The

ITEA Journal

September 2008
Volume 29, Number 3

Published quarterly by
the International
Test and Evaluation
Association

Improving Suitability Through Test

CALCULEX
Electronics, Systems, Software and Support

What is YOUR data worth when it...

...gives fire fighters the location of hot spots that have the potential to destroy thousands of family's homes?

...is used to determine if a warfighter is ready to be deployed to protect and defend millions of people worldwide?

...provides the critical information to border security in apprehending drug smugglers and curtailing criminal activities?

...is captured by the most reliable, flexible and best valued, solid state recorders?

MONSSTR®

info@calculex.com www.calculex.com (575) 525-0131

One test is worth a thousand expert opinions

Trust BAE Systems to provide instrumentation systems that exceed the demanding test and training requirements of today's weapons and tactics.

- Monopulse TSPI Radar
- CW Doppler TSPI Radar
- Wideband Imaging Radar
- Electro-Optical Tracking Systems
- Electronic Warfare Simulators
- Life Extension and Performance Upgrades

BAE SYSTEMS

For additional information contact: ESD Business Development
715 Hollywood Blvd., NW. • Ft. Walton Beach, FL 32548-4090
Phone 850.664.1334, Fax 850.244.6494, Operator 850.244.7711 ESDINFO@baesystems.com www.esdradar.com

The ITEA Journal
September 2008
Volume 29, Number 3

Board of Directors
John Smith, *President*
Russell "Rusty" L. Roberts, *Vice President*
Dr. John B. Foulkes, *Secretary*
Stephanie H. Clewer, *Treasurer*
George M. Axiotis
Scott P. Foisy
Robert T. Fuller
Kirk S. Johnson
Thomas J. Macdonald
George R. Ryan
Brian M. Simmons
Drexel L. Smith
Mark E. Smith
James R. Tedeschi
John L. Wiley

Senior Advisory Board
Charles F. Adolf
Robert J. Arnold, *Chair*
Brent M. Bennitt
John V. Bolino
Edward R. Greer
George B. Harrison
Dr. Charles E. McQueary
Dr. J. Daniel Stewart
Dr. Marion L. Williams

Committee Chairs
Awards
Denise De La Cruz
Chapter & Individual Membership Development
Mark E. Smith
Communications
Lawrence Camacho
Corporate Development
Drexel L. Smith
Education
Scott P. Foisy
Elections
Gary L. Bridgewater
Events
Brian M. Simmons
Historian
Dr. Michael Gorn
Publications
Dr. J. Michael Barton
Technology
George J. Rumford
Ways and Means
Michael A. Schall

Staff
Executive Director
Lori Tremmel Freeman
Assistant Director
Eileen G. Redd
Manager, Exhibits and Corporate Development
Bill Dallas
Managing Editor, ITEA Journal
Rita A. Janssen
Office Manager
Jean Shivar
Coordinator, Office Support and Services
Bonnie Schendell

CONTENTS

TECHNICAL ARTICLES

DEPARTMENTS

ON THE COVER: Test and evaluation has always focused on verifying system performance, but to perform when needed a system must demonstrate reliability. Reliability is an essential component of a system's worth and must be examined throughout design and development. This issue focuses on defense system reliability today and efforts to systematically improve it. The cover is from a 1942 poster "God Help me if This is a Dud" provided courtesy of the National Archives. It was commissioned by the US government's Office for Emergency Management, Office of War Information as part of a series of World War II posters and was created by John Vickery.

■ ITEA Headquarters: 4400 Fair Lakes Court, Suite 104, Fairfax, Virginia 22033-3899; Tel: (703) 631-6220; Fax: (703) 631-6221, E-mail: *itea@itea.org;* Web site: *http://www.itea.org.*

■ ITEA is a not-for-profit international association founded in 1980 to further the development and exchange of technical information in the field of test and evaluation.

■ *The ITEA Journal* (ISSN 1054-0229) is published quarterly by the International Test and Evaluation Association at 4400 Lakes Court, Suite 104, Fairfax, Virginia 22033-3899. Single issue cover price for *The ITEA Journal* is $20. ITEA membership dues are $45 for individuals, $25 for full-time students, and $745 for corporations. Annual dues include a one-year subscription to *The ITEA Journal*. The annual subscription rate for libraries and other organizations providing timely reference material to groups is $60. All overseas mail (air mail or AOA) requires an additional $20. *The ITEA Journal* serves its readers as a forum for the presentation and discussion of issues related to test and evaluation. All articles reflect the individual views of the authors and not official points of view adopted by ITEA or the organizations with which the authors are affiliated.

© Copyright 2008, International Test and Evaluation Association, All Rights Reserved. Copyright is not claimed in the portions of this work written by U.S. government employees within the scope of their official duties. Reproduction in whole or in part prohibited except by permission of the publisher.

POSTMASTER: Send address changes to: ITEA, 4400 Fair Lakes Court, Suite 104, Fairfax, Virginia 22033-3899.

President's Corner

ITEA Journal 2008; 29: 209
Copyright © 2008 by the International Test and Evaluation Association

When the year began, we were concerned with the impact reduced budgets would have on the ITEA workshops and symposiums. Travel budgets were being drastically reduced, affecting potential event attendees. Reduced budgets forced corporate members to be more judicious in determining where to invest their advertising dollars. The success of the workshops and symposiums is heavily dependent on the attendance of its members and support from corporate members through sponsorships and exhibits. As we have completed most of the year, I congratulate the chapters and their committees who have put together substantive programs that continue to attract attendees and corporate sponsors. While event titles have remained the same, the construct of the events have been modified to reflect the reality of the challenges we face in the test and evaluation (T&E) community. Corporate members have continued to see value in the workshops, demonstrated by their continued support of the events. Because of the professionalism and technical quality of the workshops and symposiums that ITEA hosts, the association continues to maintain its reputation as the leading T&E education organization. As a team of volunteers and headquarter staff working together, we can continue to experience the success of the past. I implore you to exploit the headquarter staff in the planning and execution of your events. While the primary mission of the headquarter staff is to ensure the organization is in compliance with the laws that govern not-for-profit organizations, they are also valuable resources available to you in the planning and execution of your event as required.

As we look ahead toward the end of the year, hopefully it is not a surprise to anyone that our annual symposium is nearing. Twenty-five years ago, the founding members of ITEA decided to convene an annual symposium. The annual symposium is set aside to not only address the technical challenges of T&E, but also to obtain information as to the impact of policy changes on the T&E infrastructure and processes. Since the annual symposium is not focused on a single topic, it tends to attract more of our international members. As such, the event serves as a superb opportunity for the international T&E community to network, sharing common practices, and offering the potential for the cross utilization of resources. Needless to say, the annual symposium is the pinnacle event for our association.

This year's symposium is a shift from the normal paradigm. In the past, we have discussed broadening the focus of ITEA events beyond the Department of Defense. For this year's annual symposium, we have consciously incorporated other organizations that have T&E resources and conduct T&E as part of their acquisition processes. These organizations are the Federal Aviation Administration, the National Aeronautics and Space Administration, and the United States Department of Homeland Security. The common thread through all of these organizations is aviation, including the interoperability of ground platforms with avionic systems. The annual symposium will be held November 10-13, 2008 in Atlantic City, New Jersey. If you have not already done so, I strongly encourage you to mark your calendar to attend and be a part of this grand event.


John Smith

As system engineers, developers and users, we are all too familiar with the terms reliability, availability, and maintainability (RAM). The terms are more affectionately referred to as the "ilities." The "ilities" are measures of how well the system performs; whether it's available for use when required; and how long it takes to maintain, or repair when it breaks. When all of these measures are within the desired intervals, we believe, and rightfully so, that we have designed and produced a system that is value-added to the warfighter. What if you have this superb system, but it is not the appropriate or right system for the mission? What if we tested for all of the "ilities" and the system passed, but we forgot to test it for the appropriateness/aptness of the system to the warfighter's mission? What we then have is a system that may not be suitable for the warfighter's intended use. In other words, do we have a system that is as suitable as it is reliable, available, and maintainable? The only way to have creditable confidence in the suitability of a system is to incorporate system suitability testing as an integral part of analytical test procedures. This volume of *The ITEA Journal* will look at the contribution of test, the impact of suitability, as well as methods for improving suitability through test. I am confident that articles printed in this edition will help each of us explore test methods for improving/reporting on the suitability of the systems we test.

THE ENEMY IS SURPRISE.
Multiple threats appear out of nowhere and put the mission at risk. What's next? What are the options? DRS test and training systems help simulate the heat of battle so aircrews gain vital experience. The fog of war is lifted. They are prepared for any situation, and surprise is on their side.

DRS is the clear leader in combat training, testing and evaluation systems with over 3,500 pods and 200 ground threat simulators in service. That's more instrumentation in the field than all other competitors combined. DRS systems facilitate training on the ground and in the air to support evaluation of both aircrew and vital onboard systems' performance in test and training environments. The result: aviators and weapon systems are as well prepared as possible. Bring us your toughest challenges. We're always looking for new enemies to conquer.

Working together
we'll get it done

DRS TECHNOLOGIES

www.drs.com

Guest Editorial

ITEA Journal 2008; 29: 211–214
Copyright © 2008 by the International Test and Evaluation Association

Improving the Suitability of Systems

Charles E. McQueary, Ph.D.
Director, Operational Test and Evaluation,
Washington, DC

I want to begin by commending ITEA for selecting "improving the suitability of systems" as the theme of this issue of the Journal. This is the latest example of ITEA's long history of serving the needs of the association's membership, and of ITEA's continued dedication to focusing on issues that are important to the defense test and evaluation (T&E) community.


Charles E. McQueary

For several years now, the Department of Defense has been trying to better control the costs and schedules associated with new systems. This increased focus on cost and schedule is due to the realization that we simply cannot afford the kind of cost overruns that have become all too familiar with major systems acquisitions. When program development costs rise, we are often forced to procure fewer units, thus driving the unit cost even higher.

During the late 90's, the Department's efforts to reform the acquisition process were often focused on the "quicker, faster, cheaper," mantra popular in industry. Acquisition strategies such as evolutionary and spiral acquisition were adopted to get essential capabilities into the field quickly, deferring additional capabilities to subsequent blocks. This focus on agility and near-term capability did improve the timeliness of some programs; however, as Chairman of the Joint Chiefs of Staff, Admiral Mike Mullen, pointed out when he was Chief of Naval Operations:

> *"We have a tendency to look at what it takes to get a program out the door. We don't think too much about what the life cycle [cost] is. It's 'Can I build it?' I would like us all to be mindful of what it costs to operate whatever we are building for whatever its life is going to be because I have to pay that bill every single year."*

In an effort to improve timeliness, activities that help ensure suitable systems with high reliability get into the hands of the warfighter—including operational T&E (OT&E)—were often either bypassed or reduced. This added major expense to operations and sustainment, because inadequate suitability and reliability results in increased logistics footprints, large sustainment costs, and reduced system effectiveness. Taken to the extreme, poor reliability not only greatly increases operating and support (O&S) costs, it undermines warrior confidence and adversely impacts our Nation's warfighting capability. This shift had become so pronounced, that in 2007, Dr. Hank C. Dubin, who was then the Army's T&E Executive, reported that for the first time in its history, the Army spent more on maintenance than it did on acquisition.

But with the help of Congress, the pendulum has begun to swing back toward more formal and traditional oversight and toward those mechanisms that can prevent immature technologies from entering the development process—where they can result in significant cost and schedule delays, and increased sustainment costs due to low reliability.

Helping maintain the momentum toward more formal and traditional oversight, have been several major T&E studies and policy initiatives that have called for sound systems and reliability engineering—and which will improve the suitability and reliability of systems we send our warfighters. I will briefly describe these T&E studies and policy initiatives and greatly encourage you to read more about them in this issue of the ITEA Journal, where several are described in detail.

The first T&E study I would like to discuss is the Section 231 report to Congress mandated in the fiscal year 2007 National Defense Authorization Act. The report required the Office of the Secretary of Defense, Operational Test and Evaluation Directorate (DOT&E) and the Office of the Under Secretary of Defense for Acquisition, Technology and Logistics (AT&L) to review, reaffirm or modify T&E policy as appropriate.

The Congressional concern that prompted the Section 231 report requirement was the proliferation of nontraditional acquisition strategies that we have seen in recent years, such as evolutionary, spiral, and rapid acquisition strategies, and uncertainty about what those strategies implied about how a program was being managed and whether adequate testing was being conducted. The end result of the Section 231 report requirement was that on Dec. 22, 2007, John Young (USD/AT&L) and I signed a T&E policy revision letter, which brought all T&E under one set of policies, and which was designed to improve the suitability and reliability of systems we send our warfighters.

The main provisions of the policy revision are:

- Developmental test (DT) and operational test (OT) activities shall be integrated and seamless throughout the system life cycle.
- Evaluations shall include a comparison with current mission capabilities using existing data, so that measurable improvements can be determined.
- T&E should assess improvements to mission capability and operational support based on user needs and should be reported in terms of operational significance to the user.
- To maximize the efficiency of the T&E process and more effectively integrate developmental and operational T&E, evaluations shall take into account all available and relevant data and information from contractor and government sources.
- Operational evaluators will continue to fulfill their statutory roles in providing assessments of operational effectiveness, operational suitability, and survivability to the Milestone Decision Authority (MDA). In addition, program managers shall report the results of completed developmental testing to the MDA at milestones B and C.
- To realize the benefits of modeling and simulation, T&E will be conducted in a continuum of live, virtual, and constructive system and operational environments.

The policy revision also enforced the central tenet that a system will be tested in accordance with how it will likely be used, and not just according to how we bought it, which helps to move the entire emphasis from acquisition strategy to effectiveness and suitability.

There have been several other major T&E initiatives, studies, and policies completed in the past year that have contributed to the momentum toward sound systems and reliability engineering—and which will improve the suitability and reliability of systems we send our warfighters. Several of these are described in this issue of the ITEA Journal.

In a special invited article in this issue, Mr. Pete Adolph, the former director, Test, and Evaluation, describes the Defense Science Board Study on DT&E that DOT&E and AT&L co-sponsored and which he led. The study provided convincing evidence that systems engineering along with a robust reliability growth program is an essential and missing element of assuring system suitability. Pete will outline the study's recommendations for revitalizing DT&E in the department—a key strategy and component of the effort to improve suitability and reliability.

In this issue's *Inside-the-Beltway* column, Mr. Chris DiPetto, the deputy director of DT&E in AT&L, will describe the efforts of the Reliability Improvement Working Group (RIWG), which Dr. James Finley, DUSD/AT&L/A&T, and I established in February to implement recommendations of the Defense Science Board Taskforce on DT&E, to increase the reliability of the systems we send our warfighters.

Chris will detail RIWG actions needed and underway to:

- Ensure programs are formulated to execute a viable systems engineering strategy from the beginning, including a reliability, availability, and maintainability (RAM) growth program, as an integral part of design and development;
- Ensure government organizations reconstitute a cadre of experienced T&E and RAM personnel, and finally;
- Implement mandated integrated DT and OT, including the sharing and access to all appropriate contractor and government data and the use of operationally representative environments in early testing.

We'll also hear from Randy Fowler of AT&L about life cycle metrics designed to improve Office of the Secretary of Defense (OSD) oversight, and from Pete Nolte, also from AT&L, on the new manual for developing the Sustainment Key Performance Parameter (KPP).

The Sustainment KPP is comprised of one KPP (materiel availability) and two key system attributes (KSAs) (reliability and total ownership cost). The manual requires the requirement's generation community to document expected system operational conditions, tempo, failure definitions, operator, and maintainer crews—all of which have a bearing on the system design. The development contractor must have this type of information. The Operational Test Agency (OTA) commanders have agreed to support the Sustainment KPP by providing data to Service databases that will enable OT&E to inform estimates of system ownership cost. In a related action to assure T&E data is available to support these estimations, DOT&E worked closely with both the OTAs and the cost community, which resulted in the OTA commanders agreeing to provide test reports to the Service offices who estimate projected ownership costs.

In this issue we'll also get a report from Dr. Jim Forbes and Mr. Andy Long, of the Logistics Management Institute, on the studies DOT&E and AT&L co-sponsored on the relationship between reliability engineering investment and O&S costs. Based on the findings of the studies that increased reliability engineering investment can decrease O&S costs because of reliability improvement, LMI has developed a reliability investment model that will assist in determining how much investment in reliability is needed to achieve a desired amount of reliability improvement.

Using data from 17 projects, LMI developed an estimating relationship that appears to be valid across technologies, different types of weapon systems, and a wide range in complexity extending from components to subsystems to complete platforms. The 17 programs in the sample spanned from the early 1980s until after the turn of the Century.

In another article, Dr. Mike Cushing of ATEC's Army Evaluation Center (AEC) and Ms. Margaret Hockenberry of Army Materiel Systems Analysis Activity (AMSAA) will describe their efforts as part of the core team working with the Government Electronics and Information Technology Association (GEIA) and members from industry, to develop a reliability standard based on industry best practices. The standard will address those key activities that impact the reliability of systems, and will promote an integrated, systematic, and well planned reliability program. The Draft GEIA-REL-STD-0009 underwent formal review and balloting by the GEIA G47 Systems Engineering Committee for approval as a GEIA standard. The balloting process lasted 30 days and ended on June 30, with all members of the committee giving the "thumbs-up" to the standard, and one member abstaining.

We will also hear from Mr. David Nicholls of the Reliability Information Analysis Center (IAC) who will describe the capabilities of the center.

The studies and efforts described in this issue of *The ITEA Journal* are not the only ones that have capitalized on the momentum toward more traditional acquisition and T&E oversight, and the return to sound systems and reliability engineering concepts—but they are certainly among the most important. In fact, if you want to take a closer look at some of the key policy memorandums outlined in this issue of *The ITEA Journal*, as well as some other important policy memorandums that will improve the suitability and reliability of systems we send our warfighters, please go to http://www.acq.osd.mil/sse/dte/policy.html which AT&L has agreed to host on its website.

The key is to remember where we are headed—to the improved reliability of our defense systems—and that we can get there only if the acquisition and test communities work together to support sound systems and reliability engineering activities and those policy initiatives designed to get us there.

If we complete the journey toward greater suitability and reliability, we will see several new realities, including:

- Robust reliability programs that will be the standard practice on all major acquisition programs;
- Fewer Nunn-McCurdy breaches;
- Increased warfighter confidence as a result of increased availability, which will lead to increased mission success with reduced costs of ownership and logistics footprint;
- More effective government contracting through the use of incentives to achieve reliability and the sustainment needs of the warfighter;
- Industry recognition of high reliability as a competitive advantage and an incentive to employ reliability best practices to ensure customer success as well as market share;
- Institutionalized use of integrated developmental and operational reliability testing by Service and Joint programs, which will identify reliability problem areas early when redesign is cheaper and more effective;
- OSD and Service staffs with reconstituted reliability and maintainability positions to facilitate accomplishment of strong requirements and systems engineering at the front end of a program to produce better, more sustainable systems, more quickly and at reduced costs.

