

Implementation of ATML Test Results on the USAF Versatile Depot Automated Test Station

Kevin Coggins¹, Mark Skiba²

¹402 Software Maintenance Group, United States Air Force, Robins AFB, GA 31098, kevin.coggins@us.af.mil

²Vektrex, 10225 Barnes Canyon Road A213, San Diego, CA 92121, mskiba@vektrex.com

Abstract – The VDATS (Versatile Depot Automatic Test Station) is a Department of Defense (DOD) approved Family of Testers for the Air Force. Conceptualization began in 2006 and VDATS became a reality shortly after with the first article tester demonstrated at IEEE AUTOTESTCON 2007. Forward looking in their approach, the USAF VDATS team made the determination early in the program to employ the DOD ATS Framework where possible to standardize both hardware and software interfaces. ATML (Automatic Test Markup Language) is emerging as an IEEE standard that defines a universal framework for ATS software interfaces. ATML Test Results supports the storage of results of an individual Test Program Set (TPS) execution cycle in a standardized format utilizing Extensible Markup Language (XML). ATML shows great promise toward reducing support and maintenance costs and improving interoperability. The VDATS program has enabled the Air Force to rehost scores of test programs from unsupported legacy test systems, each containing test report formats unique to each legacy test system. Incorporating a universal model for test data capture and reporting is one a key benefit of ATML Test Results. In addition, for VDATS, the Air Force is seeking to maximize the value of ATML Test Results by taking advantage of the inherent analysis capabilities, including the detection of problem areas in test systems, TPS design optimization, and monitoring of UUT (Unit Under Test) component failure rates. The VDATS ATML Initiative leverages the SBIR program to integrate ATML Test Results into VDATS and develop some useful analysis tools. This paper will present an overview of the VDATS software, Test Results implementation, key decisions, and actual test and analysis results based on real UUTs.

Keywords – VDATS, ATML, Test Results, Versatile Depot Automated Test Station, DOD ATS Framework, OSS&E.

I. INTRODUCTION

The accurate capture, evaluation and recording of test results on an automated test system (ATS) are vital steps in the maintenance and sustainment of weapons systems. The USAF has introduced the Versatile Depot Automated Test Station (VDATS) as an official DOD Family of Tester in 2007¹. VDATS, as the Air Force ATS of choice, is a key component of effective weapons systems sustainment as required by DOD Instruction 5000.2². IEEE introduced IEEE Std. 1636.1-2007 to “provide for a standard format for the transport or storage of both quantitative (measured values) and qualitative (pass/fail determination) test results.”³

Historically, the format of the majority of test reports is text-based, and lack standardization between not only test

environments but also test systems. These text-based test results formats do not segregate or isolate the individual data elements, effectively eliminating the possibility of commonality or interoperability. Custom parsers have to be developed to extract the test results in order to achieve a context mapping. A section of a text-based test report is shown in Fig. 1.

```
--- Running Test Block 4 Powerup
Test Name      Measured  High Limit  Low Limit  Units
-----
PASS 4.1 +15.0 VDC Voltage  15.001    15.1      14.9      VDC
PASS 4.2 +15.0 VDC Current  0.2684    1.0       NA        A
PASS 4.3 -15.0 VDC Voltage  -15.0034  -1.0     -15.1     VDC
PASS 4.4 -15.0 VDC Current  0.1142    1.0       NA        A
PASS 4.5 + 5.0 VDC Voltage  4.9831    5.1       4.9       VDC
PASS 4.6 + 5.0 VDC Current  0.0072    3.0       NA        A
--- Completed Test Block 4 Powerup  Status: Passed
```

Fig. 1. Text-Based Test Report

Creating the test report in the XML-based ATML provides semantic meaning to the report elements; in other words, it allows a numerical representation to have a meaning, and even attributes such as data type and measured characteristic. The ATML schema also provides structure. The combination of structure and semantics allows for the interpretation of the test results data by automated tools and processes, which are discussed in this paper.

