6A.4   TESTING STRATEGIES FOR THE NEXT GENERATION AWIPS

Kenneth L. Stricklett*
National Weather Service, Office of Operational Systems, Silver Spring, MD

Jason P. Tuell, Ronla K. Henry, Peter K. Pickard
National Weather Service, Office of Science and Technology, Silver Spring, MD

John D. Lawson and Frank P. Griffith
Raytheon Company, Silver Spring, MD

1. INTRODUCTION

The software supporting the National Weather Service (NWS) Advanced Weather Information Processing System (AWIPS) is undergoing an extensive re-architecture and conversion to a Service-Oriented Architecture (SOA) (Lawson 2007). NWS has tasked Raytheon Information Solutions (Raytheon) to implement the new system and will be converting the AWIPS software infrastructure to a Service-Oriented Architecture over the next year. The current baseline functionality will be migrated to the new architecture by fiscal year 2009 (FY2009) and deployment of the next generation, AWIPS II, software is expected to be complete in FY2010. At deployment, AWIPS II will provide a "black box" conversion of the AWIPS I software and will be functionally equivalent to the legacy software.

This paper provides an overview of the testing strategies employed to validate the AWIPS II software. These strategies, developed in partnership with Raytheon and in coordination with NWS stakeholders, rely on multiple levels of review and both formal and informal tests and evaluations of the system software.

Testing activities typically consume between 30% and 50% of software development costs (Ramler 2006), and this figure can be even higher for safety-critical applications such as those employed by the NWS. A key issue leading to the success of any software project is the effective allocation of testing resources. In general, early discovery of defects is more valuable. Defects found in the end stages of development are often difficult or impossible to correct and are costly: Finding and fixing a defect after delivery can be 100 times more expensive than during the early stages of a project. Strategic placement of testing activities in the software design and development cycle is critical to the effective use of resources. The testing strategies employed for AWIPS II are designed to provide early and frequent user input and thus to provide early discovery of defects.

AWIPS II testing will touch nearly every organization within the NWS (field offices, National and Regional headquarters, and National Centers) and will culminate in the first comprehensive field test of AWIPS since the Commissioning Operational Test and Evaluation (1999).

2. TESTING ROADMAP

The AWIPS II software project is structured as a series of incremental Task Orders (TOs) culiminating in TO11, which will deliver the full AWIPS II software functionality. Task Order deliverables will be subjected to series of tests from the following:

- Code and Unit Tests
- Delivery Tests
- Independent Validation and Verification (IV&V)
- User Functional Evaluation–Headquarters (UFE-HQ)
- User Functional Evaluation–Regional (UFE-R)
- User Evaluation (UE)
- Information Technology Certification and Accreditation (C&A)
- Local Application Tests
- Operational Test and Evaluation (OT&E)

Task Order deliverables from 8 to 10 will be submitted to the sequence of tests shown schematically in Figure 1.
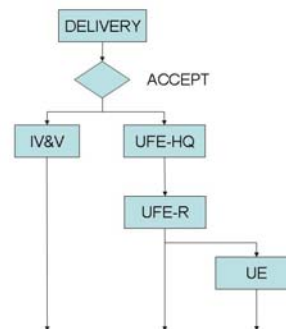

Figure 1. The Task Order testing sequence.

The sequence of tests followed for TO11 is shown in Figure 2. C&A will be added for TO11 and a period of bug fixing is planned before the start of OT&E.

Corresponding author address: Ken Stricklett, NWS, W/OPS24, 1325 East West Highway, Silver Spring, MD, 20910 (Ken.Stricklett@noaa.gov)
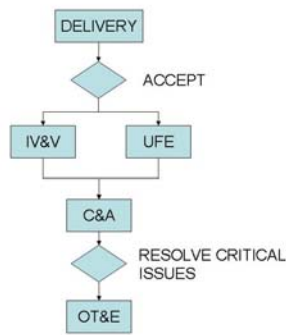
Figure 2. The TO11 testing sequence.

3. DEVELOPER TESTING

Raytheon will perform all Code and Unit Tests and other pre-delivery tests.

Raytheon will perform a Delivery Test for each Task Order to ensure the established acceptance criteria are met. Delivery Tests are conducted according to a Delivery Test Plan negotiated between the NWS and Raytheon. The Delivery Test Plan includes detailed Test Case Procedures to be used during Delivery Tests. Delivery Tests are performed on baseline AWIPS hardware and must be completed prior to acceptance of the Task Order deliverable. All Delivery Tests are witnessed by the NWS.