In the final analysis, we have little choice but to pursue these future realities, because we are operating in an increasingly constrained economic environment where costs are becoming more important, and because we have warfighters actively engaged in combat. Both conditions make increased suitability and reliability a "must have" and not an optional luxury. ❑

CHARLES E. MCQUEARY, Ph.D. was sworn in as director of Operational Test and Evaluation on July 27, 2006. A presidential appointee confirmed by the U.S. Senate, he serves as the senior advisor to the Secretary of

Defense on testing of Department of Defense weapon systems, prescribing policies and procedures for the conduct of operational and live fire test and evaluation. Prior to this appointment, Dr. McQueary was the first Under Secretary for Science and Technology at the Department of Homeland Security and confirmed by the U.S. Senate in March of 2003. In this position, he led the research and development arm of the department, utilizing the nation's scientific and technological resources to provide federal, state and local officials with the technology and capabilities to protect the homeland. Dr. McQueary is a former president of General Dynamics Advanced Technology Systems, Greensboro, North Carolina. He also has been president and vice president of business units for AT&T, Lucent Technologies, and director for AT&T Bell Laboratories. Early in his career at Bell Laboratories, Dr. McQueary served as head of the Missile Operations Department for the SAFEGUARD Antiballistic Missile Test Program, based at Kwajalein Atoll in the Marshall Islands. He later headed Bell Laboratories' Field Operations Department in Great Britain in support of a Navy oceanographic research station.

He also served as the director of Undersea Systems Development Lab. Dr. McQueary is a former executive board member of the National Security Industrial Association and the American Defense Preparedness Association (ADPA) (combined to form the National Defense Industrial Association [NDIA]). He is a past chairman of the Undersea Warfare Systems Division of ADPA and a former member of the Navy League Industrial Executive Board, the Navy Submarine League, the Electronics Industries Association, the American Society of Mechanical Engineers and the American Association for the Advancement of Science. He is also the recipient of the NDIA Homeland Security Leadership Award. Dr. McQueary is a graduate of The University of Texas, Austin, where he earned a Bachelor of Science degree in mechanical engineering; Master of Science degree in mechanical engineering; and a Ph.D. in engineering mechanics as a National Aeronautics and Space Administration (NASA) Scholar (Master of Science and Ph.D.) and member of five academic honor societies. The University of Texas has named Dr. McQueary a Distinguished Engineering Graduate.

Sponsored by the International Test and Evaluation Association

Register On-Line www.ITEA.ORG

SHORT COURSE

Introduction to the Net-Ready KPP, IA, E3 and Spectrum Management in Support of Systems Acquisition and Test

October 7 – 8 – 9, 2008 | Fairfax, Virginia

INSTRUCTOR: **Mr. Pete Christensen**, MITRE Corporation

This course is designed to provide an introduction to the Net-Ready Key Performance Parameter (NR-KPP) and Information Assurance (IA) as they relate to Information Technology (IT) and National Security Systems (NSS) and the DoD Systems Acquisition Process. The course will cover the NR-KPP, DOD IA Guidance, Security Certification and Accreditation (C&A), Electromagnetic Environmental Effects (E3) and Spectrum Management (SM) as they relate to assured information exchange. The course provides some background on the importance of information to warfare and how it has evolved into current concepts surrounding IO and IA. The course addresses the fundamentals of NR-KPP, IA, US Law and DoD Policy guiding requirements acquisition considerations. Test methodologies and metrics for evaluating NR-KPP, IA, E3 and SM as employed by operational test and evaluation organizations within the Department of Defense will be addressed in detail. The course will discuss NR-KPP, IA, E3 and SM considerations for systems throughout the entire lifecycle including requirements development, systems acquisition, Security Certification and Accreditation, developmental and operational testing. The class will also provide instruction for developing Test and Evaluation Master Plans, supporting test concepts and detailed test plans for IT and NSS. The last day will be used to provide a methodology for the system tester to be able to execute test concepts developed during the first two days.

WHO SHOULD ATTEND: The course is designed for contractor and government acquisition and testing professionals involved with the acquisition and test of Information Technology Systems, National Security Systems and Weapons Systems.

This course qualifies for Continuous Learning Points for acquisition workforce members.

More short courses will be available throughout the year. For more details visit **www.itea.org**

Invited Article
ITEA Journal 2008; 29: 215–221
Copyright © 2008 by the International Test and Evaluation Association

Defense Science Board Task Force Developmental Test and Evaluation Study Results

Pete Adolph
Task Force Chairman
Christopher DiPetto
Office of the Under Secretary of Defense for Acquisition,
Technology and Logistics,
Task Force Executive Secretary
Ernest Seglie, Ph.D.
Office of the Director,
Operational Test & Evaluation,
Task Force Executive Secretary

This article summarizes the results of a Defense Science Board (DSB) Task Force study of Developmental Test and Evaluation (DT&E) (Department of Defense, May 2008), which was conducted in 2007 and early 2008. The purpose of the study was to investigate the causal factors for the high percentage of programs entering Initial Operational Test and Evaluation (IOT&E) in recent years which have not been evaluated as both operationally effective and operationally suitable. The following is a summary of the specific issues which the Task Force was asked to assess:

- *Office of the Secretary of Defense (OSD) organization, roles, and responsibilities for Test and Evaluation (T&E) oversight. Recommend changes that may contribute to improved DT&E oversight, and facilitate integrated T&E;*
- *Changes required to establish statutory authority for OSD DT&E oversight. Recommend changes to Title 10 or other U.S. statutes that may improve OSD authority in DT&E oversight;*
- *Many IOT&E failures have been due to lack of operational suitability. Recommend improvements in DT&E process to discover suitability problems earlier, and thus improve likelihood of operational suitability in IOT&E.*

Key words: Acquisition Reform; acquisition workforce; developmental testing; integrated testing; operational reliability; suitability failure.

In recent years, there has been a dramatic increase in the number of systems not meeting suitability requirements during Initial Operational Test and Evaluation (IOT&E). Reliability, availability and maintainability (RAM) deficiencies comprise the primary shortfall areas. Department of Defense Initial Operational Test and Evaluation (DoD IOT&E) results from 2001 to 2006 are summarized in *Figures 1* through *3*. These charts graphically depict the high suitability failure rates during IOT&E resulting from RAM deficiencies.

Early in the Defense Science Board (DSB) study, it became obvious that the high suitability failure rates were the result of systemic changes that had been made to the acquisition process; and that changes in developmental test and evaluation could not remedy poor program formulation. Accordingly, the Task Force study was expanded to address the broader programmatic issues, as well as the above issues identified in the Terms of Reference (TOR).

A number of major changes in the last 15 years have had a significant impact on the acquisition process. First, Congressional direction in Fiscal Year (FY)

Program	Service	ACAT	IOT&E Result		Reason
FY 2001					
F-15 TEWS	USAF	II	Effective	Not Suitable	Reliability, Maintainability, Availability
V-22 Osprey	Navy	1D	Effective	Not Suitable	Reliability, Availability, Maintainability (RAM), Human Factors, BIT
Joint Direct Attack Munitions (JDAM)	USAF	1C	Effective only with legacy fuses	Not Suitable	Integration with delivery platforms
M2A3 Bradley Fighting Vehicle	Army	1D	Effective	Suitable	
FY 2002					
Joint Primary Aircraft Training System (JPATS)	USAF	1C	Effective with deficiencies	Not Suitable	RAM, Safety, Human Factors
Cooperative Engagement Capability (CEC)	Navy	1D	Effective	Suitable	
Multiple Rocket Launcher System (MLRS)	Army	1C	Effective	Suitable	
MH-60S	Navy	1C	Effective	Not Suitable	RAM, excessive administrative and logistic repair time impacted RAM
FY 2003					
B-1B Block E Mission Upgrade Program	USAF	1D	Effective	Not Suitable	16% decrease in weapons release rate, reduction in accuracy of Mark 82 low drag weapons, 14% hit rate on moving targets
Sea wolf Nuclear Attack Submarine	Navy	1D	Effective	Suitable	Several requirement thresholds were not met but overall system effective and suitable

Figure 1. DoD IOT&E results FY 2001–2003

1996, 1997, 1998, and 1999 Defense Authorization Acts reduced the acquisition workforce (which includes developmental test and evaluation). Several changes resulted from the implementation of Acquisition Reform in the late 1990s. The use of existing commercial specifications and standards was encouraged, unless there was justification for the use of military specifications. Industry was encouraged to use

Program	Service	ACAT	IOT&E Result		Reason
FY 2004					
Evolved Sea sparrow Missile	Navy	II	Effectiveness unresolved	Suitable	Testing was not adequate to determine effectiveness.
Stryker	Army	1D	Effective	Suitable	
Advanced SEAL Delivery System (ASDS)	Navy	1D	Effective with restrictions	Not suitable	Effective for short duration missions; not effective for all missions and profiles. Not suitable due to RAM.
Tactical Tomahawk	Navy	1C	Effective	Suitable	
Stryker Mortar Carrier-B (MC-B)	Army	1D	Effective	Not Suitable	RAM and safety concerns.
FY 2005					
CH-47F Block I	Army	1C	Effective	Not Suitable	RAM; communications system less suitable than CH-47D; did not meet Information Exchange Requirements for Block I.
F/A-22	USAF	1D	Effective	Not Suitable	RAM; needed more maintenance resources and spare parts; BIT
Joint Stand-Off Weapon-C	Navy	1C	Not Effective		Not effective against moderately hardened targets; mission planning time was excessive.
Guided-MLRS	Army	1C	Effective	Suitable	
High Mobility Attack Rocket System (HMARS)	Army	1C	Effective	Suitable	
V-22 Osprey	Navy	1D	Effective	Suitable	
EA-6B (ICAP III)	Navy	II	Effective	Suitable	

Figure 2. DoD IOT&E results FY 2004–2005

Program	Service	ACAT	IOT&E Result		Reason
CY 2006					
Common Missile Warning System (CMWS)	Army	1C	Effective	Suitable	Effective and suitable in the OIF/OEF environment but needs further testing outside of the OIF/OEF environment.
Deployable Joint Command and Control (DJC2)	Navy	1AM	Effective	Not Suitable	Operational Test Agency, COTF, reported effective, not suitable. BLRIP not complete.
Integrated Defensive Electronic Countermeasures	Navy	II			Test suspended due to reliability problems.
Surface Electronic Warfare Improvement Program (SEWIP) Block 1A	Navy	II	Not Effective	Not Suitable	Block 1A Upgrade does not make the AN/SLQ-32 EWS operationally effective and suitable but does enhance ability to protect ships
C-130J	USAF	1C	Effective single ship; Not effective in formation	Suitable with shortfalls	Effective single ship; not effective in formation air land / air drop; not effective in non-permissive threat environment. Shortfalls in suitability due to maintainability issues
Small Diameter Bomb (SDB) Increment 1	USAF	1D	Effective with limitations	Suitable with limitations	Limited effectiveness and suitability due to bomb rack reliability and deficiencies in software used to predict optimum fuzing solutions. Oct 2006 flight operations suspended

Figure 3. DoD IOT&E results for 2006

commercial practices. Numerous military specifications and standards were eliminated in some Service acquisition organizations. The requirement for a reliability growth program during development was also de-emphasized, and in most cases, eliminated. At the same time, systems became more complex, and systems-of-systems integration became more common. Finally, there was a loss of a large number of the most experienced management and technical personnel in government and industry without an adequate replacement pipeline. The loss of personnel was compounded in many cases by the lack of up-to-date standards and handbooks, which had been allowed to atrophy, or in some cases, were eliminated. It should be noted that Acquisition Reform included numerous beneficial initiatives. There have been many programs involving application of poor judgment in the last 15 years that can be attributed to acquisition/test workforce inexperience and funding reductions. It is probable that these problems would have occurred independently of most Acquisition Reform initiatives.

All Service acquisition and test organizations experienced significant personnel cuts, the magnitude varying from organization to organization. Over time, in-house DoD offices of subject matter experts (who specialized in multiple areas, such as promoting the use of proven reliability development methods) were drastically reduced, and in some cases, disestablished. A summary of reductions in developmental test personnel follows. The U.S. Army essentially eliminated their military developmental testing (DT) component and declared the conduct of DT by the government to be discretionary in each program. The U.S. Navy reduced their DT workforce by 10 percent but no shift of "hands-on" government DT to industry DT occurred. The trend within the U.S. Air Force gave DT conduct and control to the contractor. Air Force test personnel have been reduced by approximately 15 percent and engineering personnel supporting program offices have been reduced by as much as 60 percent in some organizations. The reduction of DT personnel in the Services occurred during a time when programs have become increasingly complex (e.g., significant increases in software lines of code, off-board sensor data integration, and systems of systems testing).

Principal findings and recommendations

RAM

As a result of industry recommendations in the early 1970s, the Services began a concerted effort to implement reliability growth testing as an integral part of the development process. This implementation consisted of a reliability growth process wherein a system is continually tested from the beginning of development, reliability problems are uncovered, and corrective actions are taken as soon as possible. The Services captured this practice in their reliability regulations, and the DoD issued a new military standard on reliability, which included reliability growth and development testing as a best practice task. The goal of this process from 1980 until the mid-1990s was to achieve good reliability by focusing on reliability fundamentals during design and manufacturing rather than merely setting numerical requirements and testing for compliance towards the end of development.

The general practice of reliability growth was discontinued in the mid to late 1990s, concurrent with the implementation of Acquisition Reform. This discontinuance may not be a direct result of Acquisition Reform, but may be related instead to the loss of key personnel and experience, as well as shortsighted attempts to save acquisition funds at the expense of increased life cycle costs. With the current DoD policy, most development contracts do not include a robust reliability growth program. The lack of failure prevention during design, and the resulting low initial Mean Time Between Failure (MTBF) and low growth potential are the most significant reasons that systems are failing to meet their operational reliability requirements.

Findings:

Acquisition personnel reductions combined with acquisition system changes in the last 15 years had a detrimental impact on RAM practices:

- With some exceptions, the practice of reliability growth methodologies was discontinued during system design and development (SDD);
- Relevant military specifications, standards, and other guidance were not used;
- Suitability criteria, including RAM, were de-emphasized.

- Improved RAM decreases life cycle costs and reduces demand on the logistics system.
- The deficiency report can be a valuable tool for early identification of RAM-related suitability problems, when used in conjunction with an adequately resourced deficiency correction system.

Recommendations:

The single most important step necessary to correct high suitability failure rates is to ensure programs are formulated to execute a viable systems engineering strategy from the beginning, including a robust RAM program, as an integral part of design and development. No amount of testing will compensate for deficiencies in RAM program formulation. To this end, the following RAM-related actions are required as a minimum:

- Identify and define RAM requirements during the Joint Capabilities Integration Development System (JCIDS) process, and incorporate them in the Request for Proposal (RFP) as a mandatory contractual requirement.
- During source selection, evaluate the bidders' approaches to satisfying RAM requirements:
  - Ensure flow-down of RAM requirements to subcontractors,
  - Require development of leading indicators to ensure RAM requirements are met.
- Make RAM, to include a robust reliability growth program, a mandatory contractual requirement and document progress as part of every major program review.
- Ensure that a credible reliability assessment is conducted during the various stages of the technical review process and that reliability criteria are achievable in an operational environment.
- Strengthen program manager accountability for RAM-related achievements.
- Develop a military standard for RAM development and testing that can be readily referenced in future DoD contracts.
- Ensure an adequate cadre of experienced RAM personnel are part of the Service acquisition and engineering office staffs.

Roles and responsibilities of government test and evaluation organizations

The traditional role of the government during the DT planning phase included the identification of the test resource requirements and government test facilities, the development of the test strategy and detailed test and evaluation plans, as well as the actual conduct of T&E. When a program moved from the planning phase to the test execution phase, the government traditionally participated in test conduct and analysis; performing an evaluation of the test results for the program office. With some exceptions, this is no longer the case. Until recently, it was recognized that there should be some level of government involvement and oversight even when the contractor has the primary responsibility regarding planning and execution of the DT program.

Findings:

The changes in the last 15 years, when aggregated, have had a significant negative impact on DoD's ability to successfully execute increasingly complex acquisition programs. Major contributors include massive workforce reductions in acquisition and test personnel, a lack of up-to-date process guidance in some acquisition

The ITEA Journal of Test and Evaluation jite-29-03-15.3d 21/8/08 12:28:34 218

organizations, acquisition process changes, as well as the high retirement rate of the most experienced technical and managerial personnel in government and industry without an adequate replacement pipeline.

- Major personnel reductions have strained the pool of experienced government test personnel.
- A significant amount of developmental testing is currently performed without a needed degree of government involvement or oversight and in some cases, with limited government access to contractor data.

Recommendations:

- As a minimum, government test organizations should develop and retain a cadre of experienced T&E personnel to perform the following functions:
  - Participate in the translation of operational requirements into contract specifications, and in the source selection process, including RFP preparation;
  - Participate in Developmental Test and Evaluation (DT&E) planning including Test and Evaluation Master Plan (TEMP) preparation and approval;
  - Participate in technical review processes;
  - Participate in test conduct, data analysis, and evaluation and reporting; with emphasis on analysis and reporting.
- Utilize red teams, where appropriate, to compensate for shortages in skilled, experienced T&E domain and process experts.
- Develop programs to attract and retain government personnel in T&E career fields so that the government can properly perform its role as a contract administrator and as a "smart buyer".

Integrated test and evaluation

Integrated testing is not a new concept within the Department of Defense, but its importance in recent years has been highlighted, due in part to the growth of asymmetric threats and the adoption of net-centric warfare. The December 2007 Office of the Secretary of Defense (OSD) Test and Evaluation Policy Revisions memorandum reinforces the need for integrated testing. Implementation of integrated test concepts has been allowed to evolve on an *ad hoc* basis. The time has come to pursue more consistency in integrated test planning and execution.

Collaboration between developmental and operational testers to build a robust integrated test program will increase the amount of operationally relevant data that can be used by both communities. DT and Operational Test (OT) planning is separate and this inhibits efforts by the Services to streamline test schedules, thereby increasing the acquisition timeline and program test costs.

DoD policy should mandate integrated test planning and execution on all programs to the extent possible. To accomplish this, programs must establish a team made up of all relevant organizations (including contractors, developmental and operational test and evaluation communities) to create and manage the approach to incorporate integrated testing into the T&E Strategy and the TEMP.

Findings:

- Service acquisition programs are incorporating integrated testing to a limited degree through varying approaches.
- Additional emphasis on integrated testing will result in greater T&E process efficiency and program cost reductions.

Recommendations:

- Implement OSD and Service policy mandating integrated DT&E/OT&E planning and execution throughout the program:
  - Require sharing and access to all appropriate system-level and selected component-level test and model data by government DT and OT organizations, as well as the prime contractor, where appropriate;
  - Integrate test events, where practical, to satisfy OT and DT requirements.

Operational test readiness review (OTRR)

Each Service has an operational test readiness review (OTRR) process. Although it varies from Service to Service, the process generally results in in-depth reviews of readiness to undergo an IOT&E event.

Findings:

- Shortcomings in system performance, suitability, and RAM are usually identified during the OTRR.
- In most cases, the operational test readiness certifying authority is well aware of the risk of not meeting OT criteria when major shortcomings exist.
- Because of funding constraints, the low priority given to sustainment, as well as the urgency in recent years to get new capabilities to the warfighter, major suitability shortcomings have rarely delayed the commencement of dedicated IOT&E.

Recommendations:

- Conduct periodic operational assessments to evaluate progress and the potential for achieving

predetermined entrance criteria for operational test events.
- Conduct an independent assessment of operational test readiness (AOTR) prior to the OTRR.
- Include a detailed RAM template in preparation for the OTRR.
- Require the Command Acquisition Executive (CAE) to submit a report to OSD that provides the rationale for the readiness decision.

OSD test and evaluation organization

The Task Force was asked to assess OSD roles and responsibilities for T&E oversight. T&E has been a visible part of OSD since the early 1970s, reporting to the Research and Engineering command section when it was in charge of acquisition oversight and subsequently to the Under Secretary of Defense for Acquisition (now AT&L). The early T&E office was responsible for all T&E, ranges, resources oversight, and policy. In 1983, Congress established an independent Director, Operational Test and Evaluation (DOT&E) organization, reporting directly to the Secretary of Defense (SECDEF), responsible for operational test and evaluation policy, budget review, and assessments of operational effectiveness and suitability. The Live Fire Test (LFT) oversight function was created and added to the DT&E office responsibilities in the mid-1980s. Later, the LFT oversight function was moved to the DOT&E organization.

