INITIAL ENVIRONMENTAL STRESS/LIFE TESTING OF BOEING 777 AVIONICS

Edward O. Minor, Robert W. Deppe and Sherwood S. Stolt
The Boeing Company (M/S 88-54)
P.O. Box 3707
Seattle, WA 98124-2207

AUTHORS' BIOGRAPHIES

Mr. Minor received BSEE and BSGS degrees from Seattle University, and has more than thirty years experience in missile and aircraft system integration with the Boeing Company. He is currently a member of the Quality Assurance Engineering (QAE) organization of the Boeing Commercial Airplane Group (BCAG) for the 777 Airplane Program. His assignment includes facilitating proactive quality initiatives applicable to the 777 airplane avionics. Ed was responsible for planning, conducting, and evaluating the initial pilot program of STRESS/LIFE testing and BCAG QAE test administration within the team.

Mr. Deppe joined The Boeing Company after receiving his BSAE and MSAE (Reliability and Quality Control) from the University of Arizona in 1972 and 1974 respectively. He has more than twenty years experience in the field of reliability and maintainability. His current responsibilities are as manager of the reliability group for the Commercial Avionics System (CAS). Previously, he provided reliability support for a variety of missile programs. Bob is the manager of the CAS team responsible for adaptation and implementation of the Reliability Enhancement Testing (RET) concept for the 777 avionics produced by Boeing.

Mr. Stolt received a BSEE and MSEE from the University of Wyoming in 1975 and 1977 respectively. He has worked as a reliability engineer on several military and commercial programs, including the air-launched cruise missile and the 777 cabin management system (CMS). His current responsibilities include test director for CMS STRESS/LIFE testing.

Abstract

The Boeing Company included a requirement to perform RET on avionics for the 777 airplane. A rigorous application of this program was implemented for the 777 electronic avionics to be designed and manufactured by Boeing. The program is based on applying increasing levels of thermal, vibration, and electrical stress to precipitate failures of weak links in parts, design, and processes. The first phase of testing was a pilot program focused on determining the correlation between failures found during RET testing and historical field failures. The tests show that a high degree of correlation exists. This paper addresses the pilot program and the initial testing of 777 units.

KEYWORDS

Avionics, combined environment, environmental stress, field failures, latent defects, quality, reliability enhancement testing, stress margin, thermal stress, vibration.

INTRODUCTION

The Boeing 777 twin jet airplane was designed to introduce a new standard of performance, quality, reliability, and passenger accommodation. A number of innovative approaches to development and production were implemented to facilitate and assure the accomplishment of these imperatives.

A requirement for a program of RET was established for all suppliers of avionics for the 777 airplane. The main criteria for this testing are summarized as follows:

1. The tests will be performed as part of the development process, for minimizing the in-service time required to reach mature reliability levels, by eliminating deficiencies other than defects normally detected during Environmental Stress Screening (ESS) and Qualification testing.

2. The implementation requirements include combined stress such as temperature, vibration, power, and other factors. It is required that equipment be in operation with normal loads during test and be monitored continuously.

3. The testing will be completed and any corrective actions incorporated prior to delivery of units for the first airplane.

To establish an application of the RET objectives for the avionics developed by Boeing, several potential alternative approaches were investigated. The approach taken employs high rates of change in thermal environment and six-degrees-of-freedom (6DOF) vibration. This type of stimulation has been shown to cause latent defects to be revealed in a variety of electrical, electronic, and electro-mechanical devices and
systems. Available data reviewed dealt primarily with products required to function only in a stationary installation, such as a laboratory or office environment. One of the early questions to be addressed was to determine the applicability of this approach to the RET program objectives for the 777 avionics.

Boeing conducted discussions and on-site field surveys with several other industrial companies having previous experience with this type of testing. Similar approaches were being used in each case, but with application differences for each product.

Initial Objectives

A series of preliminary tests was planned, with electronic avionics items representative of existing Boeing product lines to be used as the test articles.

Broad objectives were laid out for preliminary testing, primarily to verify and optimize the testing planned for the 777 avionics items. The following objectives and guidelines were implemented:

1. Conduct structured tests of incrementally increased stress severity to precipitate failures that identify characteristic weak link/strong link relationships in the device tested.

2. Perform root cause analysis of failures to determine the immediate failure mechanism and contributory circumstances.

3. Review historical files of field anomalies and subsequent resolutions. Determine the extent to which failures precipitated during the RET testing correspond with failures documented in the factory and field service records.

4. Determine options for alternative designs, parts, packaging, or processes that could eliminate the weak link that resulted in the failure.

5. Implement corrective action and perform retesting to verify effectiveness of the changes.

6. Apply the lessons learned to develop and optimize criteria for RET testing of the 777 avionics.

7. Document the test criteria, test processes, and results as a resource of lessons learned data.

8. Develop and implement a follow-up process to assure that the lessons learned are appropriately incorporated into the 777 avionics and future product lines.

A key aspect of RET is that the product under test is in operation during the application of stress, and monitoring is performed to detect the occurrence of any failures. This facilitates real-time detection of failures under operating conditions. An additional benefit is the potential for detection of latent defects which are manifested only during exposure to the stress stimulation. This is characteristic of a class of intermittent field failures that are reported as an in-flight anomaly but cannot be duplicated in the static maintenance environment.

Test Article Definition

The initial phase of testing was performed on seven units of a printed circuit card identified as the Universal Logic Card (ULC). The parts and manufacturing technology of the ULC are primarily typical of the production baseline configuration for previous Boeing product lines. The salient features of this technology are as follows: (1) both surface mount and plated-through-hole mounting techniques are used for smaller size integrated circuit devices, resistors, and capacitors; (2) larger capacitors, transistors, crystals, programmable logic devices, and other large-scale, integrated-circuit devices are mounted in plated-through-hole configuration.

Testing of the ULC was performed with a breakout box (BOB), with control/monitor signals connected to the ULC by cables. This allowed the ULC to be mounted in the test chamber, with the test support equipment remaining on an adjacent test bench. The ULC was mounted in a fixture that provided a structural mounting similar to that in which the card would ordinarily be placed. A timing device provided a signal to cause the ULC to execute its built-in-test (BIT) every 30 seconds. Commands could be issued to the ULC, and real-time responses retrieved throughout the period of active testing.

Environmental Test Chamber/Configuration

The test chamber used is a combination thermal/vibration system, with the capability of thermal or vibration stress, either separately or combined. The thermal capability provides means to vary the chamber temperature as much as 30 degrees Celsius (C) per minute, from minus 60 to plus 150 degrees C. The vibration is applied simultaneously in three axes by pneumatic impact actuators, which results in a randomized 6DOF excitation. Power spectral density (PSD) is distributed primarily over the higher portions of the range of 5 to 2000 Hz, as illustrated in Figure 1. The level of vibration can be controlled at fixed levels over a range of approximately two to thirty Grms.

Planned Test Parameters

The first series of tests performed also were planned to verify that the test chamber could be controlled to execute the test...
conditions to be applied for the Boeing 777 electronic avionics. A thermal test sequence, a vibration sequence, a combined thermal/vibration test sequence, and an extended duration combined thermal/vibration test to establish ESS levels were included.

The thermal test sequence consists of a series of high rate-of-change excursions, between high and low temperatures, with a dwell of from 1 to 15 minutes at each extreme. The minimum temperature rate of change used is 15 degrees Celsius per minute. Four thermal profiles are defined, based on the following high and low temperature extremes used.

1. Temperature range: -15 to +70 degrees C
2. Temperature range: -30 to +85 degrees C
3. Temperature range: -40 to +100 degrees C
4. Temperature range: -55 to +115 degrees C

A stabilization period of 10 minutes is provided at the beginning of each extreme temperature dwell. During the remaining time of the dwell, power to the unit under test (UUT) is cycled on and off. The UUT is cycled through each of the four temperature profiles eleven times. For each of the eleven cycles, a different combination of power and signal input conditions is applied, as indicated in Figure 2.

Vibration is conducted in a series of 15-minute periods at a constant level of Grms at ambient temperature. The initial levels used are 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, and 22 Grms. The UUT is removed from the test chamber for visual examination following each of the vibration test periods.

The combined environment test regimen consists of a five-cycle thermal profile with vibration during each of the extreme temperature dwell periods. The thermal range is -50 to +100 degrees C. The vibration level is 10% less than the maximum achieved during the vibration sequence.

A 30-cycle, combined environment test is performed to ESS criteria. This is intended to determine whether the planned ESS regimen will significantly consume useful product service life.

![Figure 1](image)

**Figure 1**

Typical Power Spectral Density of Test Chamber

**Initial Testing/Results**

The preliminary test effort was carried out with few deviations from the original plan. Latent defects discovered were subjected to rigorous root-cause analysis to determine the nature of the failure and the immediate cause. The results were compared to records of failures occurring in product use relative to failures recorded in manufacturing, airplane integration, and airline utilization. This was intended to provide an indication of whether testing of this type could proactively identify conditions that would otherwise eventually result in field failures.

**Thermal Test/Results**

Thermal cycling was performed, alternating between increasing levels of high and low temperature extremes, until the maximum capability of the chamber was reached. The card functioned throughout the thermal test, with no indication of malfunction or degraded operation.

After several passes through the profile with increasing temperature extremes, the maximum thermal range (from -60 degrees C to +150 degrees C), and a 20 degree C per minute
ramp rate, were consistently attainable. This thermal profile was used with variations in the high and low temperature dwell times in all of the remaining experimental thermal and combined environment testing of the ULCs.

Vibration Tests/Results (First ULC Tested)

Vibration testing at ambient temperature was initiated on the first ULC, starting at 2 Grms, and increasing by 2 Grms steps. Vibration was maintained for 1 to 10 minutes at each level, and the ULC was operated continuously. Self-test was initiated at approximately every 30 seconds. After each vibration step test period, the ULC was removed from the test chamber and examined visually (with unaided vision) to look for evidence of damage. After completing the 8 Grms vibration step, it was noted that the two ICs that were installed in sockets were starting to withdraw. The ICs were extracted and reinstalled with an epoxy-based bonding cement to avoid further occurrences of this failure mode. No other evidence of deterioration was noted, until self-test failure occurred at the 28 Grms level. Testing was suspended to determine the cause of failure and appropriate resolution. Examination at 30X with a microscope revealed several instances of stress fatigue that apparently occurred at lower vibration levels, but were not evident when examined without magnification.

Correlation Between Lab Test Results And Field Failures

Virtually all of the component failures that were precipitated in the lab were found to have corresponding multiple instances documented in the field failure records. Four major groups of components failed during the testing on every card tested and were prominently represented in the factory/field failure cases. In all these instances, the components were of considerable mass, supported only by a few interfacing wire leads. These were characterized as follows:

1. Cylindrical electrolytic capacitors, mounted parallel to the plane of the card.
2. Metal-encased cylindrical suppression diodes, mounted parallel to the card.
3. Rectangular slab ceramic capacitors, mounted upright, perpendicular to the card plane.
4. FET output drivers in TO-92 cases, mounted standing on three leads.

The primary failure was broken lead wires. This is due in part to a gap left between a component and the board, primarily to allow solder connections to be inspected visually. A component so installed is supported entirely by its interface lead wires. As motion is imparted to the board, the resulting pendulous response tends to break the leads. The failure analysis identified such fractures in several of the components where a lead had been welded, brazed, or otherwise attached directly to a metal surface of the part. Epoxy-based bonding cement was used to fill the inspection gap and secure the part to the board to prevent relative motion between the part and the board. This temporary approach allowed testing to proceed to higher levels as each weak link was identified and stabilized in turn.

Initial Testing Of The 777 Electronic Equipment

Testing was initiated on the 777 units using test regimens based on levels estimated to be consistent with the technology baseline. One change was the addition of a test sequence to determine the threshold of destruction.

The 777 Avionics General Configuration

The testing was initiated on a Line Replaceable Unit (LRU) of the Cabin Management System (CMS). The CMS is a system of integrated units developed to operate, control, and monitor the passenger comfort and entertainment services.

Testing the Overhead Electronics Unit (OEU)

The first 777 unit tested was the Overhead Electronics Unit (OEU). The OEU provides switching and power for cabin and passenger lights and information signs and is installed in the passenger cabin overhead area. Typically there are 103 OEUs per 777 airplane.

The OEU consists of a vented aluminum enclosure housing a printed wiring assembly (PWA) and a power supply transformer. The PWA is supported by aluminum spacers, and the transformer is mounted directly on the housing of the OEU. The power supply capacitors are mounted on the PWA. The design includes a power supply, a processor, digital interface capabilities, solid-state relay lamp drivers, and extensive BIT capabilities.

The test setup included a personal computer used as a function controller, the appropriate leads and power supplies, means to vary the power and signal levels, and means to accomplish continuous monitoring of the OEU.

Testing performed on the OEU was established as indicated in Figure 2:

The initial exposure conditions were maintained within the estimated technology limits of the unit.

PROCEEDINGS—Institute of Environmental Sciences
Eleven thermal cycles were executed with power/signal variations under each of the four thermal conditions. The variations of power and signal levels applied in each of the cycles are as listed. Power was applied and removed four times during each high and low temperature dwell period.

2. A number of additional thermal cycles were run to evaluate still-air conditions and various failures related to thermal stress.

3. Vibration at ambient temperature was performed in 15-minute periods at 6, 10, 14, 18, 22 and 24 Gms.

4. Combined environmental stress consisted of 13 cycles of varied conditions as indicated in Condition 1 and Condition 2 of Figure 3, (thermal stress over the range of -45 to +120 degrees C, and vibration at 20-22 Gms.).