4. GOVERNMENT TESTING

The NWS will test, evaluate, and review each Task Order deliverable leading to national deployment of the AWIPS II software. NWS tests and evaluations include IV&V, UFE, UE, C&A, and OT&E. The NWS is responsible for both planning and executing each of these tests and evaluations. Raytheon will provide technical support for fixing work stops, processing trouble tickets, providing the disposition of discrepancy reports, and assistance in defining performance test procedures. A complete description of each of these tests is beyond the scope of this document; however, a few points on each are provided below.

4.1 Independent Validation and Verification

IV&V will provide a formal, independent assessment of each Task Order deliverable. IV&V activities started with the delivery of AWIPS Development Environment (ADE) release 0.1 and will continue for the duration of the project.

IV&V assesses the quality of the software, ensures the internal documentation is complete and accurate, ensures the proper functionality of the software, ensures mathematical calculations are carried out to the required accuracy, and ensures the software meets all applicable performance requirements. IV&V activities include simulated NWS operations, which are intended to subject the software to realistic field conditions.

Early discovery of defects is vital to the success of the project. Defects discovered during IV&V will be entered into a database, adjudicated to assign their impact and priority, and tracked to ensure all critical defects are fully resolved and correctly implemented in the AWIPS II software.

Raytheon technical support for IV&V consists of resolution of defects, recommending performance metrics and tests, answering technical questions and consulting, and providing the disposition of discrepancy reports.

4.2 User Functional Evaluation

UFE will be conducted after acceptance of each Task Order deliverable. UFE is similar to the Pre-Integration Test (PIT) done today for new AWIPS functionality, and is intended to demonstrate the correct implementation of forecaster functionality, i.e., the UFE will verify that the functionality provided in the AWIPS II software adequately duplicates that provided in the legacy software. UFE will be performed by NWS forecasters, hydrologists, and other expert AWIPS users. UFE will first be conducted at NWS National Headquarters facilities (UFE-HQ) and will be conducted shortly thereafter at the NWS Regional Headquarters (UFE-R). UFE will be performed on baseline AWIPS hardware or baseline compatible hardware.

Raytheon technical support during UFE will consist of fixing work stops and providing the disposition of discrepancy reports.

4.3 User Evaluation

UE complements UFE by allowing experienced AWIPS users the opportunity to try out the new system under less formal conditions. UE will provide early user feedback on the "look and feel" of baseline applications, localization and customization, system administration, local applications migration and testing, and unique interfaces and data feeds. A key component planned for UE is the facility for side-by-side comparison of the two systems, i.e., the user will have access to systems running the legacy AWIPS I software and the AWIPS II software and can directly compare the functionality of each system. The implementation of the AWIPS II software for UE will be independent of the operational AWIPS and will thus have no impact on site operations.

UE will be conducted at selected NWS sites. Sites under consideration include Central Region Headquarters and the NWS Training Center, Weather Forecast Offices, a River Forecast Center, and the National Oceanic and Atmospheric Administration's Earth System Research Laboratory, Global Systems Division.

4.4 Certification and Accreditation

AWIPS is scheduled to undergo a C&A update in early 2008. Deployment of the AWIPS II software to NWS field sites will require revision of the C&A plan to address the specific changes implemented in the AWIPS II software since the update. The scope of the

AWIPS II C&A is expected to be limited primarily to the technical controls, e.g., authentication and authorization, of AWIPS II within the security architecture of AWIPS. As such, AWIPS II C&A is not expected to be a major departure from the established plan. C&A should be substantially completed before the start of OT&E. However, if significant configuration changes are required due to the findings of IV&V and UFE, portions of the C&A may be completed in parallel with OT&E.

The NWS is responsible for completion of all C&A activities.

4.5 Operational Test and Evaluation

OT&E is the final test of the AWIPS II software undertaken prior to national deployment. OT&E will provide a comprehensive system evaluation, including installation and de-installation of the AWIPS II software, training of NWS personnel, system documentation, support services provided by the AWIPS Network Control Facility (NCF), AWIPS communications, message handling, and local applications.

The OT&E will be conducted on baseline AWIPS hardware, including operational Weather Forecast Offices and River Forecast Centers selected from each of the NWS Regions, National Centers for Environmental Prediction, Center Weather Service Units, and Alaska Aviation Service Units.

OT&E is nominally planned for a period of six months, however, OT&E is structured around the verification of specific technical objectives rather than calendar dates. The technical objectives that must be demonstrated during OT&E include: 30 consecutive days of reliable operation at each of the test sites and successful completion of all assigned Test Case Procedures.

OT&E will be overseen by a Test Review Group comprised of user representatives and subject matter experts selected from the NWS Regions and NWS Headquarters. The OT&E Plan is currently under development by an NWS Integrated Working Team and will be submitted to the NWS Regions and to NWS management for approval prior to commencing OT&E.