In 1999, the DT&E organization was dismantled by DoD. Many functions were moved to DOT&E, including test ranges and resources, and joint T&E oversight. Some of the remaining T&E personnel billets were eliminated to comply with a congressionally mandated (AT&L) acquisition staff reduction. The residual DT&E policy and oversight functions were separated and moved lower in the AT&L organization.

A 2000 DSB Task Force Study on Test and Evaluation Capabilities recommended that DoD create a test and evaluation resource enterprise within the office of the DOT&E to provide more centralized management of T&E facilities. This recommendation ultimately led to removing the test ranges and resources oversight from DOT&E, abandoning the notion of centralized management, and the establishment of the Test Resource Management Center (TRMC) in AT&L (as directed by the National Defense Authorization Act for Fiscal Year 2003).

Findings:

Current policy as of December 2007 mandates that developmental and operational test activities be integrated and seamless throughout the system life cycle. There must be enough experts in OSD with the ability to understand and articulate lessons learned in early testing and the ability to execute the new T&E policy. That policy is to "take into account all available and relevant data and information from contractors and government sources" in order to "maximize the efficiency of the T&E process and effectively integrate developmental and operational T&E."

- Currently there is not an OSD organization with comprehensive DT oversight responsibility, authority, or staff to coordinate with the operational test office:
  - The historic DT organization has been broken up and residual DT functions were moved lower in the organization in 1999, and lower yet in 2002;
  - Programmatic DT oversight is limited by staff size and often performed by generalists vice T&E experts;
  - Recruitment of senior field test personnel is hampered by DT's organizational status;
  - Existing residual organizations are fragmented and lack clout to provide DT guidance;
  - System performance information and DT lessons learned across DoD have been lost;
  - DT is not viewed as a key element in AT&L system acquisition oversight;
  - Documentation of DT results by OSD is minimal.
- Access to models, data, and analysis results is restricted by current practice in acquisition contracting, and by the lack of expertise in the DT organization.
- TRMC has minimal input to program-specific questions or interaction with oversight organizations on specific programs.
  - Organizational separation is an impediment.

Recommendations:

- Implementation of integrated and seamless DT and OT will require, at a minimum, greater coordination and cooperation between all testing organizations.
- Consolidate DT-related functions in AT&L to help reestablish a focused, integrated, and robust organization:
  - Reestablish program oversight and policy, and Foreign Comparative Test (FCT);
  - Have Director, DT&E directly report to Deputy Under Secretary of Defense, Acquisition and Technology (DUSD[A&T]);
  - Restore TEMP approval authority to Director, DT&E.

- Integrate TRMC activities early into DT program planning:
  - Make TRMC responsible for reviewing the resources portion of the TEMP.
- If such an organization is established and proves itself effective, consider as part of a future consolidation moving LFT back to its original DT location.

The LFT change requires the concurrence of DOT&E and a legislative change to Title 10 because of the change in reporting official. All the other recommendations made throughout the report can be implemented within current DoD authority.

Other issues

Several other issues were addressed as a part of the study. A discussion of each of the following topics, along with findings and recommendations, may be found in the body of the report.

- Program Structure
- Requirements Definition
- Contractual Performance Requirements
- Alignment of DoD Technology with Systems Engineering Procedures
- Commercial Off-The-Shelf
- Systems of Systems

Summary and implementation status

In summary, the single most important step required to remedy the high suitability failure rates is to insure that programs are formulated to execute a viable systems engineering strategy from the beginning, including a robust RAM program, as an integral part of design and development. A second and related priority is to ensure that government organizations reconstitute a cadre of experienced T&E, engineering and RAM personnel to support the acquisition process. A third priority is to integrate developmental and operational testing to the extent practicable. A Reliability Improvement Working Group was established in March 2008 to address these three issues.

The reliability subgroup worked on developing a reliability acquisition policy and framework that includes consistent, concise sample RFP language that will encourage developers to plan for and resource a reliability growth program as a part of design and development; Phased templates to evaluate RAM activities throughout program reviews; and Standard evaluation criteria to provide a consistent way to evaluate an acquisition program's reliability health throughout the development process. On July 21, 2008 the Undersecretary of Defense for Acquisition, Technology, and Logistics signed a policy memo on Reliability, Availability, and Maintainability which implements the key RAM recommendations in the DSB report.

The personnel subgroup addressed four major issues: first, a policy to enable workforce reconstitution; second, a plan to reconstitute RAM and T&E personnel where necessary, third, training and education for RAM and T&E personnel; and fourth, establishing and staffing Centers of Excellence and expertise.

The integrated testing subgroup developed guidelines for early involvement in requirements and RFP development; contractual language for data access and sharing; and synchronization of the Test and Evaluation Master Plan (TEMP) and Systems Engineering Guide.

Further implementation of these and other recommendations in the report will not be easy, but will pay large dividends in improvements to the acquisition process and reduced life cycle costs. ❑

Pete Adolph has over 45 years experience in test and evaluation and systems acquisition. Following three years as an Air Force officer, he held a variety of positions with the Air Force from 1960 to 1987, advancing to technical director at the Air Force Flight Center. From 1987 to 1994, he held several positions in the Office of the Secretary of Defense (OSD). For most of that period, he was director, Test and Evaluation, Acquisition and Technology. He also served as interim director of Operational Test and Evaluation and interim director of Defense Research and Engineering. He was a senior vice president for SAIC from 1994 to 2000 and served as the manager of the SAIC test and evaluation group. He is currently a consultant.

Chris DiPetto is deputy director, Developmental Test & Evaluation (DT&E), Office of the Under Secretary of Defense for Acquisition, Technology and Logistics.

Dr. Ernest Seglie is science advisor, Office of the Director, Operational Test and Evaluation (DOT&E), the Pentagon, Washington, D.C. He provides scientific and technical guidance on the overall approach to Department of Defense (DoD) evaluation of the operational effectiveness and suitability of major DoD weapon systems; provides technical review of test reports; and serves as chief technical advisor to the Director, DOT&E.

References

Department of Defense. May 2008. "*Report of the Defense Science Board Task Force on Developmental Test and Evaluation*." Office of the Under Secretary of Defense for Acquisitions, Technology, and Logistics. The complete DT&E report is available on the Defense Science Board website at: http://www.acq.osd.mil/dsb/reports/2008-05-DTE.pdf.

2009 ITEA Journal *Themes*

The ITEA Publications Committee has established themes for the 2009 issues of *The ITEA Journal* and invites articles in the following areas:

Test and Evaluation of Highly Complex Systems (March issue). Complex systems embody the idea that the whole is greater than the sum of the parts – the large scale behavior cannot be predicted from knowledge of the individual constituents. Many of these systems are network enabled and must be tested in the presence of the network effects, including interactions with other networks, with services, and with applications. Service-oriented architectures are being integrated with legacy systems and sorting out behaviors and performance requires insight beyond the system under test. The human element – the cognitive domain – is an essential piece and is the place where perception, awareness, understanding, beliefs and values reside. This issue casts a broad net that our current test environment is being asked to accommodate: joint and distributed testing, chaos and complexity theory, emergent behaviors, virtual testing, modeling and simulation, cognition and autonomy, and assessing mission effectiveness from individual human, component, and system performance. ***(Manuscript deadline: December 1, 2008)***

The Future of Test Facilities (June issue). Shrinking RDT&E budgets signal a coming reduced workload for test facilities. To remain viable, commercial organizations seek to expand the use of facilities by redefining their markets, for example, aerospace wind tunnels are used for flow around cars, buildings, downhill skiers, and bobsledders. Government programs seek economies by partnering with industry and academia, overlapping developmental and operational testing, and attempting a peaceful co-existence between test and training. Opinions are sought on both sides of the test and training issue – what can and should be done and what can never be – as well as lessons learned from past efforts. Ideas are welcome on the synergy and benefits of other common use of test facilities: non-Department of Defense testing, non-traditional testing, extending the customer base, inclusion of experimentation, expanded use of commercial and academic test capabilities, and collaborating with international partners. ***(Manuscript deadline: March 1, 2009)***

Integrating Test and Evaluation (September issue). In December 2007 the Department of Defense issued new policies that represent a shift in emphasis from test to evaluation and promote an emphasis on integrated test and evaluation throughout the system life cycle in a seamless continuum. This issue addresses implementation and follows up the new policies to examine integrating contractor, developmental, operational, and live fire testing and the renewed role of developmental evaluation. The issue also examines the ethics and obligations of test and evaluation. Test and evaluation exist to serve the customer and must coach the customer to ask the right questions, must report the truth, report in a timely manner, and report assertively to ensure appropriate attention is paid. Questions arise from using data from Iraq as test data: how do we instrument systems to get data? How do commercial entities get customer data? How do non-military government organizations get data in areas for which they are responsible? The issue also asks: what are the impediments to realizing integrated test and evaluation and what are the limitations of doing so? ***(Manuscript deadline: June 1, 2009)***

Air & Space (December issue). The year 2009 marks the 40th anniversary of the first moon walk by Neil Armstrong. The first powered flight at Kitty Hawk, North Carolina occurred 106 years ago in December. Today space is more than exploration and air is more than airplanes. Earth is blanketed by countless satellites viewing, recording and communicating; the international space station is an orbiting laboratory; government and private organizations are pursuing commercial access to space. The national and international airspace is pushed to record densities at a time when unmanned air vehicle use is booming, and the Federal Aviation Administration is being asked to incorporate these unpiloted and remotely piloted flying creatures into the soup. This issue takes a retrospective look at how we arrived here, where technology is taking us, and the demands that will be placed on test and evaluation. Air and space constitute the realm of rockets, missiles, weapons, satellites, aircraft of every pedigree, transportation, intelligence, sensors, communications, hypersonics, and so much more. ***(Manuscript deadline: September 1, 2009)***

In addition: T&E articles of general interest to ITEA members and *ITEA Journal* readers are always welcome. Each Issue includes specialty features, each 2-3 pages long: **"Featured Capability"** describes unique, innovative capabilities and demonstrates how they support T&E; **"Historical Perspectives"** recall how T&E was performed in the past, or a significant test or achievement, often based on personal participation in the "old days" of T&E.; **"TechNotes"** discusses innovative technology that has potential payoff in T&E applications or could have an impact on how T&E is conducted in the future. **Interested authors**: should submit contributions to the **ITEA Publications Committee Chairman (itea@itea.org, attn.: Dr. J. Michael Barton)**. Detailed Manuscript Guidelines can be found at **www.itea.org** under the ITEA Publications tab.

ITEA Journal 2008; 29: 223–224
Copyright © 2008 by the International Test and Evaluation Association

Developmental Test and Evaluation Role in Assessing System Reliability

Christopher DiPetto
Deputy Director,
Developmental Test and Evaluation, Washington, D.C.

In April 2007, the Under Secretary of Defense for Acquisition, Technology, and Logistics [USD (AT&L)] requested the Chairman of the Defense Science Board (DSB) to establish a task force on developmental testing. A subject of concern was that although several initiatives had been implemented within the Department to stress better system engineering practices, recent history shows that poor performance during initial operational test and evaluation (IOT&E) "suggests deficiencies in developmental test & evaluation (DT&E) processes." One specific problem identified was the lack of focus on reliability, availability, and maintainability (RAM) during system design. The Under Secretary requested the DSB address this problem and develop recommendations to improve the Department's DT&E processes. The goal is to identify suitability problems early enough to change the system design while in development versus retrofitting it after it has performed poorly in IOT&E.

The Defense Science Board (DSB) Task Force finished their study in early 2008. The DSB's findings concluded that systemic changes to acquisition processes and a lack of a disciplined systems engineering process have resulted in the high failure rates in suitability. In February 2008, the Deputy Under Secretary of Defense (Acquisition and Technology) and the Director, Operational Test and Evaluation (DOT&E) established the Reliability Improvement Working Group to implement three specific recommendations of the DSB: (a) ensure programs are structured with a viable systems engineering strategy to include a reliability, availability, and maintainability (RAM) growth program as an integral part of design and development, (b) reconstitute a cadre of personnel within the Department and the Services with training and experience in test and evaluation (T&E) and RAM, and (c) implement the Office of the Secretary of Defense (OSD) policy to integrate developmental and operational testing.

Mr. Christopher DiPetto

In July 2008, the RIWG recommended, and the USD (AT&L) issued, a new RAM policy to ensure RAM requirements are incorporated into development contracts and system designs, and evaluated in each phase of the acquisition life cycle (available at http://www.acq.osd.mil/sse/dte/docs/USD-ATLMemo-RAM-Policy-21Jul08.pdf). In addition to OSD policy, there are several initiatives within the Services to address RAM during the development phases of acquisition. In late 2007, the Army published a new policy mandating programs establish a reliability threshold before entrance into Milestone B. This new policy requires the threshold be incorporated into the system design and development (SDD) contract. Additionally, the system is expected to meet or exceed the threshold value for reliability by the conclusion of the first system-level test in SDD. The other Services are also assuring the proper policies are in place to focus on reliability during system development. Currently, both the Air Force and Navy require the system developer to address the requirements for reliability during system design as part of the SDD contract.

Another area addressed was the reduction in personnel with experience in T&E and RAM backgrounds. In the late 1990s, Congress directed several cuts to the military's acquisition workforce. These reductions, according to the DSB report, "put the DoD acquisition workforce on a precipitous path" to losing vital technical expertise, while at the same time, our weapon systems are becoming more complex. As one DSB member put it: "We went from Insight, to Oversight to Out of Sight."

The Services are reassessing their acquisition manpower allocations to ensure there is the proper focus on growing the experience levels of the T&E and RAM workforce. Additionally, the Defense Acquisition University has taken the initiative to examine their course curriculums to ensure our workforce is properly trained to employ sound T&E and RAM principles during system development.

The final focus area is the implementation of an integrated test policy within the Department. Integrated testing is the process of collaborative planning and execution of test phases and events to ensure the objectives of the stakeholders (both operational and developmental) are addressed. One of the primary purposes of T&E is to ensure that the system, as designed, will meet the warfighter's requirements in the operational environment. Operational Test and Evaluation (OT&E) evaluates the operational effectiveness and operational suitability of the design. By the time a system is ready for OT&E the design is pretty much fixed—it's too late to make major changes. One of the fundamental focuses of developmental test & evaluation (DT&E) is to test and evaluate *the system design* to ensure it will meet the warfighter's requirements. System developers and the DT&E community use the Joint Staff validated capability requirements documents as the source for system requirements. However, as is often the case, the Concepts of Operation (CONOPs) is not made available (or is not used) during SDD to help define the scope of developmental tests. Therefore, the system is not properly stressed during SDD when there is still time to make design changes. Integrated testing brings a "mission-oriented" approach during DT&E by getting all the team players involved in system development (contractor, program office, user, developmental test, and operational test) to incorporate the mission context into the developmental test strategy. That way, the system design can be tested based on how the system will be employed. Not taking the operational environment into account during development is akin to an automobile manufacturer building a half-ton pickup not knowing that the user needs a four-wheel drive truck. Both vehicles carry the required tonnage, but how it will be used (off-road) was not taken into account when the truck was on the drawing board. By using the CONOPs during the development of the statement of work, the system developer and manufacturer would have known the truck would be employed on rough terrain and would have incorporated a robust suspension system, four-wheel drive, etc. into the design. The vehicle would have then been tested in an off-road environment before being delivered to the customer. Although much more complicated, the same principle applies to weapon systems. Additionally, by infusing a mission-oriented approach during DT&E, data that is operationally representative can be used to reduce the scope of initial operational test and evaluation. Integrated Testing should not only save time and money in the test program, but the real savings will be in the dollars and time saved by less redesigns and retrofits after the system is in production. It's a "win-win" proposition. ❑

ITEA Journal 2008; 29: 225–227
Copyright © 2008 by the International Test and Evaluation Association

The Simulated Infrared Earth Environment Laboratory, an Infrared Signature Resource to Support Test & Evaluation

Jeffrey S. Sanders, Ph.D.
Trideum Corporation, Huntsville, Alabama

Thermal sensors are highly prevalent on the modern battlefield and test and evaluation (T&E) of these sensors requires thermal infrared (IR) signature representations of threat systems. These signature requirements range from user training and model inputs for digital sensor simulations, to full-scale targets that are deployed to test ranges for destructive testing of weapon systems. IR signature data for threat systems is not something that is normally readily available to a tester or evaluator, but there are resources for this type of data for the T&E community.

One such resource is the Simulated Infrared Earth Environment Laboratory (SIREEL) website. The SIREEL website was created for support of gunner training. The SIREEL website contains extensive infrared signature data on numerous threat and friendly vehicles and the site is designed to provide country-specific vehicle identification training in support of U.S. military deployments. However, the information on this website and the models used to generate content for the website can be utilized to support T&E of emerging sensor technologies. *Figure 1* shows an example of the SIREEL website structure and sample content.

It is important to note that SIREEL is not just a website, it is also an IR signature modeling program. The core component of the SIREEL website is its System Index which organizes IR signature data for over 200 foreign systems. This online encyclopedic IR signature reference system contains both measured and synthetically generated IR signature data. The synthetic data is generated with predictive, physics-based IR signature models that allow a user to generate a vehicle signature under any operational and environmental state.