Vektrex was tasked via an SBIR⁴ to develop the capability to allow the VDATS system software to capture test results compliant with IEEE Std. 1636.1-2007 and to develop some initial proof-of-concept analysis tools. This paper demonstrates how ATML Test Results are being integrated into the VDATS, and describes how the analysis tools provide real-world benefits. Some issues with the trial-use schema that were discovered during the implementation on VDATS are also presented.

The VDATS system was developed prior to the creation of the ATML Test Results standard. To add support, an ATML TestResults “plug in” with supporting framework components was developed by Vektrex and integrated into the existing architecture by the 402 Software Maintenance Group.

II. SYSTEM OVERVIEW

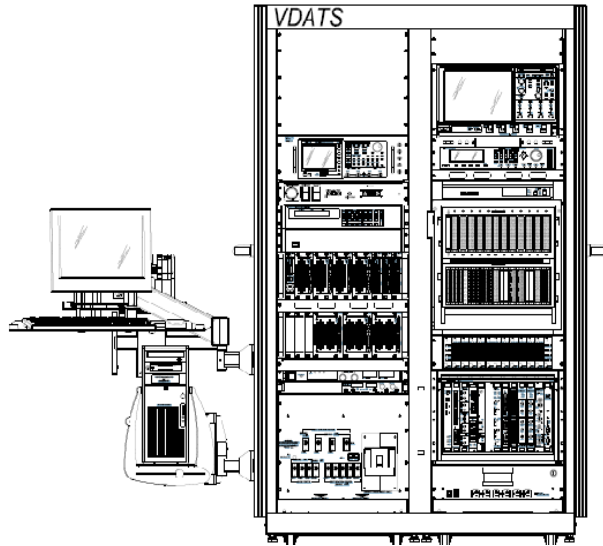


Fig. 2. VDATS System

An integrated plug-in architecture was selected to provide the support needed for existing VDATS TPSs to generate ATML TestResults compliant XML files while at the same time shielding the operator from ATML complexity and any changes to the schema that may occur. Integration of the plug in was seamless. Once written, these ATML TestResults files are then available for viewing and analysis.

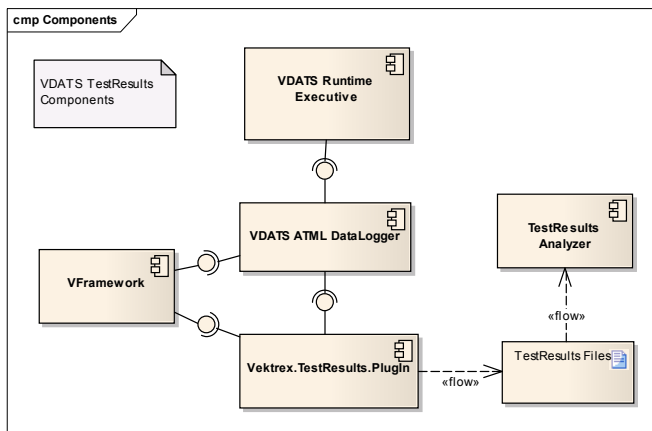


Fig. 3. VDATS TestResults components

After the TestResults “plug in” was integrated into the VDATS system, TestResults files may be generated automatically when a TPS is run. These TestResults may then be viewed and analyzed by the Test Manager or TPS Engineer. A primary benefit enabled by the usage of XML for TestResults is the ability to enhance the presentation of the test data.

III. TEST RESULTS USE CASES

A. Viewing ATML Files

Instead of basic character formatting (Figure 1), rich formatting is possible (Figure 6). Tables can be used for organizing data, color can be used to highlight failures, and graphics can be added to identify the organization. To enable such rich presentation, a custom XML style sheet was developed for the VDATS program that allows for viewing test results locally in web browsers such as Internet Explorer or Firefox. By default on most Windows based systems, double clicking on XML files will invoke Internet Explorer to view the file. The report generated from ATML test results is shown below.

