

Thermal HALT— A Tool for Discovery of Signal Integrity and Software Reliability Issues

Kirk Gray

Accelerated Reliability Solutions, L.L.C.
Lafayette, Colorado
kirk@acceleratedreliabilitysolutions.com

Abstract—Failures in the execution of software can have many different causes. Errors in coding, incorrect sequence of software execution, and other factors result in software “bugs” that when discovered must be corrected for reliable system operation. This paper, however, focuses on another cause of software failures; namely, those that come from variations in the many tiers of component manufacturing and system assembly. This paper shows how the application of thermal HALT on electronic circuits and systems help skew parametric performance at the circuit board and system levels and increase the probability of discovering marginal signal quality and integrity which can lead to software operational failures.

Index Terms - HALT, signal integrity, software reliability

I. INTRODUCTION

HALT (Highly Accelerated Life Testing) is a reliability development tool that discovers the strength of operating electronics by applying measured and increasing stress to an empirical (observable, not theoretical) stress operational and destruct limit to find and improve the strength and robustness of the product during the design phase. It has been well established that HALT can rapidly precipitate and detect a large percentage of the weaknesses and latent defects in electronic and electromechanical hardware, by using thermal stepped stress, vibration and voltage margining, as well as other stresses to operational and destruct limits. A more useful working definition of HALT would be: a highly accelerated “limit” test (as that is what it fundamentally discovers) of the empirical stress operational and destruction limits.

Historically, HALT has been used to find mechanical weaknesses in solder, interconnects, and component packaging in order to increase product strength and robustness. Thermal cycling of circuit boards and assemblies creates expansion and contraction of materials and the different thermal coefficients of this expansion (TCE) in the material bonds and interfaces results in fatigue damage during each cycle. HASS (Highly Accelerated Stress Screens) are applied before shipment to precipitate and detect latent defects (hidden defects that will cause failure over time) that may result from manufacturing errors.

When Greg Hobbs, PhD., P.E. created the HALT and HASS methods back in the 1980s, digital systems were not as prevalent and bus speeds were much slower than in today’s electronics. By comparison, today’s electronics have

increased clock and bus speeds that are up to 1000 times faster than 20 years ago; however, despite these advancements, the materials used in circuit board design has not changed significantly in the same timeframe. FR4 is still being used for PWBA (Printed Wiring Board Assembly) at much higher speeds than were ever expected when it was initially developed. As data bus speeds increase, effects causing errors in data transmission that were not significant at the time (when bus frequencies were in the megahertz range) became dominant at the gigahertz frequencies of today’s digital systems. Interconnect resistance, capacitance, and inductance are frequency-dependent and, as bus speeds increase and geometries continue to shrink, these variables may prove difficult— if not impossible— to accurately model.

The materials and methods of fabrication of components, circuit boards, and systems assemblies affect the quality and speed of propagation of digital signals in a system. In some instances, this may lead to a race condition in signal transmissions. A race condition (or race hazard) is the behavior of software in a digital system where the output is dependent upon the sequence or timing of software instructions to the logic devices. The term originates from the idea of two signals racing each other to influence output first. A race condition becomes a software bug when events do not happen in the order the programmer intended. If the signal quality or signal propagation is skewed enough due to hardware variations, solder defects, or fluctuating temperatures it may result in timing errors or false binary logic states.

The continual decrease in metallization dimensions and increase in bus frequencies will result in increased sensitivity to fabrication variations [1]. Studies have demonstrated that crosstalk in a 0.8 μm CMOS device can increase the circuit delay 100% from the mean due to process variations [2]. “Modern bus designs have become so fast that the designer must calculate the voltage and timing numbers to a resolution as small as a few millivolts and a few picoseconds. This degree of resolution was unheard of in computer designs just a few years ago” [1]. These new high frequency bus designs will lead to more SI errors in manufacturing processes when there are process variations at all tiers of assembly. Marginal signal integrity (SI) of a digital system can result in bit errors that, in turn, cause either intermittent operational software failure or degraded performance due to error correction processes. Frequently, the returned hardware may be

considered good when tested on the bench due to the intermittent nature of the failure. This increases the costly churn of sending out new circuit boards and subsystem parts to replace parts returned that have no detectable failures.

