# Keysight Technologies
# ATE System Level Calibration and Its Impact on Cost, Quality and Schedule

White Paper

Lee Nichols, Senior Systems Engineer
Duane Lowenstein, Business Development Manager
Joe LaGrotta, Business Development Manager
Keysight Technologies

This white paper was first published in the
IEEE AUTOTESTCON 2014 proceedings.

KEYSIGHT TECHNOLOGIES
Unlocking Measurement Insights

# Abstract

It has been shown that ATE systems which are targeted at families of products can have a significant positive impact on the cost of test and capacity availability. In addition it has been shown that the quality (accuracy and repeatability) of the test system reduces the overall cost of test and rework. The historical challenge has been that test systems are traditionally targeted at an individual product and thus have acceptance criteria constrained to that individual product specification and not the performance envelope of the tester. This has two major unintended consequences when a test station is intended to be used for more than one product. First, the system specifications are typically developed around a *golden unit* methodology. Second, the specifications are not easily extended to other products of a similar type.

In this paper we will propose an alternative method to provide system level specification and system measurement plane calibration which will extend the system's ability to be used for multiple products. This approach is measurement-centric, not device specification-centric. This philosophy has several major impacts on the cost, quality and schedule.

# I. Introduction

As companies continue to look at ways to minimize costs, reduce schedules and simplify the efforts it takes to introduce new products, new approaches to test and measurement philosophies need to be explored. Over the years many companies have invested in new test strategies from Built-in-Test, to distributed test, to multiple head testers, with success in reducing the cost of capital for test and measurement equipment. Although these have also removed some of the complexities in managing high mix low volume manufacturing test strategies, they still have not addressed one of the fundamental philosophies of test. That is, how do you determine the accuracy of your measurement, is it the capacity of passing a *golden unit* or a metrology verification at the measurement plane.

Though this may seem like a *so what* question, the overall impact on cost, quality and schedule is huge both from the impact on current production but just as much on future production. Getting away from *golden unit* verification and moving to a measurement plane approach allows ATE systems to become product agnostic and measurement-focused allowing a host of advantages. Many of these advantages can be realized through both new product introduction and instrument upgrades in the ATE station due to obsolescence or new measurement capabilities. New products can be designed with a prior knowledge of the system accuracy specification. Instrument upgrades, required over time, will be determined by the system level performance and the historical requirement to recertify Test Program Sets (TPS) will no longer be an obstacle in light of the system level specifications.

The concept of a *measurement plane* for setting the set of specifications will be introduced and considered for typical RF and microwave measurement systems. By logical extension, these concepts can be used for lower frequency systems and for optical test systems. This concept will be illustrated in terms of both vector and scalar calibrations to show that a more robust way to ensure measurement accuracy can be achieved while reducing cost, increasing quality, and shortening schedules. This ultimately leads to greater affordability and better product specification.

# II. Designing A Test System to Cover Families of Products

There are numerous starting points to design a test system for a family of products. Traditionally, one starts with the most complicated UUT (unit under test) and then tests that design against all other UUTs in the family for coverage and applicability. An alternative method is to consider all of the measurements which must be performed on each UUT and then the specific performance limits that must be achieved. By doing this two key questions can quickly be answered. Do the UUTs constitute a proper family for testing? What parameters must be tested and to what accuracy? An organized method to summarize this information is set forth below[1].

In Figure 1, the measurement specifications are noted across the top two lines of the matrix. A generic name is given for each measurement which is independent of the specification and in the second line a worst case specification is given for each measurement. As mentioned before the table identifies the measurement and not the method by which the measurement will be made. As an example, spur search can be performed on a network analyzer or a spectrum analyzer. The optimal choice of which instrument is the best choice can be determined by intuition or by calculations. If the frequency band is extremely broad and it is necessary to find narrow spurs very close to the thermal noise of the test system, a spectrum analyzer (SA) is probably the correct choice. In some circumstances the location of the spurs are known based on the design and there are only a few spurs which must be measured, and the fact that a network analyzer (NA) must be used for some of the other measurements, a NA may be the correct choice.

| Measurement | | Insertion loss/gain | VSWR | 1dB compression power | Saturated output power | Noise figure | Spurious signals | Radiated emissions | Switching speed | Switch command status | Temp sensor | Voltage verification |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Measurement specification/requirement | PNA input powers of -10 to +20 dbm 1 GHz to 18 GHz, 0.1 dB | Reflected powers of 0 to -65dbm | -10 to +27 dbm 0.1 dbm resolution 6 GHz to 18 GHz | -10 to +30 dbm 0.1 dbm resolution 6 GHz to 18 GHz | Min Noise Figure of 1.15 dB 6 GHz to 18 GHz | 100 Khz Res BW, 10 KHz Video BW -80dbm, 1 GHz to 18 | -60 dbm, 1 GHz to 18 GHz | TTL logic levels Speed > 1 usec RF Levels -5 to +20 dbm | 0-5 V, mV increments | 0-5 V, mV increments | 0-30 V, mV increments |
| DUT 1 | Gross tune | X | X | | | X | | | | | | |
| | Integration test | X | X | | X | | | | X | | | |
| | Final test | X | X | | | X | X | X | | | | |
| DUT 2 | Preliminary test | X | X | X | | X | | X | | X | | X |
| | Post vibe | X | X | | | | | | | | | |
| | Final test | X | X | X | | X | | X | | X | X | X |
| DUT 3 | Preliminary test | X | X | | X | | | X | X | | | |
| | Final test | X | X | | | | X | X | | | | |
| DUT4 | Preliminary test | X | X | | | X | | | | | | |
| | Final test | X | X | | | X | | X | | X | | X |