VDATS

ATML Test Results For : CPIN 81B-2ALQ155-SP-00A REV 000 13 Mar 2007

Station ID	VDATS DA-1 P/N 200625630-10 S/N 0001
UUT ID	ALQ-155 Signal Processor (123456-000)
UUT Serial Number	123456-000
Operator	NA
Test Started	2007-12-07T15:55:11-08:00
Test Ended	2007-12-07T15:55:37-08:00
Test Result	Failed

Test Group Id	Start Date Time	Description	OutCome
T10	2007-12-07T15:55:11-08:00	Running Test Block 1 T10 Crowbar Test - Checking for Required Instruments	Passed

Test Name	Start Date Time	Outcome	Measured	High Limit	Low Limit	Unit
	2007-12-07T15:55:11-					

Fig. 6. Viewing TestResults in Internet Explorer

Note that viewing TestResults data does not require an Internet connection. This capability is fully supported locally provided that the style sheet file is located in the same folder as the TestResult XML file. This is important for supporting remote maintenance depots that do not have connectivity.

B. Test Analysis Cases

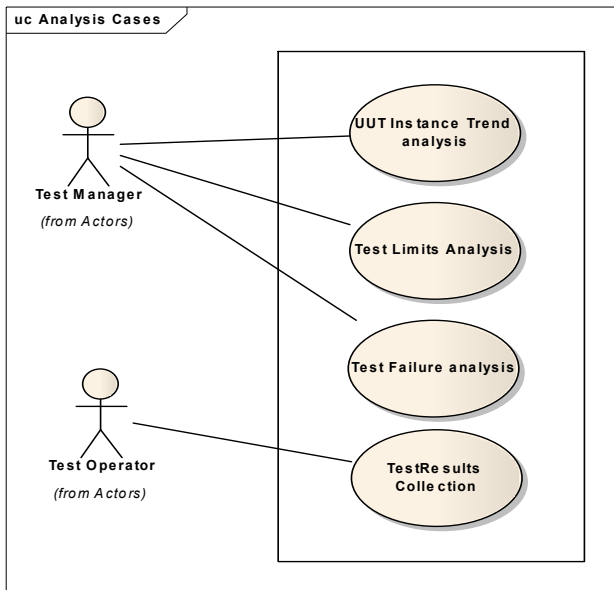


Fig 7. Test Analysis Usage Cages

The TestResults Analysis tool supports the usage cases shown in Figure 7 via a set of simple Windows wizards. Once the TestResults have been imported into the tool, the user selects the TPS, the set of TestResults files, and other information special to the type of analysis. Once the selected TestResults are analyzed, the results are presented graphically.

C. Trend Analysis

“Trend Analysis” is intended for tracking the results of a specific serial number of a UUT. Common cases for this usage include monitoring calibration drift in self-tests of instruments as well as monitoring changes over time in specific UUTs.

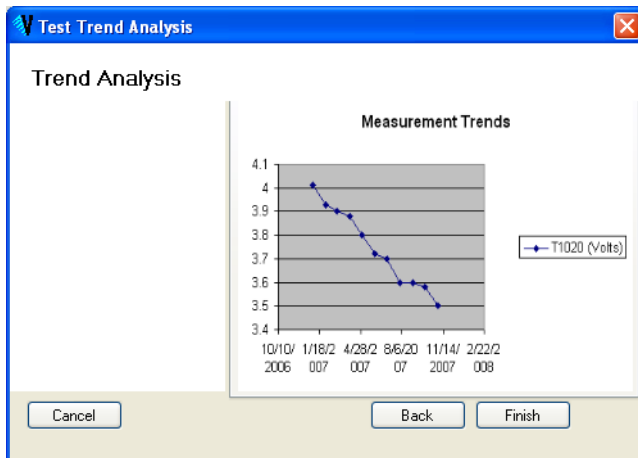


Fig. 8. Trend Analysis Wizard

In addition to the TPS and TestResults selection, this wizard requires a UUT serial number and Test ID. Using this query information, a line plot is generated showing the test measurements over time for the given serial number unit (see Figure 8). This information can be useful in determining proper calibration schedules as well as illustrating component deterioration in UUTs.

D. Limits Analysis

“Limits Analysis” is intended to see how well the test limits correlate with actual test results measurements.