5. Proof of ESS was 30 cycles of reduced level, combined vibration, and thermal stress as indicated. (Thermal stress over the range of -55 to +100 degrees C, and Vibration at 9 Gms.)

6. An additional combined environment test was conducted to identify the effects of additional vibration.
near to the thermal threshold of destruction of the unit. This test was conducted under the following criteria:

Thermal stress over the range of -60 to +120 degrees C, and vibration at 18 Grms with a higher duty cycle (Figure 3, Condition 3).

Thermal Testing/Results

The first failure occurred during the -15 degree C period of Thermal Condition 1 (per Figure 2). The increased resistance of the aluminum electrolytic capacitors at that temperature delayed power-up, causing the power-up timing monitor to initiate shutdown. Analysis indicated the increased power-up time was acceptable, so the monitor circuit was modified to allow more initialization time.

Shutdown again occurred during the -30 degree C portion of Thermal Condition 2 (per Figure 2). This was due to the increased switching speed of a zener diode at the lower temperature, which allowed a noise-triggered activation of the shutdown circuitry. This was corrected by resistive-capacitive compensation, reducing the frequency response and sensitivity of the detector.

A number of failures were noted during the -55 degree C dwell periods, arising from an out-of-tolerance test measurement. This was due to a voltage reference problem on an Application Specific Integrated Circuit (ASIC). This situation has been addressed in a later design version of the ASIC, which was not available at the time of testing. A re-test is planned to verify the effectiveness of the new design.

Lamp illumination failure resulted from inoperative solid-state relays (SSRs) during the 85 degree C portion of Thermal Condition 2 (per Figure 2). This was found to be caused by weakened response of the light-emitting diode drive circuit of the SSR at the high temperature. The drive current specification on the devices was 6 milliamps. The supplier had been unable to meet this requirement, and had supplied parts that needed 10 milliamps drive current to function at elevated temperatures. Modification of the drive circuit to provide adequate control current was implemented to facilitate testing. A final resolution of this will involve consideration of the other design criteria that were based on the 6 millamp requirement. These include power supply capacity, hold-up time, and heat dissipation.

Software Problems Found

Considerable effort was expended in the initial installation and checkout of the test set-up, and in verification that the support equipment and the OEU were functioning as planned in the test environment.

The microprocessor of the OEU experienced an interface processing problem that resulted in lockup, due to a stack underflow. This was resolved by a software change.

A failure to cope with SSR turn-on time variations at high and low temperatures was encountered. This function is controlled by software and was corrected by revising the thermal compensation logic. A re-test was conducted during later qualification tests.

Vibration Testing/Results

Vibration was conducted in 15 minute exposures at the levels of 6, 10, 14, 18, 22, and 24 Grms, the highest achievable by the test chamber, with no failures. The OEU was removed from the test chamber and inspected visually under magnifications up to 30x after each level. Significant deterioration in the housing was noted, particularly in the fasteners and bolt holes.

Combined Environment Testing/Results

Combined environment testing was performed in three phases, using the stress levels of Figure 3. The OEU indicated a failure mode at the first high temperature level. This was found to be due to an internal failure of aluminum electrolytic capacitors in the power supply. X-ray pictures revealed that internal conductors were broken. Upon dissection, it was apparent that movement of the foil rolls within the container resulted in fatigue of the fragile conductors that connect the foils to the external lead wires. Testing was completed with a capacitor connected to the UUT by wires that allowed the capacitor to be outside the test chamber. Capacitors capable of withstanding higher levels of stress were obtained several months later, and the unit was retested at 18 Grms and 100 C for 2.5 hours. Dissection of the capacitors revealed no internal damage.

Failure of one corner pin of a microprocessor chip was detected in the first low temperature portion of the test cycle. Visual examination at 50X indicated that this pin had a slight twist, apparently present at the time of installation. The pin spacing on this device is 40 pins per inch, and the pins are easily bent in handling because of their small size.

Visual examination under magnification was performed following the completion of the combined environment testing, revealing the presence of a number of cracked solder joints. The cracked joints were found between the component pin and the circuit board, primarily on corner pins, and on devices generally located toward the center of the board. It was observed that variation in the gap between the lead and the PWA pad was a common factor in the failure mode. This suggests that the solder deposition and control of the variability of the spacing between the pin and pad are highly significant.
Summary of Test Profiles And Flow Time

Figure 4 outlines the tests performed and the overall flow times as the test sequence was planned. The flow times are indicated in increments of eight-hour shifts. Initial test setup and verification required from two to three shifts.

Summary of Lessons Learned

1. Generic failures found in the RET testing of the ULC were highly indicative of documented field failures. Correlation was virtually 100%.

2. Failures were primarily in the connecting lead technology.

3. Internal failures of components were extremely rare. None were found on the ULC. A notable internal failure in the 777 OEU was the failure of the internal connections of aluminum electrolytic capacitors under combined environment stress.

4. Failure modes were attributable to weak links associated with factors of design, processes, parts, and handling.

5. Corrective actions consisted primarily of stabilization by bonding pendulous parts to the circuit board to reduce relative motion due to thermal or vibration stress.

6. Compensation for circuit performance variations attributable to thermal stress was a significant area of improvement in design margin in the 777 equipment.

7. If the accelerated life testing had been applied to the ULC in its development cycle, a very significant reduction in field failures would have been realized.

8. The latent weak links already found and eliminated in the 777 avionics would have been more costly to correct as field failures than the cost of the test program.

9. Removal and dissection of redundant parts, such as large power supply capacitors, may be required to detect intermittent failures.

Rationale for Testing at Levels of Stress Far Beyond the Design Limits

The rationale for using stress levels above the intended environmental limits to represent an accelerated life test is similar to that associated with qualification testing. Qualification testing is conducted at levels intended to be representative of an equivalent lifetime of stress in a much compressed time. RET testing is in a sense an extrapolation of this intent. The focus of RET testing is not on conditions near the qualification criteria, but rather on the use of elevated levels of stress that are intended to precipitate failure of the product by deliberately exceeding expected tolerances. By maintaining constant operation and monitoring of the functioning of the test article, the detection of failures provides for immediate contributors to the observed failures. Evidence of lead twisting, lack of planarity, or other distortion were typically involved. It is believed that these defects are precipitated primarily by board flexure, due to vibration and thermal stress. Process control measures are being studied to find means of improvements in these areas.

Proof of ESS Results

The proof of ESS test was intended to demonstrate that a combined environment ESS test that would not consume a significant fraction of the OEU's useful life could be designed. This was complicated by the desire to screen both flight test hardware that did not yet include all RET-related corrective actions and, later, production hardware. Initially, the RET unit was rebuilt by replacing the capacitors (with the lower vibration-rated parts) and all large integrated circuits. Some cracked solder joints and lifted pads produced faults during the 90 cycles. These were believed to be residual RET damage, so a second OEU was obtained. Three months after the completion of the Proof of ESS, the original unit was found to be non-functional. The aluminum electrolytic capacitors had failed the same as during RET, even though no fault was detected during test or during the subsequent complete functional test. A retest was conducted on both units for 60 cycles with the vibration reduced to produce 9 Grms measured at the top of the fixture (rather than on the test chamber shaker table as was done previously) and the maximum temperature was reduced to 71 C.

The RET unit continued to deteriorate, and the numerous solder joint failures were ignored. The new unit (which had no previous ESS) had two infant mortality failures due to solder defects, one in the first cycle and one in the seventh. The RET unit failed to start on one attempt on the 48th cycle, and later dissection revealed both units to have failed capacitors, and some pads had lifted on one hand-installed IC on the new unit.

A third test was run with the vibration reduced to 7 Grms. After 24 cycles, dissection revealed no damage to the capacitors. The new ESS now planned for other flight test hardware will be six cycles at 7 Grms and 71 C maximum. For production hardware, it will be three cycles at 8 Grms.
empirical identification of weak links in the product. By application of stresses in incrementally increasing steps, weak links can be detected sequentially, as thresholds of susceptibility are exceeded. As the process is continued, the threshold of destruction of the test article eventually will be indicated by a general deterioration and multiple failures at nearly the same intensity of stress. Of the products tested, the identification of latent weak links appears to have been readily discernible against the background of the general technology capability. By application of reasonable changes, a number of unforeseen weak links were detected and addressed. It is believed that this will make a significant improvement in the quality, reliability, service life, and environmental stress tolerance of the products.

<table>
<thead>
<tr>
<th>PROCESS</th>
<th>SHIFT</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
<th>27</th>
<th>28</th>
<th>29</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Install Thermocouples</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Setup &amp; Checkout</td>
<td>1-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thermal Profile Tests</td>
<td>2-8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional Test</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Install Accelerometers</td>
<td>8-9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vibration Step Tests</td>
<td>9-12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional Test</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined Environment Tests</td>
<td>12-14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional Test</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proof of Environmental Stress Screen (ESS)</td>
<td>14-19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Threshold of Destruction Testing</td>
<td>19-21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remove from Lab</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4 - Approximate End-to-End Flow Time (8 Hour Shifts)
The Benefits and Costs of Overstress Testing

by Nahum Meadowsong and Edmond L. Kyser, Cisco Systems

The partnership of design and manufacturing is central to the process of bringing a product to market. The impact of problems in either of these areas can increase exponentially if they go unnoticed until after the product reaches the customer.

Overstress test, using stresses beyond the design limit of the product, is successful at uncovering such faults in both the product design and the manufacturing process and ensures the overall robustness of the product. The benefits of overstress test include the following:

- Rapid design and process maturation.
- Less total engineering time and cost.
- Reduced production and warranty costs.
- Earlier and mature product introduction (yields stabilized).
- Higher mean time between failures (MTBF).
- Reduced manufacturing screening costs.
- Faster corrective action for design and process problems.
- Satisfied customers.

Highly accelerated life test (HALT) is a step-stress-to-fail destruct test that gradually increases the environmental stresses to determine the operational limits and find any design faults. The process is one of test, fail, and corrective action to prevent possible field failures.

HALT is not a compliance test and not limited by component or product specifications. All products are candidates for HALT.

In the case of systems, HALT finds defects in the weakest subsystem. Production should be delayed until HALT results are satisfactory.

Stress Overview

The following stresses, both alone and combined, can identify product weaknesses. When available, each of these stresses should be used in HALT design.

Temperature

Operation under prolonged elevated temperatures can uncover marginal design, bad components, or process problems inherent in the product. Temperature cycling detects weak solder joints, IC package integrity, a temperature coefficient of expansion (TCE) mismatch, and PCB processing and mounting problems, all of which will show up over time once a product is in the field.

Vibration

Vibration is useful for testing poor solder connections and a product’s robustness during shipping. Cold and insufficient solder joints can be stressed to failure with vibration levels that will not harm good connections.

Voltage Margining

Voltage margining can be useful in identifying marginal components and marginal design, especially when used in conjunction with temperature.

Frequency Margining

Frequency margining is not always an option, but if the circuit under test allows for it, it can be useful in identifying marginal components.

Functional Stresses

The most abstract and sophisticated aspect of the HALT design is functional test. The product must endure the combination of environmental and electrical stress while operating at peak processor utilization and bandwidth. Functional test should simulate this worst-case real world as accurately as possible to ensure that no product functionality goes untested.
Hot Step

Begin with two hours at each temperature step, one hour with a high-voltage margin and one hour with a low-voltage margin (Figure 1a). Power cycle after each temperature step. The first step is 60°C, and the last meaningful step is 90°C. The temperature may continue to be stepped; however, failures become less and less meaningful after a certain temperature. The point of diminishing returns is around 90°C.

Cold Step

This is similar to the hot step (Figure 1b). The first step is at -10°C, and last meaningful step is -40°C.

Vibration

Vibration is stepped 5 grms/h until the maximum capabilities of the vibration table are reached (Figure 1c). For the QualMark OVS4 Chambers, the value is 60 grms. Accelerometers should be used to determine the amount of vibration incident on the UUT. Generally, and especially when the UUT is part of a larger system, only a fraction of the table vibration is transmitted to the test subject.

Determinants of HALT Success

For HALT to be successful, the closed-loop corrective-action process must be followed. If failure analysis is not carried through to root cause, the benefits of HALT are lost. To be effective, the results must be:

- Fed back to design to make a circuit change, select a different supplier, or improve the existing supplier's process.
- Fed back to manufacturing for a process change.
- Used to determine the production test profile.

Another key for success is the intimate involvement by members of several departments within the company.

Management

Management must allocate sufficient resources, time, and funds for HALT to take place. Support must be provided during the failure-analysis phase to get closed-loop corrective action in a timely fashion.

Design Engineering

Design engineering must be immediately available to troubleshoot failures that sometimes can be beyond the scope of the HALT test engineers.

HALT Profile

There is no industry-standard HALT profile; it should be tailored to the needs of the program. Our experience with testing telecommunications products at Cisco has lead to the design of the following three-phase profile.

www.evaluationengineering.com
Suppliers

Suppliers must be willing and able to provide component failure analysis to obtain root cause.

Other key factors to success include the placement of HALT in the product development timeline, sample size, the perceived relevance of failures, and failure reporting. HALT should begin as soon as the hardware and software are available and stable.

Test as many units as possible because the probability of uncovering a defect increases with sample size. Failures that occur during testing should be treated as relevant and pursued to root cause. Finally, failure reporting must be visible enough so the failures and solutions gleaned from HALT are not overlooked.

MTBR Prediction Using HALT Margin

Can HALT be used to predict a product's mean time between returns (MTBR)? If so, then we are in a good position to estimate the benefits and cost-effectiveness of HALT.

The data that follows was collected from multiple HALT procedures of similar Cisco telecommunications products. MTBR and return material authorization (RMA) require that this data be tracked for a year past the completion of HALT.

