test categories and test coverage

The cylinders in the center denote Boundary-Scan Test (BST), Processor-Controlled Test (PCT), and Intel® High-Speed I/O (HSIO) and/or FPGA-Controlled Test (FCT). As may be seen, BST provides structural test coverage (PCOL/SO); PCT provides a superset of structural and functional test coverage (PCOL/SO/FA); and HSIO and/or FCT provide a superset of structural, functional and performance-based test coverage (PCOL/SO/FA/M).

These detection technologies constitute a complementary set of solutions for providing high levels of test coverage and fault detection/isolation. Their inherent capabilities are presented in the next section.

DETECTION TECHNOLOGIES

Boundary-Scan Test (BST)

Boundary-scan test, based on the IEEE 1149.1 Boundary Scan Standard, was conceived in the 1980s and adopted as a standard in 1990. Its adoption was accelerated in the mid-1990s as a result of pins disappearing under the silicon die in ball grid array (BGA) packages. Boundary scan tests are applied to a circuit board through a connector and the four-wire serial interface on boundary scan's Test Access Port (TAP). When implemented, this interface is commonly referred to as the 'JTAG port', which comes from the informal name of the working group that began development of the standard, the Joint Test Action Group.

Since its development, the boundary-scan standard has been adopted extensively by the industry and it is now deployed in chips, on circuit boards and in systems. Because of its widespread acceptance, the boundary-scan infrastructure has been appropriated by other applications and related standards. It is used to program logic and memory devices in-system, for example, and it provides the basis for the IEEE 1149.6 standard for testing high-speed differential and AC-coupled interconnects as well as other standards.

Boundary-scan test really shines when applied to high-speed serial I/O, such as PCI Express, and SDRAM memory. Let's examine a couple of case studies that apply the technology.

PCI Express

The differential, AC-coupled nature of PCI Express allows this bus to be somewhat self-healing, whereby some structural defects will allow the bus to transparently run, albeit at a degraded performance. Due to this, these short-circuit and open-circuit defects may be completely masked from conventional functional test. But such defects are important to detect, because they will affect the throughput of the port. Boundary scan can be used to detect these defects, subject to the implementation of IEEE 1149.1 and IEEE 1149.6 in the chips.

As we saw summarized above, some shorts and opens will cause a link to run at a degraded performance, with lower margins and higher bit error counts. For example, a single open due to a missing capacitor, or a short between Tx1+ and GND or between Tx1+ and Tx2+ will have this effect. That's why these defects are so nefarious; they may be invisible to conventional

manufacturing functional or system test, and then only cause problems subsequently out in the field. It is reasonable to inquire whether boundary scan might detect such defects.

Whether boundary scan will be effective will depend on the implementation of IEEE 1149.1 and 1149.6 (AC-JTAG) within the associated chip. Considering a hypothetical processor chip with PCI Express gen3 out to an unpopulated PCI Express connector:

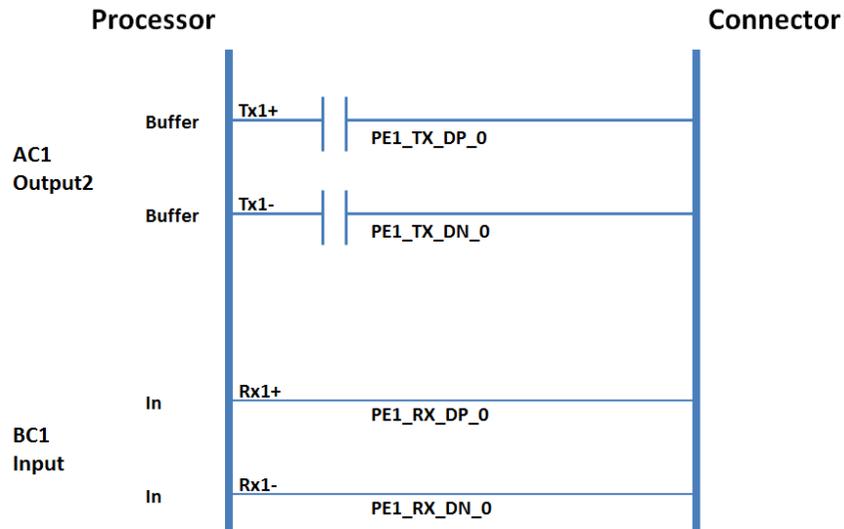


Figure 19: PCI Express lane between processor and PCIe slot

It is worthwhile to note that the PCI Express specification requires that the AC coupling capacitors be as close as possible to the transmitter buffers. They will be on-board the printed circuit board where the processor resides in this diagram. The other coupling capacitors for the processor's receive buffers will be on the PCI Express add-in card.