The IR signature of a ground vehicle is highly dimensional, and predictive, physics-based modeling offers an approach to providing IR signatures to support a wide range of T&E activities. It is simply not possible to measure the IR signatures of threat ground vehicles in sufficient quantity and fidelity to support all of the needs of the T&E community. Some reasons for this are:

- many threat systems required for T&E are not available for signature measurements or range testing;

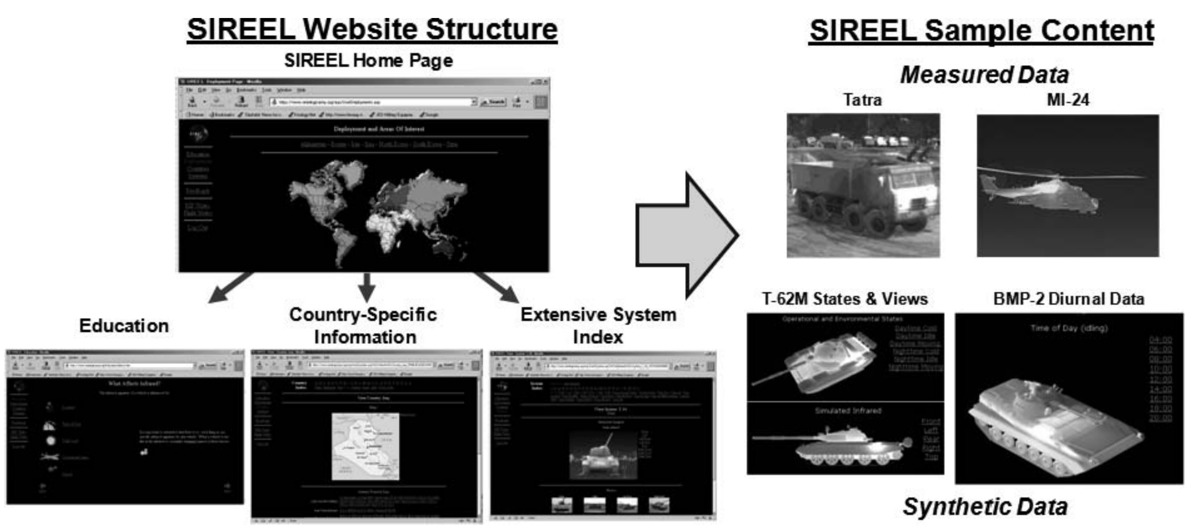


Figure 1. Simulated Infrared Earth Environment Laboratory (SIREEL) website structure and sample content

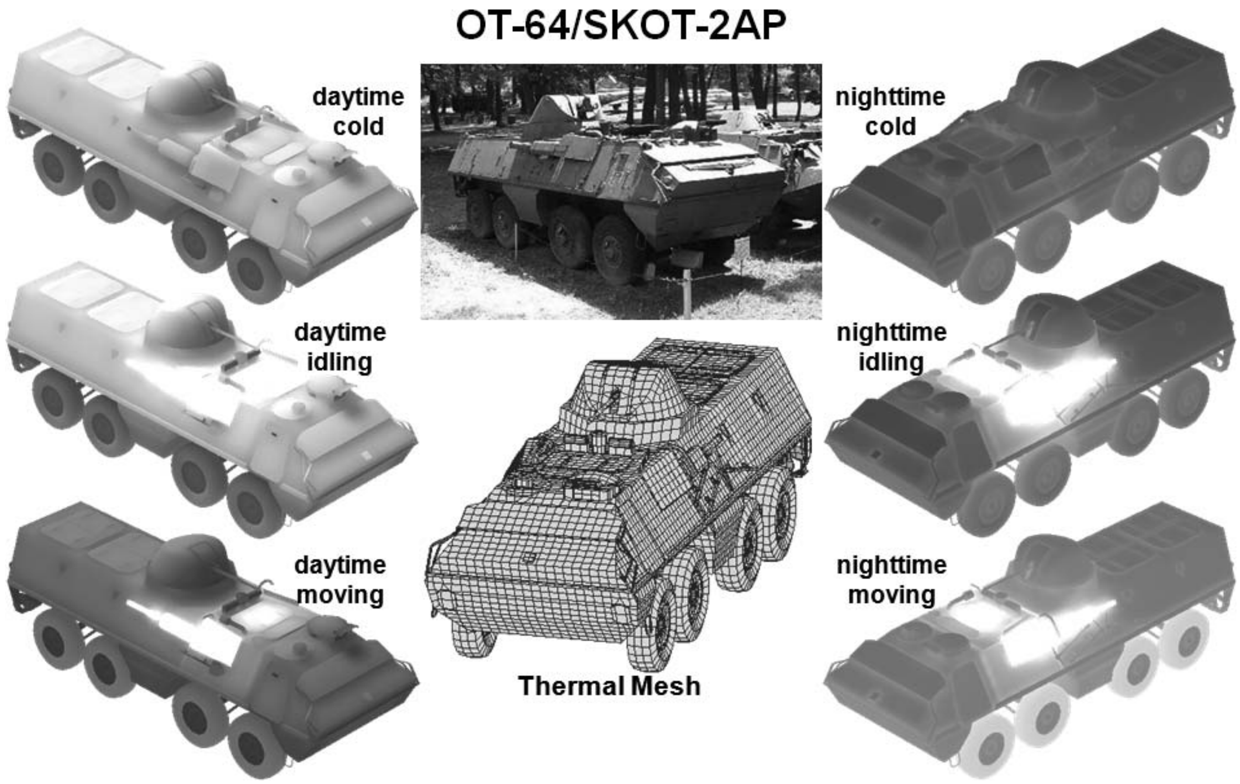


Figure 2. Example Simulated Infrared Earth Environment Laboratory (SIREEL) signature model outputs. Daytime signatures are shown on the left, nighttime signatures on the right, and cold, idling, and exercised signatures going from top to bottom. Hot, engine-related target components are apparent in the idling and exercised cases and tire heating and hull cooling is shown in the exercised signatures.

- IR signature data supporting simulations must be environmentally correlated (same place & time) for multi-target scenarios and measuring data in this manner is challenging and expensive;
- the IR signature of a ground vehicle is affected by so many different factors that it is simply not possible to capture *every* signature state.

Predictive IR signature modeling can fill in these signature requirements gaps and the SIREEL program has an archive of over 200 signature models that are available to support T&E needs. The signature data on this website and the models used to create the data have been used to provide IR signature design support for multiple range target development programs. These include the Threat Vehicle Surrogate Target program, the Realistic Low Cost Targets program, and the Precision Target Signatures programs. Literally thousands of users have gained access to the SIREEL website since it was launched. While it is not practical to track the end use of the IR signature data obtained by these users from the website, it is known that multiple T&E target-related programs in addition to those mentioned above have benefited from the SIREEL program.

Example SIREEL signature model outputs are shown in *Figure 2*. This figure demonstrates the six fundamental signature states of a ground vehicle, the environmental states of daytime and nighttime, and the operational states of cold (engine off), idling, and exercised. This is a *gross simplification* of the dimensionality of ground vehicle IR signatures, but it demonstrates the majority of a target's critical IR signature features.

Creating this type of signature model requires the following steps:

- building a geometry model in a computer-aided design program;
- converting the geometry into a thermal mesh suitable for accurate temperature calculations;
- segmenting and attributing the thermal mesh with thermal properties;
- developing the algorithms to calculate engine-related heat transfer;
- executing a scenario specific simulation in an IR signature code such as the Multi-Service Electro-Optic Signatures (MuSES) code.

This is a limited example of the power and flexibility of predictive IR signature modeling to support T&E by filling in the IR signature gaps for emerging IR sensor technologies. The SIREEL Program has a large archive of these models and they are available to end users for generating signature data for a wide range of T&E applications. ❑

JEFFREY S. SANDERS is a principal engineer at Trideum Corporation in Huntsville, Alabama. He has managed the IR signature modeling component of the SIREEL program since its inception. Dr. Sanders received a Bachelor of Science in electrical engineering, Master of Science in electrical engineering, and a Ph.D. in engineering from Memphis State University in 1987, 1990, and 1994, respectively. E-mail: jsanders@trideum.com

Acknowledgments

The SIREEL point of contact is Daniel J. Rinald (daniel.j.rinald@us.army.mil). Access to the SIREEL website, which is for official Department of Defense (DoD) users only, can be requested at https://sireel.ngicarmy.org. (A .gov or .mil domain is required to request a login and for access.)

Sponsored by the International Test and Evaluation Association

Register On-Line www.ITEA.ORG

SHORT COURSE

Fundamentals of the Test and Evaluation Process

October 15 – 16 – 17, 2008 | Fairfax, Virginia

INSTRUCTORS: **Lawrence M. Damman** and **Matthew T. Reynolds**

This three-day intensive course will describe how to plan and execute a major test and evaluation program. Industry and government experience over the last two decades has resulted in a set of guidelines and process that are vital to ensuring the success of T&E as a timely indicator of progress and the achievement of system performance objectives during a development program. Much of this experience is being applied in military systems development programs today through updated Defense Department policy, procedure, and guidance documents. Although this course will focus on application of this experience in the Defense Department, the principles apply equally well to their complex systems and products such as commercial aircraft, automobiles, and communications and control networks. Past course participants have included representatives of the Departments of Energy and Transportation, including the Coast Guard and Federal Aviation Administration.

This course covers the role of T&E in system development, the determination of appropriate test requirements, developmental and operational T&E, the T&E Master Plan, writing T&E requirements in contracts, and the role of modeling and simulation in T&E. Taught by seasoned professionals with up-to-date knowledge, the course will consist of lectures and case studies, as well as individual and group exercises. Each student will receive a notebook that includes the presentation material with a narrative text, excerpts from selected publications, and a certificate of completion.

WHO SHOULD ATTEND: The course is designed to meet the needs of professionals who want a current overview of T&E with sufficient insight into major technical and organizational issues. The lectures are designed for executives and managers responsible for testing by government agencies, engineering centers, research laboratories, industry, and academia; for project managers; for test directors and analyst; and for those responsible for managing test facilities.

This course qualifies for Continuous Learning Points for acquisition workforce members.

More short courses will be available throughout the year. For more details visit www.itea.org

Will it work for the Warfighter?

Ask SAIC. We take it personally.

Supporting the Warfighter is a job our 44,000 employees take to heart every day.
For more information, call

SAIC T&E Support
Military Utility Assessment, Operational Test, Distributed Test,
Joint Test & Evaluation Support, T&E Infrastructure
Contact: Mark Smith
mark.e.smith-2@saic.com
(505) 830-6757

SAIC.com

© 2007 Science Applications International Corporation. All rights reserved.

SAIC®
From Science to Solutions

ITEA Journal 2008; 29: 229–231
Copyright © 2008 by the International Test and Evaluation Association

Test, Evaluation, and Pogo Problems on Saturn V[1]

J. D. Hunley, Ph.D.

Invariably, rocket engineers seek to discover and fix potential problems before their vehicles reach the flight-testing stage. Typically, they have done so by examining reports about previous rockets and by extensive ground testing.

In the case of the Saturn V rocket (Figures 1 and 2) (that ultimately launched 12 astronauts to the Moon), the von Braun team at NASA's Marshall Space Flight Center was aware that the huge vehicle was potentially subject to longitudinal oscillations (known as the pogo effect). This phenomenon had come "into general attention in the early days of the Gemini program," when it manifested itself on the Titan II missile used to launch spacecraft and astronauts into orbit. The von Braun team tested and analyzed the Saturn V for possible pogo problems before its initial flight test (AS-501) and found "an acceptable margin of stability," suggesting there would be no significant oscillations. On November 9, 1967, the giant Saturn V did, however, exhibit some low-level pogo oscillations on this first flight test, but each of the first-stage F-1 engines experienced them "at slightly different points in time" so that the overall effect was not severe and tended to confirm the von Braun team's earlier analysis of stability. (Hunley 2008a and 2008b)[2]

Overall, AS-501 was a success. Yet, despite the von Braun team's ground testing, the second Saturn V flight test—AS-502 on April 4, 1968—proved to be a technical failure. Among its problems was a much more severe pogo effect than had occurred on AS-501. Toward the end of the first-stage burn, the oscillations exceeded 5 hertz (cycles per second), beyond the design specifications and a reason for alarm (Bilstein 1980).[3]

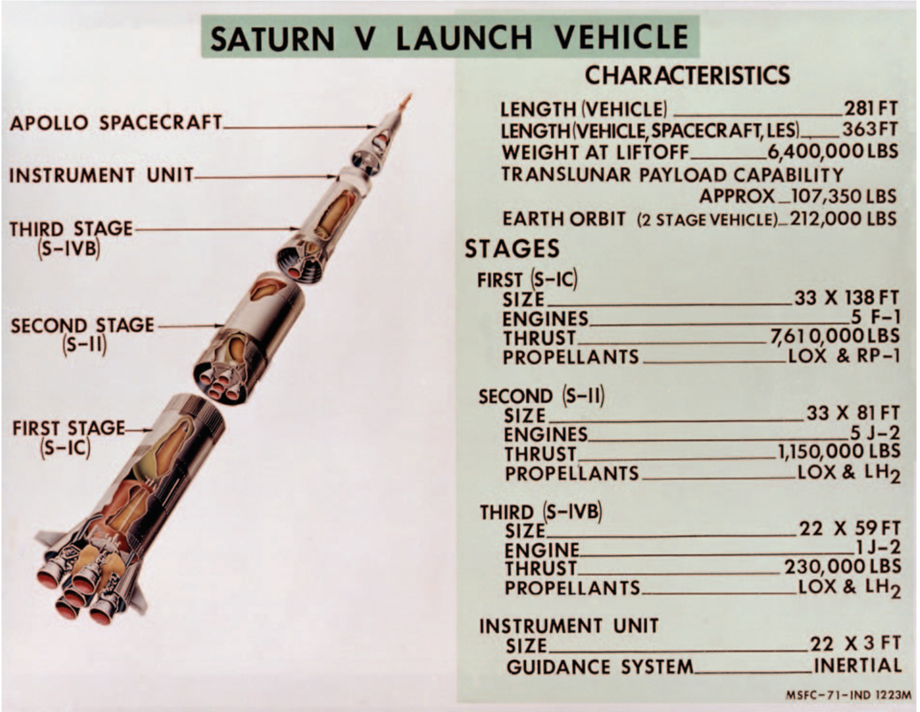


Figure 1. Diagram of a Saturn V launch vehicle's components and characteristics. Courtesy of NASA

Figure 2. A Saturn V launching from Kennedy Space Center, Florida. NASA Historical Reference Collection, Washington, D. C. Courtesy of NASA

On that flight, the pulsations of all five 1.5-million-pound F-1 first-stage engines "came into a phase relationship" where "the engine pulsation was additive." At the same time, the entire vehicle developed a bending frequency that increased as it consumed propellants, reaching 5.25 hertz about 125 seconds into the flight. The engine vibrations traveled longitudinally up the vehicle structure, with their peak occurring at the top, where the spacecraft would carry astronauts on future flights. By themselves, the engine vibrations would not have posed a problem, but coupled with the vehicle's oscillations (bending frequency), which moved in a lateral direction, they became a serious issue. When the two types of vibrations intersected, with both at about the same frequency, their effects combined and multiplied. In the words of Apollo Program Director Samuel Phillips, the "complicated coupling" of the two frequencies was "analogous to the annoying feedback squeal you encounter when the microphone and loud speaker of a public address system are coupled." The combined effect could have interfered with an astronaut's performance, a significant concern (Bilstein 1980)

To analyze and correct this problem, NASA created a pogo task force that included experts from Marshall, other NASA organizations, contractors, and universities. The solution this group recommended was to detune the five engines, changing the frequencies of at least two of the five so that they would no longer produce simultaneous vibrations. Engineers did this by inserting liquid helium into a cavity formed in a liquid oxygen prevalve in which a casting bulged out and encased an oxidizer feed pipe in each F-1 engine. The bulging portion was only half filled with the liquid oxygen during engine operation. The helium would absorb pressure surges in oxidizer flow—the cause of the oscillations—and thereby reduce the frequency of the oscillations to 2 hertz, lower than the frequency of the structural oscillations. Engineers eventually applied the solution successfully to all four outbound F-1 engines that encircled a single central engine in a cluster of five (Figure 3). Marshall, Boeing, Martin, TRW, the Aerospace Corporation, and North American's Rocketdyne Division contributed the technical talent that solved the dilemma (Bilstein 1980) (Dunar and Waring 1999).[4] This incident illustrates the large number of people and organizations needed to solve technical problems like this one (Figure 4). It also shows how difficult it was for rocket designers to predict when and how phenomena such as pogo might occur, even while aware of and actively testing for them. The von Braun team was well known for its extensive ground testing, (Bilstein 1980) (Neufeld 2007) yet it was still unable to predict the extent of the Saturn V pogo problem. Like rocket design itself, testing and evaluation was not a clear-cut science. It lacked firmly established rules capable of eliminating problems before actual flight testing. Rather, it was a complex part of an engineering culture involving analysis and testing that required a good deal of art, coupled with scientific methods. This culture was successful in solving the pogo problem on the Saturn

Figure 3. Saturn V first stages in NASA's Michoud Assembly Facility near New Orleans, showing the five clustered F-1 engines

Figure 4. Key mission officials for NASA's Apollo program in the Launch Control Center. The four in the middle foreground of the photograph are, from left to right, Charles W. Mathews, deputy associate administrator for manned space flight; Dr. Wernher von Braun, director of Marshall Space Flight Center; Dr. George E. Mueller, associate administrator for manned space flight; and Lt Gen Samuel C. Phillips, director of the Apollo program. Courtesy of NASA.

V, but it took an expensive flight test to reveal its full extent. Until the data from the flight test were available, engineers were simply unaware of the extent of the problem and the steps that would be needed to fix it. ❑

DR. J. D. HUNLEY was chief historian of NASA Dryden Flight Research Center before his retirement in 2001. He has written a two-volume overview of U.S. missile and rocket technology, the first entitled Prelude to U.S. Space-Launch Vehicle Technology: Goddard Rockets to Minuteman III; and the second, U.S. Space-Launch Vehicle Technology: Viking to Space Shuttle (both, Gainesville, Florida: University Press of Florida, 2008) among many other books and articles on aerospace history. Besides working as a historian for the U.S. Air Force and NASA during his lengthy career, Hunley has also taught at Allegheny College and served as a Ramsey Fellow at the Smithsonian National Air and Space Museum. E-mail: dillh@roadrunner.com

Notes

[1]This article and its illustrations are adapted from J.D. Hunley, *U.S. Space-Launch Vehicle Technology: Viking to Space Shuttle* (Gainesville, Florida: University Press of Florida, 2008), esp. pp. 195, 210–212, 214, 216, 219. Reprinted with permission of the University Press of Florida.

[2]U.S. Congress, Senate, "Summary of the Problems Encountered in the Second Flight of the Saturn V Launch Vehicle," Hearings before the Committee on Aeronautical and Space Sciences, 90th Cong., end sess., April 22, 1968, pp. 6, 16–17, including quotations from the testimony of Samuel C. Phillips; NASA, "Saturn V News Reference," MSFC, KSC, and contractors, August 1967 (portions changed December 1968), <history.msfc.nasa.gov/Saturn_apollo/saturnv_press_kit.htm>, pp. 12-1 to 12-2; U.S. President, *United States Aeronautics and Space Activities, 1967: Report to the Congress from the President of the United States* (Washington, D. C.: GPO, 1968), p. 131. For background on the pogo problems in the Titan II, see , J. D. Hunley, *Preludes to U.S. Space-Launch Vehicle Technology: Goddard Rockets to Minuteman III* (Gainesville, Fla.: University Press of Florida, 2008), pp. 269-73.

[3]Senate, "Summary," 6, 16–17, including first two quotations; Samuel C. Phillips, draft article for *New York Times*, 1969, Samuel C. Phillips Collection, box 118, folder 4, "Manned Lunar Landing Program; Writings by Phillips," Library of Congress, Manuscripts Division (LC/MD), for last two quotations; NASA, "Saturn V Tests Completed," News Release 68-128, July 18, 1968, Folder 010450, "Saturn V, General (1965–1969)," NASA Historical Reference Collection, NASA Headquarters, Washington, D.C.

[4]Senate, "Summary," pp. 18, 22–27; NASA, "Saturn V Tests Completed," pp. 2–4, 5A; "Minutes of the Saturn V Semi-DCR/AS-502 Evaluation," April 21, 1968, p. 3, George E. Mueller Collection, Box 92, folder 8, LC/MD.

References

Bilstein, R. 1980. *Stages to Saturn: A Technological History of the Apollo/Saturn Launch Vehicles*. NASA SP-4206, 1980. Washington, D. C.: NASA, pp. 5, 16–17, 27–28, 183, 360–363.

Dunar, A. J. and Waring, S. P. 1999. *Power to Explore: A History of Marshall Space Flight Center, 1960–1990*. NASA SP-4313, 1999. Washington, D. C.: NASA, p. 97.

Hunley, J. D. 2008a. *U.S. Space-Launch Vehicle Technology: Viking to Space Shuttle*. Gainesville, Fla.: University Press of Florida, pp. 195, 210–212, 214, 216, 219.

Hunley, J. D. 2008b. *Preludes to U.S. Space-Launch Vehicle Technology: Goddard Rockets to Minuteman III*. Gainesville, Fla.: University Press of Florida, pp. 269–273.

Murray, C. and Bly, C. 1989. *Apollo: The Race to the Moon*. New York: Simon and Shuster, p. 312.

Neufeld, M. J. 2007. *Von Braun: Dreamer of Space, Engineer of War*. New York: Alfred A. Knopf, p. 388.

U.S. President. 2968. *United States Aeronautics and Space Activities, 1967: Report to the Congress from the President of the United States*. Washington, D. C.: Government Printing Office (GPO), p. 131.