Figure 2 shows the correlation between MTBR, the actual field performance of the product, and the HALT margin. In this data, the HALT margin is the smallest margin in degrees C between operating specifications and any HALT failure. Vibration failures are not considered.

The correlation is strong and intuitive. MTBR can be predicted based on the following least-squares fit:

$$MTBR = \left(0.0131HM + 0.0876\right)(NF)$$

where: $HM$ = HALT margin for the product
$NF$ = normalization factor used for Figure 2

Economic Justification

Having obtained the relationship between the HALT margin and MTBR, we can assess the cost-effectiveness of HALT. For HALT to be cost-effective, the cost must be less than the anticipated benefits.

HALT requires destruction of at least one prototype at a critical stage, and prototype build is a leading cost item in product development. Also, there are costs such as the manpower required to conduct HALT, the depreciation of test equipment, consumable costs, and corrective-action costs.

Cost Justification: Improve Reliability

As can be seen in Figure 3, if the operating margin is increased $n^\circ$C, the normalized RMA rate is reduced 0.0192 n. Also, the cost of each RMA approximates the cost of producing the board, which is termed whole product-cost (WPC) dollars.

*Continued on page 66*
A HALT program with a few tests under its belt does not have the data necessary to correlate HALT performance to field performance. That is not to say that a correlation does not exist. Quite the contrary.

A seasoned HALT program that has conducted multiple tests on similar products possibly could predict MTBR more accurately than current predictors in use such as those based on component count and individual component field data. Current predictors do not factor any aspects of the actual product design into their calculations where HALT margin is specific to the product under test.

Consider the comparison of a reliability determination test (RDT) using traditional methods and RMA as predicted by the HALT margin in Figure 4. It shows the required test time for an 80% confidence level in an MTBF prediction greater than 75,000 hours. The blue line indicates that RDT requires 40 boards tested for 10 weeks, assuming one failure and Arrhenius acceleration due to a 50°C test temperature. RDT is a very time-consuming test and needs equipment and labor that otherwise would be used for production.

The same prediction using the HALT margin requires as little as one board for one week. In Figure 3 showing the RMA rate vs. the HALT margin, the dashed line illustrates an 80% confidence for an operating margin of 30°C indicating a normalized RMA rate below 0.55. The traditional MTBF predictor may be improved by factoring the HALT margin into the MTBF calculation.

Conclusions
Several benefits can be obtained from a well-designed HALT program:
• The HALT process is adept at finding and correcting design faults and determining design margins.
• The costs of HALT can be justified in terms of improved margin.

Figure 4. Required Test Time as Predicted by HALT Margin
A Method of Reliability Improvement Using Accelerated Testing Methodologies

by

Kevin Granlund
ESS Engineering Manager
EMC Corp.

Abstract

Reliability growth is a major goal of many new product introductions. There are many ways of achieving this growth in a new product. One of the new methods, HASS ¹, will be discussed in this paper. Utilizing Highly Accelerated Stress Screening along with highly focused root cause analysis will help reliability grow rapidly which is highly desirable in today’s fast moving time to market driven companies. An example of how this process works based on the experience gained in the introduction of a new power system will be discussed. HASS is part of an overall iterative process of forcing failures, analysis, correction, and retesting. HASS itself needs to be continually monitored for its effectiveness and corrections must be made for missed defects. Case studies, with reliability data, will be presented to demonstrate how HASS can be a major part of a reliability growth process.

Introduction

In today’s commercial electronics environment, the winners are the companies who come out with the products first and grab the market before the competition. Although this is good financially, it makes the job of evaluating and growing reliability in a product much more difficult, especially when you also add to this that the expected reliabilities are far greater than they were 10 to 15 years ago, in many cases an order of magnitude higher. The challenge is to quickly turn designs into full scale production with as few defects as possible both in the design as well as the assembly process and ship only a highly reliable product.

This paper is a focus on a HASS process and its impact on reliability, what is not being discussed here is the design evaluation phase of HALT (Highly Accelerated Life Test) testing which was also performed on this product successfully prior to release to production. All major design failure modes exposed were root caused and corrected. The HASS process utilized thermal cycling, 6dof vibration, power cycling, load variations, etc. (see fig. 2).
The traditional methods, RGT, ORT, Duane, PSRT, etc., run into a wall when trying to beat the competition with the 1st product out. You simply don’t have enough time to gain confidence that the product is reliable before shipping it. Mil Std 781 defines a host of methods to statistically demonstrate reliability, but who can afford to perform one of these tests today! MTBF is no longer a value in the 10’s of thousand of hours but instead the 100’s of thousands and in many cases now millions of hours of MTBF.

Growing Reliability

How can you quickly grow reliability? First, we must understand what the reliability of a product is based on. The traditional bathtub curve (Fig. 1) is the standard method to model electronic reliability throughout product lifecycle. Standard reliability theory says that there are 3 major states in electronic reliability: infancy, steady state, and wearout. Responsibility for these 3 states falls on manufacturing and design for infancy, while design owns steady state and wearout. Manufacturing is responsible for a controllable, reproducible process, while design is responsible for a well margined design that can accept variation in production that won’t degrade its reliability. This is what many companies now call Design For Manufacturability or sometimes Design Robustness.

The mechanisms we are talking about today are also different from decades past as well. Silicon level Arrhenius reaction based defects are no longer the dominant issue in reliability of assemblies. Assembly or design margins are now predominant. Arrhenius models are still valid for some silicon based defects, but even that is varied among technologies like CMOS where the dominant mechanism is oxide defects. A voltage acceleration model is the proper model to work from in this case. Assemblies are dealing with mechanically or electrically based failure mechanisms and Arrhenius is not valid here.

![Bathtub Curve](image)

Figure 1. Bathtub Curve
Design Robustness has a direct link to the reliability of a product. Terms like Design For Manufacturability are now commonplace, but not always understood. Designing for Manufacturability means to account for manufacturing variability in a design - Design Robustness.

Product Infancy defects will be the focus of this paper and how they can be influenced by using HASS as a tool to measure the robustness of the design/manufacturing process, and ultimately its reliability. This process is meant to be a compliment to good DFM methodology, not a substitute.

Problem

The basic problem is having enough samples to “measure” the reliability of a product. A program utilizing a PSRT method could be set up to measure the initial product reliability. What if it fails? Does the program slip by some significant amount of time? Never mind that the samples needed today may be impractical to produce before a product needs to be released. This type of failure “counting” is an open looped process that adds little or no value.

An Example:

A 1000 watt switching power supply with 5/12V outputs and battery backup capability has been predicted to demonstrate 50,000 hrs MTBF. Using plan VC in Mil-Std-781, one would need to run 3.75 times the expected MTBF if 0 failures are detected. That equates to 187,500 hours of failure free running time, which would mean building and running 260 units for a month of time, 24hrs/day, to demonstrate it, and this assumes there are NO failures. Assuming a cost of $1500 each, that equates to $390,000 in inventory alone, not including other costs associated with the test. Many companies will apply Arrhenius acceleration rates and run the units at elevated temperature for a shorter period of time and equate a longer run time, to save cost, but the assumption is flawed since it doesn’t apply to today’s predominant failure mechanisms. Then, as all reliability engineers know, the next 260 units produced will be different. The next logical step is to develop an ORT program, but this is simply too slow and too small a sampled process to be effective against lot variations as well.

It is not economically feasible to take this type of approach.

“An effective reliability test program should aim at generating failures, since they provide the information on how to improve the product”.
How can one generate failures through a large enough sample size to quickly improve the reliability? A HASS program that will utilize a highly effective root cause corrective action process can meet this requirement. The surest way to have large sample sizes is to be in a production line performing HASS on 100% of the product. While this means having a production process that makes some production managers cringe, it has proven to be a very cost effective method as long as the value added step of Failure Analysis/Corrective Action (FRACAS) is taken. Otherwise, this is only a cost added measure.

Since one is screening 100% of the product, a direct benefit will be seen by the customer with a higher level of reliability. For example, based on another system's field performance at EMC, Before HASS then After HASS, the 1st 90 days operation failure rate went from approximately 3% to a 0.5%. Eventually even higher reliability's will be seen as corrective actions are applied.

Case Study

During the development phase of a new Integrated Cached Disk Array, including the power system, it was decided to implement a HASS program on the power system. The goal was to screen failure mechanisms before shipment to the customer as well as to provide a feedback tool for the quality of the design/manufacturing process to drive correction action. This power subsystem was subcontracted to a power supply manufacturer. The subcontractor performed a 24 hour burn-in and a final functional test before shipment. The supplier was involved in the decision making regarding this process.

This system consisted of three (3) 1200 watt switching power supplies with 208, single phase input and 4 outputs, 56, 24, 12 and 5 volts DC. The system was an N+1 redundant configuration with dynamic load sharing of the 5 and 12 volt outputs. Battery backup was also utilized with the 56 volt bus. The 24 volts is a low current output. Each supply also had a series of TTL inputs and outputs for monitoring and control.

An earlier study done at EMC to determine the effectiveness of a dynamic (powered on, fully monitored) process vs. a static (powered on, unmonitored) process had been done. A static process is only 10% effective. In other words, given a defective population, 90% of the defects would escape a static process. Therefore all HASS processes are dynamic. A tester was developed that would control the UUT, loading conditions, AC line levels on/off cycles, temperature chamber, vibration controller, DC battery source. The system would log all the failures with their conditions.
Failure replication is an important step due to the intermittent nature of most mechanically based failures. Elimination of fixture errors (cables, software, A/D boards, etc.) was important to prevent a high no fault found rate and to clarify the failure modes for easy diagnosis by the supplier. This would aid in faster resolution. For instance, it had been seen that some plug and play failures (no stress needed) would not replicate and
return for retest and fail the 3rd cycle for the same symptom and eventually be analyzed to an intermittent electrical interface.

A database is maintained for each failure by serial number, failure mode, conditions, cycles to failure, etc. Root cause analysis was logged into the system linked to the original failure as part of determining the effectiveness of the process. Analysis of the effectiveness of the corrective action taken for a particular mode could also be analyzed since sometimes corrective actions don’t work (because root cause was not fully understood) and need to be readdressed or may not affect 100% of the population due to technology constraints.

Don’t develop tunnel vision if you can. Continue to look at errors and faults found in follow on test stages as well to determine if your screen is as effective as you think it is and to determine if other correction actions need to be taken. Adjustments need to be made to either add other stresses or just plug test holes. The profile shown is actually had a number of changes made to it to either gather further information regarding a particular failure mode or to tighten the screen against certain defects. Also, don’t think this is a panacea. Even though every major failure mode was detected in this process, not all failure mechanisms are 100% screenable either in a chamber. Some would take serious amounts of overstress in order to accelerate effectively. This adds further to the argument regarding fast root cause to prevent further fallout. Or other more effective methods of screening could be done, such as leakage current measurements of bulk capacitors as an example. A proof of screen should always be redone on all major screen parameter changes.

Profile Utilized

Fig. 2
Fig. 4

1200W/ASS Failures By Test Cycle.

Fig. 3

- Count
- Punctured Insulator
- Poor Solder Joint
- Foreign Object
- Cables/Connectors
- Improper Torque
- Shorts
- Optocoupler
- Backward Component
- Capacitor
A Look at the Failure Mechanisms

Certain failure modes were strictly due to design tolerance, some due to strictly manufacturing control, and some, like PCB layout are a mix. If component leads have to be deformed to fit the spacing layout on a PCB, then this is uncontrollable for manufacturing. Fig 3 is a sample pareto of defects found for a specific population of production units. What at first appears to be solely workmanship issues, in actuality was a combination of design and manufacturing since a further breakdown of Poor Solder Joints and Shorts was related to board layout (DFM) and workmanship in rework (manufacturing). Punctured insulators again could be broken down to work area cleanliness (manufacturing) and a simple change in the drawing (design) calling out a punch operation in the wrong direction leaving burs. Connector issues included incomplete insertion in assembly (manufacturing) and when properly inserted, higher impedance connections (design). Corrective action results from this testing varied significantly from PCB layout (DFM), to component changes (design), to process flow (manufacturing).

Fig 4 is presented as an example of determining test time. This would vary, as well as the mode of precipitation would (vibration, temperature, line, etc.) as different lot quality problems would surface. So diligent monitoring of the process is needed to ensure its comprehensiveness.

The Bottom Line

While at first this looks like an expensive process and might frighten a few businesses away, the reality is it is only adding about 4% to 5% to the actual total cost of each unit when amortized over the life of the program. This includes capital and operating expenses. This can be even lower if the corrective action cycle can be enhanced to prevent further capital investment as the volumes increase. As for the benefit to actual MTBF, the prediction for this design was between 45K hrs. to 75K hrs. depending on models and assumptions used. The actual MTBF is running about 126K hrs. for the 1st uncorrected population (screened, but uncorrected) and can be expected to grow by 30% to 60%, maybe a little more or less as a guestimate, as the population with changes begins to carry the weight.

Conclusion

The benefits of taking this type of approach are clear. Reliability can be directly benefited by screening, but ultimately it is the corrective actions that count. The goal should always be to work towards elimination of the screen by eliminating failures. Depending on how well your suppliers are under control will determine your ability to take the next step, like HASA (Highly Accelerated Stress Auditing), when you have reached a point of diminishing return and are only concerned about lot to lot variations.
The cost of such an investment considered over the life of the product is easily justified when considering the level of quality generated. While this is “testing in the Quality”, which severely goes against the grain of today’s quality training, it is sometimes a necessary step to get to that next level. What is learned by both the design and manufacturing community will be incorporated into the next project, yielding even higher reliability in the next design. Certainly, this type of investment yields far better results since it is a proactive step to finding problems and fixing them, rather than waiting for them to happen and simply “counting” them. Predictions also fall short as well, since it sets expectations, leading to complacency when they are met, when clearly more needs to be done. If your customer, like most, are in a business that requires very high levels of reliability, the added value of high quality products is well worth the investment of an extra 5%.