Higher processor and bus clock frequencies, the higher density of IC's and systems and their impact on operational reliability lead to/introduce challenges of finding and correcting marginal SI issues quickly during the development phase: before market release. As digital clock and bus speeds increase and circuit features get smaller, thermal HALT offers potentially significant benefits for the discovery of marginal SI issues during new product development. Thermal stress skews electronic parametric operational conditions of components and interconnects and stimulates the future intrinsic parametric variation that may occur during mass manufacturing or occur over time from fatigue damage or material aging. There is currently a dearth of published data or studies on how the variations in manufacturing materials and methods at all tiers of the circuit board assembly affect SI and software failures. Therefore, thermal HALT is an excellent tool and opportunity for the electronics industry to discover issues with interactions of hardware and SI that lead to marginal software reliability earlier, during product development.

This paper discusses how thermal step stress methods, aka HALT, use empirical high and low temperature operational limits to discover marginal SI reliability issues. By stressing systems to thermal operational limits, the weaknesses in signal timing and propagation have a higher probability of discovery before market release. In doing so, it suggests that, as a tool, thermal HALT is highly effective in discovering operational reliability as digital data transmission speeds continue to increase.

II. BASICS OF DIGITAL SIGNAL PROPAGATION

It is widely known that in digital systems the basic function is to transmit binary information with electrical signals that represent a 1 or a 0. Ideally this involves sending and receiving an electrical square wave with the higher voltage being a "1" and the lower voltage a "0". As the speed of signal transmissions increases the ideal square wave becomes a trapezoid shaped with wide boundaries due to signal timing variations which cause skews and jitters in the signal.

Every conductor has a capacitance, inductance, and frequency-dependent resistance. At a high enough frequency, none of these things is negligible. Thus a wire is no longer a wire but a distributed parasitic element that will have delay and a transient impedance profile that can cause distortions and glitches to manifest themselves on the waveform propagating from the driving chip to the receiving chip. The wire is now an element that is coupled to everything around it, including power and ground structures and other traces. The signal is not contained entirely in the conductor itself but is a combination of all the local electric and magnetic

fields around the conductor. The signals on one interconnect will affect and be affected by the signals on another. Furthermore, at high frequencies, complex interactions occur between the different parts of the same interconnect, such as the packages, connectors, vias, and bends. All these high-speed effects tend to produce strange, distorted waveforms that will indeed give the designer a completely different view of high-speed logic signals [1].

A digital signal can be observed on a logic analyzer. The signal shown in figure 1 is referred to as an eye diagram; derived from the fact that the space between high and low signal waves is similar to the shape of a human eye. A compliance mask overlay on the signal, also shown in figure 1, indicates the area that the signal must not cross, or the receiver may misread the one or zero being transmitted.

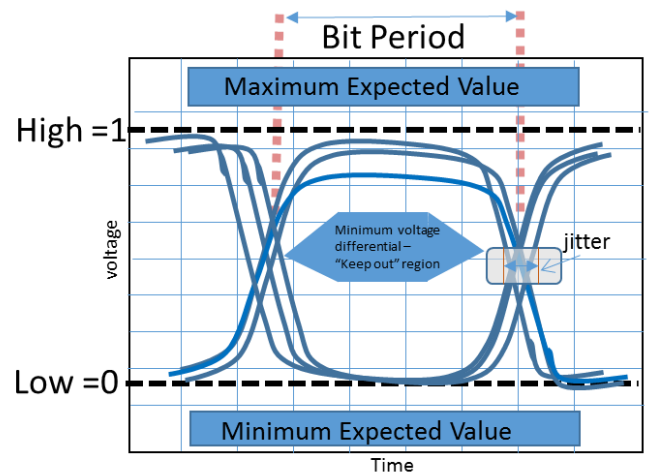


Fig. 1. The eye diagram with Compliance Mask [adapted, [3]]

III. THERMAL STRESS EFFECTS ON SIGNAL PROPAGATION

Thermal stress affects electrical characteristics and speed of signal propagation in materials used for circuit boards and systems. As the temperature of electrical conductors increases, the physical vibration of atoms in the material increases thereby reducing the charge carrier mobility and therefore increases electrical impedance. Doped semiconductors have a more complex response to temperature, but generally will increase resistance as temperature increases at typical use conditions [4]. Temperature can be a significant factor in the energy band gap, current density, leakage current, interconnect resistance, and mobility of charge carriers combined with the complex interplay of scattering parameters due to surface roughness, phonon, bulk charge, and Coulombic scattering [4]. In conductors such as aluminum and copper, the wire resistance can change as much as 72% and 77% (respectively) over a military specified temperature range of -55°C to 125°C [4]. Throughout the process of mass manufacturing circuits and systems, variations in material quality, semiconductor and other component fabrication processes can shift the effect of

temperature on signal propagation during the production or manufacturing life cycle that causes robust SI margins to decline to a level that causes errors in data transmission.