ITEA Journal 2008; 29: 232–239
Copyright © 2008 by the International Test and Evaluation Association

The Reliability Information Analysis Center (RIAC)—The Department of Defense Center of Excellence in Reliability Engineering

David Nicholls
RIAC Operations Manager, Utica, New York

The Reliability Information Analysis Center (RIAC), formerly known as the Reliability Analysis Center (RAC), serves as a Department of Defense (DoD) information analysis center (IAC). As an IAC, the RIAC is a Center of Excellence and technical focal point for information, data, analysis, training, and technical assistance in the engineering fields of reliability, maintainability, quality, supportability, and interoperability (RMQSI).
Chartered by the DoD, the RIAC also undertakes a variety of other government and industrial support projects each year. The team is led by Wyle Laboratories and includes Quanterion Solutions Incorporated (QSI), the University of Maryland Center for Risk and Reliability, the Pennsylvania State University Applied Research Laboratory (ARL), and the State University of New York Institute of Technology (SUNYIT). Resources at the nationwide locations include over 2,000 employees of various technical backgrounds and expertise. The RIAC operation serves as a model of successful government–industry–academia cooperation and collaboration in addressing the technical needs of both the defense community and commercial industry.

Key words: Failure; maintenance; reliability analysis.

November 14, 2007 marked the ribbon-cutting at the State University of New York Institute of Technology (SUNYIT) to celebrate the "Grand Reopening" of the Department of Defense (DoD) Reliability Information Analysis Center (RIAC). The event recognized the 40th anniversary of this critical Information Analysis Center (IAC). Prior to 2005, the RIAC was known as the Reliability Analysis Center (RAC), but the name was changed to emphasize its role as part of the IAC family of Defense Technical Information Center (DTIC)-sponsored centers. Since its start in 1968, the Center's technical mission has evolved from the reliability of emerging microelectronics devices to addressing full system RMQSI. In the course of doing so, the Center's products and services have expanded from raw failure data compilations to a variety of state-of-the-art RMQSI-based reports, guidance publications, software tools, training courses, and technical and administrative consulting capabilities. Table 1 indicates some of the major milestones in RIAC/RAC history.

The RIAC team

Wyle Laboratories

The Huntsville, Alabama facility of Wyle Laboratories serves as the prime contractor for the RIAC program. The main focus of the Wyle personnel in operation of the RIAC is to provide contract management and to market, distribute to the team, and perform technical area tasks (TATs) and subscription accounts (SAs). Wyle Laboratories has provided comprehensive technical services to government and industry since 1949. Headquartered in El Segundo, California, it has grown to over 4,500 employees with annual revenues in excess of $700 million. The company provides research, testing, engineering, and technical support services to the aerospace, defense, telecommunications, consumer electronics, energy, transportation, and other industries. Wyle ranks among the world's leading independent engineering and testing organizations, operating over 24 engineering and test facilities in Huntsville, Alabama; China Lake, Camarillo, San Diego, San Bernardino, El Segundo, Santa Clara, and NASA Ames Research Center, California; Hampton, Crystal City, and Arlington, Virginia; Lanham and Lexington, Maryland;

Table 1. Major milestones in Reliability Information Analysis Center (RIAC)/Reliability Analysis Center (RAC) history

Years	Focus	Milestone events
1965–1970	Component reliability	1968 RAC started
		1971 RAC becomes Department of Defense (DoD) Information Analysis Center
1971–1975		1971 Failure databases
1976–1980	EOS/ESD	1976 Reliability design handbook published
		1976 First training course presented
		1978 NPRD Databook published
1981–1985	System reliability	1983 VZAP Databook Published
1986–1990	Quality/TQM	1987 Technical Advisory Group (TAG) Established
1991–1995	Productization	1995 "Reliability Toolkit: Commercial Practices Edition" published
1996–2000	Acquisition reform and SIDAC integration	1996 "RAC Blueprints for Product Reliability" published
		1999 DoD "PRISM" automated system reliability assessment tool released
2001–2005	RAC to RIAC	2005 Wyle team of QSI, University of Maryland, Penn State ARL, and SUNYIT brings broadened capabilities to RIAC (Renamed from RAC)
2006–	Total system RMQSI	2006 "217Plus" system reliability assessment methodology and tool released as DoD "PRISM" replacement
		2006 "System Reliability Toolkit" published in collaboration with the Data and Analysis Center for Software (DACS), a sister DTIC IAC
		2007–08 RIAC participation in joint Govt/industry GEIA reliability program standard development
		2008 RIAC participation in DoD RIWG and MIL-HDBK-217G update

EOS/ESD, electrical overstress/electrostatic discharge; TQM, total quality management; SIDAC, Supportability Investment Decision Analysis Center; RAC, Reliability Analysis Center; RIAC, Reliability Information Analysis Center; RMQSI, Reliability, Maintainability, Quality, Supportability, Interoperability; NPRD, nonelectronic part reliability data; GETA, Government Electronics & Information Technology Association; RIWG, Reliability Improvement Working Group; VZAP, electrostatic discharge susceptibility data.

Dayton, Ohio; Houston and San Antonio, Texas; Cedar Falls, Iowa; and NASA Kennedy Space Center and Jacksonville, Florida.

Quanterion Solutions Incorporated

Quanterion Solutions Incorporated (QSI) is responsible for operating the core operations of the RIAC out of its Utica, New York office located on the SUNYIT campus. QSI was formed in 2000 by professionals with over 80 years of combined reliability engineering analysis experience. Key personnel within the company have had significant roles in the successful operation of the RIAC (as well as its predecessor, the RAC) and at the U.S. Air Force Rome Laboratory (renamed Air Force Research Laboratory—Rome Site) when it was the principal DoD organization responsible for reliability and maintainability (R&M) prior to acquisition reform in the mid-1990s. A partial list of contributions to RAC/RIAC includes:

- Performing independent reviews of competitive proposals, business operations, and product designs for a variety of organizations;
- Developing R&M program approaches for complex AF systems and evaluating request for proposal (RFP) responses;
- Auditing state-of-the-art equipment developers' reliability approaches and practices;
- Participating in the development of the GEIA-STD-0009 Reliability Program Standard;
- Participating in the OSD(AT&L)/OSD (DOT&E) Reliability Improvement Working Group (RIWG);
- Participating in the MIL-HDBK-217G Update Working Group;
- Managing, and developing component models for, MIL-HDBK-217 "Reliability Prediction of Electronic Equipment" for the DoD over a 10-year period;
- Serving as content editor for the Defense Acquisition University (DAU) Acquisition Community Connection (ACC) R&M Special Interest Area (SIA);
- Developing an expert web-based system/knowledge management tool (PROTOCOL) for the Army to define and tailor reliability programs;
- Providing R&M expertise to the ASR-9 Radar, ARSR-4 Radar, Voice Switching and Control System (VSCS), Mode-S Radar, Advanced Tactical Surveillance Radar, 60K Aircraft Loader, and JTIDS programs;
- Performing reliability assessments, FMECAs, FTAs, and reliability critical item identification on Army and commercial anti-chemical-biological warfare equipment;
- Performing reliability assessments on Navy submarine Towed Array sonar, ship decoy, and tactical military communications systems;

- Developing and implementing tailored reliability qualification test plans for Air Force and FAA systems;
- Supporting definition of degradation characteristics over time of electronic timing assemblies for an Army nuclear projectile;
- Developing Weibull Analysis tools for tracking and reporting the field reliability of specialized Air Force systems;
- Defining accelerated reliability test plans and environmental stress screening (ESS) options for multiple customers;
- Developing the RAC PRISM automated system reliability prediction tool and its replacement, the RIAC 217Plus methodology and tool;
- Authoring R&M guidance documents such as "Program Managers' Guide to Reliability and Maintainability", "Reliability Growth Testing Effectiveness", RAC and AFRL "Reliability Toolkits" (several versions), the RAC "Blueprints for Product Reliability" and the current RIAC/DACS "System Reliability Toolkit";
- Coauthoring widely used reliability databooks such as the RIAC/RAC "Nonelectronic Part Reliability Data (NPRD)" and "Failure Mode/Mechanism Distributions (FMD)";
- Presenting Reliability 101, Failure Mode Effects and Criticality Analysis (FMECA) and Fault Tree Analysis (FTA) training to military and commercial organizations;
- Providing on-line reliability training structured after the ASQ Certified Reliability Engineer Body of Knowledge (REPERTOIRE).

University of Maryland Center for Risk and Reliability

The Reliability Engineering Program at the University of Maryland started within the Nuclear Engineering Department in 1984, offering a few courses mostly related to system safety and risk. The Center for Reliability Engineering (renamed the Center for Risk and Reliability) was established in 1985, and the Certificate, M.S., and Ph.D. degrees in Reliability Engineering were formally approved in 1989. The program is currently administered by the Mechanical Engineering Department and is an internationally recognized leader among academic institutions in the areas of reliability education and research. The program focuses on design, development, operation, and management of engineered systems and processes that increasingly demand the formal integration of reliability and risk methods to ensure that such systems and processes are designed and operated with highest possible levels of reliability and safety while minimizing cost and other risks to their developers and users.

More specifically, UMD supports RIAC in the development and application of methods and tools to (a) understand why and how components, systems, and processes fail; (b) measure, track, and predict levels of reliability during the system life cycle; (c) improve reliability by applying science and engineering to remove failure causes; and (d) provide input to decisions regarding system design and operation. The risk analysis focus is on the development and application of methods and tools to determine potential undesirable consequences of developing and operating systems and processes, identify how such consequences might occur, assess the likelihood of such consequences, and provide input to decision makers on optimal strategies to reduce risk. Specific technical areas that UMD contributes to RIAC include:

- Probabilistic risk assessment
- Structural reliability
- Reliability of complex systems
- Accelerated testing
- Reliability of microelectronics
- Laser probing techniques
- Physics-of-failure
- Human reliability
- Data analysis
- Software reliability
- Software testing
- Software fault tolerant systems
- Electronic packaging materials
- Thin-film semiconductors
- Electronic system reliability
- Reliability optimization
- Design optimization
- Uncertainty modeling

The Pennsylvania State University Applied Research Lab (ARL)

The Applied Research Laboratory at Penn State is an integral part of one of the leading research universities in the nation and serves as a university center of excellence in Defense science and technologies, with a focus on naval missions and related areas. As a DoD-designated, Navy UARC (University Affiliated Research Center), ARL maintains a long-term strategic relationship with the Navy and provides support for the other services. ARL provides science and technology for national security, economic competitiveness, and quality of life through: education, scientific discovery, technology demonstration, and transition to application.

The Penn State ARL Complex System Monitoring (CSM) Department implements advanced diagnostic technologies for machinery health monitoring and supports basic and applied research in CSM technologies related to diagnostics and prognostics for electromechanical systems. These systems include rotating components, weapons system platforms, and machinery networks. The CSM Department uses proven analysis techniques, such as artificial intelligence, neural networks, and signal processing. The Accelerated Capabilities Initiative (ACI) for machinery diagnostics and prognostics was a seven-year program sponsored by ONR to develop hierarchical health management systems that provide real-time condition assessment information for Navy and Marine Corps assets. The program was comprised of several efforts whose objective was to demonstrate the application of advanced diagnostics to facilitate increased reliability and availability of critical assets, efficient maintenance planning, reduced maintenance activities, and increased safety related to the avoidance of catastrophic failure. The Demonstration of Hydraulic System Diagnostics and Condition Based Maintenance for Heavy Tactical Trucks project was a successful two-year project that involved the development and implementation of a hydraulic system health monitoring capability to the HEMTT LHS A2+ technology demonstration vehicle.

The Penn State ARL also has several unique test beds that have been designed and built by ARL for the development of Health Management technology. They include two full-scale military vehicles (Marine Corps Light Armored Vehicle and the Army Heavy Expanded Mobility Tactical Truck) on loan from the DoD for the implementation and demonstration of advanced diagnostic and prognostic technology. These test beds provide the ability to train, test, and validation and verification (V&V) diagnostic and prognostic technology that has been developed for mechanical, electrical, electrochemical, and structural applications. These test beds and vehicles demonstrate the ability to manage and operate small to large scale test bed systems and equipment.

The State University of New York Institute of Technology (SUNYIT)

SUNYIT brings a tradition of information technology-related education and research to the RIAC through its programs in computer science, computer information systems, engineering technologies, telecommunications, and its new jointly registered program in electrical engineering with Binghamton University. The SUNYIT campus provides the RIAC with access to campus computing and laboratory facilities with high speed networking, fully supported information technology services and distance learning capabilities, and growing wireless networking capabilities. The campus supports a variety of computing platforms, operating systems, collaborative computing software, and software development and simulation tools.

SUNYIT faculty members and graduate students actively engage in research and scholarship in areas such as systems theory and science, data mining, probabilistic databases, high performance computing, web archiving, nanotechnology, digital forensics, computer gaming, optical and wireless networking and wireless system security, and data security and cryptography. The SUNYIT faculty participates in RIAC research projects and is available to test and evaluate new information and knowledge management technologies and strategies. Students have the opportunity to engage in research and capstone projects under the supervision of faculty members in RMQSI (reliability, maintainability, quality, supportability, and interoperability) related topics and in information knowledge and management areas. Students also participate in internships at the RIAC.

Additionally, SUNYIT helps to foster collaborative research partnerships within the RIAC and to bring SUNY technology developments of interest to the RIAC community. The 64 campuses of the SUNY system participate in significant research programs in areas such as nanotechnology, electronic packaging, wireless systems, and advanced sensors and biometrics. Further, through the auspices of the Griffiss Institute for Information Assurance in Rome, NY, SUNYIT, other SUNY campuses, the SUNY Research Foundation, and leading New York State universities and industry partners participate in computer and network security education, research, and outreach activities.

RIAC success stories

☑ ***F-15 200-to-400 Hour Reliability Center Maintenance (RCM) Phase Extension Analysis***—Warner Robins Air Logistics Center, 830th Aircraft Sustainment Support Squadron (830 ACSG/GFEA), is tasked to perform maintenance, serviceability and sustainment for the F-15 fleet *(Figure 1)*.

The RIAC was tasked by F-15 Engineering to analyze all aircraft/weapon systems reliability and scheduled maintenance procedures by performing an RCM analysis—a 4-year effort that requires the review and analysis of 10 years of maintenance information and in-flight data. RCM is an analytical process to determine failure management strategies that must be implemented to ensure safe, reliable, and cost-effective operations. It evaluates scheduled maintenance, hidden or evident failures, safety/environmental and economic/operational impacts, preventive maintenance tasks, other actions, age exploration, or run-to-failure. The

Figure 1. F-15 success story

process ensures that the right maintenance is performed at the right time interval by the right people with the right tools—increasing aircraft availability and mission readiness while lowering the overall cost of ownership.

Based on the results of the F-15 analysis and a hazard risk assessment on failure modes associated with the phase inspection package, it was determined that the current 200-hour phase could be extended to 400 hours without impacting safety or reliability, reducing scheduled maintenance downtime. The RCM analysis also allowed the 400-hour work cards to be rewritten, further reducing the "look" phase of the inspection by approximately 2 days. After V&V, the 400-hour phase was implemented by an F-15E unit deployed to Afghanistan and reduced the mobility footprint for the number of aircraft and maintenance personnel required at the forward operating location.

The new 400-hour phase inspection is scheduled for fleet-wide implementation in 2008. Projected gains include effectively increasing fleet size by 7 aircraft due to increased aircraft availability; mission capable/mission ready rate increase by approximately 1 percent fleet-wide, approximately 40 percent reduction in phase maintenance, approximately 10 percent reduction in unscheduled maintenance with approximately $70M+ in cost avoidance/savings annually, and approximately 400K in annual man-hour savings.

☑ ***Nuclear Vehicle for Rocket Vehicle Applications (NERVA) Failure Modes and Effects Analysis (Figure 2)***—The RIAC recently completed a detailed Failure Modes and Effects Analysis (FMEA) on the National Aeronautics and Space Administration (NASA) NERVA Reactor.

The NASA Marshall Space Flight Center (MSFC) is responsible for the propulsion system on all NASA space vehicles. MSFC is developing a design tool for nuclear thermal propulsion (NTP) engine systems to generate conceptual designs and perform trade studies on one potential propulsion source for the Manned

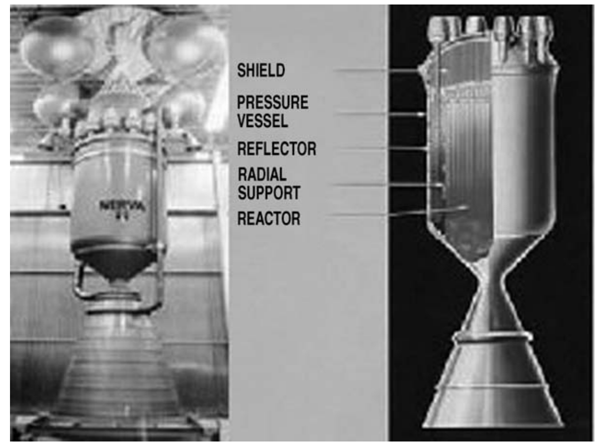


Figure 2. Nuclear Vehicle for Rocket Vehicle Applications (NERVA) success story

Mission to Mars. The NTP design tool contains modules for the major system components to assess the performance, weights, and sizing of the components and the overall system. RIAC subject matter experts completed a detailed FMEA for the nuclear reactor and the associated subsystems. This reactor, NERVA, has the largest thermal output of any nuclear system ever developed. This new analysis will guide NASA in the design and construction of the advanced NERVA, which will be a critical element in the first manned space flight to Mars.

☑ ***Significant Reduction in Joint Council on Aging Aircraft Document Review Times***—The Joint Council on Aging Aircraft (JCAA) was created to coordinate the efforts of DoD Services and government agencies in optimizing the service life of aging aircraft fleets to maximize aircraft material safety, minimize aircraft operating costs and duplication of effort, and facilitate and conduct exchanges of information on these efforts. The RIAC JCAA Team is responsible for managing the JCAA operations for this office.

One special JCAA focus area was to identify general series publications that are candidates for a joint development/update. The Services use individual publications in numerous nearly identical areas. RIAC subject matter experts (SME) coordinated and distributed metrics that showed the costs for maintaining these individual agency publications versus the savings realized through joint publications. As a result, individual publications were selected as candidates for consolidation. Supporting schedules were also prepared to develop common use documents across the Services. In the past 3 years, 115 Navy publications have been reviewed in concert with the other Services, with consolidation into 11 common documents.

A reduction in review times is a very significant benefit gained from this project. The close scrutiny of

The ITEA Journal of Test and Evaluation jite-29-03-09.3d 21/8/08 12:25:59 236

the selected publications cut the technical refresh time nearly in half over the period 2001 through 2007. Two of the most well-known impacts on technology refresh are obsolescence and understanding when the benefits of new technology (in this case, more up-to-date documents) outweigh the time and cost associated with rewriting a document to implement that technology. By taking the "best" chapters or work packages from the Services' publications and integrating them into one, research and writing time was substantially reduced. Best of all, this process "dusted the document" and ensured that reviews are conducted more frequently.

☑ ***Improved Sustainment of Army Missiles***—The Aviation and Missile Research, Development, and Engineering Center (AMRDEC) at Redstone Arsenal, Alabama, is responsible for the operational state and reliability of all missile systems used by the U.S. Army. Historically, the Army has relied on the Stockpile Reliability Program (SRP) to test samples of fielded assets to ensure that procured missiles are aging predictably. This program provides the Army with confidence that these systems will operate as designed when needed. While this management approach is effective, it is also very expensive because actual production hardware is often destroyed during testing

AMRDEC is now developing sophisticated, prognostics-based methodologies that will monitor missiles while they are affixed to aircraft and powered-up in "captive carry" mode *(Figure 3)*. Tiny sensors employed within missiles will monitor component and subsystem operation, and their output will be analyzed by prognostics algorithms to determine when operational anomalies occur. When subsystem operation falls outside of the design specification, system failure is presumed to be imminent. The missile will then be removed from service and returned to the depot for repair before it actually fails. Development of this capability will increase missile readiness and enhance the capabilities of our warfighters by ensuring that fielded assets will always operate as designed.