3 Harry McLean, “Highly Accelerated Stressing of Products With Very Low Failure Rates”, Institute of Environmental Sciences Proceedings 1991 pp 443-450
Combining Team Spirit and Statistical Tools With the H.A.L.T Process
by Larry Edson
from
The Cadillac Luxury Car Division of General Motors

ABSTRACT:

H.A.L.T., Highly Accelerated Life Testing, offers three unique facets of opportunity for the reliability engineer:

1. Provides the shortest possible test time required for the engineer to gain durability insight into the product while keeping it within the human attention span for retained enthusiasm.

2. The H.A.L.T. process is best utilized in a cooperative workshop between the customer and the supplier. The advantages from this cooperative effort extend beyond just the product on test. H.A.L.T. testing is enough like "survival on foreign soil" that prior prejudices and misconceptions are neutralized. Customer and suppliers can work together as a team in the exploration and development of the product. Profound knowledge of the failure mechanisms is available for all to learn. The attitude that exists during this process is akin to a group of students testing a class project for the first time.

3. The massive time compression of the H.A.L.T. process, in combination with the reduced data variability, provides the opportunity to use "life" as a response variable in a statistical "Design of Experiments" (DoE). The DoE would be intended to find optimized product configurations or process parameters that would extend the life and robustness of the product. Typically, life testing takes such a long time that using "life" as a response variable has been out of the question.

"... through to Z is for Zebra. I know them all well,"
Said Conrad Cornelius o' Donald o'Dell.
'So now I know everything anyone knows
'From beginning to end. From the start to the close.
'Because Z is as far as the alphabet goes.'

Then he almost fell flat on his face on the floor
When I picked up the chalk and drew one letter more!
A letter he never dreamed of before!
And I said, 'You can stop, if you want, with the Z
'B because most people stop with the Z
'But not me!'

'In the places I go there are things that I see
'That I never could spell if I stopped with the Z.
'I'm telling you this 'cause you're one of my friends.
'My alphabet starts where your alphabet ends!'

-Dr. Seuss, *On Beyond Zebra* (1983)

**GENERAL DISCUSSION**

**THE TIME NEEDED TO TEST**

The automotive industry has longed for the H.A.L.T. test as the answer to all of the durability testing that goes into the development and validation of a new automobile. The idea of finding some way to highly accelerate time has always been the key issue. The American Automobile Industry would like to be able to bring a new vehicle from conception to birth within a three year time interval. This process would then be staggered within a company such that several new models would appear every year. This allows technological advances to keep pace with the market place and fosters growing enthusiasm in potential customers. The steps required in bringing forth new vehicles is well modeled using a four phase vehicle development process. Everyone pretty much knows what they are supposed to do, its just that some steps take too long. The majority of the steps that take too long have something to do with developing and validating the systems that go into the vehicle. In particular, it is the "life" related tests that have been the road block. Many of our tests take months to complete and cost hundreds of thousands of dollars. This scenario is played out in many different areas of the car as the vehicle is made up of so many diverse and complex systems. Considering this timing and cost dilemma, the entire mindset going into these tests is one of sincere hope and prayer that nothing goes wrong. The time needed to invent a fix and retest when problems are encountered, delays not just the
system, but the entire vehicle program. Well, you might say, "why don't you just do it like the Japanese; develop and test systems and put them on the shelf." That way when you want a new car you just pick the pieces off the shelf." The problem with this approach is that as we try to merge the age of the technology with the time of car sales, there is very little time left for systems to be sitting on shelves.

The H.A.L.T. test has begun to revolutionize this process at The Cadillac Luxury Car Division. You are probably wondering how we have been able to stuff a whole car into one of these little chambers? We haven't. We are using the H.A.L.T. process on a system by system basis. However, we have not constrained its use to just electronic circuit boards! I have brought with me some of the systems that we have used H.A.L.T. on so far. You can see that it includes lamp assemblies, power mirror systems, sunroof systems, automatic transmission cooler lines, antenna systems, instrument panel clusters, and entertainment systems. And this is only the beginning! (H.A.L.T. at The Cadillac Luxury Car Division began in November of 1995) We have also learned to adapt the H.A.L.T. approach to other tests and test equipment. We are applying the H.A.L.T. strategy to exhaust systems and various brackets while using existing non-H.A.L.T. test equipment.

Please realize that there is more here than just test equipment. The "In Search of Weakness" strategy provides a robustness in overcoming the age old problems of sample size and product variability.

A NEW MENTAL ATTITUDE

There is a unique psycho-social phenomenon that occurs when we team up with our suppliers to H.A.L.T. test a product. We begin by agreeing that this will be a time void of demeaning paperwork, that no one will be contractually obligated to change their product and that it should be "fun". I realize that this may not exactly seem "business like" but it does elevate personal integrity, respect and sincerity to a new working level. The Cadillac Luxury Car Division has taken the approach of "Guided Discovery" in working side by side with the supplier while running H.A.L.T. tests. The joint effort allows The Cadillac Luxury Car Division to develop a growing product knowledge base that becomes applicable to products that follow. It also fosters a unique relationship between working engineers from both companies. In some instances, the working team has included several suppliers, some of which were competitors to each other. In two test episodes, the working engineers did not speak the same language, yet they were able to communicate because they so desperately wanted to!
In every “Guided Discovery” HALT test we have insisted that each engineer in attendance take an active role in running the test. This participation has included chamber control, attending to monitoring equipment and being good visual spotters (more important than you realize). Once the participation has begun then everyone wants to assist in developing fixes and performing the actual hands-on labor leading to a retest.

I believe that the HALT process, in conjunction with this “way of doing business”, is foreign enough and exotic enough, to allow engineers to approach the test with a child-like open mind, ready for the first day of class. The “Guided Discovery” process then takes on the role of “teacher” and all attention is devoted to “WHAT IS GOING TO HAPPEN INSIDE THIS CHAMBER?” All prior knowledge of testing and what to expect is suspended and each person obtains confidence (don’t forget now we are on foreign soil!) from working with every other person. It’s a little difficult to describe, and may even sound comy, but this process when carried out this way invokes a kind of enthusiastic energy that is almost startling (certainly startling for the automotive industry). I think the “instant gratification” that occurs during HALT testing, as a result of test time compression, allows this energy to stay alive.

A NEW OPPORTUNITY: DoE MEETS LIFE TESTING

Reliability people have long lived with the curse of “waiting” for results from life tests. The “wait” is also accompanied with extreme variability when adequate sample sizes are tested. Weibull analysis is usually used to predict from these non-Gaussian (asymmetrical and long tailed) distributions of life. The “wait”, in combination with the variability, tends to diminish the enthusiasm for the test and has generally precluded the use of Design of Experiments (DoE) as a tool in developing long product life.

The HALT test not only reduces the “wait” drastically, but also compresses the variability in the life data.

The compression in life data variability is one of the most profound aspects of HALT testing for the statistically minded engineer. One can begin to appreciate the magnitude of this compression in variability by reviewing the S-N graph of plastic tensile fatigue in figure 1.

The horizontal lines with little vertical bars represent the confidence intervals on the mean. The mean in this case is the average number of cycles to failure at different stress levels. The shortening of the length of the lines as you move from the right side of the graph to the left side of the graph gives a sense of the compression of variability that is taking place.
When the compression of data variability is combined with the compression in actual test time, the reliability engineer's time curse is removed and the door is open for integrating statistical tools for life enhancement.

**TENSILE FATIGUE OF A PC/ABS MATERIAL**
**CYCLIC LOADING FROM ZERO TO SPECIFIED STRESS CONDITIONS: 3 HZ AND 73 DEGREES F.**

![Graphical representation of stress vs. cycles to failure](image)

Figure 1.

Let's take just a minute to review the basic concept of a designed experiment. The reason we give this idea such a fancy title is that it has a unique advantage in efficiency, and a robustness in drawing conclusions. The following discussion describes a basic weakness in the undergraduate college education process and attempts to start us thinking within the framework of the designed experiment.

Most engineers graduating from college have never been taught how to design an experiment efficiently. This topic would be taught as a graduate course under advanced statistical methods. Many engineers believe that they must hold everything constant and change only one variable at a time when a number of variables are to be evaluated. This seems intrinsically correct, after all this should give one "control" over variation in test results. We will call this strategy "one factor at a time" (Ofat). This strategy is fundamentally flawed and brings on various forms of "depression" as one ponders how to form the number of combinations that must be run to obtain valid results. I can remember the first time this happened to me. I was in my last year of college taking a powertrain lab class. We were to evaluate the effects of air-fuel ratio, spark timing, engine compression ratio, fuel type and water injection on power output. While trying to figure out what combinations we would need to run, it started to dawn on us how many combinations there were. There would not be enough lab time to run all of those combinations and so we gave up in frustration. This same situation occurs in many different businesses every day to an extent that would surprise all of us.
The designed experiment alleviates this frustrating situation and offers a "guided pathway" for the experimental process. Let's begin by first considering a three dimensional cube (please review figure 2.) Each axis of the cube represents a factor that we believe will effect the outcome of the experiment (three factors for this example). The reasonable extremes of each factor (high level and low level) are located at the eight corners of the cube. The Ofat approach would have us traveling through the interior of the cube taking several data values at each stop-over point. The corners of the cube, under Ofat, would be the least explored regions. The designed experiment approach would have us take a single data point at each of the eight corners (the corners will now be the regions of focus). Contrasts are then drawn between opposing sides. The average of the four data values, one from each of the four corners on any one side are averaged together. The same is done for the opposing side. The difference between these two averages reflects the effect of the variable whose axis connected the two sides. This same process is carried out for the sides that were at right angles to our original sides. When we are done we will have the three contrast values for our three axis factors. When you carefully consider what we have done you will realize that we used each data point (at the corners of the cube) three times. You can also see that we have used an average for each side that is a perfect balance of the other two factors making up that side. We can continue establishing values for the averages of planes through the corners of the cube. The difference between the diagonal planes will provide the contrast values for the interactions between the factors that make up those planes.

Now that you have this three day course under your belts, we will apply it to a H.A.L.T. test.

(at the time of writing this proposed test was only in the planning stages. I felt that the product represented the kind of diversity, yet had an ease of appreciation that made it a good candidate for this explanation)

The product will be the crimp attachment on the end of a transmission cooler line. The three factors we'll explore are the depth of crimp, the length of crimp and the shape of the crimp. Two levels of each factor will be tested and these levels will represent the reasonable extremes possible in manufacturing. The cube plot of our test is shown in figure 2.
This test will require 8 part configurations and all 8 will fit inside the chamber at the same time (the paired testing concept is important). The response variable will be the time needed to bring on a leak. We have soaked the inside of the hoses with a dyed oil that is easily picked up by a black light. If and when this oil is forced out the crimp, it will be easy to see with the black light. We are pressurizing the hose with air and have a pressure gage attached at the other end of the hose. The volume of air under pressure is kept small so that we can detect small changes in pressure (a leak). Thus we have two ways in which to detect a leak. The failure mechanism is the loss of intimate contact inside the crimp sandwich. This results from changes in temperature in combination with vibration and is exactly what the product sees in the car. The oil in the transmission line in the car takes on the consistency of peanut butter at minus 38 °C. As the oil in the transmission heats up it eventually pushes this plug of "peanut butter" through the line and hot oil quickly reaches the crimp. The vibration from the engine/transmission is directly affecting this joint while this rapid temperature change is taking place. The test will be conducted using repetitive temperature swings from -100 °C to +200 °C while applying a maximum vibration level of 55 Grms. Please note that we are not using an explorative step stress approach. We are trying to invoke a failure as fast as possible. We will test until all designs begin to leak. We will plot the contrasts in "time to leak" on normal probability paper and will use the noise from the unimportant contrasts to pick out important signals. This will allow us to decide which factors are most important and if there are important interactions between these factors.

If we had more variables to begin with, we would probably chose to run a fractional factorial (most similar to a Taguchi method). This would allow us to still test all the parts at the same time in the chamber but we would give up our ability to discern interactions in the first test. We would use a follow up test that would provide the full interaction effect between the most important factors. If we had four factors, our cube plot will become two cube plots split on the high and low levels of the fourth variable.

The important concept here is that, although there is no correlation to actual miles of use, the paired comparison approach retains integrity in the final data. The best design combination will be chosen from the highly accelerated test results and this should provide the best design for use by the customer.

EPILOGUE

I have been in the Reliability/Validation field a long time (25 years) and have endured the debates between the following good intentioned engineers: those who are trying to decide how many parts to run to a test
“bogey”, those who want to run multiple lives to a test “bogey”, and those who are trying to figure out what to do with their six sample Weibull plot when they have a mixture of failure modes. Oh yes, I almost forgot the guy who is trying to correlate an accelerated test to a test “bogey”. While this discussion is entertaining, experience has shown that it does not provide what the product really needs (results from warranty usually highlight this fact), nor is it capable of doing so in a timely manner. These reenactments of generations old methodology are beginning to change as the H.A.L.T. approach takes hold. The H.A.L.T. approach breaks through many of the statistical barriers that have “muddied up” the data involved in life testing. **It is also important to realize that these ideas can be transplanted into testing environments that cannot take advantage of an existing “H.A.L.T. chamber”**.

I have tried to help you “read between the lines” of what is generally described in a H.A.L.T. test. While the H.A.L.T. test has its roots in engineering, the opportunities to utilize good statistical tools and develop good inter-company cooperation are wonderful. I hope through the video and experiences I have described here today, you can begin to appreciate how great these opportunities are.