Also, since the PCI Express slot in this example is unpopulated, the two Rx1+ and RX1- lanes are open; that is, they are in a high-Z, or tri-stated high impedance condition. In other words, they are not being driven to any defined logic level.

Examining an excerpt from the hypothetical BSDL file for this processor for PCIe lane 1:

(from the port statement):

```
PE1_RX_DN_0;           in bit;
PE1_RX_DP_0;           in bit;
PE1_TX_DN_0;           buffer bit;
PE1_TX_DP_0;           buffer bit;
```

(from the port grouping statement):

```
"Differential_Current ((PE1_TX_DP_0, PE1_TX_DN_0)), "&
```

(from the boundary register statement):

```
"121 (BC_1 PE1_RX_DN_0,           input,           x           ), "&
"122 (BC_1 PE1_RX_DP_0,           input,           x           ), "&
"123 (AC_1 PE1_TX_DP_0,           output2,        x           ), "&
```

(from the AIO pin behavior statement):

```
"PE1_RX_DN_0           :           HP_time=8.0e-9           ; "&
"PE1_RX_DP_0           :           HP_time=8.0e-9           ; "&
"PE1_TX_DP_0           :           ; "&
```

The boundary scan internals of the processor chip, per this BSDL, can be seen diagrammatically:

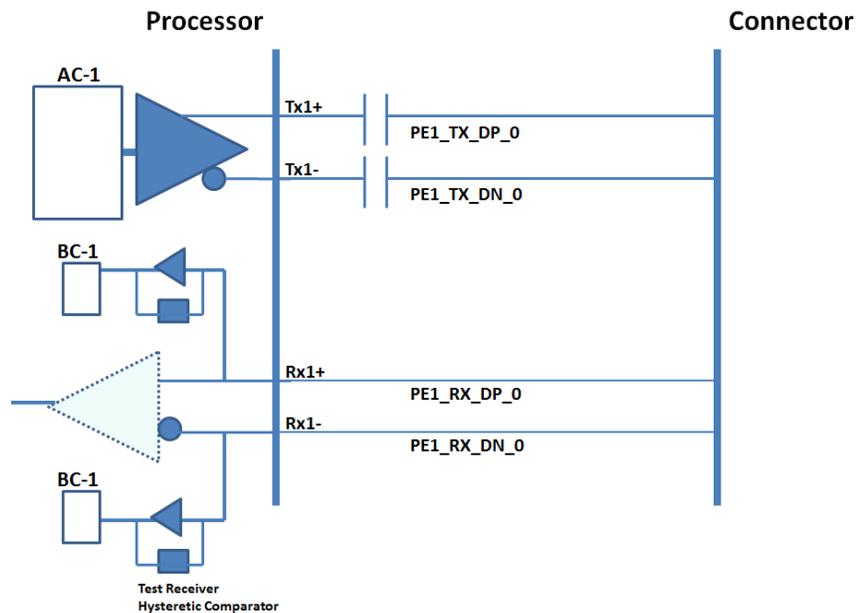


Figure 20: PCIe boundary scan cell implementation

There are several useful things to note here regarding a static DC-only level-sensitive boundary scan implementation for IEEE 1149.1:

- The positive and negative transmit nets are driven by a two-state output-only buffer, so these nets can drive but cannot sense. So shorts between these two nets could not be detected. Also, with DC stimulus only, opens on these nets cannot be detected (of course, these nets

are open because there are capacitors on the nets which in turn connect to an open connector).

- Shorts between either of these two transmit nets and a separate net with a bidir cell could be detected.
- The receive cells are input-only and can't drive.
- Since the receive cells are open and high-Z, these nets don't have a known state at all test steps, so shorts between these nets and any other nets are not covered. And of course opens cannot be detected because these nets are explicitly open.

So the boundary scan test coverage on PCIe is categorized as Class 3: some coverage on stuck-at 0 or 1, but not much else.

Things get interesting when a passive loopback card (which connects transmit nets to receive nets) is placed in the PCIe slot and some of the features of IEEE 1149.6 (AC-JTAG) are used on these nets. This provides shorts and opens test coverage on both the component side and connector side of the bus.