IV. SOFT FAILURES AND THE NFF (NO FAULT FOUND) PROBLEM

There are many causes of electronics systems that, when returned to the manufacturer, seem to have no defects or are determined to have No Fault Found (NFF). There are many underlying reasons for NFF conditions, including errors in test equipment, misapplication of the product, and the lack of re-testing under field stress conditions. Another contributing factor is the simultaneous replacement of many subsystems, in which only one subsystem has failed, with the goal being to return equipment to operation as quickly as possible and not spend time isolating the specific device failure [5]. Another cause of intermittent or marginal operational reliability comes from poor SI and the errors it creates in binary data. Poor quality of digital signal transmissions or SI can lead to bit errors. The signal quality, timings, skew and jitter and electronic noise levels are affected by the variations in materials and manufacturing processes, along with environmental conditions such as temperature, voltage and humidity [6]. Software failures due to these variations are usually “soft failures;” where a system can be reset and operate normally potentially. Depending on the frequency of these operational failure events, the user may or may not tolerate their occurrence. When a customer determines the resettable fault occurs too frequently it may result in returning the device or product to the manufacturer.

When a system is returned due to field failure it may be disassembled and each subsystem will be sent for failure analysis. If the system is returned due to reported soft failures as a result of SI marginality, each subsystem may test as functional. The marginal SI may be due to tolerance stack-up of cabling, connectors, or thermo-mechanical stresses on the circuit board that may be included in the failure analysis testing. The returned parts that test good may be used for warranty repair and may be less marginal in another system. This can result in the manufacturer never discovering the cause of a high NFF rate in warranty returns for a product resulting in an expensive cycle of functionally good parts being replaced with other good parts.

The vast majority of new electronic systems being produced today are digital systems. In the mass-manufacturing phase of digital electronic systems there are many electrical parameters that will be affected by the stack-up of manufacturing process variations. There are process variations in the many steps of fabrication of silicon die that will affect parametric performance from die to die on the same wafer. In deep submicron processes, variation of transistor threshold voltage can produce over 30% sheet resistance, and variation of poly-silicone resistors can reach 40% for some technologies [7].

In the manufacturing of CMOS semiconductors, the main sources of variation come from gate oxide thickness, which may consist of single digit numbers of atoms thick, random doping fluctuations, device geometry from lithography in the nanometer region, and transistor threshold voltage. The

variations can be a range of 100% for threshold voltage across a chip, 30% speed variation across a wafer and 100% leakage current variation in a wafer manufactured with 130 nm transmission line widths [8]. There is a distribution of values within the minimum and maximum required electrical performance parameters for semiconductor devices because IC fabrication is an imprecise process. For some semiconductor devices the measured parametric deviations from process variations can be used for product ‘binning’. Binning is the selection process for IC die that IC manufacturers characterize finished products for different markets by measuring thermal and frequency differences and using specific algorithms for an IC’s performance. The IC manufacturer does not generally disclose the deviation or margin in parametric between the units contained in each bin category.

During the process of PWBA (printed wiring board assembly) manufacturing there will exist variations in the metallization thickness, between lamination layer in the substrate, and surface roughness of the interconnects among other variables. Dimensional variations in PWB can affect impedance crosstalk, noise, and EMI issues in the system. Expansion and contraction of the structures and material interfaces in a PWBA is what induces the thermo-mechanical fatigue damage during thermal cycling that has been a primary focus of HALT and HASS methodology, but the dimensional variations impact SI quality as well.