**Larry Edson**

“I led him around and I tried hard to show
There are things beyond Z that most people don’t know.
I took him past zebra. As far as I could.
And I think, perhaps, maybe I did him some good.....

Because, finally, he said:
‘This is really great stuff!
‘And I guess the old alphabet
‘ISN’T enough!’

**NOW the letters he uses are something to see!**
Most people still stop at the Z...
But not HE!”

- Dr. Seuss, On Beyond Zebra (1983)
ACHIEVING PHENOMENAL RELIABILITY GROWTH

Clifton J. Seusy
Hewlett-Packard Company
Disc Memory Division
Boise, Idaho, USA

ABSTRACT
Reliability does not come easily. Nor is it free. But armed with adequate tools, techniques, knowledge, time, and an ample budget, project managers can meet their reliability goals. Achieving rapid growth requires exceptional means for exposing failure modes and an absolute, unswerving resolve to permanently eliminate them. Many prototypes are needed, and each one must be elicited for failure information. Then each failure mode must be diligently pursued from analysis through implementation of a permanent solution. Issues addressed are:
- How many units must be tested to meet a specific reliability goal?
- How is this number of units justified?
- How are these units acquired?
- How is failure information extracted?
- How does a manager track the failures and their status?
- How are the failures treated?

The tools presented are derived from statistical theory and from experience. Numerous case studies are presented to illustrate the successful application of these principles.

RELIABILITY GROWTH is often modeled as a straight line on a log-time versus log-reliability plot (Figure 1). The figure illustrates that to improve reliability more test hours must be accumulated. Hundreds of thousands of test hours are expensive but may still be necessary to reach a reliability goal. To reach reliability goals more quickly, increase the reliability improvement rate; i.e., increase the slope of the line in Figure 1. Simple project management concepts and attitudes can be employed to increase this rate for free. These concepts merely streamline and optimize the test, analyze, and fix (TAAF) process.

Even with rapid growth rates, numerous units have to be tested for many hours. The testing should be designed to extract all possible failure mode information from each unit. Failure modes must then be quickly analyzed to find the fundamental reason, or root cause, behind the failure. The analysis must be correct and the corrective design must be perfect to maximize reliability growth. To maximize growth rate, the turnaround time for this process should be as short as possible without jeopardizing the quality of either analysis or design.

Finally, all failure modes must be addressed. Opinions and feelings must be cast aside in favor of facts and evidence. This is not an easy task, given the disposition of human nature. Reliability growth challenges are psychological as well as technical, even among presumably logically-minded engineers.

HOW MANY UNITS MUST BE TESTED?

Determining how many units to test should be mathematically calculated based on reliability goals. If the number is chosen according to budget constraints or, worse yet, intuitive judgments, the reliability goal may not be met. If the budget and reliability goals are mu-
tually exclusive, one or both must change. In today’s
marketplace, unreliability is extremely costly, so an
investment in an adequate number of prototypes can
pay big dividends.

**BINOMIAL PROBABILITIES** - The binomial
probability distribution is used to calculate the number
of test units required to detect one specific failure mode.
If the failure mode has a particular probability of existing
in each unit and if it can be assumed that the test
will uncover an existing defect, the number of units
required to ensure detection with a certain probability
can be easily calculated.

The binomial probability distribution is

\[ P(y) = \binom{n}{y} p^y (1-p)^{n-y} \]

where \( f(y) \) is the probability of having \( y \) bad units in a
sample of \( n \) units when the probability of each unit
being bad is \( p \). The total probability of finding the
problem is the sum of the probabilities of observing exactly
one, two, three, or more failures, up to \( n \) failures. This is equal to one minus the probability of
observing exactly zero failures \((y=0)\):

\[ P(D) = 1 - f(0) = 1 - \binom{n}{0} p^0 (1-p)^{n-0} \]

where \( P(D) \) is the probability that a certain defect will
be detected, or at least exist, in a sample of \( n \) units if
the probability of any unit having the defect is \( p \). Solving Eq. (2) for \( n \):

\[ P(D) = 1 - (1-p)^n \]

1 - \( P(D) \) is\( = (1-p)n \]

\[ \log[1 - P(D)] = \log[(1-p)^n] = n \log(1-p) \]

\[ n = \frac{\log[1 - P(D)]}{\log(1-p)} \]

If more than one failure is required to accept there is
a problem, or to establish the root cause, the number
of units required can be recalculated ignoring the terms
for exactly zero and exactly one failure. Again, it is
simpler to sum the probabilities of observing exactly
zero and exactly one failure and subtract them from one.

\[ P(D) = 1 - [f(0) + f(1)] \]

\[ = 1 - [(1-p)^n + np(1-p)^{n-1}] \]

Substitute the desired probability of detection for
\( P(D) \) and solve for \( n \) numerically.

To illustrate, if a particular failure mode exists in 5
percent of the population and if it must be found with
90 percent probability — or 90 percent of all 5 percent
failure modes must be found — then 45 units are needed
to find the defect. This assumes it can be detected with
only one occurrence. If two occurrences are required for
detection, 77 units must be tested. It becomes clear how
expensive it is to ignore the first failure(s). With minimal
test units, it is obvious that every failure mode
must be assumed to be significant. Each failure mode,
therefore, has to be addressed.

The number of units to be tested can also be deter-
mined graphically using Figure 2. The horizontal axis
is the probability that the defect exists in any one unit.
Follow the appropriate line until it intersects the curve
indicating the desired probability of detection. Then
draw a horizontal line over to the vertical axis and read
the number of units which must be tested. This nomo-
graph shows the probability of having one or more
defective units in the sample.

**TWO-FOR-THREE OBSERVATION** - Tripling
the number of units on test will approximately double
the average number of unique failure modes found.
This has been observed to be true for several different
products (Figure 3), and will be further substantiated
after discussing failure mode distributions.

Failure Mode Distribution - Corvin Kuklinski,
Hewlett-Packard's statistician at Disc Memory Divi-
sion, modeled Pareto lists of failing subassemblies. He
discovered that the distribution can be reasonably ap-
proximated by a geometric series\(^1\) (Figure 4). In Pareto
order, each subassembly has a failure rate equal to the
previous subassembly’s failure rate times a constant
ratio:

\[ F_i = F_{i-1} r = F_1 r^{i-1} \]

where: \( F \) is the defect probability or failure rate
\( i \) is the failure mode number, the largest
mode being 1
\( r \) is the geometric series ratio

Kuklinski observed that the ratio \( r \) is about 0.7 for
new products and approaches 0.9 for mature products.
It is not known for certain whether this model describes
actual failure mode distributions, but only that it fits
some field replaceable units (FRU) data. It may be very
close, especially considering that the geometric series
has a finite area. The area equals the largest element
divided by the quantity one minus the ratio. This area
corresponds to the total product failure rate.

\[ F_{total} = \sum_{i=1}^{\infty} F_i = \frac{F_1}{(1-r)} \]

The Pareto principle states that a small proportion
of causes account for the majority of effects. In a Pareto
list of failure modes distributed according to a geo-
metric series with a ratio of 0.8, the top seven failure modes
(out of infinite failure modes?) account for 80 percent
of the total failures. If the ratio is 0.7, the top five failure
modes contribute 80 percent of the problem; and if the
ratio is 0.9, 15 failure modes contribute 80 percent.

Mathematical Substantiation - A closed form bi-
nomial approximation which assumes applicability of
the geometric series shows a two-for-three relationship
under certain conditions. Table 1 shows the average
numbers of failure modes found for different numbers
of units, failure rates, and geometric series ratios. The
relative value of testing more units is not very sensitive
to the geometric series ratio. At low failure rates with
few prototypes, the return is closer to three-for-three. At
Figure 2 - Binomial Probability of Failure Mode Detection

Figure 3 - Two-for-Three Observation

Figure 4 - Geometric Series
Table 1 - Average Number of Failure Modes Found
(numbers in parentheses are the ratio of the value on the right to the value on the left; i.e., the ratio of failure modes when n is tripled)

<table>
<thead>
<tr>
<th>Probability</th>
<th>Total Failure Fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Failing n=10 30 90 270</td>
</tr>
<tr>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>10%</td>
<td>3.51 (1.82) 6.38 (1.48) 9.43 (1.33) 12.5</td>
</tr>
<tr>
<td>5%</td>
<td>0.93 (2.55) 2.37 (2.08) 4.94 (1.62) 7.99</td>
</tr>
<tr>
<td>1%</td>
<td>0.48 (2.77) 1.33 (2.41) 3.20 (1.89) 6.05</td>
</tr>
</tbody>
</table>

Geometric Series Ratio = 0.7

<table>
<thead>
<tr>
<th>Probability</th>
<th>Total Failure Fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Failing n=10 30 90 270</td>
</tr>
<tr>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>10%</td>
<td>3.96 (2.05) 8.11 (1.60) 13.0 (1.38) 17.9</td>
</tr>
<tr>
<td>5%</td>
<td>0.95 (2.71) 2.57 (2.32) 5.95 (1.80) 10.7</td>
</tr>
<tr>
<td>1%</td>
<td>0.49 (2.84) 1.39 (2.58) 3.59 (2.15) 7.60</td>
</tr>
</tbody>
</table>

Geometric Series Ratio = 0.8

<table>
<thead>
<tr>
<th>Probability</th>
<th>Total Failure Fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Failing n=10 30 90 270</td>
</tr>
<tr>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>10%</td>
<td>4.43 (2.42) 10.7 (1.88) 20.1 (1.49) 30.0</td>
</tr>
<tr>
<td>5%</td>
<td>0.97 (2.86) 2.77 (2.60) 7.21 (2.14) 15.4</td>
</tr>
<tr>
<td>1%</td>
<td>0.49 (2.94) 1.44 (2.78) 4.00 (2.46) 9.82</td>
</tr>
</tbody>
</table>

Geometric Series Ratio = 0.9

- Two for Three Region

high failure rates with many prototypes, the return is less than two-for-three. In the region where most commercial products operate, the return is approximately two-for-three, which agrees with the empirical observation. This lends credence to the geometric series as being a reasonable failure mode distribution.

KUKLINSKI CURVES - Kuklinski extended the binomial probabilities to not only include one failure mode, but to include all failure modes that must be caught. Referring to Figure 5, defects with low probability of occurrence can be allowed, but all the larger defects must be found and eliminated. With n units on test, there is a certain binomial probability \( P(D) \) that each of these large failure modes will be detected. The product of these probabilities yields the probability of finding all the unacceptable defects. (When the probability of finding all the defects is over 0.90, only the first three failure modes to the left of the dividing line in Figure 5 contribute significantly to the result.)

The discussion so far has been limited to defect probabilities. The result of Kuklinski's work is a series of curves plotted as the desired fraction failing during warranty versus the number of units to test. An understanding of a product's time-to-failure distribution is prerequisite to convert defect probability into fraction failing during warranty. The time-dependent nature of defect rate will be ignored for the moment.

Figure 5 shows how many units to test to achieve a desired defect rate with a particular probability. To use this figure, one must assume that testing will reveal any existing defect. In terms of failure rates, testing must simulate the operating period of interest, e.g., a warranty period.

Different sets of curves can illustrate the effect of various probabilities (Figure 6), geometric series ratios (Figure 7), and numbers of failures that must be observed to ensure the problem gets solved (Figure 8).
If the warranty period is 2000 hours and the acceptable fraction failing is 0.06:

\[ F(2000) = 1 - e^{-2000/\alpha} = 0.06 \quad (8) \]

Rearranging to solve for \( \alpha \):

\[ \alpha = \frac{2000}{\ln(1-0.06)} = 32,323 \quad (9) \]

Substituting the value for \( \alpha \) into the failure function:

\[ F(t) = 1 - e^{-t/32,323} \quad (10) \]

Now the fraction failing can be determined for any other period. For instance, the fraction failing by 1500 hours is 0.0453. If testing for a shorter period, the failure modes will each be smaller and therefore harder to find. The modes for this example would be scaled by a ratio of 0.0453/0.06 = 0.76 before calculating the binomial probability. This ratio will be different for every failure rate, because the term \( \alpha \) in the failure function would change.

Figure 9 illustrates the effect of different test times for the exponential failure distribution in the example.

The curves assume that all the failure modes observed are corrected so they do not cause field failures. In other words, they show the best defect rate obtainable with a given number of prototypes.

Here is a summary of the assumptions made when using the Kuklinski curves:
- The testing sample is representative of the population.
- All failure modes are independent.
- The failure mode distribution can be reasonably approximated by a geometric series.
- The test has the same probability of producing failures as field use for equivalent time periods.
- All failure modes discovered are permanently eliminated and no new failure modes are introduced.

In addition, a geometric series ratio based upon experience with similar products must be assumed. Study
the effect of error in this assumption and choose a conservative value; i.e., a larger number of prototypes. Refer again to Figure 7.

It is evident from mathematics that a goal of high reliability necessitates multitudinous prototypes, that is until products can be designed that need no testing, analyzing, and fixing. As a point of reference, one of Hewlett-Packard's recently released system storage products had 610 prototypes built prior to introduction. Another system storage product still under development is budgeted for over 1000 prototypes.

**Empirical Data** - Figure 10 shows empirical correlation between the number of prototypes built and product reliability at introduction. Initial reliability is used rather than current reliability in order to remove the effect of product maturity.

Included in Figure 10 is a Kaklinski curve for a geometric series ratio of 0.7, a probability of 0.50, and assuming one occurrence is adequate for detection.

**HOW ARE THESE UNITS ACQUIRED?**

**Prototype Production** - With new understanding of reliability growth concepts, prototypes can no longer be built in batch mode because process and design changes which improve reliability cannot be tested until the next batch. Also, batch mode processes rarely achieve full process control or even an understanding of what input parameters need to be controlled.