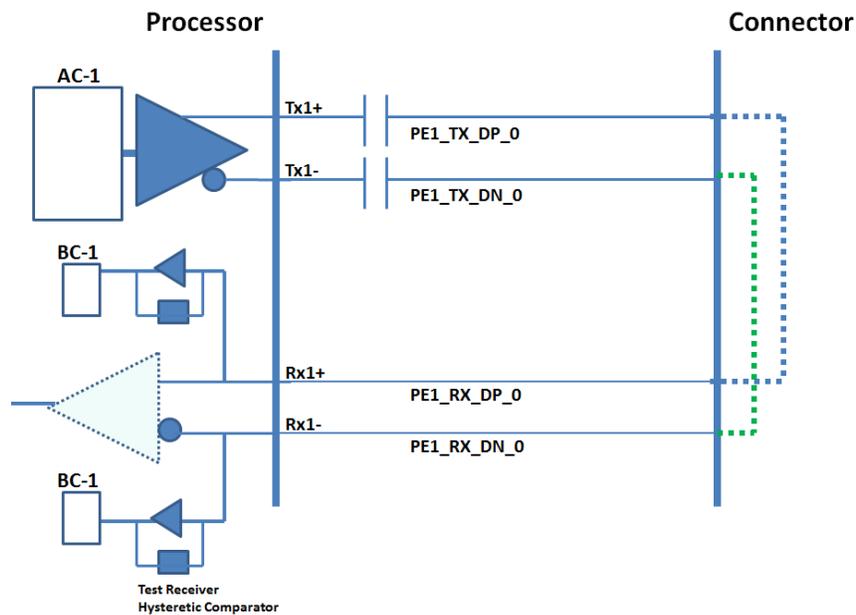


Figure 21: PCIe boundary scan implementation with passive copper loopback

Note that Tx1+ is now looped back to Rx1+ and Tx1- is looped back to Rx1-. So shorts between these nets are now explicit as part of the loopback and they will not be detected by boundary

scan. But, these defects will be detected by either processor-controlled test or HSIO using an active loopback card, as will be seen later.

Any opens on the board, connector or add-in loopback card due to missing BGA balls, connector plated through-hole issues, etc. will be detected as normally by IEEE 1149.6.

Of course, if there are shorts between Tx1+ and Tx1-, or between Rx1+ and Rx1-, no signal will go out on these differential pairs, and the lanes will fail. A clever 1149.6 interconnect test which detects edges on the receive side of the looped pair will now see the lack of an edge and isolate the fault.

It can be seen that using edge detection on the virtual receivers to the BC_1 cells using EXTEST_PULSE or EXTEST_TRAIN allows detection and diagnosis of shorts between Tx1+ to Rx1+, Tx1+ to Rx1-, Tx1- to Rx1+ and Tx1- to Rx1-.

So a prescribed boundary scan test sequence for PCI Express would consist of:

1. With nothing in the PCIe slot, run an 1149.1 static test to get Class 3 coverage on shorts.
2. Insert a passive loopback card into the PCIe slot and then run an 1149.6 test for opens and shorts.

The only defects of the Tx1+ to Rx1+ and Tx1- to Rx1- type that might escape from this methodology could then subsequently be detected with either processor-controlled test and/or HSIO.

DDR3 and DDR4 SDRAM

As was summarized above, it is critically important to detect structural defects on memory nets. Some shorts and opens (such as a DQ stuck high or open) will cause a memory channel to be mapped out or even a catastrophic error; other defects (such as two DQ shorted together or a DQS stuck at GND) may be invisible to conventional functional test, but may cause high error counts and memory performance issues after the product is shipped.