Subject matter experts from the RIAC team worked together to analyze a single missile component, a gyroscope, and determine potential ways it might fail during service. The team then developed notional prognostics algorithms that could interrogate appropriately designed sensors so that insight into missile aging can be obtained while they are still in service. The information obtained through these efforts helps form the basis for the future initiatives needed to implement and mature prognostics-based approaches to managing the Army's missile stockpile.

Figure 3. Improved sustainment success story

☑ ***Review of Reliability Centered Maintenance Programs on the CH-47 Chinook, UH-60 Blackhawk, and AH-64 Apache Helicopters for the U.S. Army Aircraft Engineering Directorate***—The U.S. Army Aviation Engineering Directorate is responsible for the airworthiness of Army developed aircraft and provides matrix-based engineering support to the Army aircraft crew, passengers, and maintainers that operate the Army aviation systems. Their customers include the Program Executive Officer Aviation Program project/product managers (PMs) and the U.S. Army Aviation and Missile Command (AMCOM). The U.S. Army Aviation Engineering Directorate Sustainment Engineering Division contracted with the RIAC to conduct RCM evaluations on three helicopter platforms: the AH-64 Apache, the UH-60 Blackhawk, and the CH-47 Chinook *(Figure 4)*.

The results of the RCM analysis that RIAC performed were as follows:

- The RCM analysis that is currently in progress on the CH-47 is approximately 70 percent complete with all of the aircraft systems. This effort equates to roughly 51 percent of the overall analysis being completed based on the size and level of effort of the remaining systems
- A component mapping effort is under way for the UH-60. FMEA developed for the Navy H-60 are being used to identify failure modes and their existing PM tasks. The stated output of the component mapping process will be a list of prioritized candidates for the PM to consider for further development.
- A baseline RCM analysis for the current AH-64 fleet is in progress.

The results of RIAC RCM support have resulted in significant O&S cost savings on other weapons system platforms.

☑ ***AN/ALQ-184 (V) Electronic Attack (EA) Pod Reprogrammable Low Band Processor Printed Wire Assembly (RLB PWA) Failure Analysis Study***—Warner Robins Air Logistics Center (542 CBSSS/

Figure 4. Reliability Center Maintenance (RCM) success story

GBEAA) is tasked with sustainment of the AN/ALQ-184 (V) EA Pod. The mission effectiveness of the AN/ALQ-184 (V) EA Pod is degraded due to the frequent failures of the Reprogrammable Low Band Standard Processor Printed Wire Assembly (RLB PWA). The purpose of the work is to provide 542 CBSSS/GBEAA with detailed analysis of the underlying causes of the failures associated with the RLB PWA. There were two phases to the analysis; Causal and Failure Elimination and Control.

Six RLB PWAs and one card cage were evaluated by the RIAC Team during the Causal Analysis Phase. The analysis included visual inspections, modeling, environmental, vibration, maintenance, and electrical testing. Though environmental conditions may affect long-term life of different components, they are not root cause failure mechanisms. Visual inspections of the sample PWAs determined board composition or production flaws are not functional failures. Maintenance practices performed in the field and at depots are not directly responsible for PWA failure. Better shipping and handling procedures, new tools for easier removal/insertion of the RLB PWA, and a thorough review of Technical Orders and Troubleshooting Manuals to eliminate known discrepancies will reduce maintenance time and effort saving money and increasing reliability.

The results of the electrical analysis however, represent an unusual root cause failure mechanism. Very high voltage pulses (spikes) cause time domain reflections that generate unwanted over-range voltage spikes resulting in U1 IC chip bond wire failure. This causes either Vcc node failure (spike on the supply path); or I/O ESD protection circuit coupled failure (current path will still be through each of the Vcc wire bonds and ground). The U1 IC chip bond wires act as high power fuse protectors, actually protecting the chip from catastrophic burnout. The multiple spikes that cause the bond wires to melt are random and cannot be filtered at the source. It is also cost prohibitive to filter out all random pulses.

The RIAC Team's Failure Elimination and Control Phase recommendations provided three options for government engineers to consider for root failure mitigation and long term sustainment of the system; implement an LC network to carry the noise pulses to ground; add diffused or thin film resistors to bond pad wire in order to ensure noise spike pulses dissipate rapidly to ground; or a combination of both in the form of an LCR network on a separate chip. Follow-on work is planned to either redesign the U1 chip or redesign the PWA to eliminate the failure. This will ensure that the USAF can continue to fly the approximately 1,000 AN/ALQ-184 Jammer Pods well into the next decade.

Summary

The RIAC works with government and industry customers on both short term and long term Reliability, Maintainability, Quality, Supportability, and Interoperability projects. The RIAC helps clients identify problems and implement solutions that positively impact their most important organizational goals: safety, mission success, market share, life cycle cost, profitability, and productivity. It offers an independent, unbiased perspective in all its work, performed by a competitively-selected team chosen on the basis of the depth and breadth of its technical experience and expertise:

- Reliability Program Planning and Development
- Reliability Audit/Maturity Assessment
- Parts Control Program Development
- Proposal/Source Selection Support
- Design Reliability Evaluation (Hardware, Software, Human Factors)
- Critical Item Identification and Management
- Reliability Growth Planning and Implementation
- Probabilistic Risk Assessment
- Closed-Loop Root Failure Cause Analysis and Corrective Action Implementation and Verification (e.g., FRACAS)
- Reliability Modeling (Component and System Level)
- Reliability Assessments/Predictions
- Component Obsolescence Planning
- Maintainability Analysis
- Reliability-Centered Maintenance
- Condition-Based Maintenance (CBM)
- Application and Use of COTS/NDI/CFI
- System Lifetime Extension
- Failure Mode, Effects and Criticality Analysis (FMECA/FMEA)
- Fault Tree Analysis (FTA)
- Finite Element Analysis (FEA)
- Derating Analysis
- Test Planning and Evaluation
- Environmental Stress Screening Plans

It is easy for the DoD/Government to benefit from RIAC services because of its unique status as a competitively awarded government contract. Short term support can be arranged through a Subscription Account (SA) in a matter of a few days, and longer term Technical Area Task (TAT) support in a matter of a few weeks. Using either of these two vehicles, a relatively simple modification is made to the RIAC contract. The process starts by mutual agreement on a scope of work with the Center. A Military Interdepartmental Purchase Request (MIPR) or other government funds transfer document is then sent to the government management organization for RIAC. The government management of the RIAC contract also encourages the industrial and commercial use of the Center's technical services. Commercial customers simply provide a purchase order to the RIAC contracting team and agreement on the scope of the work. ❑

David Nicholls currently manages the Core Operations of the Reliability Information Analysis Center. He has over 30 years experience in the reliability engineering discipline, serving in project lead and management positions on numerous defense programs within GE Aerospace (now Lockheed-Martin). He also has over 15 years of technical and management experience with the Department of Defense-funded, DTIC-sponsored Reliability Information Analysis Center (formerly the Reliability Analysis Center), including Deputy Director and his current role as operations manager. He has served as deputy director of the Data and Analysis Center for Software (DACS). His publishing credits include the RIAC "System Reliability Toolkit", the DACS "Software Reliability Sourcebook", and the RAC "Reliability Toolkit" and "Reliability Blueprint" series. He is also a key contributor to the RIAC 217Plus system reliability assessment methodology. David Nicholls was an active participant in the development of the GEIA-STD-0009 "Reliability Program Standard for Systems Design, Development and Manufacturing", and the joint OSD(AT&L) and OSD(DOT&E) Reliability Improvement Working Group (RIWG), and is a member of the Navy Crane Working Group which is updating MIL-HDBK-217. E-mail: dnicholls@theRIAC.org

ITEA Journal 2008; 29: 240–243
Copyright © 2008 by the International Test and Evaluation Association

Tropic Regions Test Center

Lance VanderZyl
Director, Tropic Regions Test Center
U.S. Army Yuma Proving Ground, Yuma, Arizona

A key element of U.S. military strategy is the ability to rapidly deploy, employ, sustain, and redeploy capabilities in geographically separate and environmentally diverse regions around the world. To support that strategy, U.S. military systems must be able to operate effectively and safely in the full range of natural environments found around the globe. The ability to operate in a tropic environment is especially critical as approximately three-fourths of all regional conflicts have taken place in tropical areas.

While some aspects of the natural environment can be recreated in test chambers, it is only through testing in the natural environment that the synergistic effects of all of the challenges posed by nature can be fully understood and evaluated. The detailed, realistic testing ensures U.S. military personnel from all services that their systems will function as intended in these environments.

The U.S. Army's Yuma Proving Ground (YPG), the Department of Defense (DoD) lead for natural environment testing, operates and maintains sites for testing the full range of military systems in cold, desert, and tropic environments. The Tropic Regions Test Center (TRTC) is a subordinate command of YPG (a subordinate command of the Developmental Test Command and the Army Test and Evaluation Command) and has the mission for planning and conducting testing of military systems in a tropic environment.

This article provides information on TRTC's capabilities as well as a review of the history of tropic testing and a listing of the key environmental challenges systems may encounter in a tropic environment.

Key words: Natural environment testing; Panama; realistic conditions; tropics.

While the origins of tropic testing by the U.S. military can be traced back to the early 1920s, efforts to establish a permanent tropic test capability did not begin until the early 1950s. At that time, the Panama Canal Zone was selected as the most appropriate site for tropic testing. However, due to funding problems, these plans were laid aside until 1962 when the U.S. Army established the U.S. Army Research and Development Office, Panama. The name was subsequently changed to the U.S. Army Tropic Test Center in 1964 and it evolved over the following years into what is now known as the Tropic Regions Test Center (TRTC).

Following the withdrawal of U.S. forces from Panama in 1999 (in accordance with the Panama Canal Treaty) and the relocation of TRTC to Yuma Proving Ground (YPG), the primary focus of TRTC's efforts has been to rebuild the Department of Defense's (DoD) tropic test capabilities and to restore tropic testing workload to a sustaining level.

To support rebuilding these tropic test capabilities, YPG commissioned an independent scientific review of potential tropic test sites by a panel of experts assembled by the U.S. Army Research Office (ARO). The studies concluded that a suite of sites would offer the best technical approach. Based on this input, as well as the direction of higher headquarters and specific customer requirements, subsequent TRTC efforts focused on regaining access to sites in Panama and the establishment of test sites in Hawaii, Honduras, and Suriname (*Figure 1*).

TRTC's efforts have been extremely successful, as evidenced by ready access to tropic test sites in Panama, with additional sites in Hawaii, Honduras, and Suriname. Work at these sites has restored tropic test

Figure 1. Tropic Regions Test Center test sites

workload to levels in excess of historical averages, reaching an all time high in FY 06.

Tropic environment—Characteristics and challenges

Two of the more common challenges found in the tropics are:

- Dense, old-growth forests with double or triple canopies that block sunlight and retain moisture from frequent, heavy rainfalls, while keeping the temperature nearly constant at 75 degrees Fahrenheit and the humidity level at 95–100 percent around the clock, all year long.
- Open areas of short or 20-foot tall grasses, mangrove swamps and coastal environments with year-round temperatures in the 80–95 degrees Fahrenheit range coupled with relative humidity in the 80–100 percent range.

The combination of climatic, biotic and geographical features found in the tropics can often lead to the following problems with military systems:

- Degradation in target detection and acquisition due to dense tropic vegetation, terrain, and climatic conditions.
- Blockage or degradation of tactical communications due to rugged terrain, dense foliage, and climatic conditions.
- High levels of heat stress due to temperature and humidity levels.
- Impeded foot and vehicle mobility due to steep slopes, frequent heavy rainfall, and weak soils.
- Obscuration of targets and degradation in weapon effectiveness due to jungle canopies and vegetation.
- Corrosion of metals, clogging or rusting of air ducts and tubing, shorting of electrical components and slippery surfaces due to heavy rains, high humidity, and fungal growth.
- Degradation of materials such as polymers, rubber, plastics, and paints due to solar radiation and high temperatures, which can cause cracked surfaces and allow moisture intrusion.
- Insect and microbiological damage and corrosion.

TRTC capabilities

TRTC operates and maintains an array of test sites in a variety of jungle and tropical open lands and coastal environments. Testing is accomplished by a combination of Army civilian personnel, support contractors, and (on occasion) borrowed military personnel. Key tropic test sites are described in the following paragraphs.

Panama provides a near "ideal" tropic environment and is the primary site for the testing of individual soldier systems and small caliber weapons. It is the primary location for support contractors. Key capabilities include:

- **Cerro Tigre Test Site** is a secure general purpose test area which includes a 700-meter small arms firing range, small arms storage bunkers, a manpack portability course, and a MOUT complex; it is located on a training academy operated by the Panamanian National Police (PNP) about 30 minutes from Panama City (*Figure 2*).
- **Pacora Test Site** is a leased site in a nearby area east of Panama City. It is used for the testing of sensor and communication systems in a triple canopy jungle environment with a nearby helipad.

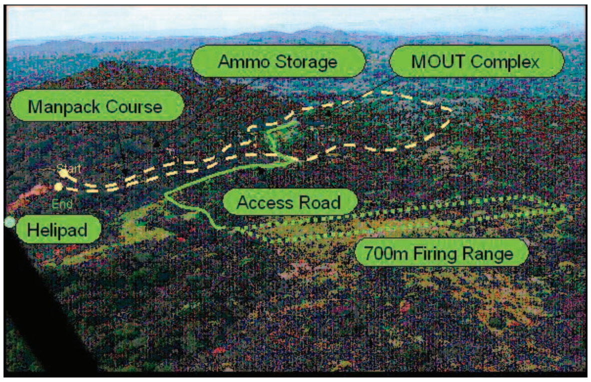


Figure 2. Cerro Tigre test site, Panama

- **Llano Carti Test Sites** are located on leased lands in a remote area further east of Panama City for testing of sensors in a triple canopy jungle environment with a nearby helipad.
- **North Coast Area Test Sites** include breakwater, coastal, open-inland and jungle test sites for the exposure testing of materials and systems across the full range of conditions (humidity, temperature, rainfall, solar radiation, salt spray, etc.) found in coastal and inland tropic areas. Testing at these sites is conducted through a Cooperative Research and Development Agreement (CRADA) with the Technical University of Panama.
- **Panama City Support Complex** is located in downtown Panama City and features office space, conference rooms, telecommunications capabilities, medical personnel and other general administrative support capabilities for test customers. It also serves as the operations center for the TRTC support contractor and the organization's approximately 35 test support personnel.

Tropic Test Site—Hawaii is located on the U.S. Army's Schofield Barracks on the island of Oahu. It is the primary site for the testing of mine, countermine and smoke obscurant systems and serves as a back-up test site for individual soldier systems and small caliber weapon systems (*Figure 3*). Its colocation with the 25th Infantry Division provides enhanced access to soldiers and the firing and training ranges operated and maintained by the 25th Infantry Division.

Tropic Test Site—Honduras is located near Mocoron in a remote area of eastern Honduras on an outpost manned and secured by the Honduran Army's 5th Infantry Battalion. Due to minimal electromagnetic "clutter" in this area, as well as its remoteness, it is ideal for the testing of sensitive systems, C4I systems and sensors (in partnership with the U.S. Army's Electronic Proving Ground and the Redstone Technical Test Center). A C-130 capable airstrip is located on the Honduran Army outpost.

Figure 3. Man Pack Course, Hawaii

Tropic Test Site—Suriname is located on the Suriname Aluminum Company, L.L.C. (a subsidiary of Alcoa) bauxite mining complex near Moengo in the northeastern part of Suriname. This site is used for the testing of heavy vehicles and leverages the existing support infrastructure (repair shops, maintenance facilities, housing, medical clinic, deep water port, etc.) associated with the bauxite mining operations.

Because TRTC operates in a number of foreign countries in the SOUTHCOM area of operations, access to test sites is closely coordinated with the command and U.S. embassies in each country. The relationships and agreements that have been developed combined with the experience of TRTC personnel in dealing with the nuances of conducting testing in a foreign country help to streamline and facilitate the test process for TRTC's customers.

Figure 4. Campo Victoria, Honduras

Summary

TRTC has successfully re-established itself as the premier Department of Defense (DoD) activity for testing in the challenging natural tropic environment. The test sites TRTC operates and maintains, combined with the skill and experience of its workforce, ensure realistic testing of equipment and systems. This has led to an enhanced understanding of system and equipment performance, reliability, and safety when subjected to the unique climatic, biotic, and geographical challenges found in the tropics, and will continue to do so in the future. Further information on TRTC, as well as points of contact, can be found at http://www.yuma.army.mil/tc_trtc.shtml. ❑

Sponsored by the International Test and Evaluation Association

Register On-Line www.ITEA.ORG

SHORT COURSE

Radio Spectrum Management for Test and Evaluation

October 21 – 22 – 23, 2008 | Fairfax, Virginia

INSTRUCTOR: **Mr. Darrell Ernst** is a Principal Space Systems Engineer with the MITRE Corporation

This course introduces the student to the electromagnetic spectrum, with an emphasis on the radio band. The curriculum covers the discovery of electromagnetism and the development of radio, the basic physical principles of the spectrum from an intuitional standpoint rather than a rigorous scientific or engineering standpoint. It describes how the use of the spectrum is managed, the environment in which radios operate, the applications of radio, testing of radio systems, using radios for testing and relevant Department of Defense (DoD) policy. The course will address the management of spectrum at test facilities and such current issues as encroachment on spectrum used for testing and the use of wireless systems for testing as they relate to spectrum. The course is intended for non-radio professionals who need to understand the issues associated with radios and the radio spectrum in the acquisition and testing of systems. The course is oriented towards the DoD, but much of the material applies equally to non-DoD applications.

WHO SHOULD ATTEND: The course is designed for contractor and government acquisition and testing professionals involved with the acquisition and test of systems involving radio frequency (RF) communications, or systems that must be tested using RF systems. The course is appropriate for program managers, systems engineers, test planners, and other personnel who desire a general understanding of the practical aspects of the radio spectrum, particularly as it applies to test and evaluation. The course will be useful to radio engineers who want to know how the radio spectrum is managed or used for T&E, but the initial material on the fundamentals of radio is presented in a very simplified form.

This course qualifies for Continuous Learning Points for acquisition workforce members.

More short courses will be available throughout the year. For more details visit **www.itea.org**

ITEA Journal 2008; 29: 244–246
Copyright © 2008 by the International Test and Evaluation Association

Reliability:
A Look at Four Decades of Reform Within the Department of Defense

Larry Crow, Ph.D.
Crow Reliability, Madison, Alabama

Developing weapon systems in the Department of Defense (DoD) is a particularly challenging task for many reasons. The most obvious and well known difficulties are that the systems will almost always push the state of the art in technology, there are always schedule constraints and the need to deploy as soon as possible, and of course, budget considerations. What is not so obvious to many people is the especially difficult challenge of meeting reliability requirements. DoD data over the period 1997 to 2006 show that a very low percent of DoD systems attained their reliability requirements during an Operational Reliability Demonstration Test at the end of the program. Systems that fail to meet the reliability requirements during this test are either deployed with the low reliability or subjected to a reliability improvement program as an attempt to increase the reliability of the system. Both alternatives are costly.

Key words: Reliability; DoD acquisition reform; demonstration tests; reliability growth.

An unreliable weapon system has a clear impact on the war fighter and will always have a major impact on lifecycle sustainment costs. Because of this situation there have been many studies of DoD reliability practices over the past several years by the National Academy of Sciences and the Defense Science Board.