A continuous prototype-building process affords the advantage of instantaneous process and design change implementation. A continuous process also makes sense when considering how many prototypes need to be tested. With a continuous process, the design and process can be proved simultaneously, process control parameters can be identified and quantified, training can be an ongoing process rather than a crash course, and production rates can be altered according to changing needs.

**Final Product Production** - Production yield problems are a valuable source of failure information because there is strong correlation between failures occurring in the factory and those occurring in the field. Table 2 is one product's Pareto list of failing subassemblies and their contributing fraction of all failures in the factory and in the field. These numbers have a statistical correlation coefficient of 0.98!

Figure 11 shows a product's predicted field failure rate — which is derived purely from factory yield data — and its actual field failure rate. The excellent correlation of field failure types and rates to similar factory data indicates that failure modes and rates observed in the first few hours of life are indicators of reliability performance in the first few months in the field. Others' experience is similar.*

John Smethurst, a Hewlett-Packard manufacturing manager, was fed up with poor yields and, therefore, determined to increase his production line's yield to 100 percent. Several months after reaching his goal, it was noticed that the warranty failure rate for Smethurst's product dropped to essentially zero. Again, this points to the same failure modes existing in the factory and in the field. A logical exception might be early wearout modes, but even some of these have been observed in the factory.

This correlation is important for at least two reasons. First, design and process problems should be caught in the factory, rather than by customers. Fac-

---

### Table 2 - Pareto List of Failing Subassemblies

<table>
<thead>
<tr>
<th>Part</th>
<th>Fraction of All Factory Failures</th>
<th>Fraction of All Field Failures</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>37.1%</td>
<td>36.7%</td>
</tr>
<tr>
<td>B</td>
<td>6.0%</td>
<td>4.1%</td>
</tr>
<tr>
<td>C</td>
<td>5.1%</td>
<td>4.0%</td>
</tr>
<tr>
<td>D</td>
<td>4.6%</td>
<td>3.7%</td>
</tr>
<tr>
<td>E</td>
<td>4.5%</td>
<td>5.9%</td>
</tr>
<tr>
<td>F</td>
<td>4.0%</td>
<td>4.1%</td>
</tr>
<tr>
<td>G</td>
<td>3.9%</td>
<td>6.7%</td>
</tr>
<tr>
<td>H</td>
<td>2.7%</td>
<td>3.5%</td>
</tr>
<tr>
<td>I</td>
<td>2.6%</td>
<td>2.4%</td>
</tr>
<tr>
<td>J</td>
<td>2.5%</td>
<td>2.7%</td>
</tr>
<tr>
<td>K</td>
<td>2.5%</td>
<td>0.1%</td>
</tr>
<tr>
<td>L</td>
<td>2.2%</td>
<td>1.4%</td>
</tr>
<tr>
<td>M</td>
<td>1.2%</td>
<td>2.2%</td>
</tr>
<tr>
<td>N</td>
<td>0.9%</td>
<td>1.3%</td>
</tr>
<tr>
<td>O</td>
<td>0.6%</td>
<td>1.4%</td>
</tr>
<tr>
<td>P</td>
<td>0.3%</td>
<td>0.1%</td>
</tr>
</tbody>
</table>

*Willis J. Willoughby, Jr. stated at the 1986 Institute of Environmental Sciences Annual Technical Meeting that factory failures are a perfect representation of field failures. This statement is made from his many years of reliability management experience with the Navy. Brigadier General Frank S. Goodell of the Air Force said he concurred with Mr. Willoughby.
Some stresses, such as actuated stresses, continually point to more significant weaknesses than a defective one. They should be forced or accelerated with stress

Factory failures are an invaluable source of feedback for reliability growth.

FORCE FAILURES

THE SUCCESS OF FAILURES - Product and process weaknesses must be found before they can be fixed. To quickly improve reliability, any threat to reliability must be quickly uncovered. An attitude of wanting to observe failures is foreign to many, but reliability growth occurs only in the presence of failures. If there are no failures for feedback, it is difficult to improve reliability and impossible to measure improvement.

Every device has its weak links, and their manifestation must be actively sought. Every tool should be employed to coax narrow margins to reveal themselves. Failures may appear in time, but time is costly; failures should be forced or accelerated with stress.

GENERIC STRESSES - Some stresses, such as temperature cycling, vibration, power cycling, and humidity, can be applied to almost any product. These are convenient because they require minimum test design and no product specific test equipment. Generic stresses can be very effective in precipitating failures, depending upon what failure modes are present. These may or may not reveal the weakest link in a product but generally reveal significant problems and are extremely cost effective. They should be used in combination with product specific stresses.

PRODUCT SPECIFIC STRESSES - Greater product knowledge facilitates designing more focused internal stresses, such as DC voltage variation, clock frequency variation, and even component value variation. These potentially point to more significant weaknesses than a generic stress, but results again depend upon failure mode makeup.

A product specific stress is more likely to reveal narrow margins in expected areas, whereas generic stresses usually turn up a few surprises. Both kinds of tests are necessary: product knowledge must be used for product specific "rifle-shot" tests; generic "shotgun" tests can reveal overlooked areas.

STRESS LEVELS - Stress should be applied beyond the published product specifications. The test levels are dynamic; they should be cranked up continually to force an approximately constant failure rate. Depending upon the product strength distribution, the levels may go far beyond the product specifications. For example, HP's San Diego Division tests electronic assemblies from -55°C to 125°C at a ramp rate of 20°C per minute. These assemblies are not going to the battlefield, but only to offices. Their products' temperature specifications are 0 to 55°C.

When failures are no longer produced at these extremes, using a different stress will usually reveal weaker links than those found by increasing temperature further.

RELEVANCE OF STRESS FAILURES - Failure modes precipitated by stress outside of product specifications are statistically relevant field failures. Rarely does an unaddressed stress-testing failure not become a field problem.

Many feelings and opinions surround a failure occurring at 80°C in a product designed for only 55°C. The recurring tendency to gloss over a failure is a psychological barrier to reliability growth. Facts must override opinion here because evidence and statistics often point in the opposite direction of intuition.

Failures occur only when stress exceeds strength (Figure 12). Strength is generally broadly distributed. Product strength varies with time, and it generally gets worse rather than better (Figure 13). Applying stress merely simulates what happens with age (Figure 14). For strengths that don't decrease with age, the increased stress increases the overlap area, thereby improving the chances of a small failure mode occurring with a smaller sample of units than would otherwise be required. Stress testing amplifies unreliability so it can be detected.

It is extremely rare for an increased stress to cause a foolish failure, a failure mode which would never occur.
demanded to be a foolish failure because it had not been observed within the operating limits. To further prove the failure's irrelevance, these same drives which had previously failed were restored to operating condition and then retested within the specified operating range. No failures were noted. The problem was ignored and so became this product's number one field failure mode for the next two years!

The same product also had a digital flip-flop that failed at -5°C but returned to normal at 0°C. Because it was only specified to 0°C, no action was taken. What was actually uncovered was the tail of a very wide strength distribution which extended far into the operating range. This IC became the product's most frequently failing electronic component.

Failure modes precipitated in stress testing are relevant, almost without exception. Attempts to prove otherwise only delay the redesign process and, therefore, reduce the reliability growth rate. Stress reduces the time required to find failures; even more time is saved by immediately initiating failure analysis and then immediately redesigning.

ADDRESS ALL FAILURE MODES

One early wearout mode in a Hewlett-Packard product cost the responsible division over ten million dollars. This failure mode, it was later discovered, had been observed during product development but was not addressed because its several isolated occurrences were not understood, believed, documented, or made visible.

DESIGN DEFECT TRACKING - All failure modes must be documented and addressed. A project manager has the responsibility to fight against the human shortcomings of defensiveness, self-justification, and mistake suppression.

A tool to assist in accomplishing this goal was developed by Larry Brunetti of Hewlett-Packard's Boise Division. Called Design Defect Tracking (DDT), it is to be used by project managers to manage failure analysis and solution implementation.

The tracking procedure is initiated by completing a defect report, an activity which must be well rewarded.

DDT provides discrete classifications for defect status which facilitate understanding and simplify visual communication. Making the status of each defect highly visible to all project team members and to upper management is the best way to get action. People respond to visible priorities.

The six defect status classifications are listed in Table 3.

The number of defects in each classification are plotted in a stacked bar chart with zeros at the bottom (Figure 15). After a period of time, the tops of the bars, representing the sum of all defects found to date, will asymptotically approach the maximum number of inherent defects, at least for a given stress level. The number of unsolved problems will approach zero. These two signs are indicative of a maturing design.

In addition to the bar chart, the defects and their current status and activity should be posted in clear view of all team members.
Table 3 - DDT Defect Status

<table>
<thead>
<tr>
<th>Status</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Cause unknown.</td>
</tr>
<tr>
<td>1</td>
<td>Root cause has been isolated.</td>
</tr>
<tr>
<td>2</td>
<td>Solution has been designed.</td>
</tr>
<tr>
<td>3</td>
<td>Solution has been implemented in all prototypes.</td>
</tr>
<tr>
<td>4</td>
<td>Solution has been verified.</td>
</tr>
<tr>
<td>5</td>
<td>Problem and solution have been recorded in a lessons learned database so that it will never recur.</td>
</tr>
</tbody>
</table>

Psychological and organizational roadblocks must be cleared so a failure mode can move through these states as quickly as possible. However, avoid the tendency to move too quickly from state 3 to state 4; lots of accumulated test time is necessary to verify a fix.

As a rough point of reference, Jim Bobroff, Quality Assurance Engineering Manager at HP's Personal Office Computer Division, tracked the average time to go from state 0 to state 1 at 37 days for electrical problems and 15 days for mechanical problems. The average time to go from state 1 to state 4 was 7 days for electrical problems (too fast for proper verification) and 36 days for mechanical problems.

FAILURE ANALYSIS - Achieving DDT state 1 is most critical. When the root cause is known, the solution is often obvious. It has been said, "A problem well stated is a problem half solved." Good trouble-shooting technique as well as good statistical experiment design must be employed in order to operate at a level above amateur exorcism.

Katsu Yoshimoto of HP's Japanese subsidiary, YHP, suggests asking "Why?" five times to get to the source of the problem. Consider the following example:

<table>
<thead>
<tr>
<th>Engineer</th>
<th>Manager</th>
</tr>
</thead>
<tbody>
<tr>
<td>The 8510 failed.</td>
<td>Why?</td>
</tr>
<tr>
<td>Bad microprocessor board.</td>
<td>Why?</td>
</tr>
<tr>
<td>EPROM died.</td>
<td>Why?</td>
</tr>
<tr>
<td>Electromigration on buried metalization layer.</td>
<td>Why?</td>
</tr>
<tr>
<td>Violation of current density design rule.</td>
<td>Why?</td>
</tr>
<tr>
<td>Chip designer didn't catch the violation.</td>
<td>Why?</td>
</tr>
</tbody>
</table>

Each step will require engineering analysis and then the question is repeated. A variation of Yoshimoto's words might be to ask why $n$ times, where $n$ is large.

Every failure mode must be tracked carefully and adequately addressed. There is never scientific reason to ignore a failure mode before DDT state 1, which is when the root cause of a failure is precisely understood.

RESULTS

Of the many different reliability growth models, the Duane model is used at Diac Memory Division because of its inherent simplicity, broad acceptance, and robust nature.
The Duane model uses cumulative data, avoiding the abhorrent tendency to disregard failure data when problems are fixed. It is helpful in predicting future reliability based on early results and also yields an idea of current MTBF.

The slope of the Duane curve is a measure of reliability growth rate and can range from 0 to 1, with greater slope indicating faster growth. It can also be negative if there is negative growth. It is generally accepted that the more aggressive the TAAF program, the higher the growth rate. Typically, slopes of 0.3 to 0.6 are considered good growth rates. By implementing the concepts described herein, a growth rate of 0.7 has been observed for HP's new 571 megabyte disc drive (Figure 16).

CONCLUSIONS

Phenomenal reliability growth is achieved through a combination of many test hours, many test units, accelerated testing, good failure analysis, and a resolve to fix every failure mode found. The number of prototypes required must be determined by the reliability goal and sound mathematical modeling. If the project budget and reliability goals are mutually exclusive, change one or both.

Test the process design simultaneously with the product design by using a continuous build process.

Use stress to precipitate failures quickly. Assume all failures represent narrow margins and weak links.

Search for the root cause of all problems before attempting to fix them. Only by understanding the real failure mechanism can the fix be permanent. Shotgun fixes do not last.

Address all failure modes. Testing a product to see if it has any weaknesses and then ignoring the test results wastes time and money.

Reliability verification tests often verify that more growth is needed. Allow robust reliability models to measure reliability and use all available resources toward reliability growth rather than reliability proof.

ACKNOWLEDGMENTS

This paper represents years of accumulated experiences by many individuals. Any list of names would surely be missing several key players. As a result, it is the organization for which they all work, Hewlett-Packard Company, that deserves and receives credit. I am particularly indebted to Corvin Kuklinski for his willingness to contribute his work to this effort and to Chet Haibel, my manager, for his ongoing support. Gratitude is also extended to Barbara Helling for her excellent editing, Helen Bokman for superb typesetting, and Jeff Chrisue for the precision graphics.

REFERENCES

1) Kuklinski, Corvin, unpublished work, Hewlett-Packard Company, Disc Memory Division, P.O. Box 39, Boise, Idaho 83707. Used by permission.