Figure 1. UUT vs. measurement requirements.

Historically, intuition has played a large role in the instrument selection process. For engineers new to test or new engineers, a more mechanistic approach may be in order. This procedural approach uses additional information about the test time for each measurement for each UUT and possible choices of measurement instruments. In most cases this is a small subset of the measurement required; but, it is of significant financial value. Test time is used as the differentiator because it is directly related to the cost of test for the UUT. The estimate for time can then be used to decide if an additional instrument is worth the capital investment over the useful life of the system. For example, if a measurement requirement is for an extremely long spur search, the capital investment in a SA and the supporting infrastructure (matrix, rack space, software complexity, calibration and selftest, etc.) will usually pay for themselves.

At this point, one has a set of tests, a dedicated set of possible instruments to make the measurements and estimated test times. The selected instruments and test times can be factored into the results of Figure 1. This integration is shown in Figure 2. It should be noted that the measurement specification are still maintained in the Excel spreadsheet but are hidden for clarity. Now, one can quickly estimate the total test time required by each UUT and a preliminary cost of test can be calculated.

| Measurement | | Demand | Insertion loss/gain | VSWR | 1dB compression power | Saturated output power | Noise figure | Spurious signals | Radiated emissions | Switching speed | Switch command status | Temp sensor | Voltage verification |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Test asset | 2 port PNA | 10638 | N5245A | N5245A | | | | | N5245A | | | | |
| | Noise figure meter | 1630 | | | | | N8975A | | | | | | |
| | Spectrum analyzer | 756 | | | | | | N9030A | | | | | |
| | Synthesized source | 2082 | | | HP83650L | HP83650L | | HP83650L | | HP83650L | | | |
| | Power meter | | | | X | X | | | | | | | |
| | Power supplies | 17956 | E3631A E3645A | E3631A E3645A | E3631A E3645A | E3631A E3645A | E3631A E3645A | E3631A E3645A | E3631A E3645A | E3631A E3645A | E3631A E3645A | E3631A E3645A | E3631A E3645A |
| | Pulse generator | 1040 | | | | | | | | X | | | |
| | Multimeter | 4701 | | | | | | | | | E34401A | E34401A | E34401A |
| | Oscilloscope | 750 | | | | | | | | X | | | |
| | Total test times | | 846 | 2439 | 516 | 60 | 1340 | 466 | 1083 | 750 | 1851 | 510 | 1825 |
| DUT 1 | Gross tune | 10 | 3 | 12 | | | 10 | | | | | | |
| | Integration test | 10 | 3 | 12 | | 13 | | | | 10 | | | |
| | Final test | 10 | 3 | 12 | | | 10 | 12 | 9 | | | | |
| DUT 2 | Preliminary test | 15 | 3 | 12 | 15 | | 10 | | 9 | | 22 | | 25 |
| | Post vibe | 15 | 3 | 12 | | | | | | | | | |
| | Final test | 15 | 3 | 11 | 15 | | 10 | | 8 | | 25 | 30 | 25 |
| DUT 3 | Preliminary test | 18 | 5 | 11 | | 18 | | | 8 | 23 | | | |
| | Final test | 18 | 5 | 11 | | | | 14 | 8 | | | | |
| DUT4 | Preliminary test | 14 | 5 | 11 | | | 10 | | | | | | |
| | Final test | 14 | 5 | 11 | | | 10 | | 7 | | 24 | | 25 |

Figure 2. Test time and Instrument selection

The particular UUTs chose for this paper do not require testing at temperature (hot, cold and ambient). Using the same format, additional tests could have been added to the matrix to represent temperature testing. Mock estimates for ramp and soak times would then allow the designer to consider the impact of these elements and their impact on asset utilization problem in an expanded light. When the combination of test time and soak time approaches the test time then alternative dual fixturing should be considered. For example if the total test time is 2 hours and the combined soak and ramp time for each transition approach 2 hours a dual fixture strategy may have a significant impact on the total asset utilization and the cost of test. This strategy makes better use of the more expensive instruments in the system with a small investment in additional infrastructure. Typically the following elements must be duplicated or expanded: power supplies, fixtures, matrix, temperature environment and temperature controllers. The cost of these elements is typically small for an RF test system when compared to the cost of the primary RF measurement equipment.