As noted by the most recent (2007–2008) Defense Science Board Task Force addressing reliability, we need to put the problem in historical perspective so we will have a better understanding of what it is we need to do. That is: What policies work and what policies do not work? Are we repeating past mistakes? What are proven historical DoD reliability best practices? To get an understanding of the answers to these questions and the various DoD approaches in the past, we need to go back about four decades.

Era 1969 to 1980

During this period the main DoD guidance document on reliability was Mil Std 785 A released on March 28, 1969 and entitled "Reliability Program For Systems and Equipment Development and Production." The DoD policy during this period was to only specify a reliability requirement and subject the system to an Operational Reliability Demonstration Test at the end of the program. Depending on the number of failures during the test, the system either passed the test or failed the test. Unfortunately, in the earlier part of this era most programs failed their Operational Reliability Demonstration Test. The situation improved with the use of reliability growth testing. Beginning in the 1960s companies such as General Electric (GE) were taking a different approach to reliability testing than the DoD demonstration test method. Instead of waiting until the end of a development program and subjecting the system to a fixed reliability demonstration test, why not test the system earlier, and incorporate corrective actions when reliability problems were discovered. This approach not only improved the reliability but J. T. Duane of GE was also able to show empirically that the reliability improvement trend had a predicable learning curve pattern. This observation is called the Duane Postulate.

This predicable pattern could be used for reliability management and J. D. Selby and S. G. Miller of GE recommended reliability growth management to the DoD in the early 1970s. At the same time (1972) the author, at the Army Materiel Systems Analysis Activity (AMSAA), showed that the Duane postulate could be expanded and formulated into a statistical framework, called the Crow (AMSAA) model. This model provided the necessary groundwork for valid statistical methods

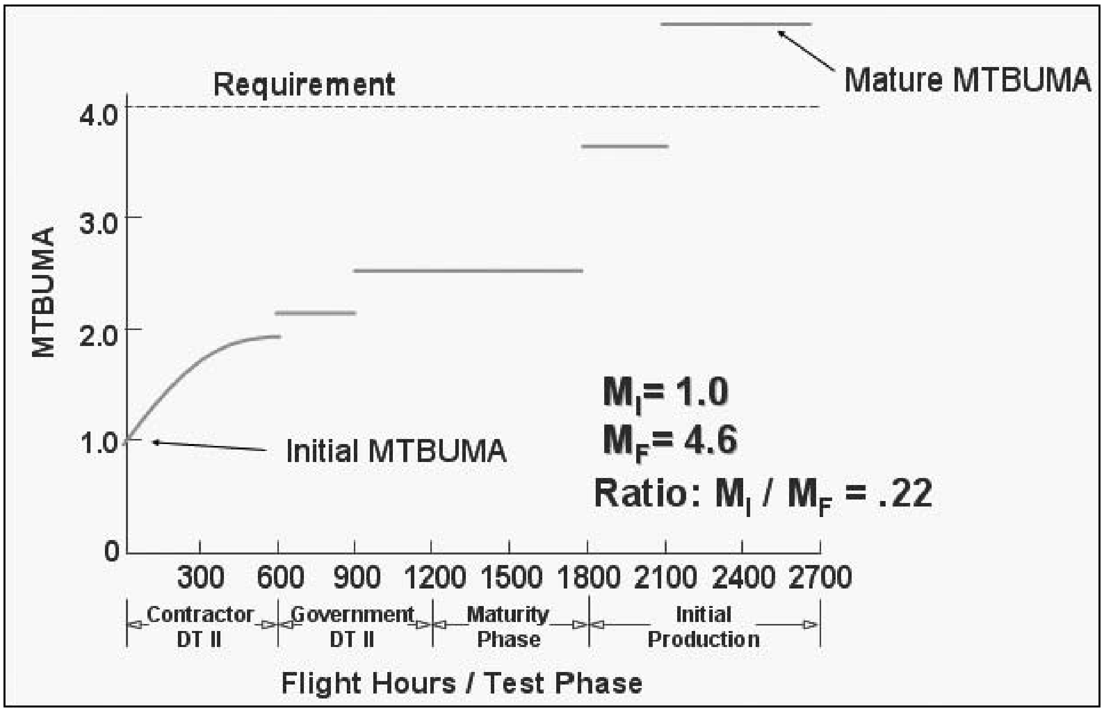


Figure 1. The Blackhawk helicopter reliability growth measured in meantime between unscheduled maintenance action (MTBUMA) per flight hours per test phase

to evaluate the current mean time between failures (MTBF) of a system while the reliability is improving during reliability growth testing.

This first major DoD application of reliability growth was on a U.S. Army program, the Blackhawk helicopter (1972–1978). The Blackhawk had a reliability requirement for mean time between unscheduled maintenance action (MTBUMA) of 4 hours. When the Blackhawk was first tested the MTBUMA was 1 hour. If this was the usual demonstration test, the system would have failed. Instead, the system was subjected to reliability growth testing where the reliability was improved and the Crow (AMSAA) model was used for assessment of the progress. This process increased the Blackhawk MTBUM from 1 hour to 4.6 hours, exceeding the requirement. For this system the initial MTBUMA was 22 percent of the final value actually attained, with an improvement of 4.6, *Figure 1*.

During the 1970s reliability growth was applied to many major systems within all the services, for example, the Army's M1 Abrams main battle tank, the U.S Navy's Tomahawk Weapon System, and the U.S. Air Force's F16 Fighter aircraft. With the success of this new approach to DoD reliability the Joint Logistic Commanders in 1976 directed that a new Mil Handbook be written. This document, "Reliability Growth Management," MIL Handbook 189, was released in 1981. In the meantime the DoD made no changes in its reliability policies. However, each service did make policy changes which made reliability growth and reliability engineering a best practice through various regulations and documents such as the Army's 702-3, the Air Force R & M 2000, and the Navy Willoughby Templates Best Practice NAVSO P-6071.

A lesson that was loud and clear was that the old practice of simply specifying a reliability requirement, not requiring upfront reliability design engineering, and then subjecting the system to a demonstration test was a failed and flayed approach. A demonstration test is a go-no go management test. On the other hand reliability growth testing is an engineering tool where failure information is used to improve the system reliability.

The use of reliability growth and test-analyze-and-fix (TAAF) testing became widespread within the DoD to complement and, in some cases such as the Blackhawk helicopter, a substitute for formal reliability demonstration testing. Despite the success of reliability growth there was still concern by many reliability experts within the DoD that it was rewarding contractors for sloppy initial designs. This concern was addressed in 1980.

Era 1980 to 1998

In 1980 the DoD released Mil Standard 785 B, "Reliability Program for Systems and Equipment

Development and Production." This revision contains fundamental changes from Mil Std 785 A. The document states that "Increased emphasis has been placed on reliability engineering tasks and tests. The thrust is toward prevention, detection and correction of design deficiencies, weak parts, and workmanship defects." Mil Standard 785 B consisted of reliability tasks in three main areas: management, design, and testing. The management and design tasks were directed toward increasing the reliability of the initial design, and reliability growth was now a complementary reliability best practice task to further increase the reliability toward the requirement. The goal of DoD reliability policies from 1980 until the mid to late 1990s was to achieve good reliability by focusing on reliability fundamentals during design and manufacturing rather than merely setting numerical requirements and testing for compliance toward the end of development. DoD data shows that during this period the reliability growth best practice increased the initial reliability MTBF by an average factor of about four times, similar to the results of the Blackhawk helicopter. Most of the systems used in the First Gulf War were developed using reliability growth testing, Mil Std 785 B, and other design best practices, and as the results showed, these systems were generally very reliable during this conflict.

DoD policy changed again in 1998.

Era 1998–present: Acquisition reform

In 1998 DoD implemented Acquisition Reform. The DoD noted that "DoD's transition to a Performance-Based Business Environment, maximizing the use of commercial items and practices, is a key step toward achieving civil military integration." The consequence of Acquisition Reform on reliability was that now the only reliability requirement was that the system passes an operational test at the end of the development program. Reliability growth testing and upfront design reliability engineering were no longer best practices and were not required. Mil Std 785 B was cancelled, although Mil handbook 189 is still active. This is basically the same approach taken by the DoD in the early 1970s which did not work then and is clearly a major factor affecting the low percent of systems passing reliability operational tests today. Unfortunately, the DoD had repeated history in terms of its policy and direction regarding reliability. This policy is not likely to change. However, some good news is that there are pockets within the DoD development community that never gave up key reliability best practices and reliability growth. This is particularly true of aviation systems, both fixed wing and rotary. A big part in effectively addressing this issue is recognizing that DoD has gone full circle over the past four decades and learning from the experiences in between. For example, although Mil Standard 785 B provided a structure for reliability engineering, we know today that many of these tasks were actually not that effective. We can do better. Therefore, issuing a new Mil Standard 785 B-like document is not the answer. The Defense Science Board Task Force addressing reliability has noted this issue in its report and steps are already underway in the DoD to attempt to correct this problem by issuing a new industry-government reliability standard with a more modern approach, and implementing comprehensive reliability training programs. Although we have made a full circle in policy, DoD now has an opportunity to provide leadership in the reliability community in order to improve the number of systems passing operational reliability demonstration tests.

Dr. Larry H. Crow is president of CRR. Previously Dr. Crow was VP, Reliability & Sustainment Programs, at the Illinois Institute of Technology Research Institute, which became Alion Sciences. From 1985 to 2000 Dr. Crow was director, Reliability, at Bell Labs Advanced Technology Systems (later named General Dynamics ATS). From 1971-1985, Dr. Crow was chief of the Reliability Methodology Office at the U.S. Army Materiel Systems Analysis Activity (AMSAA). He developed the Crow (AMSAA) reliability growth model, which has been incorporated into U.S. DoD handbooks, and national & international standards. He chaired the committee to develop Mil-Hdbk-189, Reliability Growth Management and is the principal author of that document. Dr. Crow is a fellow of the American Statistical Association, and the Institute of Environmental Sciences and Technology. He is a Florida State University Alumni Association Distinguished Alumnus and the recipient of the FSU "Grad Made Good" Award for the Year 2000, the highest honor given to a graduate by Florida State University. E-mail: Crowrel@knology.net ❑

ITEA Journal 2008; 29: 247–250
Copyright © 2008 by the International Test and Evaluation Association

Sustainment Key Performance Parameter and the RAM-C Manual

Peter Nolte, John Quackenbush, and Grant Schmieder
Systems and Software Engineering,
Office of the Under Secretary of Defense for Acquisition
Technology and Logistics,
Department of Defense, Washington, D.C.

This article addresses an increasing trend within the Department of Defense of systems not achieving the required reliability during developmental testing and subsequently being found unsuitable during Initial Operational Test and Evaluation. It introduces a Department systems engineering initiative to help requirements managers and program managers develop balanced and measurable sustainment requirements of reliability, availability, and maintainability (RAM) requirements with the development of a RAM-Cost Rationale Report Manual. This manual, in a coordination draft, will assist program managers and requirements managers to infuse robust systems engineering activities early in the program so that informed RAM trades are made throughout the life cycle. Thus better reliability will be designed into systems, validated through testing, and presumably resulting in systems that are more reliable and maintainable long term.

Key words: Maintenance; materiel availability; materiel reliability; mission requirements; ownership cost; sustainment requirements.

The Department of Defense spends billions of dollars acquiring, maintaining, and operating a wide variety of equipment. In doing so, the Department expects to acquire reliable and maintainable products that are of high quality, readily available, and able to satisfy user needs with measurable improvements to mission capability and operational support at a fair and reasonable price.

Not only is this equipment becoming more expensive to purchase, but the cost to operate, sustain, and maintain some new equipment is becoming much higher than anticipated or the equipment is not meeting the expected level of reliability, availability, and maintainability (RAM). Both of these ramifications have a negative impact on operational suitability and acceptable cost.

Sustainment requirements

In an effort to change this trend, the Department and subordinate Services are focusing more closely on sustainment requirements. In 2006, the Chairman of the Joint Chiefs of Staff defined three mandatory sustainment requirements to be articulated in requirements documents throughout a program's life cycle. These sustainment requirements include a Materiel Availability Key Performance Parameter (KPP) and two supporting Key System Attributes (KSA), Materiel Reliability and Ownership Cost. Although these are the three mandated sustainment requirements, maintainability and supportability analysis also underpin the systems engineering efforts that involve designing and developing these traits.

The Department understands that "sustainment" is a key component of performance and must be planned for during the first stages of concept refinement and system development, with the planning for all other mission capabilities. If these sustainment attributes are not adequately designed into the system, programs may breach Nunn-McCurdy thresholds[1], cost more than anticipated to own, or fail to achieve availability expected by the warfighter. Developing reasonable sustainment requirements will help ensure other mission performance requirements are achieved while balanced against the system life cycle cost for new and fielded systems.

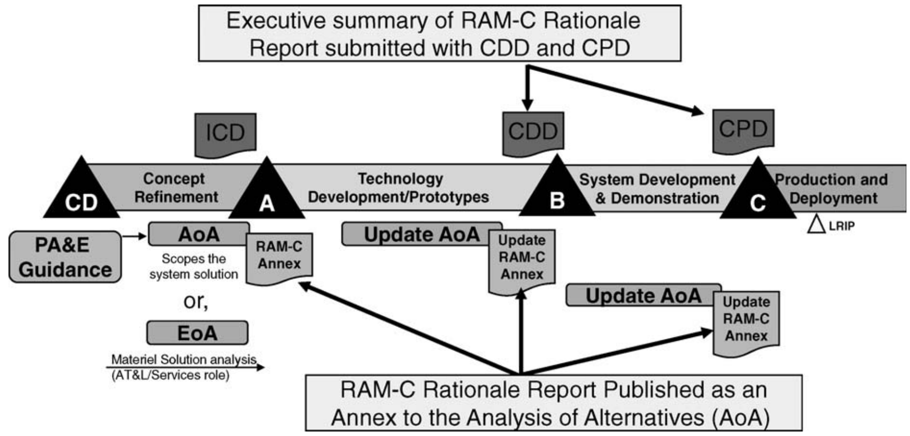


Figure 1. Expected timeline and associated documents with which the RAM-C Rationale Report will be submitted

The value of the sustainment KPP is derived from the operational requirements of the weapon system, assumptions for its operational use, and the planned logistical support to sustain it. In order for the program manager to develop a complete warfighting system with appropriate sustainment attributes, programs must establish both mission capability and sustainment requirements and measure the performance of the entire system against those requirements. The mandatory KPP and two supporting KSAs establishing that framework are summarized as follows:

- Materiel Availability KPP—Measures the percentage of the total inventory of a system that is operationally capable (ready for tasking) of performing an assigned mission, at a given time, based on materiel condition. Materiel availability also indicates the percentage of time that a system is operationally capable of performing an assigned mission and can be expressed as (end items ready for tasking)/(total end items procured).[2]
- Materiel Reliability KSA—Measures the probability that the system will perform without failure over a specified interval under specified conditions.
- Ownership Cost KSA—Provides balance to the sustainment solution by ensuring that the Operation and Support (O&S) costs associated with materiel readiness (e.g., maintenance, spares, fuel, support, etc.) are considered in making program decisions. The Ownership Cost KSA is ultimately based on O&S Cost Estimating Structure elements as specified in the OSD Cost Analysis Improvement Group (CAIG) "Operating and Support Cost-Estimating Guide."

RAM-C rationale report

To reinvigorate RAM systems engineering, the Department is developing a process, originally emphasized by the Army in the 1980s, requiring programs to develop a Reliability, Availability, Maintainability, and Cost (RAM-C) Rationale Report at the beginning of a program or during early concept development before it officially becomes a program.

The writers of requirements (the combat developers) and the program sponsors (or managers) must work together early to develop and understand mission and sustainment requirements that facilitate achieving the objective of affordable, suitable, and available systems. The logical process of developing sustainment requirements involves well-defined activities to arrive at values that are realistic, achievable, measurable, documented, and therefore defendable.

By documenting the rationale behind the development of the sustainment requirements with underlying assumptions, requirements writers and program managers will understand the basis for decisions made early in the program and will be better informed when trades need to be made later in the program.

The activities required to develop the RAM-C Rationale Report are detailed in a draft manual intended to be published in the October/November timeframe. Figure 1 shows the expected timeline and associated documents with which the RAM-C Rationale Report will be submitted.

The Service sponsor for concept refinement or system development of a new system will first document the sponsor's requirements rationale as early as the Analysis of Alternatives (AoA) during concept refinement. It is during this analysis that maintenance and supportability assumptions are first made and documented as part of the AoA. The Service sponsor conducting the AoA compares the various alternatives to determine the best potential materiel solution for the government, including how sustainment attributes and their sensitivities will be traded over the entire life cycle.

Once a materiel solution is selected, the combat developer will begin establishing the sustainment

1.1 Executive Summary
1.1.1 Summary of RAM Goals and Constraints
1.1.2 Description of Sustainment Requirement Element Values
1.1.3 Summary of Program Manager Analysis
1.1.4 Summary of Combat Developer Analysis including updated RAM-C Goals as appropriate
1.1.5 Summary of Sustainment System
1.1.6 Information for obtaining full RAM-C Report
1.1.7 Approval Signatures for Mid-Phase Updates to Sustainment Requirements
1.2 Program Summary Introduction
1.3 Predecessor System
1.4 Reliability, Availability, Maintainability, and Cost Goals and Constraints
1.4.1 Materiel Availability
1.4.1.1 A_M Requirement
1.4.1.2 A_M Rationale
1.4.1.3 Assumption Rationale
1.4.1.4 Relevant Facts Known
1.4.1.5 Supporting Analysis (Combat Developer and/or Program Manager)
1.4.2 Materiel Reliability
1.4.2.1 R_M Requirement
1.4.2.2 R_M Rationale
1.4.2.3 Assumption Rationale
1.4.2.4 Relevant Facts Known
1.4.2.5 Supporting Analysis (Combat Developer and/or Program Manager)
1.4.3 Ownership Cost
1.4.3.1 OC Requirement
1.4.3.2 OC Rationale
1.4.3.3 Assumption Rationale
1.4.3.4 Relevant Facts Known
1.4.3.5 Supporting Analysis (Combat Developer and/or Program Manager)

Figure 2. Outlines the suggested report format

requirements as part of the development of the Capability Development Document (CDD) and Capability Production Document (CPD). When the AoA is updated prior to each milestone decision, the RAM-C Rationale Report will be updated as well, to include validating or invalidating assumptions made earlier and to include an explanation of trades made as a result of lessons learned. This process, articulated in the manual, will require an executive summary of the RAM-C Rationale Report be attached to the CDDs and CPDs.

The draft manual will be a worthwhile tool for combat developers and program managers to use in developing their requirements and documenting their rationale. It provides a suggested Rationale Report format and walks through the steps to develop sustainment requirements and rationale using an example notional gun system. Figure 2 outlines the suggested report format.

The manual describes the process of developing and refining sustainment requirements. It begins with the combat developer using the Operational Mode Summary and Mission Profile (OMS/MP) along with the Failure Definition and Scoring Criteria (FD/SC) to conduct the analysis required to determine the maintenance and support concepts. This information is used to draft initial materiel availability, materiel reliability, and ownership cost goals. It is also used in supporting rationale and assumptions, including the levels of maintenance and the maintenance activities to be conducted at each level.

The program manager takes the above information and works with the combat developer to determine what is achievable based on technology maturity and other factors in order to make appropriate trades. Once the combat developer and program manager have reached agreement on a balanced solution with acceptable trades based on what is technologically possible, the combat developer needs to identify the appropriate sustainability requirements for inclusion in the CDD/CPD.