USE OF HALT TECHNIQUES DURING PRODUCTION OF HIGHLY RELIABLE MILITARY ELECTRONIC EQUIPMENT

Bruce A. McAfee
Hughes Defense Communications

Abstract
The use of Highly Accelerated Life Test (HALT) during product design and development is well documented as an efficient process for increasing the field reliability of military avionics. As production run rates increase, parts from new and unproven suppliers of electronic components are used in the manufacturing of hardware. The majority of today's newly designed semi-conductor parts are plastic encapsulated and available only as commercial or industrial grade components not certified to typical military environments. Qualification of individual part suppliers by the OEM is cost prohibitive and few commercial suppliers are willing to perform the testing to satisfy an ever shrinking military demand. HALT techniques provide a fast cost efficient methodology to verify that new suppliers of parts are indeed meeting the military reliability requirements. Using HALT periodically during production will not only certify new vendors, but also verify that current suppliers have not degraded the reliability of their parts by modifying their part building process. This paper will address the successes and shortcomings of HALT concerning part supplier verification.

Introduction
The necessity of supplying a reliable product to the military is becoming extremely important as the DOD focuses on field reliability of products and longer warranty time periods. This change in emphasis by the DOD, along with the shift in the semi-conductor business to essentially a commercial (non-military) business, has placed a great burden on military avionics suppliers. Being cost competitive in an ever shrinking military budget mandates the use of new technology components that may not be "qualified" for traditional military environments. There is no future in developing tomorrow's products utilizing yesterday's technology.

The use of HALT, as a reliability verification tool, during production can be just as beneficial as the use of HALT during product design. Typically, as the production run rates increase, parts from new and unproven suppliers of electronic components are used in manufacturing. The majority of today's newly designed semi-conductor components are plastic encapsulated and available only as a commercial or industrial grade not certified for military applications. In particular most new technology semi-conductor components are non-hermetically sealed and specified for temperature ranges much less than those required for military operation. The use of these components, in applications beyond the stated temperatures ranges, is necessary to produce a cost competitive military product.

The question arises as to how to evaluate the reliability of various component manufacturers during the life of production. Simulated field environments are normally used to evaluate the reliability of a product. Significant time compression is typically gained by
eliminating field environments with low stress levels and concentrating on the most severe environments. Time compression from years of field data into several months of simulated testing is not uncommon for traditional reliability tests. HALT takes time compression a step farther by increasing the stress levels significantly beyond expected field environments. This acceleration further decreases time from months of testing to days or even hours. HALT is a cost effective process that can be performed during production to verify that current part manufacturers maintain an acceptable level of quality and new part manufacturers are qualified to a similar standard.

Hardware Background

Hardware Description

The avionics equipment developed by Hughes Defense Communications for the United States Air Force is a 10 Watt VHF-FM and VHF-AM Receiver-Transmitter (R/T) utilizing a Remote Control Set (RCS) for cockpit operation. The entire system consists of approximately 2700 parts. A majority of these parts are commercial quality and surface mount technology. Less than 10% of the parts are "through-hole" technology. The operational environment for the avionics system is a rigorous military environment including a -54 to +71°C temperature range with severe vibration requirements.

Hardware Development

HALT was successfully performed on the individual modules and units during design and development of qualification hardware. High levels of random vibration along with rapid temperature transitions were applied to hardware samples to determine weak links in product design. A total of 23 faults were identified during accelerated reliability testing. Six workmanship problems were discovered and feedback was provided to the appropriate manufacturing area. Seventeen design related faults were identified, deemed relevant to performance, and corrective action implemented. The hardware successfully completed environmental qualification and performed flawlessly during Air Force flight testing.

Prior to starting production, several design modifications were implemented to enhance functional performance and increase product producability. The basic design remained the same, however, the production process was modified and new process variables were introduced to the hardware. These new internal variables along with a few surprises from our component suppliers provided a large amount of frustration during the first few months of production. Because of the success of HALT during design and development, HALT was applied to the initial production hardware with similar success.

HALT Process

Description

The HALT Process consists of multiple environmental stimuli applied at levels well beyond those expected during product use. Traditional environmental testing programs utilize well-defined field environments over a relatively long period of time to identify design weaknesses. HALT attempts to reduce the time needed to precipitate these same defects.
The goal of HALT is to improve the product's design to a point where manufacturing variances and environmental effects have minimal impact on performance and reliability.

Mechanical stress (random vibration and temperature) and electrical stress (voltage, frequency, power cycling, etc.) are utilized to precipitate faults in traditional environmental tests. Highly Accelerated Life Testing applies these same stimuli in a very aggressive manner to precipitate and detect weak links in the product's design. Other environments, humidity, shock, salt fog, etc. may also be used to stimulate faults depending upon the intended use of the product.

Step stress testing with multiple environments begins with the least destructive environment and proceeds to the most destructive (this is assuming assets are limited). Typically, electrical stressing is performed first at ambient conditions, then low temperature testing is performed, high temperature next and finally random vibration. Once the separate environments are complete, mechanical stress and electrical stress are combined for a thorough exercise of the test item. During all stress applications complete monitoring of the test item is necessary to detect any abnormality or degradation in performance. Figure 1 describes the general HALT Process.

![Figure 1: HALT Process Diagram](image-url)
General HALT Results

**Transformer**
- Visual examination revealed a broken lead from the transformer at the PWB termination point. The break was caused by movement of the coil during vibration. The coil is attached to the PWB with an RTV adhesive that allows a small amount of movement during vibration or shock. The transformer leads had been trimmed to a minimum length during installation. The small gage leads were stretched taught and could not support minimal coil movement during mechanical vibration. RECOMMENDATION: Add a service loop between the coil and PWB termination point on the PWB to allow for movement of the transformer coil during vibration. Verification of the service loop took less than two hours in HALT.

**Linear Microcircuit**
- Examination via a scanning electron microscope of five failed parts revealed areas of questionable die metallization coverage. Metallization runs over steps exhibited thinning, voiding and possible cracking or separation (Figure 2). Stress from rapid temperature transitions caused fractures to surface at thin metallization areas. RECOMMENDATION: Evaluate the device manufacturer's die metallization process. Step height, cold substrate evaporation or a poor material source are potential causes for improper metallization.

![Figure 2: Linear Microcircuit Internal Scan](image)

**Band Pass Crystal Filter**
- Microscopic examination revealed a broken lead from the coil near the PWB termination point. The failure occurred at the apex of the soldered area where the wire meets the solder on the PWB (Figure 3). The break was caused by movement of the coil during vibration. Internal examination of good parts revealed inadequate lead length to allow coil movement during mechanical vibration or shock. Also, the solder joint of the failed lead was perpendicular to the coil instead of parallel to the coil as found upon internal examination of good parts. RECOMMENDATION: Add a service loop between the coil and PWB termination point on the internal PWB.
RF Oscillator

Internal microscopic examination of three failed devices revealed identical fracture locations. The failure occurs on the bus wire lead connection from the crystal oscillator to the PWB (Figure 4). The fault was caused by vibration fatigue fracturing. The device manufacturer had eliminated the RTV substance used during qualification of the device. RECOMMENDATION: Return all stock for RTV adhesive inclusion and retest for compliance. Requalification of the RF oscillator parts with RTV applied to the crystal took less than four hours with HALT.
**Ceramic Filter**

Visual examination revealed multiple fractures in the PWB solder joints. The fractures were induced by movement of the ceramic filter during vibration. Further investigation revealed that solder beneath the component leads had not reflowed during PWB assembly. A spectral analysis of the component lead finish indicated the part manufacturer had used silver based solder to coat the leads. The reflow temperature of silver solder is higher than tin-lead, therefore complete reflow beneath the leads did not occur. Thus, the solder joints that were visually acceptable were in fact extremely weak. **RECOMMENDATION:** Change the lead finish to tin-lead based solder. Re-verification of the new parts took less than four hours in HALT.

**Toggle Switches**

The switch becomes very difficult to operate at -54°C. Internal examination revealed binding in the pivot area at the O-ring seal of the switch handle. Further investigation revealed the part manufacturer had increased the size of the O-ring to improve the moisture sealing characteristics of the switch assembly. **RECOMMENDATION:** Return to the original O-ring design for operation at cold temperatures. Re-verification of the original O-ring took less than two hours in HALT. Subsequent testing of the toggle switch seal revealed no moisture problems.

**Digital Microcircuit**

Testing of the device at room ambient conditions revealed no anomalies. However, applying pressure to the device lid caused breakdown characteristics to change. Scanning acoustic microscopy revealed poor die attachment over most of the die area (Figure 5A). Microscopic examination during and after cross-sectioning revealed the die attach material had delaminated from the die paddle (Figure 5B). The delamination is the result of moisture absorption and subsequent assembly induced temperature stress, commonly referred to as popcorn effect. **RECOMMENDATION:** Store the parts in a dry atmosphere and bake prior to assembly.

![Figure 5: Digital Microcircuit Sonogram (A) and Micrograph (B)](image)

**Discussion Of Results**

The failures described above are the abnormalities of the production process. Many other second source manufacturers were evaluated with HALT and had successful results.
There are some very good manufacturers of commercial parts and a few not so good that are willing to improve their assembly process to meet the demands of the military. And of course, some part manufacturers exist that are not willing to improve their product for military applications.

The descriptions above include part manufacturers of each category. The two in-house caused failures (improper transformer lead length and moisture absorption within the digital microcircuit) have been corrected and no further failures have occurred. The part manufacturer of the linear microcircuit with improper metallization has discontinued production of the part. The RF oscillator vibration sensitivity has been eliminated by adding the RTV adhesive between the crystal and the PWB. The new parts were successfully tested in HALT upon receipt. The new ceramic filter supplier is using tin lead based solder on the part leads. The new parts passed HALT and the manufacturer is now an approved source for this part. The toggle switch manufacturer has returned the O-ring to the original design per our request.

Summary

The use of commercial grade components for military applications require the parts to be used beyond the published temperature range. Commercial grade components are normally specified for 0 to 70°C, industrial grade for -40 to 85°C, while traditional military avionics require -55 to 125°C. Our experience with commercial and industrial grade parts is all easily exceed published data and most will perform satisfactorily beyond military requirements. The question still exists as to why part manufacturers do not advertise the extended performance capability. One can only speculate that the relatively low demand for military avionics is the main reason.

Other issues arise when using parts outside of the published temperature range. How can the user ensure the parts will continue to perform beyond specified environments for years to come? The device manufacturer should notify users if a change occurs in form, fit or function of published data. However, if use of a part is occurring beyond published limits, then the manufacturer is not obligated to notify users. Periodic testing of hardware with HALT is a cost effective tool to verify part performance during production.

Another variable that occurs during production is the expansion of part manufacturers by material procurement to obtain the best possible price for the end product. These new suppliers are potentially introducing material that has not been verified to perform under the rigors demanded by the military. Previously the government published a Qualified Parts Listing (QPL) that theoretically identified part manufacturers of equal qualifications. Unfortunately a similar listing does not exist for plastic integrated circuits. Each OEM must develop its own approved part manufacturer listing or risk putting questionable product into the field. Development of a comprehensive internal QPL is cost prohibitive. Use of HALT is to verify second and third source component vendors is fast and cost effective.

The use of distribution centers for part purchases is another variable that brings multiple manufacturers at irregular intervals to the production process. Again, performing HALT periodically during production will identify whether the distribution center has selected a quality manufacturer or one that is not desirable.

The benefits from applying HALT periodically during production are profound. Experience has shown the most important variable to controlling warranty cost is the ability to contain the quantity of abnormal parts in the deliverable hardware. The HALT process is effective in identifying part manufacturers of marginal or unreliable parts.
Analysis of concurrent Tri-axial Random Vibration and Thermal Cycling Applied to Computer Circuit Cards

Dennis.P.Pachucki

Sun Microsystems Computer Corporation
SPARC Technology Business
410 N. Mary Avenue
Sunnyvale, CA 94086

Biography

The author is presently a Program Manager for the SPARC Tech Modules Division and previously held the position of Staff Engineer in the Advanced Test Process Research Department at Sun Microsystems. He has been instrumental in the introduction and development of the ESS processes and equipment presently used in manufacturing. During his 7 year tenure at Sun he has also worked in the quality and technology research departments, responsible for new product introduction quality, the development of an on-going reliability test process and facility, and more recently the establishment and management of an environmental stress facility used for ESS experimentation and accelerated screening tests. Previous to his work at Sun, he has held test, manufacturing engineering, R&D, and management responsibilities in other commercial and government industries. He has a B.S in Industrial Engineering from San Jose State, and an A.A. in Electrical Engineering from M.I.T.

Abstract

Sun Microsystems has strived to be a leader in the workstation market. To maintain and advance in this leadership role, manufacturing process improvements which increase productivity, decrease test process time, and improve customer satisfaction are being pursued. The application of environmental stress screening is a method of achieving these improvements.

This paper describes in detail the process, implementation, and results of 3 independent experiments directed solely at applying random vibration or concurrent application with temperature to Sun Microsystems’s computer circuit cards.

The goal of the experiments is to establish if Sun’s circuit card manufacturing could replace their thermal cycling stress screening process with a much shorter ambient temperature random vibration screening process - code named The Simple Plant. The experiments consists of stressing 600 circuit cards, divided into 3 lots of 200: Two hundred are vibrated at ambient for ten minutes, another 200 are vibrated at 0C and 55C for 1.5 minute durations at each of the 8 dwell periods included in a 4 thermal cycle profile. The final 200 are super stressed for one cycle by vibrating at 85C and -55C for 1.5 minute at each dwell then followed with vibrating at 0C and 55C for 3 thermal cycles for 1.5 minutes at each dwell period. All circuit cards are functionally test monitored during the thermal and random vibration stress process, except during the super stress cycle. The cards are then tracked through the standard manufacturing process.