Now the accuracy of the system, as opposed to the accuracy of the individual instruments can considered. Many times, engineers use the instrument accuracy as a proxy for the system accuracy. In some cases this estimate is reasonable and in others it is not. The cases where the assumptions fails most often is when a large switch matrix is use to connect the test instruments to the UUT. The uncertainty in the matrix, additional cabling, switches and connections, increase the measurement uncertainty beyond acceptable limits. In section *III. Establishing Measurement Fidelity Using Product Independent Calibration*, we will consider the system level calibration and uncertainty calculation.

# III. Establishing Measurement Fidelity Using Product-Independent Calibration

The idea of a *measurement plane* is an extension of the idea of a *reference point* to an entire set of configured stimuli and/or measurements implemented for the purpose of testing a family of products. In the common scenario of applying an RF stimulus (perhaps a local oscillator) to a UUT, it is apparent that the signal generator output power will only resemble its configured value at the output port of the instrument. In an automated test system, there will typically be significant power loss through the cable and connection infrastructure on the way to the UUT. This system loss must be taken into account when configuring the generator. In the simplest analysis, one must set the generator's power level (in dBm) to $P_{gen} = P_i + L$, where $L$ is the system loss in dB determined via system calibration and $P_i$ is the desired stimulus power (in dBm) to the UUT. In this scenario, the UUT port into which the RF stimulus is injected may be considered the reference point. Using the measured system loss, one relates configuration of a stimulus at another point in space to the specified stimulus at the reference point. If the system loss is stored in a lookup table and retrieved by driver-level routines in order to configure $P_{gen}$ appropriately, then the specified stimulus can be set at the reference point by a test developer without regard to the details of this process (sometimes called *normalization*); it becomes a calibrated system stimulus at the point of interest.

One can think of establishing a reference point for a power measurement in a similar manner. Whereas a bench top RF power meter or spectrum analyzer might be used to determine the output power from a device by hooking the device directly to the measurement instrument, in an automated test system, one must establish a reference point at the end of a combination of cables and interfaces which incur loss $L$. The measured power must be corrected as $P_o = P_{meas} + L$ where $P_o$ is the calibrated output power at the point of interest and $P_{meas}$ is the uncorrected instrument measurement. Again, the system losses would be stored in a lookup table. Of course, in a test system designed to address the qualification needs of more complicated devices, there will be multiple stimuli and measurement reference points, adding to the dimensionality of the lookup table. From this notion, we generalize the idea of a reference point to a reference plane or measurement plane so that an automated test system can be thought of as a single calibrated instrument with a set of capabilities at that plane.

If power and loss values could be determined exactly under all relevant conditions, a sufficiently large correction table would theoretically be sufficient. Practically, we know that the measurement fidelity (or accuracy) at the plane must be quantified in terms of the uncertainty of the measurement. In a system designed to measure a family of products, one source of uncertainty is the finite set of frequencies used to characterize the system: the lookup table contains loss values over this set of frequencies and the driver that retrieves the calibration value for a specific frequency of interest may do so using an interpolation algorithm. Depending upon the nature of the connectivity infrastructure, good engineering judgment has to be applied in determining which frequencies need to be calibrated and which do not; for instance, losses at frequencies below a few MHz may not be variable enough to require many calibration points versus spectral regions having greater variation considering the required accuracy of measurement. Also, considering the design of UUTs, RF power measurements may be considered acceptable with lower test accuracy ratios (the ratio of the test limit range to the $\pm k\sigma$ accuracy of the measurement, where $\sigma$ is the standard deviation and $k$ tends to be 2 or 3) at least over some range of frequencies. If there are a large number of UUT interfaces, over such frequencies, one can save time and cost (in process complexity) by not applying rigorous RF calibration to all of them.

Power linearity may also be a factor in the uncertainty of a measurement. If there is a nonlinear characteristic to the signal source, there is an error associated with assuming that the generator will emit 1 dB more power when it is configured to $P_{gen}$ + 1; this uncertainty can be reduced by treating the specified power as an independent variable similar to the treatment of frequency. Again, there will still be an uncertainty component associated with interpolation between power settings. In general, one wants to have normalization implementations that utilize data acquired over a multi-dimensional matrix of defined parameter values (frequency, power, temperature) and interpolate for values in between. Thus, common calibration datasets and software implementation can be used for a variety of UUTs and the implementation can be leveraged even across multiple system designs if isolation is maintained between the calibration drivers and the test sequences. The key benefit is that one isn't developing catered calibration procedures and datasets for specific UUTs with each new UUT requiring additional development time and cost. The datasets reflect the variability of the test system, not the variability of the UUTs.