Raytheon support during OT&E will consist of normal trouble ticket processing and priority resolution of any critical defects. Raytheon developers will also provide "rapid response" to NWS field sites as needed.

5. SUBTASKS

The NWS has identified several highly leveraged areas for concentrated effort during the software update.

5.1 Algorithm Validation and Verification

Virtually all fundamental meteorological parameters may be calculated using multiple algorithms. The selection of appropriate algorithms for a specific application may depend on the required accuracy, computational speed, and the range of available physical parameters, for example, and is often a matter of judgment.

Significant benefits may be gained by eliminating duplicate algorithms: both the lines of code maintained and the incumbent probability of errors are reduced by eliminating duplicate code. However, rationalization of the computational algorithms must be guided by good science and by the needs of the NWS forecasters and hydrologists.

The goal of this activity is to provide forecasters and hydrologists with complete, accurate, and reliable computational tools. The meteorological algorithms implemented will be identified and compared to document their underlying assumptions, range of validity, and reliability. The code supporting these algorithms will be reduced and simplified when appropriate.

5.2 Performance Testing

System performance tests will be conducted both during IV&V and OT&E. The figures of merit considered for performance testing include the timeliness of system response and the reliability and latency of system communications. The activities required to support performance testing include defining performance metrics, obtaining baseline performance data for AWIPS I, and testing the performance of AWIPS II. These tests will rely, in part, on side-by-side tests of AWIPS I and AWIPS II running on baseline AWIPS hardware. The performance of AWIPS communications will be assessed using the Product Availability Monitoring System (Stricklett 2007), which has been extensively used during earlier AWIPS OT&E.

5.3 Local Applications

One of the key strengths of AWIPS is the ability to configure the system to satisfy local needs, and NWS field sites have invested considerable resources in development of local applications to meet the requirements of site forecasters and hydrologists.

A recent survey conducted by the NWS Office of Science and Technology found in excess of 3200 unique local applications in use at NWS field sites. Of these applications, approximately 1250 were listed as critical to site operations (Olsen 2007). Developing effective methods for migration of local applications to the AWIPS II environment is thus essential to the success of project.

While the development of local applications is, and will remain, the responsibility of the NWS sites and Regions, NWS Headquarters must provide meaningful support to the field sites for planning, training, and facilities to enable the sites to efficiently complete the task of migrating their local applications to AWIPS II.

An NWS Integrated Work Team, comprised of Regional representatives and subject matters experts, has been formed to address this issue. The recommendations under consideration to date include implementation of a collaborative work environment and code repository for local applications development and limited technical support. The NWS Office of Science and Technology, Systems Engineering Center will play a key role in leading this transition and will interface with Raytheon when escalating technical issues. The NWS

Office of Climate, Water and Weather Services will coordinate and provide training in the new software architecture.

6. AWIPS II TESTING

The AWIPS II software is based on an open source, Service-Oriented Architecture, with object oriented programming and platform independence. These features will improve both the quality of the code and its flexibility and maintainability.

Implementation of the AWIPS II software is projected to have direct benefits for development of new applications and operations and maintenance. One of the objectives of testing is to critically examine the paradigm for system operations and maintenance and to refine and adapt current NWS practices to the new system architecture.

The testing paradigm for operations and maintenance of the AWIPS I software involves multiple levels of development, integration, acceptance, and field tests (Code and Unit Tests, Software Integration Test, System Integration Test, System Acceptance Test, Alpha, and Beta). AWIPS I software has evolved over 10+ years and currently consists of custom and open source architectural components. Coupling among these components and the applications they support has meant that regression testing usually consists of a comprehensive test of each new build, limiting the frequency of system updates to about one per year. Changes in AWIPS II will be able to occur much more frequently, with testing both more efficient and effective.

AWIPS II is based largely on open source software. AWIPS has been long employed open source software for its operating system, and this strategy has proven to be both reliable and effective. The introduction of open source Service-Oriented Architecture components will further enhance reliability and maintainability by replacing all custom architectural components. Platform independence will allow researchers to achieve results that can be much more easily integrated into AWIPS.

Object orientation implements a modular architecture. This has two advantages: (1) the likelihood of unwanted side effects is reduced when the code is modified and (2) the ability to reuse objects. Since the source code consists of independent modules, the need for extensive regression testing should be reduced in the new architecture. Once an object has been validated it may be reused with little risk.

The re-architecture of the AWIPS software will have the added benefit of reducing the total lines of code. AWIPS I has grown to include over four million lines of code (LOC), and the total LOC are projected to be reduced by approximately 40% in AWIPS II, with application code approximately 10% of the current total. In addition to replacing architectural code with open source, functional redundancies will be removed.