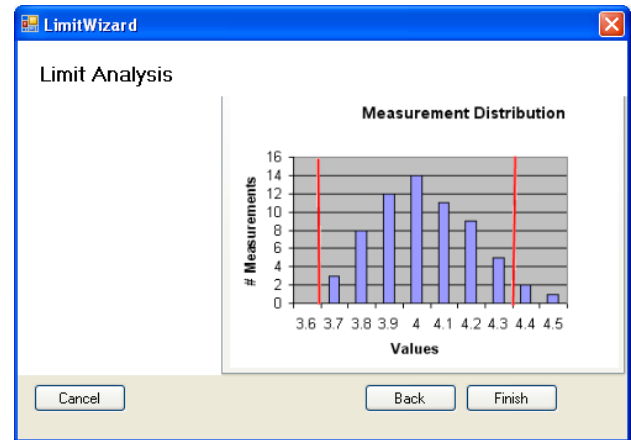


Fig. 9. Limit Analysis Wizard

After a Test ID is selected from a TPS and TestResult runs, selected measurements data is organized into a histogram chart showing the frequency of measurements by value along with the limits (see Fig. 9). This data typically takes on the form of a Gaussian curve and a curve fitting is provided. The results of this analysis can be used to feedback new limits or highlight test system issues. For example, a skewed distribution curve could point out how resistance in a test adapter needs to be taken into consideration in the limits. This analysis can result in reduced false failures and a better understanding of the system.

It is also possible to feedback new limits to an ATML TestDescription file to describe the modified limit. If the runtime system is setup to use the limits in a TestDescription file, a limit optimization can be done in a close loop without modification of the TPS.

E. Failure Analysis

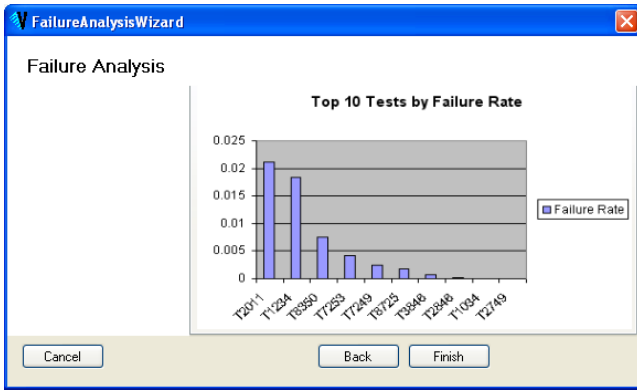


Fig. 10. Failure Analysis Wizard

“Failure Analysis” is intended to show which tests are most likely to fail based on failure rate. After selecting the TPS test runs, and the maximum number of top failures to show, a sorted histogram chart is displayed with the tests most likely to fail by test id. This information can be used to identify problem UUT components as well as optimize future TPS operation to capture failures sooner.

Similar to the “limits analysis” case, this information can be fed back to the corresponding ATML TestDescription file. If the test executive is designed to follow test sequences described in an ATML TestDescription file, close loop TPS optimization can be achieved.

IV. TPS OPTIMIZATION USING TEST RESULTS

A significant portion of the sustainment of weapons systems is maintenance and modification (M&M) of existing test programs. The M&M phase of a test program’s lifecycle is the prime opportunity to utilize historical test results information to optimize TPS performance. Additionally, Operational Safety, Suitability and Effectiveness (OSS&E), which is intended to assure engineering rigor and discipline, is applied during weapons system sustainment, benefits greatly from the commonality and interoperability of test results data. Operational Safety is the condition of having an acceptable risk to life, health, property or environment caused by a system in an operational environment. Operational Suitability is the degree to which a system can be placed satisfactorily into field use. Operational Effectiveness is the overall degree of mission accomplishment of a system. The achievement of the combination of these three is absolutely vital to the success of a weapons system in a real-world scenario throughout its operational lifecycle.