Now, as mentioned previously, the notion of establishing power at the measurement plane via adjustment of a generated or measured power level by a path loss value is an oversimplification. As we know from microwave theory, the power delivered to an input of UUT A (or, for instance, calibration device A) from a signal generator within the ATE will not be the same as that delivered to an input of UUT B unless the two inputs share the same complex input impedance $Z = R + jX$, where $R$ and $X$ are the resistance and reactance, respectively. In general, some portion of the incident power at the measurement plane will be reflected at the interface according to the standard formula for the complex reflection coefficient $\Gamma_l = (Z_l - Z_0)/(Z_l + Z_0)$, where $Z_0$ is the source impedance looking back into an ideal generator. With a well-defined impedance mismatch, the delivered power is $P = (1 - |\Gamma_l|^2) P_0$, where $P_0$ is the generator's output into a matched load.

However, this is still oversimplified: we have to replace the ideal generator's output with an incident power $P_{inc}$ which is a function of $P_0$ as well as both the source and load impedances which are not in general equal to $Z_0$. In fact [2], a more complete analysis using signal flow graphs shows that $P = (1 - |\Gamma_l|^2 )P_0)/|1 - \Gamma_s \Gamma_l|^2$ where $Z_s$ is the true source impedance looking back into the generator. While it may seem strange that the incident power would be a function of the source reflection coefficient, one has to keep in mind that the steady-state signal flow is a product of many reflections back and forth between the source and the load (it may help to think of inserting a lossless length of $Z_0$ between the two). The ratio $P/P_0$ characterizes a mismatch effect that can reduce or even increase the power delivered to the load (that is because $P_0$ is not in general the *maximum available power* from the generator, which is the power delivered to a conjugate-matched load[2]), depending on the relative phase and magnitudes of the coefficients.

Unfortunately, sources and loads are not often characterized as regards their exact impedance at any given frequency; instead, manufacturing test may characterize only their magnitudes. Since the numerator of the expression for $P/P_0$ is phase-independent, this well-characterized loss factor is usually termed *mismatch loss*. In contrast, the denominator will vary from $(1 + |\Gamma_s||\Gamma_l|)^2$ to $(1 - |\Gamma_s||\Gamma_l|)^2$ over a range of frequencies and is typically referred to as *mismatch loss uncertainty*. This is a potential discrepancy between the power level measured during the calibration procedure and that measured with the UUT even if the calibration sensor and the UUT share the same magnitude of reflection coefficient $|\Gamma_l|$ and must be taken into account in the overall error analysis of the system. For example, a typical microwave signal generator with a standing-wave ratio (SWR) of 1.6:1 has $|\Gamma_s| = (SWR - 1)/(SWR + 1)$ of 0.231 and a typical microwave power sensor might have SWR of 1.23:1 has $|\Gamma_l|$ of 0.103. Thus, the denominator is in the range of 1 ±5% and the delivered power varies by ±0.2 dB.

There are several ways to combat mismatch loss uncertainty in one's test system design. One is to attack the problem of mismatch reflections directly by introducing components with attenuating properties in one or both directions. In microwave network theory, such a device is one whose $S_{12}$ or $S_{21}$ parameter (the scattering parameter that governs the transference of power in the reverse or forward direction, respectively) is significantly less than unity. Some obvious examples would be a magnetic RF isolator or an amplifier, but a less expensive, more practical option is a simple attenuator. In return for a potentially-significant but well-characterized power loss hit, the component ensures that reflected signals are quickly attenuated, thereby reducing the effective magnitudes of $|\Gamma_l|$ or $|\Gamma_s|$.

A more expensive, but very effective way to reduce the mismatch component of system calibration uncertainty is to introduce a means of characterizing the phase as well as the amplitude of the generated or measured signals. This is often termed a *vector* calibration method. This is commonly achieved by integrating a vector network analyzer into the test system itself. The architecture of a vector network analyzer supports the characterization of relative power measurements with regard to phase using a combination of high-directivity couplers and narrowband digital IF sampling. When combined with sophisticated RF signal switching networks to add channel capacity without introducing too much additional measurement uncertainty, overall system calibration accuracy can be improved relative to the simple *scalar* calibration methods.

Part of establishing system measurement accuracy is ensuring the integrity of the calibration data: is the data acquired in a consistent manner; is the acquired data applied correctly. While it is natural to expect measurements to be repeatable in the course of production testing, the calibration process is often overlooked as a candidate for automation. Manual calibration procedures targeting a limited set of tests for a single product may be revered as a means to achieving the highest accuracy possible in the generation of the requisite lookup tables, but the possibility of human error and the time required to perform the procedure call into question the practical value of the accuracy achieved. The cost of developing automated calibration routines that address the entire measurement plane rather than the ports utilized by a single product can be offset by the elimination of errantly-performed procedural steps, the broader applicability of the data acquired (in a shorter time) to multiple products, and the leveragability of the implementation to other system designs.