So, as was done with the PCI Express case study above, a hypothetical BSDL of a commercially-available memory controller will be postulated and the boundary register cell definitions for some of the memory nets will be examined:

```

"271  (BC_0,  *,                control,  1                )," &
"272  (BC_0,  DDR3_DQS_DN_16    output3,  x  271,  1,  z  )," &
"273  (BC_8,  DDR3_DQ_63,       bidir,     x,  271,  1,  z  )," &
"274  (BC_8,  DDR3_DQ_62,       bidir,     x,  271,  1,  z  )," &
.
.
"335  (BC_0,  DDR3_MA_10,       output3,  x,  271,  1,  z  )," &
"336  (BC_0,  DDR3_CAS_N,       output3,  x,  271,  1,  z  )," &
"337  (BC_0,  DDR3_RAS_N,       output3,  x,  271,  1,  z  )," &
"338  (BC_0,  DDR3_WE_N,       output3,  x,  271,  1,  z  )," &
"339  (BC_0,  DDR3_BA_0,       output3,  x,  271,  1,  z  )," &
"340  (BC_0,  DDR3_BA_1,       output3,  x,  271,  1,  z  )," &
"341  (BC_0,  DDR3_MA_0,       output3,  x,  271,  1,  z  )," &
"342  (BC_0,  DDR3_MA_PAR,     output3,  x,  271,  1,  z  )," &
.
.
"348  (AC_1,  DDR3_CLK_DP_0,    output3,  x,  271,  1,  z  )," &
"349  (AC_1,  DDR3_CLK_DN_0,    output3,  x,  271,  1,  z  )," &
.
.

```

The BSDL reveals some very interesting things about this chip. Cell number 271 at the top denotes a control cell, which acts for all of the device's DQ, DQS, MA, CAS, RAS, WE, BA, etc. signals. In fact, this control cell controls all of the nets on the memory controller. Since the control cell provides the means to disable the driver attached to an output or bidirectional signal, from a boundary scan point of view, it is impossible to simultaneously drive on address and control, and read on the DQ. This is a huge problem for opens testing, which requires that we have separate control of address, data and control signals. If all the memory nets are driven from the processor, then there are no memory pins that are quiet enough to monitor an incoming stimulus. This is essential for opens detection since a driver that is open will still "see" what it drives. So there is no boundary-scan opens coverage on any of the memory.

As a side note: in boundary scan implementations, it is common practice for a control cell to be provisioned at least for each given logic bus. For example, different control cells will be

implemented for address, control and data buses. In FPGAs, since I/O can be grouped in an arbitrary fashion, there's one control cell behind each pin. However, in the above BSDL example, a single control cell fans out to multiple driver enables.

Looking at the BSDL further, note that the DQS cell is BC_0, which means they are not self-monitoring and there is no stuck-at coverage on the strobes.

The DQ are supported by BC_8 cells, which means they are self-monitoring. There is stuck-at coverage on the data lines. Shorts between these nets and a separate non-memory net (such as a PCIe) which had a bidir cell will also be detected. However, DQ to DQ shorts testing is problematic because of the DQ's bias toward sensing the voltage level it is driving.

So to summarize, for this hypothetical example, the boundary-scan test coverage is:

DQ to DQ Short	No
DQ Stuck	Yes
DQ Open	No
MA to MA Short	No
MA Stuck	No
MA Open	No
DQS1+ to DQS2+ Short	No
DQS Stuck	No
DQS Open	No
DQS1+ to DQS1- Short	No
Control	No

So, the boundary-scan test coverage for this example is quite poor. Opens coverage is non-existent, and shorts coverage is very limited. If there is no physical access for in-circuit test (which is quite common on today's designs), how is it possible to achieve structural test coverage on memories?

Processor-Controlled Test (PCT)

Processor-controlled test (PCT) makes use of a board's JTAG infrastructure to access and utilize the extended debug commands provided by a circuit board's processor. Control of the processor is temporarily given over to the PCT system so that the CPU can be used to read and write memory and I/O registers in addressable devices on the circuit board. In this way, PCT exercises the functionality of the circuit board and as a result detects and diagnoses structural faults.

As it is a functional test, PCT is device and bus-centric; that is, it will exercise the devices and buses on a circuit board. In addition, because it operates at CPU speeds as an “at-speed” test, it will detect faults which only manifest themselves while the board is running at operational speeds. On the other hand, static tests such as ICT, MDA, FPT and JTAG/boundary scan are effectively DC tests. And of course PCT, which reads and writes from/to targeted devices on a board, verifies the board’s functionality, which static tests cannot. Thus, PCT’s fault coverage spectrum is broader than that of static tests.

PCT Test Generation

PCT instructs the processor to sequentially test all addressable devices on the board under test. These tests are normally carried out without booting the board to its operating system, so device initialization is handled by the PCT test program in place of the BIOS. Test programming is greatly simplified by PCT’s Automatic Test Generator (ATG), which identifies the devices present on the board and then assimilates the appropriate device profiles from a built-in library into a board-specific test script.