We know from the SPC (statistical process control) that reduction of manufacturing variation is the path to making a defect-free product. When a complex high-speed digital electronics system is designed, computer simulation models are used to analyze the electrical performance such as SPICE. SPICE and other simulation software are useful tools to verify basic functionality; however, the models are limited in that they cannot include effect of impact on operation of potential variations in component manufacturing, circuit board fabrication, solder quality, and second sources of components over production period, field use, and fatigue-damage over time.

Finding marginal SI in a circuit that leads to operational failures during early product development is challenging. Early samples of a new electronics product are typically expensive and are shared by all development teams. Since early prototype samples are built with components from the same production lots, and using the same production lines, the variation between samples will be small compared with mass production. Using only a few samples during development it is difficult to determine whether the parametric variation and SI in the actual manufactured system is marginal. Most companies that have performed thermal HALT to operational limits on digital electronics almost always find/discover an operational limit only; as finding a destruct limit beyond that would come from an irrelevant field failure mode, such as solder being reflowed. It is rare to ever find a thermal destruct level in digital systems, such as IT (information technology) hardware, beyond operational levels. The change in material states, such as reflow of solder or melting of wire insulation, beyond operational limits is usually not relevant failure in operational field reliability (although it might be in severe

non-operation storage or shipping conditions). Hot and cold thermal stress causes signal propagation shifts in conductors and semiconductors resulting in “skewing” of signals throughout the system. This is the likely reason why thermal HALT on most digital systems results in finding an operational limit and not a destruct limit. Thermal HALT operation limits in digital systems often comes from a failure in SI, and a lock-up or shutdown occurs, but it can easily be reset when the stress is removed. The beneficial effect of skewing of signal propagation on a small number of system samples as a result higher temperatures and causing an is shown in Figure 2 and lower temperatures in Figure 3.

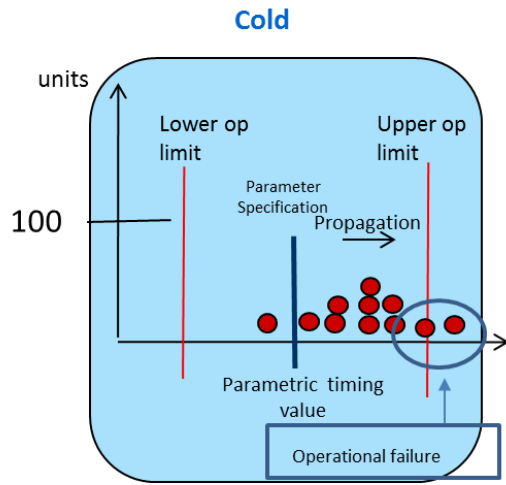


Fig.2. Low temperature HALT increases signal propagation and detection of marginal SI (adapted,[6])

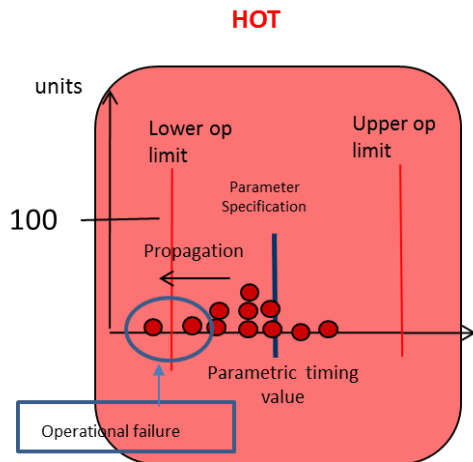


Fig.3. Low temperature HALT decreases signal propagation and increases detection of marginal SI

Marginal operational failures may be observed later in the field from worst-case stack up of parametric variations in a

smaller percentage of products when thousands or millions are produced. As manufacturing volumes increase, a wider distribution of parametric variations may then extend near or over the stable operational limit (as previously shown on the right graphic in figure 1). Of course the stimulation of timing variations using thermal stress on a system stimulates all the components to either speed up or slow down electrical signal propagation.

In the larger mass manufacturing population, the lot-to-lot and second source of components parametric variation is mixed with high and low speed distributions. Temperature affects the impedance of electronic conductors and semiconductors, which, in turn, affects the speed of signal propagation. An example of the relationship between temperature and signal propagation in a semiconductor is shown in Figure 4. The graph shows the measured low-to-high propagation delay versus case temperature in Fairchild Octal buffer MM74HC244N (rated for -40 to 85C).