The critical component of a product-independent and repeatable calibration procedure at the measurement plane is some type of calibration device that can be connected directly to the measurement plane. Unlike a particular UUT, this device has mating interfaces for all test ports and is capable of switching (repeatable and with low loss) external signals to/from embedded sensors and/or reference oscillators. In support of certain kinds of measurements, it may even incorporate a noise source. This calibration device is preferably DC-powered and uses compact remote interfaces like USB or LAN to receive control instructions and return measurement data from and to the test system. With detailed planning, it can be designed to work in the context of the same detachable test fixture used by a set of UUTs. Critical attention must be paid to the system error analysis when designing the calibration device and how it is utilized in the system calibration process. Since the device itself must be calibrated for internal losses, the overall accuracy of any system-level measurement becomes a function of the accuracy of its characterization (offline by a network analyzer, for instance) as well as the accuracy of the system instruments used in the calibration process.

In order to ensure that the acquired calibration data is applied correctly, the handling of calibration data can be enhanced through the use of *smart* calibration devices and even *smart* fixtures. Rather than storing characterization data for the calibration device or the test fixture in the test system controller, the data can be stored within memory which is embedded within the device or fixture itself. The advantage of this approach is the reduced likelihood of using an incorrect dataset in an environment with multiple calibration devices and/or fixtures. A further enhancement regardless of where the calibration data is stored is the retention of historical calibration data for use in trending. Analyzing loss values over time can highlight test system component failures, simplifying troubleshooting. It can also be used retroactively to increase system utilization by justifying longer calibration intervals for some or all measurement paths.

## IV. Developing and Verifying System-Level Specifications For Cross-Program Utilization

In the last section, we have tried to justify and outline methodologies for treating an integrated ATE as a single instrument with system-level specifications at the measurement plane. This is more than just a conceptual novelty; it is a strategy for promoting the use of test systems across product families and multiple programs. Such cross-program utilization can have a measurable positive impact on cost-of-test. However, affecting this strategy requires the ability to communicate the true system-level capabilities/specifications in a consolidated form. In part, this consists of documenting the ways in which the integrated system enhances or augments the capabilities of the bench top instruments. For instance, a digitizer connected to the measurement plane via a multi-port RF switch with high-power attenuators on some of the ports can be described generically as a quantity of non-simultaneous high-speed signal analysis channels with calculated per-channel maximum input power among their specifications. However, it is also important to acknowledge the ways in which the system constrains the capabilities of the bench top instruments: the losses within the cables, the switch, and the attenuator effectively increase the input noise floor of the channels so that the overall system dynamic range may be reduced. The details behind the measurement plane are immaterial as are the details of the products currently being tested on the station; what matters to the prospective user of the ATE is the measurement capability at the measurement plane.

In addition to analyzing parameter range specifications such as those just described, it is important to properly assess the effect of the entire system on per-channel measurement uncertainty. The discussion in the previous section of mismatch uncertainty is a valid approach to error analysis for a single interface. Unfortunately, an ATE tends to have numerous interfaces at the inputs and outputs of multiple switches and possibly a mass interconnect or even active components between the measurement plane and the signal source or measurement device. A complete error analysis has to consider all of the potential mismatches occurring at every cable connection point in between. Each interface potentially introduces reflections that can combine in or out of phase with the forward-propagating wave. The overall reflection coefficient looking back toward the source can be a complicated and, to some degree, unknown function of the coefficients at each interface, thereby increasing the overall uncertainty.

Consider a slightly more complicated signal path consisting of a generator followed by a cable, a connector, another cable, and a load. Assume that the components are lossless, that the cables are of the characteristic impedance, and that the connector can be modeled using the following simple S-matrix

$$\begin{bmatrix} \Gamma & 1 \\ 1 & \Gamma \end{bmatrix}$$

where $\Gamma$ is the reflection coefficient at the connector and we have taken the transmission coefficients to be unity since we are more interested in the impact of reflection. For this analysis, we will consider each reflection rather than trying to determine steady-state conditions directly. The generator launches power $P_0 = |V_0|^2/2Z_0$ into the first cable (not the steady-state power incident at the connector, but the power the generator would deliver to a matched load). This primary wave propagates along the first cable and then encounters a mismatch at the connector which generates a reflection of amplitude $\Gamma V_0$ back toward the source, where it experiences another reflection. This secondary forward-propagating wave has amplitude $\Gamma_s \Gamma V_0$. When it superimposes with the original wave, it results in a combined amplitude incident at the connector $(1 + \Gamma_s \Gamma e^{j\theta_1})V_0$. The phase factor accounts for the round-trip delay of the secondary wave relative to the primary wave. Now, of course there will be a tertiary reflection as well, but we will ignore this as a second-order effect.