Air Force Instruction (AFI) 63-1201⁵ states that “Organizations responsible for preserving OSS&E of Air Force systems or end items must ensure that operational use, configuration changes, maintenance repairs, aging, part substitutions, and similar activities and events do not degrade baselined characteristics of systems or end items over their

operational life.” VDATS, the single test system that serves as a common point in the lifecycle of these Air Force systems, must support OSS&E with its hardware and software architecture. ATML Test Results provides a key capability that allows the Air Force to achieve the OSS&E requirements on VDATS.

Table 1. OSS&E Metrics

Element	Examples of Attributed Metrics
Operational Safety	Air Worthiness, Mishap Risk, Loss Rate
Operational Suitability	Mission Capability Rate, Mean-Time-Between-Failure, Mean-Time-To-Repair
Operational Effectiveness	Capability, Range, Payload

Three areas to investigate under the SBIR that provide immediate benefit to the Air Force were identified. The first is an analysis of the measurement distribution of multiple occurrences of the same test on one or more test systems. The second is a correlation of performance changes on one or more test systems or UUTs when software, an instrument or a component changes; each of these are very likely during the M&M phase. The third is the identification of bad actors in a test system, such as a drifting power supply.

A. Measurement Distribution

The actual distribution of measurements for multiple occurrences of a given analog test typically follows a Gaussian distribution. A typical analog measurement has a nominal value, and an upper and lower limit. Ideally, the mean, μ , is located coincident with the nominal value, and the standard deviation, σ^2 , is small enough that the distribution curve does not intersect a limit threshold.

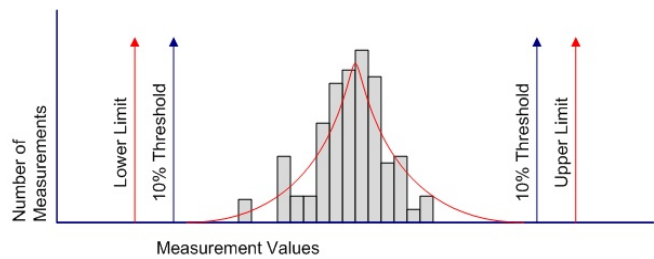


Figure 11. Measurement Distribution

Figure 11 depicts the measurement distribution of 100 measurements categorized into distinct measurement bins. The measurement distribution is Gaussian with the mean approximately centered on the nominal measurement value. The smaller the variance, the more stable and repeatable the test is likely to be.

However, there are cases where the mean is not centered on the nominal value, but rather is shifted towards one of the limit boundaries. An artificial threshold could be placed at 10% of the range from each of the limits, and the number of passing measurements between this artificial threshold and the measurement limit could be detected. This might indicate that the limits are not optimally placed, or that the UUT performance is degrading over time. In either case, this analysis can trigger an investigation to determine if the test methodology is suitable.

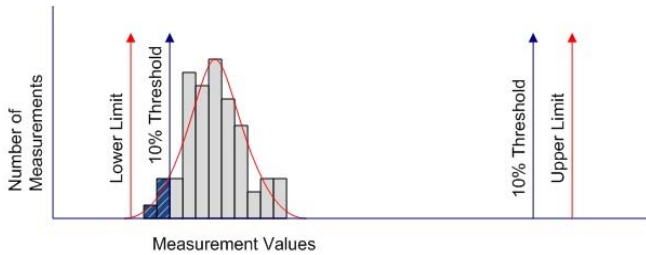


Figure 11 Measurement Values

Likewise, the extension of the probability distribution function crossing a measurement limit could signify a need to investigate limit placement.

B. Detection of Bad Actors

With all test programs using ATML Test Results, analyses can be extended to find bad actors based on a very large data set. With the integration of Test Description and Test Configuration, intelligent inferences can be determined to detect and isolate bad actors.

V. WORKING WITH THE TRIAL USE STANDARD

The current initial ATML standard for TestResults has been approved for “trial use”. It is the intention of the working group to use “trial use” feedback returned from industry to incorporate into a subsequent more permanent standard. What this means is that we can anticipate change in the standard as well as the possibility of missing elements.