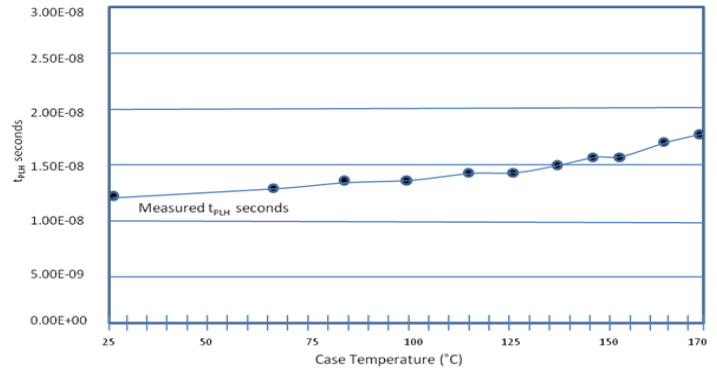


Fig. 4. Low-to-high propagation delay versus case temperature in Fairchild Octal buffer (adapted [9])

An additional benefit of rapid thermal cycling stress found in HALT chambers helps discover more potential timing variations. Low mass components have higher thermal transition rates than larger mass or high wattage components resulting in a combination of temperatures and thermal gradients across a PWBA. The temperature gradient during thermal transitions skew the impedance and mechanical stress of devices and interconnects across an operating circuit assembly. In thermal HALT the heating and cooling of a single active component is an effective way to isolate the component that has a thermal-induced parametric shift that causes an operational limit. A graphical illustration of the synergistic effects in rapid thermal transitions and the creation of thermal gradients, mechanical stresses and skewing of signal propagation throughout a PWBA assembly is shown in figure 5. The double arrows in the figure represent the thermo-mechanical stress vectors induce by thermal gradients throughout the circuit board assembly.

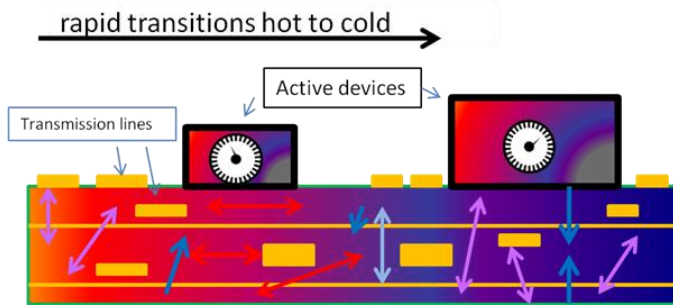


Fig. 5. The creation of thermal gradients on a PWB in rapid HALT thermal cycling [6]

V. Case Histories

Examples of the benefits of HALT techniques for finding software issues have been documented by Allied Telesis (formerly Allied Telesyn). Donovan Johnson and Ken Franks of Allied Telesis wrote and published a white paper several years ago on how the use of HALT has benefited their discovery of reliability issues due to software [10]. Some of the reliability issues Allied Telesis found were:

A. Abnormal LED activity

One fault found during cold step stress HALT at minus 10°C was a failure that caused random LED activity when the system was powering up. This fault occurred only after hard power cycling and not when a soft reset of the system was done. It was attributed to the reset pulse timing inside the PLD (Programmable Logic Device). They isolated the fault by applying heat to the suspected component. Under a hard power cycle, the PLD reset pulse duration was insufficient. After update of the code, the system had no errors at temperatures as low as minus 50°C. [10]

B. System Crash

A product that had six months of field use was used in the HALT lab to find its thermal operating margins. The first iteration of HALT found the Upper Operating Limit (UOL) to be 70°C which resulted in a system crash. By changing a register setting for a memory interface inside the boot code, the UOL was extended to over 100°C. [10]. Power Up Sequencing Electronics systems often use onboard logic to control the power sequence of the voltage rails. One product had failures of various components that after a power cycle at minus 20°C. By applying heat to the suspected they were able to isolate a PLD component that was unreliable at a lower temperature. Again the code was modified in the PLD that allowed the product to reach temperatures lower than minus 50°C without failures. [10]

In all these cases, the solution was due to a change in program coding— and not a hardware component— that extended the thermal operational margins significantly. The breakdown of the percentage of software and hardware issues found with HALT at Allied Telesis is shown in figure 6.