PCT Test Coverage

From a functional point of view, PCT can test all CPU-addressable devices, including the DDR3/4 memory and PCIe devices that boundary scan may be unable to test (due perhaps to inadequate boundary scan implementations in chips or poor BSDs). Structural faults are also detected and diagnosed as a by-product of the functional testing. PCT includes an extensive coverage reporting system, which allows fault reporting to component and pin levels. This is a unique feature of PCT. This reporting is achieved by importing the board’s netlist and then assigning parts and pins to specific device tests during the test development process.

The following is an example of a PCT test sequence:

Processor	
Initialize	Check if power OK Check JTAG Infrastructure
Processor	Take control of CPU Check CPU ID is valid
Integrated Memory Controller	Register/access test. Configure for normal operation
DDR3/DDR4	Register/access test Configure for normal operation Run ASSET's pre-programmed memory test on all RAM banks.
Chipset	Register/access test. Configure for normal operation.
PCI Express	Register/access test. Configure for normal operation. Requires suitable PCIe endpoint loopback.
Chipset	
Initialize	Check if power OK Check JTAG Infrastructure
Chipset	Register/access test. Configure for normal operation.
SATA Ports	Register/access test. Configure for normal operation. (Based upon presence of terminating SATA device).
USB Ports	Register access test. Configure for normal operation. (Based upon presence of terminating USB device).
PCIe Ports	Register access test. Configure for normal operation. (Based upon presence of terminating PCIe device).
Ethernet	Register/access test. Configure for normal operation.
Audio	Register/access test.
VGA	Register/access test.
TPM on SPI	Register/access test.
Flash on SPI	Register/access test. Configure for normal operation.
Embedded Controller on LPC	Register/access test.
Port 80/LPC Slot	Access test.

In addition, PCT works by creating comprehensive tests comprised of function calls written in the Tcl/Tk high-level language. The functions themselves provide a rich level of control over test

coverage and fault spectrum. The following figure shows a handful of the functions that would be used within a typical board test.

DownloadHexFileToMemory	Loads the UUTs memory with a hex file. For example, flash data which can be written to the device directly
ExecuteUserDiag	Runs a user diagnostic routine in memory
GPIOWrite	Writes to a discrete I/O connection
IOBusTest	Performs an I/O bus integrity check for stucks, opens and shorts
MSRRead	Reads the contents of the specified 64-bit Model Specific Register
RAMTests	Executes any of a number of different RAM tests
ReadGPR	Reads the contents of a CPU General Purpose Register
ReadIO	Reads a UUT I/O port address in 8/16/32-bit format
RunUUT	Applies a reset to the UUT and lets it run under its own BIOS
UploadFlashSector	Uploads individual sectors of a specified flash device to a specified upload file in binary format
UUT_CheckMemory	Checks a range of a memory for a specified data value

These functions provide fine control for board-specific test requirements. For example, the RAMTests function can be invoked in any of five different types:

1. RAM Bus Test - used to diagnose Address and Data Bus defects between the processor and RAM under test.
2. Basic R/W RAM Test - used to diagnose all common memory defects, i.e. stucks, opens, shorts, bad cells, etc.
3. R/W RAM Test - performs a more intensive test, which can be useful for detecting more complex types of memory problems.
4. DRAM Refresh Test - allows a user to verify that DRAM is refreshing correctly.
5. RAM Bus Test via FIFO - used to diagnose Address and Data Bus defects between the processor and RAM under test. This option is required for certain memory controllers

that contain a FIFO, which must be flushed in order for data to be written out to physical memory.

Intel® High-Speed I/O (HSIO)

Built-in Self Test (BIST) commonly refers to test mechanisms or instruments that are embedded into chips and which can be applied in non-intrusive board test applications. A particular example of this is Intel®'s Interconnect Built-In Self Test (IBIST) technology which is within the ASSET ScanWorks HSIO tool and which is being embedded by Intel, Avago and other semiconductor and IP providers into next-generation chips and chipsets. The embedded Intel IBIST functionality can be applied in a number of ways, including structural tests in non-intrusive board test applications. It can also be used in design validation applications to validate the performance of high-speed serial buses and memory on circuit boards.

Intel IBIST uses PHY-layer highly-stressful bit pattern generation-and-capture capabilities in conjunction with offsets for time and voltage to validate system operating margins and create a virtual eye diagram of memory and serial I/O performance. As such, it exposes a system to worst-case synthetic “killer patterns” and eliminates the masking effects of scrambling and encoding to deliver long strings of consecutive individual bits (CIDs). Encoding and scrambling also have the benefit of achieving DC balance in the bit stream, reducing data wander and improving error recovery; HSIO removes this as a variable. This stresses the ability of SerDes clock recovery circuits to lock and hold, and detects the effects of jitter, crosstalk, ISI and other impairments.