Our combined wave propagates onward toward the load where yet another reflection occurs, generating a back-propagating wave of amplitude $\Gamma_l (1 + \Gamma_s \Gamma e^{j\theta_1})V_0$. However, we can assume that the magnitude of all of the reflection coefficients is much less than unity so that we may approximate the amplitude as $\Gamma_l V_0$. When this wave encounters the connector, it generates a near reflection, but the (backward) transmitted wave continues back to the source, generating a far reflection. All of these combine at the load to form an incident wave of amplitude $(1 + \Gamma_s \Gamma e^{j\theta_1} + \Gamma_l \Gamma e^{j\theta_2} + \Gamma_s \Gamma_l e^{j\theta_3})V_0$. The incident power $P_{inc}$ is then this amplitude multiplied by its complex conjugate, $P_0$*(1 + 15 terms)*. Rather than compute the various products of reflection coefficients and time delay factors within the additional terms, let us recognize first of all that nine of these terms are products of four reflection coefficients and are therefore less significant. We can show that the remaining six terms combine to $2\mathrm{Re}[\Gamma_s \Gamma e^{j\theta_1} + \Gamma_l \Gamma e^{j\theta_2} + \Gamma_s \Gamma_l e^{j\theta_3}]$.

Without knowing the phase of the reflection coefficients or the time delay factors, we can see that the incident power varies roughly between $P_0(1 \pm 2|\Gamma_s||\Gamma| \pm 2|\Gamma_l||\Gamma| \pm 2|\Gamma_s||\Gamma_l|)$. Note that this is analogous to our derivation of the mismatch uncertainty for a single interface where the maximum value of $P_{inc}/P_0$ is $1/(1-|\Gamma_s||\Gamma_l|)^2 \approx 1 + 2|\Gamma_s||\Gamma_l|$ for coefficients much less than unity. Thus we see that, in the worst-case analysis, reflections from each interface have an additive effect on the overall power uncertainty of the signal path. It should be apparent that this type of worst-case analysis (assuming all reflections are in phase) can result in some pretty horrendous estimates for the measurement uncertainty associated with a more complicated test system signal path. A more practical approach is to assume that the phases of the reflected waves are randomly-distributed and that therefore the path signal amplitude uncertainty can be estimated as a root-sum-square (the square root of the sum of the squares) of the amplitude uncertainties at each interface. In such an RSS analysis, overall uncertainty will be roughly proportional to the largest individual uncertainties present in the path.

When applying error analysis to an entire system in this manner, it is important to note what this actually means in terms of the performance of multiple test systems of the same design. Suppose that the error analysis of a particular path within the test system design leads to an expected 2s uncertainty of ±1 dB. Does this mean that a proper build of the ATE with quality parts cannot result in performance outside of this range? No. As is the case with any normally-distributed performance parameter, outliers can exist even with consistent production methods. Although they are less likely, it is still important to plan for the possibility that they will occur; recall from elementary probability that one can use the binomial distribution with p-values integrated over the normal distribution to estimate the likelihood of an outlier (for a single path) within a small volume of units. Taken into account all the paths within the system and you begin to see that there is significant probability for a few under-performing paths across the entire system. One must prepare for this likelihood with alternate parts strategies.

Does an uncertainty of ±1 dB imply that measurements from a single system will vary substantially on the order of 1 dB? No. It is important here to distinguish between accuracy and repeatability as components of overall uncertainty. For a signal at a particular frequency, power level and temperature, the relevant physical characteristics of the components in the path are unlikely to be changing dynamically to any significant degree. One can expect a particular system to have significantly better repeatability than the uncertainty derived from the error analysis. The derived uncertainty is more likely to manifest across multiple systems or multiple paths that share the same component characteristics. This is why it can be misleading to use a *golden UUT* as the standard for verifying the performance of an ATE: the repeatability of a test result from a single ATE doesn't really demonstrate the accuracy of the test system. To put it another way, it is possible to get a very repeatable result which is wrong (if the normalization software retrieves the wrong correction value, for instance). It is better to use an external instrument such as a signal generator or analyzer with NIST-traceable calibration to verify system-level specifications and to do so across multiple systems if possible in order to get a better understanding of the variability. One has to keep in mind of course that a set of three test stations doesn't constitute a particularly significant sample size statistically, but it is significantly better than one.

As a final comment, it is worth pointing out that a complete test system development strategy should include an automated self-test for regular checking of the test readiness of the ATE. The purpose of this self-test, which can utilize a self-test device that connects to the measurement plane, is not to verify system-level specifications, but simply to ensure basic connectivity. It is intended to root out the kinds of ongoing system integration issues that can occur when, for instance, instruments are returned from calibration, but the integrator forgets to restore all system cable connections. It is not intended to be exhaustive performance verification. However, the self-test design should ideally cover the entire measurement plane so that operational readiness is ensured for all products and programs.

# V. Identify Measurement Parameters for Common Families

After completing the system hardware, calibration and verification designs, one needs to consider the impact of all elements including the UUT measurement on the software architecture. Traditionally, software design was done for a specific device with the parameters of the test hard coded or passed as arguments to subroutines. This process is counterproductive for two reasons: It makes the resulting code difficult to maintain and it spawns variations to support additional UUTs.

In the software architecture of a test system a number of high level design elements should be consider at the outset.