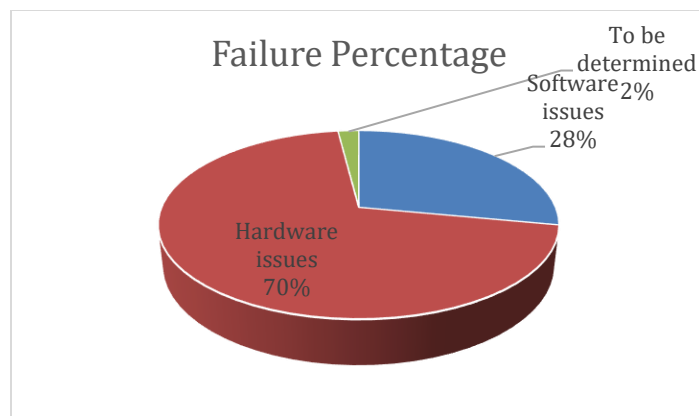


Fig. 6. The percentage of hardware and software failures found in HALT at Allied Telesis [adapted [10]]

VI. CONCLUSION

The benefits of HALT to find mechanical issues that result in catastrophic hardware failures in electronics assemblies have been well established over the last several decades. As the speed and density of electronics continue to increase, soft failures due to hardware SI issues will become more sensitive to parametric variations from manufacturing processes. Manufacturing variations in hardware designs will lead to operational failures in systems with low SI margins. In addition to the traditionally established benefits of HALT to reveal hardware weaknesses and defects, there is increasing, yet largely undocumented, evidence that suggests using thermal HALT to discover low parametric margins and signal quality that may lead to failures in software execution can improve operational reliability.

REFERENCES

- [1] Hall, Stephen H., and Howard L. Heck. *Advanced signal integrity for high-speed digital designs*. John Wiley & Sons, 2009.
- [2] Natarajan, Suriyaprakash, Melvin A. Breuer, and Sandeep K. Gupta. "Process variations and their impact on circuit operation." In *Defect and Fault Tolerance in VLSI Systems, 1998. Proceedings., 1998 IEEE International Symposium on*, pp. 73-81. IEEE, 1998.
- [3] ONSEMI, "ON Semiconductor," 2016. [Online]. Available: http://www.onsemi.com/pub_link/Collateral/AND9075-D.PDF. [Accessed 8 7 2016].
- [4] Wolpert, David, and Paul Ampadu. "Temperature effects in semiconductors." In *Managing Temperature Effects in Nanoscale Adaptive Systems*, pp. 15-33. Springer New York, 2012.
- [5] Qi, Haiyu, Sanka Ganesan, and Michael Pecht. "No-fault-found and intermittent failures in electronic products." *Microelectronics Reliability* 48, no. 5 (2008): 663-674.

- [6] K. A. Gray and J. J. Paschkewitz, *Next Generation HALT and HASS: Robust Design of Electronics and Systems*, John Wiley and Sons, 2016, pp. 190-205.
- [7] Melikyan, V., A. Durgaryan, Abraham H. Balabanyan, Eduard H. Babayan, Milena Stanojlović, and Ashot G. Harutyunyan. "Process-voltage-temperature variation detection and cancelation using on-chip phase-locked loop." In *56th Conference for electronics, telecommunications, computers, automation, and nuclear engineering–ETRAN, Zlatibor*. 2012."
- [8] Patel, Janak. "Cmos process variations: A critical operation point hypothesis." In *Online Presentation*. 2008.
- [9] Condra, Lloyd, Diganta Das, Neeraj Pendse, and Michael G. Pecht. "Junction temperature considerations in evaluating electronic parts for use outside manufacturers-specified temperature ranges." *IEEE Transactions on Components and Packaging Technologies* 24, no. 4 (2001): 721-728.
- [10] Johnson, Donovan and Ken Franks. "Software fault isolation using HALT and HASS." *Allied Telesyn*. 2005

BIOGRAPHY

Kirk Gray has over 33 years of experience in the electronics manufacturing industry. He has been teaching, consulting, and applying HALT and HASS methodology since 1992. He holds a BSEE from the University of Texas at Austin, is a Senior Member of the IEEE, and Senior Collaborator with the CALCE Consortium at The University of Maryland. He is the owner and Principal Consultant at Accelerated Reliability Solutions, L.