By margining off of zero time and voltage, it also reduces the amount of test time significantly for the detection of the cumulative effect of silicon and board design and manufacturing defects and variances, and significantly improves accuracy over conventional functional test. This is important because conventional test technologies operate at the ideal voltage and time offsets for memory and I/O, and they use conventional, but limited, ways to detect errors.

Considering PCI Express Gen3 as an example, where data is encapsulated within a TLP (Transaction Layer Packet): a CRC (Cyclical Redundancy Code) protects the entire packet (with the exception of the framing start/end bytes). A TLP looks like this:

PCIe Transaction Layer Packet (TLP)



Figure 22: PCI Express Transaction Layer Packet (TLP)

For PCIe Gen3, the link CRC (LCRC) is 32 bits wide based on the large, variable-sized payload. The end-to-end CRC (ECRC) provides some level of data integrity for different link hops. The bit error rate threshold for PCIe Gen3 is 1 in 10^{12} . So, a test would need to run on average for approximately 100 seconds, or 1.5 minutes to detect an error.

Intel QuickPath Interconnect (QPI) uses a smaller, fixed-sized payload, for which the link CRC is 8 bits. The bit error rate threshold for QPI is 1 in 10^{14} . Detecting whether the BER threshold is exceeded here would require a test lasting 10,000 seconds or 2.8 hours.

Also, CRCs are limited because they use polynomial arithmetic to create a checksum against the data they are intended to protect. The design of the CRC polynomial depends on the maximum total length of the block to be protected (data + CRC bits), the desired error protection features, and the type of resources for implementing the CRC, as well as the desired performance.

Tradeoffs among all of the above aspects of a CRC polynomial are quite common. For example, a typical PCI Express Gen3 packet CRC polynomial is:

$$x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{11} + x^{10} + x^8 + x^7 + x^5 + x^4 + x^2 + x + 1$$

Whereas for Ethernet frames, the CRC generator may use the following polynomial:

$$x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{12} + x^{11} + x^{10} + x^8 + x^7 + x^5 + x^4 + x^2 + x + 1$$

The PCI Express 3.0 CRC-32 for the TLP LCRC will detect 1-bit, 2-bit, and 3-bit errors. 4-bit errors may escape detection. Bit slips or adds have no guarantee of detection. Burst errors of 32 bits or less will likely be detected.

For QPI, the 8-bit CRC can detect the following within flits:

- All 1b, 2b, and 3b errors

- Any odd number of bit errors
- All bit errors of burst length 8 or less
 - Burst length refers to the number of contiguous bits in error in the payload being checked (i.e. '1xxxxxx1').
- 99% of all errors with burst length 9
- 99.6% of all errors of burst length > 9

Using a technology, such as Intel IBIST, which uses a combination of stressful patterns, voltage/time margining, and bit-by-bit compares provides the most robust means for checking the health of an interface.

MATRIX OF TESTING CAPABILITIES

Given the above, it is possible to create a matrix containing defects on memory and serial I/O and the detection and diagnostics capabilities of the test stack technologies. This is shown below:

		BST		PCT		HSIO		Combined Coverage		Predominant Methodology
DEFECT (Memory):	Boot/Impact	Detect	Diagnose	Detect	Diagnose	Detect	Diagnose	Detect	Diagnose	
DQ to DQ Short	Channel mapped out	No	No	Yes	Rank	No	No	Yes	Rank	PCT
DQ Stuck	Channel mapped out	Yes	Bit	Yes	Rank	No	No	Yes	Bit	BST
DQ Open	Channel mapped out	No	No	Yes	Rank	No	No	Yes	Rank	PCT
MA to MA Short	None/Hang	No	No	Yes	Bit	No	No	Yes	Bit	PCT
MA Stuck	None/Hang	No	No	Yes	Bit	No	No	Yes	Bit	PCT
MA Open	None/Hang	No	No	Yes	Bit	No	No	Yes	Bit	PCT
DQS1+ to DQS2+ Short	None/Enhanced ECC	No	No	No	No	Yes	Byte	Yes	Byte	HSIO
DQS Stuck	None/Enhanced ECC	No	No	No	No	Yes	Byte	Yes	Byte	HSIO
DQS Open	None/Enhanced ECC	No	No	No	No	Yes	Byte	Yes	Byte	HSIO
DQS1+ to DQS1- Short	Channel mapped out	No	No	Yes	Rank	No	No	Yes	Rank	PCT
Control	Channel mapped out	No	No	Yes	Poor	No	No	Yes	Poor	PCT
DEFECT (SerDes):										
Tx1+ to Tx1- Short	Lane failover	Yes	Net	Yes	Lane	No	No	Yes	Net	BST
Tx1+ to Tx2+ Short	Enhanced BER	Yes	Net	Yes	Lane	Yes	Lane	Yes	Net	BST
Tx1+ to Rx1+ Short	Enhanced BER	No	No	Yes	Lane	Yes	Lane	Yes	Lane	PCT
Tx1+ Stuck	Enhanced BER	Yes	Net	Yes	Lane	Yes	Lane	Yes	Net	BST
Tx1+ Open	Lane failover	Yes	Net	Yes	Lane	No	No	Yes	Net	BST

Figure 23: Matrix of memory and serial I/O defects versus test coverage technology

The DEFECT column relates to a postulated defect; for example, a short between two memory data lines being designated as a DQ to DQ short or a short between two adjacent positive and negative legs of a transmit differential pair being designated as Tx1+ to Tx1- Short.

The Boot/Impact column relates to what happens when a given defect exists. For example, if two memory data lines are shorted together, the system will not boot and the memory channel will be mapped out.

The Detect/Diagnose columns refer to whether the given test technology will detect the fault and, if so, to what level of diagnostics it is capable.

The Combined Coverage column identifies whether, given all of the three technologies being applied, test coverage exists on that defect and, if so, what's the best level of diagnostics possible.

The Predominant Methodology column identifies which test technology is most effective in terms of being able to detect the defect, the highest granularity of diagnostics, and the least test time.

Given the hypothetical boundary scan constraints as postulated in the earlier section, it can be seen that there is very little BST coverage on memory defects. The DQ stuck scenario is the one area where BST provides any level of test coverage and, in fact, is the dominant methodology for this kind of defect, as its test time is very short and diagnostics is to the bit level.

PCT is the dominant methodology for memory and data address defects, given its short test time, and its ability to detect and diagnose even in situations where the system will hang or the affected channel will be mapped out.

HSIO can provide no coverage where the system is hung or the affected channel is mapped out, since it requires a running BIOS to train the memory and I/O. But, its strength is in detecting defects that will escape BST or PCT, as with memory strobes; and, in fact, it diagnoses combinations of defects and variances that may be undetectable without extremely long test times at nominal time and voltage.

A combination of all three of BST, PCT and HSIO are needed in this example to detect the universe of defects described.

CONCLUSION

Defects and variances are latent in every printed circuit board design. Until recently, variances were not of concern to the test engineer, and defects could be easily detected with legacy test technologies. But now, with the advent of higher-speed memories and the ubiquity of differential signaling on I/O, the effects of defects and variances are intertwined and must be taken into consideration within the validation and test processes for circuit boards.

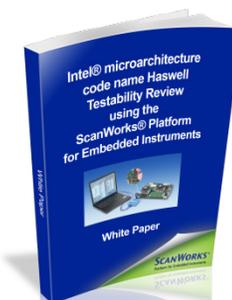
We've seen that the "self-healing" nature of high-speed signaling allows a system to function even in the presence of defects and variances that exceed what used to be an acceptable threshold. However, designs with these attributes will be prone to poorer performance, intermittent dropouts, and ultimately system failure. These effects are exacerbated over time.

With the limited access permitted on today's complex designs, conventional bed-of-nails test will often miss structural defects. And because of the above-mentioned self-healing capability, marginal systems may appear on the surface to operate normally, and issues may escape conventional functional test technology.

A new class of validation, test and debug tools based upon on-chip embedded instrumentation can assist with much greater levels of test coverage and diagnostic granularity than has been possible in the past. Such technologies include boundary-scan test, processor-controlled test, and Intel® High-Speed I/O, among others. This paper has demonstrated how these technologies can be applied to characterize a marginal design, and recommend solutions to improve its performance and reliability.

LEARN MORE

Learn more about using the ScanWorks platform when applied to the Intel® microarchitecture code name Haswell designs.



Register Today!