Those elements are as follows:

- An overall test executive with the ability to support conditional branching, test timing, a debug process, external parameter control and external data base hooks.
- A User Interface design (as simple as go, stop buttons to a full feature debug environment).
- Data collection and storage for UUT tests, System Calibration results, System verification results, System maintenance and instrument calibration.
- Resource allocation of specific assets (instruments, matrix paths ...) to support multi up UUT testing.
- Spawn multiple concurrent processes for the case of simultaneous testing of two up and multi up testing.
- Results reporting to a common database.
- Common Data formats for scalar, vector, test parameters and results. These formats should be supported by the analysis and reporting software infrastructure.
- Interface to factory control system for status reporting and coordination.
- Data analysis software which has an interface to the common data base.

Historically, each test system designer and software architect has developed unique solution to each of these elements. There are significant economic advantages that can be achieved by using a common approach to all of these elements. Clearly, the cost saving received from not recoding each of these elements in a different way can be significant. But the recoding cost pales in comparison to the cost of maintaining multiple implementations of the same functions by different methods.

Indirect advantages can also be achieved by this commonality. For example the comparison of test results for a single UUT has been possible in most historical architectures. The proposed approach allows an engineer to do comparisons work across a family of UUTs for a single measurement. This comparative work may not be focused on the UUTs relative performance but on the systems performance to make a particular measurement. Also, the same concept can be extended to multiple systems with comparative analysis extended across multiple systems.

One capability would also be to store system calibration and selftest result in the database. This method opens the possibilities of tracking calibration changes for a single system over time and self-test results for a single system and or across multiple systems. These analyses can be used to identify rogue systems, bad connectors, and bad cables and to refine the performance envelope of the architecture.

It is clear that the system software and system performance specifications must be independent of a UUT. Reliance of golden units to assess system performance is a very comfortable concept for a given UUT and for a given program. The software architecture can support such a philosophy and it also opens the door to comparative studies of *golden UUT* across multiple systems for correlation between systems. Although very useful, golden UUTs do not satisfy the objective to re-use of a system and the associated measurement libraries.

The software implement must support the measurement plane concept. Two alternatives are available for implementations: 1) The measurements can incorporate the correction factors from the calibration process 2) Raw measurements can be recorded and then post processed by a *calculation* test which immediately follows the data collection. Both options are support by commercial test executives and both have advantages and disadvantages. Ultimately this is a decision that has far reaching implications for the future. The advantage of storing raw data opens the possibly to re- correct using new calibration data or calibration techniques but is burdened by computational overhead when comparing results for large populations of measurements. Some designs have stored both raw and corrected data which takes advantage of both formats at the cost of doubling the storage requirements.

The software test environment design decisions are in the long term the most important and highest cost impact. Although RF test equipment is expensive, its cost is small relative to the support and maintenance costs associated with the software environment. The assertion that software is a very significant cost driver is supported by the US Navy's procurement for a CASS replacement platform eCASS. During that procurement, the Navy placed significant emphasis on TPS compatibility with the legacy CASS systems.

# VI. Cost, Quality and Schedule Impacts

The advantages of a measurement plane verification philosophy can easily be understood but, the paradigm will take some debate. There should be little debate on the advantages, the biggest is being cost closely followed by quality and schedule. The capability to have ATE that is designed for a measurement rather than an individual product's specifications would allow greater, utilization flexibility, and re-use. These elements make up the basis for reducing cost for an enterprise test strategy. The most important is utilization, having assets that are being utilized > 95% will almost always be a good investment assuming they are both effective and efficient. To be able to use the assets more, the ATE needs to be flexible in their capability, especially in areas that a single ATE system cannot be utilized by a single product 100%. Therefore, as described in section *II. Designing A Test System to Cover Families of Products*, the capability to find commonality for multiple products on common unified ATE platform becomes imperative. And to ensure even greater return on investment the ability to re-use all or some of the ATE (both hardware and software) for future products is a bonus. This bonus is most important to schedule, and quality and the formulation of new program bids.

The capability for re-use is the foundation for advantages in schedule and quality. It is simple to see that is an ATE system already exists that has 60-70% of the capability that you already need for a new product, that means up to 60-70% of the time to design a new ATE could be eliminated. Though it probably not a one to one correlation it is something more than zero. It also allows earlier ability to verify a design on the production test sets while still in design to ensure adequate coverage, precision and accuracy.

The last element, quality wraps up the greatest holistic value to the approach described in this paper. The quality of the ATE system used in manufacturing is its ability to produce repeatable and accurately measure specific specification. Even though it is made up of multiple individual instruments and lines of code that should have the capability, the only important aspect is that the interface accuracy between the UUT and ATE for a specific measurement. The process of re-use and system level calibration directly address this need. Measurement plane design re-use allows the capability to verify an ATE's hardware and software for accuracy and repeatability across multiple UUTs therefore statistically ensuring measurement fidelity through UUT independence. System level calibration elevates the ATE to the fundamentals of a test strategy. That is, a test strategy is to verify a specification or a capability of a UUT, and it's not how to assemble a bunch of hardware and software. Using system level calibration moves the implementation of test strategy to focus on agnostic measurements that ultimately leads to greater re-use, better utilization and fundamentally higher fidelity and quality results.