The intent of this paper is to share with the reader the resulting yields as tracked through the entire manufacturing process along with the root cause failure analysis on all failures. Conclusions and recommendations will be summarized which will point out key failure modes, profiles, time to failure averages, and the resulting failure elimination and corrections.

Keywords

Environmental Stress Screening
Triaxial Random Vibration
Accelerated Temperature Cycling
Power Cycling
Stress Profiles
Statistical Significance
Proof of Screen
Step Stressing
Introduction

The experiment is performed on the SPARC 10 circuit cards. It is a two card set. The cards consist of a 8.5 x 11-inch, double-sided, high-density mother board comprised predominately of surface-mount components that have a minimum lead pitch of 25 mils. It consists of approximately 350 components with over 200 surface mounted discrete components on the bottom and over 100 I.C.'s, ASIC devices, and crystals on the top. The second card consists of a 5.75 x 3.25 high density double sided daughter board referred to as the CPU module, consisting of SMD, ASIC, and discrete through hole devices. A sample size of 600 hundred PWAs is used for this experiment, further details on determining the proper sample size are described later in this paper.

The 600 boards are processed through the USOPS Environmental Stress Lab in Milpitas, before continuing on to production's ESS and Systems Test. The Stress Lab has a combined triaxial quasi-random vibration and temperature-cycling chamber. The Lab utilizes data acquisition software and signal conditioning hardware for temperature and vibration data acquisition. Also, four SPARC 10 chassis were modified and installed in the chamber in order to effectively apply the accelerated stresses used in the experiment.

Additionally, a custom Open Windows based program was used to control the environmental vibration chamber, as well as, the cycling and recording of the monitored diagnostic tests.

Objective

The experiment sought to determine if a 30 minute stress process consisting of 10 minutes of monitored triaxial random vibrational excitation at 25C could replace manufacturing's 7 hour temperature cycling ESS process.

Stress Overview

The stresses applied while functional testing were:

- 25GRMS Tri-axial Quasi Random Vibration
  (Frequency 20 to 2Khz)

- Temperature Cycling

The following sub-topics explain how the stresses were determined and applied for the experiments.

Temperature Screen Designs

Temperature Profiles

The experiments consisted of 3 distinct profiles applied in the Stress Lab. A quantity of four circuit cards were installed into the chamber and tested at one time, until 200 of each profile were completed. The temperatures were measured and controlled from the circuit cards.

- Profile One - All tests performed at 25C ambient temperature and ten minutes of 25 GRMS tri-axial random vibration, (see figure three).

- Profile Two - The profile consisted of four temperature and vibration cycles between the ranges of 0C and 55C, with random vibration applied for 15 minutes at each dwell temperature of either 0C or 55C, for a vibration total of 10 minutes. The profile was restricted to these functionally operating temperatures because of early development of one of the components. (see figure four).

- Profile Three - This was a two step enhanced profile. Determining, if applying a highly accelerated stress during a non operating portion of the testing, could aid in reducing the number of functionally operating temperature cycles, thereby shortening the overall stress and test times. It consists of four temperature and vibration cycles, the first cycle ranges between -35C and 85C, and 1.5 minutes vibration applied at each dwell. The circuit card is un-powered. Following this cycle the circuit card is powered up, diagnostics loaded and started at 25C before continuing to complete 3 remaining cycles of 0C to 55C with 1.5 minutes of vibration at each temperature dwell. (see figure five).

Temperature Change Rate

Product temperature is changed at a rate of 20C per minute. This is accomplished by mounting 5 muffin fans in strategic chamber locations. Enabling the air flow to be evenly directed across each of the four circuit cards installed on top of the 44 x 44 inch table. The overall chamber air velocity was approximately 600 feet / minute. The temperatures are monitored and plotted using integrated software and data-acquisition equipment developed to work on Sun workstations.

Diagnostic Monitoring and Power Cycling

The diagnostic monitoring also described as functional testing, consisted of a Power On Self Test, Unix boot, and Sundiag, a system exerciser. These tests are performed before applying any of the stresses to ensure all circuits card's functioned properly. Sundiag is applied during the
stressing. Upon completing the profile, the circuit card is power cycled once and all tests are repeated at 25C ambient before returning them to the production process. This power cycle and last battery of tests is performed to determine if the stresses precipitate any faults that may be revealed upon returning to 25C ambient condition. The chamber control, diagnostic control and monitoring, and failure recording are exercised through a custom ESS control program.

Vibration Screen Determination

The increasing step stress method determined the random vibration levels. Application of the step stress method is explained as follows: Three boards are used to determine the screen level. Two boards are each individually vibrated with the vibration responses measured at six different locations. Diagnostics are monitored during the entire process. Random vibration frequencies of 20 to 2Khz are increased in levels of 1 to 5 GRMS increments, dwelling for one to five minutes at each increment. Each failure is traced and validated. Forty GRMS was found to be the maximum functional operational limit, for the mother board and CPU card combination.

Ten minutes is selected for the experiment's vibration screen time length. In determining the vibration stress level, the proof of screen method is applied and is explained as follows: Seventy-five percent of the operational limit (.75 X 40 = 30 GRMS), is applied for two hours on the third board set with no failures occurring at 25C ambient. However, at 0C and 55C it was found that the vibration level needed to be lowered to 25 GRMS, due to SIMM socket design. Therefore, the vibration screen profile level of 25 GRMS was used through out the experiment.

The proof of screen rule of thumb in determining the experiment's screen is to apply the 75% level (~30 GRMS in this case), ten times longer than the vibration screen duration (10 X 10 minutes = 100 minutes), and for it to be error free. This validates that the stress level has not caused any flaws that were not already inherent in the product and has not significantly shortened a product's useful life.

Fixture Design

To vibrate, temperature cycle, and monitor four boards simultaneously, four SPARC 10 plastic chassis' without top covers are modified and bolted to the vibration table. Also, power, signal, disk and floppy drive cabling, are lengthened and connected to an external test bench to enable the circuit cards to be tested outside of their system enclosures. Additionally, an aluminum mounting bar and DeStaco clamps were designed and employed to ensure a firm connection to the vibration table. A 9.5 GRMS input results in ~25 GRMS response as measured at the center of each board with less than 5% variance between all four positions.

Experiment Overview

Figure 1 is the process flow diagram of the circuit card flow through manufacturing and the Stress Lab. Figure 2 shows the profile limits, cycles, and durations of manufacturing's Environmental Stress Screen and Four Corner Tests, (ESS / FCT). The process is designed to reveal and catch any test escapes during the typical production ESS / FCT process that the Stress Lab's experiments may have missed or precipitated. Following the ESS / FCT process the circuit cards are input to Systems Aging Testing, (SAT). This is a system integration test, where the circuit cards are installed in the system chassis along with their power supply and peripherals and tested with functional diagnostics at ambient for eight hours before customer delivery. The circuit card failures that occur during ESS / FCT and SAT are the screen escapes or screen induced failures directly resulting from the three different profiles.

Sample Size Selection

The sample size of 200 per experiment profile, was selected after conferring with personnel from Sun's Quality and Reliability departments. It was felt that a sample of no less than 200 could be considered significant in revealing the trend for the experiment population. The sample sizes of 200 each would then be compared to a typical production sample lot of ~1200.

Experiments Result Summaries

Table 1 - Experiment I Failure Quantities

<table>
<thead>
<tr>
<th>Failure Category</th>
<th>Stress Lab</th>
<th>ESS / FCT</th>
<th>SAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Components</td>
<td>3 ICs</td>
<td>0</td>
<td>1 IC</td>
</tr>
<tr>
<td>Workmanship</td>
<td>2 Shorts</td>
<td>1 open</td>
<td>0</td>
</tr>
<tr>
<td>Crystals</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

The Total Time to Failure, (TTF), for experiment one in the Stress Lab ranged from one to eight minutes, with 4.25 being the average TTF, during the monitored vibration testing.
Figure 1. Experiment Overall Process Flow Diagram

Environmental Stress Lab

Profile 1
(25°C Random Vibration) PASSED

Profile 2
(T/C 0°C - 55°C and Random Vibration)

Profile 3
(T/C -35°C to 85°C, 0°C to 55°C and Random Vibration)

200 circuit cards each profile for a total of 600

Figure 2. Manufacturing's ESS/FCT Profile

shows transition from 4th cycle to 4 corners

18 minutes

Power Cycle 5 secs on/off

PROCEEDINGS—Institute of Environmental Sciences

146
Figure 3. Experiment 1, Profile 1, 10 minutes of 25 GRMS Vibration at 25C.

Profile 1

200 C2's
ICT → Post/Boot/Sundiag → 10 Minutes Vib. → Power Off/On - Sundiag → ESS/FCT - SAT

@ 25C Ambient: → Power Off/On - Sundiag → ESS/FCT - SAT

Load / Testing / Stress / Unload = 1 hour

Figure 4. Experiment 2, Profile 2, 25 GRMS Vibration at 0C and 55C.

Profile 2

200 C2's
ICT → Post/Boot/Sundiag → 1.5 Min. Vib dwells → Power Off/On - Sundiag → ESS/FCT - SAT

55C

- 35 Minutes -

0C

Load / Testing / Stress / Unload = 1.25 hours

Figure 5. Experiment 3, Profile 3, 25 GRMS Vibration at -35C, 85C, 0C, 55C.

Profile 3

200 C2's
ICT → Post/Boot/Sundiag → 1.5 Min. Vib dwells → Power Off/On - Sundiag → ESS/FCT - SAT

85C

- 45 Minutes -

55C powered

0C powered

Load / Testing / Stress / Unload = 1.5 hours

PROCEEDINGS—Institute of Environmental Sciences

147
Table 2 - Experiment 2 Failure Quantities

<table>
<thead>
<tr>
<th>Failure Category</th>
<th>Stress Lab</th>
<th>ESS/ FCT</th>
<th>SAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Components</td>
<td>3 ICs</td>
<td>4 IC</td>
<td>0</td>
</tr>
<tr>
<td>Workmanship</td>
<td>7 Shorts</td>
<td>1 short</td>
<td>0</td>
</tr>
<tr>
<td>Crystals</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Stress Lab Failure Distribution = 8 failures revealed during first and second cycle, 2 failures during the third, and one during the fourth cycle.

Table 3 - Experiment 3 Failure Quantities

<table>
<thead>
<tr>
<th>Failure Category</th>
<th>Stress Lab</th>
<th>ESS/ FCT</th>
<th>SAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Components</td>
<td>15 IC's</td>
<td>5 IC's</td>
<td>0</td>
</tr>
<tr>
<td>Workmanship</td>
<td>9 heat sinks, 1 short</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Crystals</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Stress Lab Failure Distribution = 20 failures revealed during second cycle, 3 failures during the third, and two during the fourth cycle.

Experiment’s Findings and Key Points

The Step Stressing method used to determine the correct random vibration level, proved to be invaluable in finding design related issues. Five design related issues were revealed within one hour of step stressing.

- 3 discrete parts broke off
- 2 loose interconnections

These failures were identified as design robustness improvements. Product team research showed no evidence of failure observed at any customer sites. Additionally, all five failures were redesigned in the next product revision.

Experiment One Conclusions

Failure analysis indicated that not all of the latent failures were removed during 25°C ambient vibration. Failures were seen in the ESS/FCT and the SAT process. These failures proved that for this product a thermal stress was also needed for failure precipitation and elimination.

Experiment Two Conclusions

Experiment two revealed four important findings. First, there were approximately twice as many failures screened as compared to experiment one. Secondly, a power diode was found to be shorting to a via pad on the circuit card during the concurrent thermal and vibration stress process step. This was resolved by applying 4 adhesive bumps under the diode's heat sink during SMT Assembly. This simple fix added less than 0.5 seconds to the SMT process time, but eliminated potential failures. Thirdly, all latent failures were screened out. No failures were seen in the Systems Aging Test process. Lastly, 3 of the component faults observed in the Stress Lab, but dismissed as No Trouble Found's, later failed during the production ESS/FCT process. The point being, that all true intermittent failures exposed during stress screening are real and will most probably fail later in the product life cycle.

Experiment Three Conclusions

This was the most stressful of the three experiments. This experiment revealed yet another doubling of failures. It also precipitated an additional failure type, i.e., improper bonding of the heat sinks to the IC ceramic package. This failure was traced to inconsistent thermal adhesive curing process used by the supplier. Also, as with experiment two, no failures were observed in production's Systems Aging Test process step. It was also found that the greater accelerated stress cycle used in profile three, did not help in defining if a reduction in the number of thermal cycles is possible, but did however result in exposing additional failure mechanisms for this product.

Final Conclusion

The experiment revealed that the sole stress application of random vibration at 25°C is insufficient in removing all the latent failures from this specific product. This fact is supported by the evidence of failures observed in production's ESS/FCT and SAT processes, as shown in Table 1. Additionally, it revealed that the addition of random vibration to the ESS/FCT process would be significant in the removal of all SAT's circuit card failures, as shown in Table 2 and 3. It also proved that step stressing random vibration is a very successful technique for finding design related failures, in a short amount of time. Finally, it shows the value of thermal stress cycle optimization. It is believed that if adequate time and resources are allocated to empirically defining the proper number of thermal cycles and temperatures, that all circuit card functional failures could be captured in a concurrent thermal cycle and random vibration process.
References

Barker, Thomas, B. 1985. Quality by Experimental Design


Acknowledgments

I wish to thank the following USOPS West Coast departments: Product, Component, and Test Engineering for their responsive technical support. Quality Engineering for their statistical guidance. Board production for their material movement support. John Parker for his administrative, and operational help. Finally, thanks to the management team for recognizing the importance of this project and for giving me the freedom to concentrate resources and efforts as needed.