There are several caveats to these advertised benefits: Open source software is not bug-free. Fixes are typically available relatively quickly; however, it is up to the user to implement them. The user must under-

stand the desired functionality of the open source software and how to use it. Although the code itself is always available for inspection, documentation for open source projects can be uneven.

Object-orientation does not prevent bad programming. Standard class libraries and open-source patterns are well tested; however, object developers must still follow good development practices. Mitigations include control of the cyclomatic complexity (McCabe 1976) of the new software.

7. DISCREPANCY HANDLING PROCESS

NWS Headquarters; primarily the Office of Science and Technology and the Office of Climate, Weather, and Water; will administer a new process for tracking defects and user comments. The process implemented mirrors the AWIPS Discrepancy Report (DR) process. That is, suspected defects will be entered into a database tailored to the AWIPS II project and adjudicated by an AWIPS II Review Board. Each defect will be ranked by the Review Board and assigned to the appropriate development organization for resolution. Defects will be escalated to Raytheon developers as required.

Similar methods will be used during OT&E to track suspected defects and other user input. AWIPS users will enter defects into a defect tracking database and will also follow AWIPS standard operating procedures and report any problems to the AWIPS Network Control Facility Help Desk. The Network Control Facility will open trouble tickets for each problem reported. The OT&E Test Review Group will adjudicate and track all reported defects and ensure the bug tracking database is accurate and up to date. Data gathered by the Test Review Group regarding the timeliness of response and the accuracy and completeness of trouble tickets will be used to assess the quality of services provided by the AWIPS Network Control Facility and Raytheon.

E-mail list servers have been implemented to support developers and test site personnel. The developer list is intended for the free exchange of tips, hints, etc., among AWIPS II users. The administrative list is for AWIPS II site points-of-contact (POC). New software releases and other pertinent information regarding system administration will be announced via this list.

Engineers at NWS Headquarters will monitor the list servers and defect tracking database entries and provide feedback both to those posting and to Raytheon.

7.1 System-Level Problems

Problems found during performance and reliability testing will, upon confirmation, be communicated to Raytheon. In this way, possible system-level or architectural issues will be addressed with highest priority. Problems not confirmed as performance or reliability concerns will follow the normal process.

7.2 Testing Problems

Problems with test design or execution will be

communicated to the NWS Headquarters test team. Those that cannot be resolved in a timely manner will be escalated to AWIPS II management.

8. CONCLUSION

AWIPS II implements radical changes in the underlying architecture of the AWIPS software, and comprehensive system tests are required to ensure the correct implementation of AWIPS requirements. The testing strategy presented includes multiple levels of testing, facilitates early evaluation of the software by the end users, and establishes clear technical objectives that must be demonstrated prior to national deployment of the AWIPS II software.

9. ACKNOWLEDGEMENTS

10. REFERENCES

Lawson, J. D., J. P. Tuell, F. P. Griffith, and R. K. Henry, 2007: Overview of the New AWIPS SOA. Preprints, *23d Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology*, San Antonio, TX, Amer. Meteor. Soc., CD-ROM, 6.2.

McCabe, T.J., 1976: A Complexity Measure. IEEE Trans. on Software Eng., SE-2, 308−320.

Olsen, J., 2007: Results of the local applications survey, private communication.

Operational Test and Evaluation for AWIPS Build 4.2, The Commissioning Load, May 1999.

Ramler, R., S. Biffl, and P. Grunbacher, 2006: Value-based management of software testing. In *Value-Based Software Engineering*, Stefan Biffl, Aybuke Aurum, Barry Boehm, Hakan Erdogmus, and Paul Grunbacher editors, Springer-Verlag, Berlin.

Schotz, S., J. P. Tuell, S. Jacobs, D. W. Plummer, S. A. Gilbert, R. K. Henry, 2008: Integrating NAWIPS into the new NWS Service Oriented Architecture. *24th Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology*, New Orleans, LA, Amer. Meteor. Soc., 6A.3.

Stricklett, K. L., K. B. Nguyen, and M. D. Buckingham, 2007: Performance measures for the AWIPS communications network. Preprints, *23d Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology,* San Antonio, TX, Amer. Meteor. Soc., CD-ROM, P2.3.

Tuell, J. P., R. K. Henry, J. Olsen, J. D. Lawson, F. P. Griffith, 2008: Application Migration within New AWIPS Architecture. *24th Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology*, New Orleans, LA, Amer. Meteor. Soc., 6A.2**.**

Tuell, J. P., R. K. Henry, J. D. Lawson, F. P. Griffith, 2008: The Evolution of AWIPS–Progress and Future Plans. *24th Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology*, New Orleans, LA, Amer. Meteor. Soc., 6A.1.