## VI. Conclusion

Test system standardization is an economic, emotional and technical challenge. Many factors work to undermine the implementation of a unified approach to testing across an enterprise. Compelling arguments, leadership and intestinal fortitude are a prerequisite for successfully transforming an enterprise from a product focused test strategy to an enterprise focused strategy.

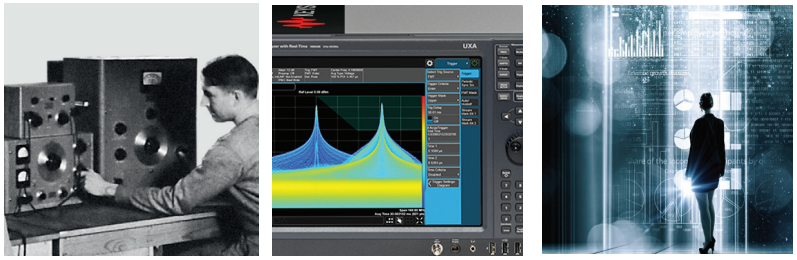
## VII. References

[1] Duane Lowenstein, Keysight Technologies and Joe Lagrotta, Keysight Technologies, *A holistic approach to reducing the life cycle costs of test*, IEEE AUTOTESTCON 2013

[2] D. Pozar, *Microwave Engineering*, Addison Wesley, Reading, MA, 1990.

## Evolving

Our unique combination of hardware, software, support, and people can help you reach your next breakthrough. We are unlocking the future of technology.



From Hewlett-Packard to Agilent to Keysight

myKeysight

**myKeysight**
**www.keysight.com/find/mykeysight**
A personalized view into the information most relevant to you.

KEYSIGHT SERVICES
*Accelerate Technology Adoption.*
*Lower costs.*

**Keysight Services**
**www.keysight.com/find/service**
Our deep offering in design, test, and measurement services deploys an industry-leading array of people, processes, and tools. The result? We help you implement new technologies and engineer improved processes that lower costs.

STANDARD 3 YR WARRANTY

**Three-Year Warranty**
**www.keysight.com/find/ThreeYearWarranty**
Keysight's committed to superior product quality and lower total cost of ownership. Keysight is the only test and measurement company with three-year warranty standard on all instruments, worldwide. And, we provide a one-year warranty on many accessories, calibration devices, systems and custom products.

CALIBRATION UP TO 10 YEARS WARRANTY
ASSURANCE PLANS

**Keysight Assurance Plans**
**www.keysight.com/find/AssurancePlans**
Up to ten years of protection and no budgetary surprises to ensure your instruments are operating to specification, so you can rely on accurate measurements.

**Keysight Channel Partners**
**www.keysight.com/find/channelpartners**
Get the best of both worlds: Keysight's measurement expertise and product breadth, combined with channel partner convenience.

www.keysight.com

For more information on Keysight Technologies' products, applications or services, please contact your local Keysight office. The complete list is available at:
**www.keysight.com/find/contactus**

### Americas

| | |
|---|---|
| Canada | (877) 894 4414 |
| Brazil | 55 11 3351 7010 |
| Mexico | 001 800 254 2440 |
| United States | (800) 829 4444 |

### Asia Pacific

| | |
|---|---|
| Australia | 1 800 629 485 |
| China | 800 810 0189 |
| Hong Kong | 800 938 693 |
| India | 1 800 11 2626 |
| Japan | 0120 (421) 345 |
| Korea | 080 769 0800 |
| Malaysia | 1 800 888 848 |
| Singapore | 1 800 375 8100 |
| Taiwan | 0800 047 866 |
| Other AP Countries | (65) 6375 8100 |

### Europe & Middle East

| | |
|---|---|
| Austria | 0800 001122 |
| Belgium | 0800 58580 |
| Finland | 0800 523252 |
| France | 0805 980333 |
| Germany | 0800 6270999 |
| Ireland | 1800 832700 |
| Israel | 1 809 343051 |
| Italy | 800 599100 |
| Luxembourg | +32 800 58580 |
| Netherlands | 0800 0233200 |
| Russia | 8800 5009286 |
| Spain | 800 000154 |
| Sweden | 0200 882255 |
| Switzerland | 0800 805353 |
| | Opt. 1 (DE) |
| | Opt. 2 (FR) |
| | Opt. 3 (IT) |
| United Kingdom | 0800 0260637 |

For other unlisted countries:
www.keysight.com/find/contactus
(BP-2-23-17)

DEKRA Certified
ISO9001 Quality Management System

**www.keysight.com/go/quality**
Keysight Technologies, Inc.
DEKRA Certified ISO 9001:2015
Quality Management System

KEYSIGHT TECHNOLOGIES
Unlocking Measurement Insights


Published in USA, March 20, 2017
5992-2234EN
www.keysight.com