A. Planning for change in the ATML standard

Since change in the ATML “trial-use” standard was a given, a “plug in” architecture was used. In this “TestResults Framework” design, the interface to various versions of TestResults data was encapsulated in a version independent framework (see Figure 3) that uses run-time loaded “plug in” components to handle different versions of the standard. This design will allow VDATS and the TestResults analysis tool to handle future versions of the standard without re-compiling.

B. The need for measurement resolution.

Most Test Requirements Documents specify a parameter to be measured and required resolution of that parameter. For example, 10 VDC, to 1 mV resolution. This means that the measurement must contain at least three decimal places of accuracy. For example, a valid measurement is 10.000 VDC.

For non-digital tests, a Test Requirements Document typically provides a nominal measurement value with high and low limits. These values specify some degree of measurement resolution, which must be measured, transferred and recorded in the test report. For example, suppose a test provided for a nominal measurement of 10.0 VDC.

The value 10.0, when extended to three decimal places, can take any value between 9.945 and 10.044, which provides for a relatively large degree of uncertainty. The following diagram shows this uncertainty for a test with nominal value 10.000 V and +/-30 mV measurement limits.

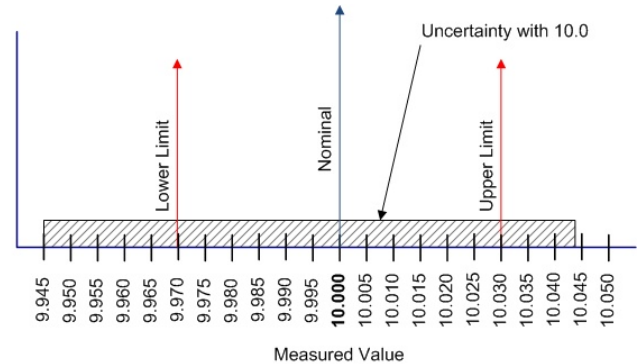


Figure 12 Measurement Resolution

An invalid measurement is 10.0 VDC, which when extended to three decimal places has an uncertainty that covers the range from 9.995 to 10.044. There are several points in a typical ATE system where this value can be corrupted, cast in data types that automatically pad (or truncate) trailing zeros, and even displayed incorrectly in a formatted string. A modification to the Test Results schema is just on part of the puzzle.

Consequently, we see that a lack of decimal precision can provide for false passes. This is a serious concern given the OSS&E mandate. In the short term, this will be addressed through the formatting of measurements in TestResults to display the significant digits to imply the measurement resolution. In the longer term, we’re pleased to receive feedback from the ATML working group that this concern will be address in the next revision of the standard.

VI. SUMMARY

The Air Force will utilize the VDATS on legacy, modern and future electronics systems for years for decades to come. The ability to synchronize the Test Results capture and analysis methodologies for these workloads will greatly enhance the ability to provide world-class support to the War Fighter.

ACKNOWLEDGEMENT

Vektrex acknowledges and thanks both the Air Force and the Navy for their support of the SBIR program and Vektrex's research related to ATML.

REFERENCES

- [1] G.F. Kilian, "DoD Automatic Test Systems Program Plan", 2006.
- [2] "DoD AUTOMATIC TEST SYSTEMS PROGRAM PLAN", http://www.acq.osd.mil/ats/DoD_ATS_Program_Plan-2006.pdf, 2006
- [3] IEEE Std. 1636.1-2007, IEEE Trial-Use Standard for Software Interface for Maintenance Information Collection and Analysis (SIMICA): Exchanging Test Results and Session Information via the eXtensible Markup Language (XML)
- [4] SBIR Small Business Innovative Research
- [5] Air Force Instruction 63-1201, "Life Cycle Systems Engineering", Secretary of the Air Force, July 2007 <http://www.e-publishing.af.mil/shared/media/epubs/AFI63-1201.pdf>
- [6] ATML Download Site, <http://grouper.ieee.org/groups/scc20/ATML/>, March 2008
- [7] Air Force Guidance on Procurement of Automatic Test Systems and Equipment", http://www.acq.osd.mil/ats/USAF_ATS_Policy_Letter_signed_26_Nov_07.pdf, November 2007