Requirements development

If done correctly, the combat developer will avoid writing requirements that include vague references to another system, such as "equal to or greater than predecessor system" or "50% less to support than the predecessor system." Such references make the requirements difficult to measure because the predecessor system requirements are often unknown or may not be compatible with the new system. In addition, such phrasing leads to side-by-side comparative testing that may be both costly and unrealistic and may not provide the results that effectively illustrate the system's ability to succeed in the field. The combat developer should write requirements that fit the operational needs foreseen while also including probability values required. For example, "the system must have a 95% chance of completing a 12-hour mission without a mission-affecting failure" or "the system should have a maximum mean time to repair of 4 hours for up to the 95th percentile repair."

The requirements development process concludes when all inputs are translated into materiel availability, materiel reliability, and ownership cost with supporting rationale so that the program manager can develop a plan to design-in and achieve the threshold values within established cost and schedule parameters. It is expected that the program managers will then contract for the appropriate activities to design-in and evaluate RAM to ensure threshold values are demonstrated in developmental test and evaluation prior to the initial operational test and evaluation. It is vital that these plans to develop, design, and measure these RAM requirements are included in the Request for Proposal and subsequent contracts.

Conclusion

In the end, the sustainment requirements must enable warfighter functional requirements and must be measurable and obtainable. All program activities must carefully balance technological feasibility with operational needs and desires. Program requirements are subject to trade-off in order to optimize all requirements, including sustainment requirements.

The manual is currently in draft and has been informally staffed within the Office of the Under Secretary of Defense for Acquisition, Technology and Logistics, the Director of Operational Test and Evaluation, the Joint Staff, and a few outside agencies. We expect the manual to be published and implemented in FY 2009. ❑

PETER NOLTE has over twenty-eight year's systems engineering and acquisition related experience in the Department of Defense. He began his career at Naval Air Systems Command where he served as the avionics logistics manager for the SH-60B SEAHAWK and SH-60F helicopters where he was responsible for developing and fielding the support system for the helicopters avionics systems. Since then, he has served as a systems engineer in various offices within the Department of Defense to include the Production and Logistics Directorate, the Directorate for Test, Systems Engineering, and Evaluation, the Directorate for Defense Systems and most recently in the Systems and Software Engineering (SSE) Directorate. In SSE, he currently serves as the senior program support team lead responsible for the Program Support Reviews (PSR) process which provides constructive independent reviews to support both program managers and senior decision makers. As the lead for Land Combat programs, he leads a team which supports program managers develop and document systems engineering, developmental test and evaluation, and acquisition plans. Mr. Nolte has a bachelor of science degree in mechanical engineering from the University of Connecticut and a masters of science in Systems Management from the University of Southern California. He is also DoD Level III certified in Systems Planning Research, Development & Engineering (SPRDE) and Acquisition Logistics. E-mail: Peter.Nolte@osd.mil

JOHN QUACKENBUSH is a retired Army officer with over 20 years of service in a variety of command and staff positions highlighted with multiple tours in tactical attack helicopter aviation units. His significant staff experience includes tours at the Army Human Resources Command and on the Joint Staff where he served in the J-3 Combating Terrorism Directorate and as the Afghanistan country director in J-5. After retiring, he worked in private industry as a systems engineer conducting requirements analysis work for companies such as General Dynamics Robotics Systems and Aurora Flight Systems on projects supporting the Department of Defense. For the past 18 months, Quackenbush has worked for the Department of Defense Acquisition, Technology and Logistics Directorate for Software and Systems Engineering as a program analyst where he helps provide oversight and support to land combat programs in order to develop systems engineering, developmental test and evaluation, and acquisition plans and documents. He has a bachelor's degree in engineering from the United States Military Academy. E-mail: John.Quackenbush.CTR@osd.mil

GRANT RICHARD SCHMIEDER is a senior reliability engineer supporting the OUSD(AT&L)SSE/AS in Arlington, VA. Mr. Schmieder has 24 years of experience in logistics and reliability engineering including 17 years in defense; a BS in engineering from the University of Santa Clara in Santa Clara, CA; an MS in systems and industrial engineering from the University of Arizona in Tucson, AZ. He spent 16 years with Hughes (now Raytheon) Missile Systems in Tucson, Arizona, working as a logistics engineer on the Maverick, AMRAAM, ESSM, AIM-9X, and Phoenix missile programs. Transitioning to reliability engineering while obtaining his masters degree in systems and industrial engineering, he served as the reliability engineer on the Stinger missile system and for a short time on Tactical Tomohawk. After a brief stint as a reliability engineer for Cubic Defense Systems in San Diego, CA, he spent 7 years working as the reliability lead for Telabs' Salix Division and BioVeris Corporation in Maryland. He's now in his first year of supporting the OUSD(AT&L). He is an American Society for Quality Certified Reliability Engineer and holds the Professional Certificate in Reliability and Quality Engineering from the University of Arizona. E-mail: Grant.Schmieder.CTR@osd.mil

Endnotes

[1]Programs may breach Nunn-McCurdy thresholds with significantly higher than projected Program Acquisition Unit Cost or Average Procurement Unit Cost in the Acquisition Program Baseline due to resulting corrective action costs.

[2]Discussion of the Materiel Availability (A_M) KPP must begin with the differences in purpose between A_M and the more well known Operational Availability (A_O) metric. The purpose of A_O is to provide a measure of a single end item readiness for use when intended. As such, the uptime and downtime calculations for A_O are related to restoring individual end items to use after a maintenance action is performed. Conversely, A_M applies to the entire inventory of a given end item and covers not only those end items in operational use but also those in a temporary non-operational state. For the A_M metric, items in a temporary non-operational state (at depot for overhaul/upgrade, held in reserve as spares, not assigned to an operational unit, etc.) are recorded as being "down" (i.e. unavailable for operational use). For the A_O metric, most temporary non-operational states are not considered as either "up" or "down" times.

ITEA Journal 2008; 29: 251–253
Copyright © 2008 by the International Test and Evaluation Association

Determining How Much to Invest in Reliability

Jim Forbes, Ph.D. and E. Andrew Long
Logistics Management Institute, McLean, Virginia

This article reports on the development of a reliability investment model that can be used to predict the investment in reliability required to achieve a given amount of reliability improvement. Earlier work by Logistics Management Institute (LMI) in 2007 provided evidence that such a model was feasible. The reliability investment model comprises a set of four submodels that include:

- *A basic model that computes reliability development effort and cost as a function of program size and desired reliability improvement;*
- *An intermediate model that computes development effort (and schedule) as a function of program size, desired reliability improvement, and a set of relevant cost drivers;*
- *A detailed model that incorporates the characteristics of the intermediate version with an assessment of the cost driver's impact on each step of the reliability engineering process;*
- *A companion production/support cost model to estimate.*

This article discusses the basic and intermediate submodels.

Key words: Cost estimating; cost of unreliability; reliability investment model; reliability values.

In his guest editorial, Dr. McQueary makes the point that rapid acquisition strategies were used to circumvent the traditional acquisition milestone system—including dedicated operational test and evaluation (OT&E). He also noted that this practice caused a shift in the major expense of systems from development activities to operations and sustainment, because inadequate suitability and reliability have increased logistics footprint, sustainment costs, and reduced system effectiveness. Assuming that the technical engineering understanding necessary to fix the problem exists (although not necessarily universally), an important issue is underinvestment in reliability engineering early in programs. The question then is: "How much investment is needed?" The purpose of this article is to address research currently underway to answer that question.

There have been efforts to answer the reliability investment question dating back to at least the late 1970s. We believe it is fair to say that there has been limited success at best. The more promising efforts developed various forms of algorithms (e.g., for least cost allocation of reliability) but then did not succeed in connecting the algorithms to empirical data in a way that could be used for cost estimating. As a consequence, reliability engineers and logisticians, when defending reliability values and budgets have been compelled to argue all to often simply from emotion.

In late 2006 the Department of Defense (DoD) Director of Operational Test and Evaluation tasked Logistics Management Institute (LMI) to look at the downstream costs that were resulting from the adverse trend coming out of operational test. In a June 2007 report,[1] we examined reliability investment and improvement data from six programs and analyzed five in some detail. Although the intent of the study was to characterize the cost of unreliability, an unanticipated result was the discovery of what appeared to be a heretofore unrecognized, systematic relationship between investment in reliability and reliability improvement. We noted that when investment in reliability was divided by average production unit cost (essentially normalizing for complexity), the logarithm taken of that ratio and then the logarithm taken of the improvement in reliability, the result was a straight line. Such log-log endogenous technological progress functions have been uncovered in other domains but, to our knowledge, this is the first such relationship for reliability.

Although intriguing, a relationship defined by five data points is an insufficient basis for cost estimating purposes. Therefore, as a follow-on to the 2007 study, the

The ITEA Journal of Test and Evaluation jite-29-03-16.3d 18/8/08 11:35:28 251

Director of Operational Test and Evaluation; Deputy Director, Assessments and Support, Systems & Software Engineering; and the Deputy Under Secretary of Defense (Logistics and Materiel Readiness) tasked LMI to develop a reliability investment model that would assist in determining how much investment in reliability is needed to achieve a desired amount of reliability improvement. The earlier work was taken as evidence that such a model was feasible. The reliability investment model is intended to comprise four submodels.

- A basic model that computes reliability development effort and cost as a function of program size and desired reliability improvement. Use of this model would be suitable, for instance, on the very front end of programs when little more is known beyond desired reliability improvement.
- An intermediate model that computes development effort (and schedule) as a function of program size, desired reliability improvement, and a set of relevant cost drivers.
- A detailed model that incorporates the characteristics of the intermediate version with an assessment of the cost driver's impact on each step (analysis, design, etc.) of the reliability engineering process.
- A companion production/support cost model to estimate
  - delta investment in production (e.g., for retrofit and spare parts) and
  - change in operations and support cost.

Readers who are familiar with the software engineering constructive cost model (COCOMO) may note a parallelism between COCOMO's submodels and the set of submodels described above. This was deliberate; the intent was a set of models spanning the range from high-level to high-fidelity—a COCOMO for reliability.

The basic model will hereafter be referred to as phase IIA (the earlier study being phase I, although not called so at the time) and the intermediate model as phase IIB. Since the fall of 2007 LMI has been developing the phase IIA and phase IIB models. The detailed model, notionally phase III, has intentionally been held in abeyance pending results of phase II. However the companion production/support cost model (for which there is no COCOMO parallel) is part of the phase II effort.

Phase IIA model

To create the phase IIA model we performed regression analysis. Using data from 17 projects, we developed the cost estimating relationship (CER) shown in *Figure 1*. So far as we have been able to determine, this is most likely the first time such a CER has been available. Required investment increases

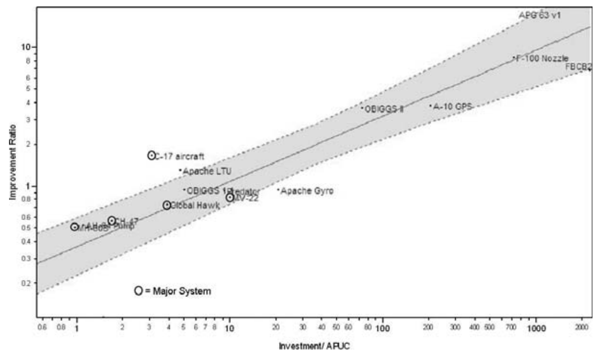


Figure 1. Phase IIA cost estimating relationship

linearly with average production unit cost (APUC) and is a power function of the reliability improvement ratio. APUC is essentially normalizing for program size and complexity. The reliability improvement ratio is (new MTBx - old MTBx)/old MTBx. MTBx can be mean time between failure, mean time between removal, mean time between system abort, mean cycles between removal, or other similar measure—as relevant to a specific program. In the case of a modification, old MTBx is the reliability prior to the modification. In the case of a new development, since there are very few programs for which a legacy predecessor does not exist, old MTBx can be taken to be the reliability of an analogous predecessor. In cases where a one-to-one analogy to a legacy predecessor is not reasonable, a synthetic predecessor may need to be constructed from subsystems of legacy programs.

The R^2 (explanatory power) of the CER is 0.81 and the CER appears to be valid across technologies, across different types of weapon systems, and across a wide range in complexity extending from components to subsystems to complete platforms. The programs in the sample spanned the timeframe from early 1980s until after the turn of the century. The reliability data are from either test or service reliability management systems. Cost data are from service budget submissions or other related sources.

Phase IIB model

The phase IIB model, in development at the time this article was written, is based on the mathematics that underlie the Army Materiel System Analysis Agency (AMSAA) Maturity Prediction Model (AMPM). Starting from the same premises as the AMPM, LMI rederived the model while incorporating terms representing cost. For purposes of development, we divided the reliability engineering process into three periods:

- A design period beginning with Old MTBF (M_0) and producing the initial reliability entering test, analyze, and fix (M_i);

- A test, analyze, and fix (TAAF) period, beginning with M_i and ending with the final reliability (M_f); and
- A validation period that includes sufficient testing to provide a desired degree of confidence that the required M_f has been achieved.

Since the design-period model was in the earliest stages of development at the time of this writing and the validation period model is exploratory, the remainder of this article will concentrate on progress with the TAAF period model.

TAAF period model

Central to both the AMPM and the TAAF phase reliability estimating model is the notion of A modes and B modes. One of the earliest descriptions of the A-model and B-model concept and an excellent overview is Larry H. Crow's article, "Achieving High Reliability" in the Fourth Quarter 2004 *Journal of the Reliability Analysis Center*

In simplified form:

- A-modes are failure modes that management decides not to address, for whatever reason (low consequence, too costly, etc.);
- B-modes are failure modes that management elects to address.

Removal of B-modes results in growth. Conversely, A-modes, since they are not removed, contribute to remaining unreliability as do B-modes to the extent they are not mitigated. The AMPM relies on this concept to project reliability growth. As noted above, LMI rederived the AMPM while incorporating terms representing cost.

After developing the needed formulas, we used 14 data points from early AMSAA estimates performed for the concept development phase of a major Army program to calibrate the TAAF period model. We inferred a non-dimensional TAAF time, τ, from the AMPM and the ratio M_f/M_i (neglecting the rate of A-mode failures, λA) for each data point. We then determined a model cost for each τ and calibrated the model by adjusting two parameters. Finally, we compared costs estimated by the model with the data points. Results are given in *Figure 2*, and show a good fit to the data. The mean absolute deviation of the data points from the line representing best fit is 19 percent. (This is not a form of linear regression so R^2 is not applicable.)

There are two important caveats regarding this result. First, although the data are from multiple platforms they are from a single program and a single service. Secondly, and probably more importantly, in contrast to the basic model where we used demonstrated reliability values, the data in this instance are estimated costs and estimated reliability improvements. Thus, additional validation of the TAAF period model is needed. ❑

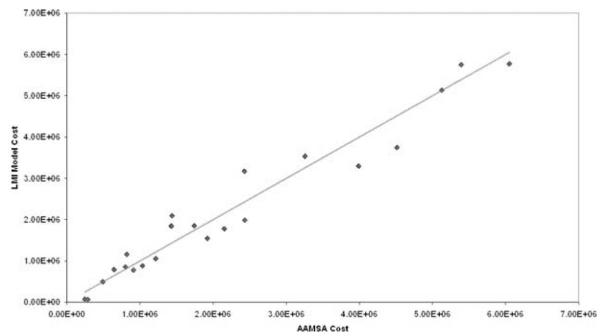


Figure 2. Phase IIB test, analyze, and fix (TAAF) period model

Acknowledgments

In this effort we have received substantial assistance from many organizations both inside and outside the Department of Defense. Primary among these are the Army Materiel Systems Analysis Agency, the Army Evaluation Center, more than two dozen program offices in all three Services, and a number of contractors who graciously provided us crucial data and insights, without which none of the results in this study could have been possible. In particular, AMSAA provided invaluable recommendations on mathematical approaches, identified prior relevant work, and generously participated in a sequence of in-progress reviews along the way.

Dr. Jim Forbes is a certified professional logistician with more than 30 years of experience analyzing policy for logistics, reliability, and maintainability; managing technology programs, acquisition programs, and total ownership costs; and managing operational logistics. At LMI, he has been a principal researcher in numerous studies of depot maintenance, functional strategies for logistics-related information systems, reliability investment, total ownership cost, data management, and knowledge management. E-mail: jforbes@lmi.org

E. Andrew Long has over two decades of experience in a broad range of systems analysis problems in performance and logistics reliability. He has participated in several efforts to help programs address supportability issues, and has performed both theoretical and applied studies of reliability and availability. Recent work includes logistics cost realism analysis for the Coast Guard's Integrated Deepwater System. He performed logistics modeling and footprint analysis for the Army's Future Combat Systems Program. Currently, he supports the OSD Director, Operational Test and Evaluation on reliability issues related to suitability. E-mail: andy.long.ctr@osd.mil

Endnotes

[1]*Empirical Relationships Between Reliability Investments and Life-cycle Support Costs*. SA701-T1, June 2007.

ITEA Journal 2008; 29: 254–262
Copyright © 2008 by the International Test and Evaluation Association

Best Practices for Reliability Assessment and Verification

Michael J. Cushing, Ph.D.
U.S. Army Evaluation Center,
U.S. Army Test & Evaluation Command, Aberdeen Proving Ground, Maryland
Margaret Hockenberry
U.S. Army Materiel Systems Analysis Activity,
U.S. Research Development and Engineering Command, Aberdeen Proving Ground, Maryland
E. Andrew Long
Logistics Management Institute, McLean, Virginia

The Department of Defense (DoD) is working closely with the Government Electronics and Information Technology Association on the development of a GEIA-STD-0009, Reliability Program Standard for Systems Design, Development, and Manufacturing, at the behest of the Defense Science Board Developmental Test Task Force. It is hoped that GEIA-STD-0009 will improve the odds that military systems will successfully demonstrate reliability requirements in both developmental and operational testing. This article provides an overview of GEIA-STD-0009 along with initial guidance regarding its application, with an emphasis on the assessment and verification of system reliability.

Key words: Developmental testing; failure mode; GEIA-STD-0009; reliability; operational load; system design; system reliability modeling.

During the past year, the U.S. Department of Defense (DoD) has been working closely with both industry and the Government Electronics and Information Technology Association (GEIA) on the development of a new standard, GEIA-STD-0009, *Reliability Program Standard for Systems Design, Development, and Manufacturing*. The DoD's motivation for this undertaking is that many systems are not achieving the required reliability during developmental testing and are subsequently found unsuitable during Initial operational test and evaluation. The Defense Science Board Developmental Test (DT) and Evaluation Task Force examined this issue and concluded that a new reliability program standard is urgently needed. The purpose of this article is to provide a brief overview of the new standard as well as guidance regarding how to best assess and verify system reliability with it.

GEIA-STD-0009 overview

Embodied in GEIA-STD-0009 is a new approach to the development, production, and fielding of reliable systems. As depicted in *Figure 1*, the standard is primarily comprised of four objectives:

1. understand customer/user requirements and constraints
2. design and redesign for reliability
3. produce reliable systems/products
4. monitor and assess user reliability

During the development of GEIA-STD-0009, the Working Group identified the essential reliability processes (termed "*reliability activities*" both in the Standard and herein) that simply must be performed in order to design, grow, build, and field reliable systems. The *reliability activities* are mandatory in nature but merely specify "what to do."

GEIA-STD-0009, at its core, is a reliability engineering and growth process that is fully integrated with systems engineering as depicted in *Figure 2*. The new standard is not a menu of reliability tasks that one may select from as with many previous reliability program standards. Readers who are concerned that their favorite reliability methods or tools do not figure prominently in the Standard should not fear. The primary mechanism for tailoring GEIA-STD-0009 is by selecting "how to" and best practices in order to implement each of the activities. Many of these methods and tools are listed in Annex A of the Standard and are essential to the implementation of

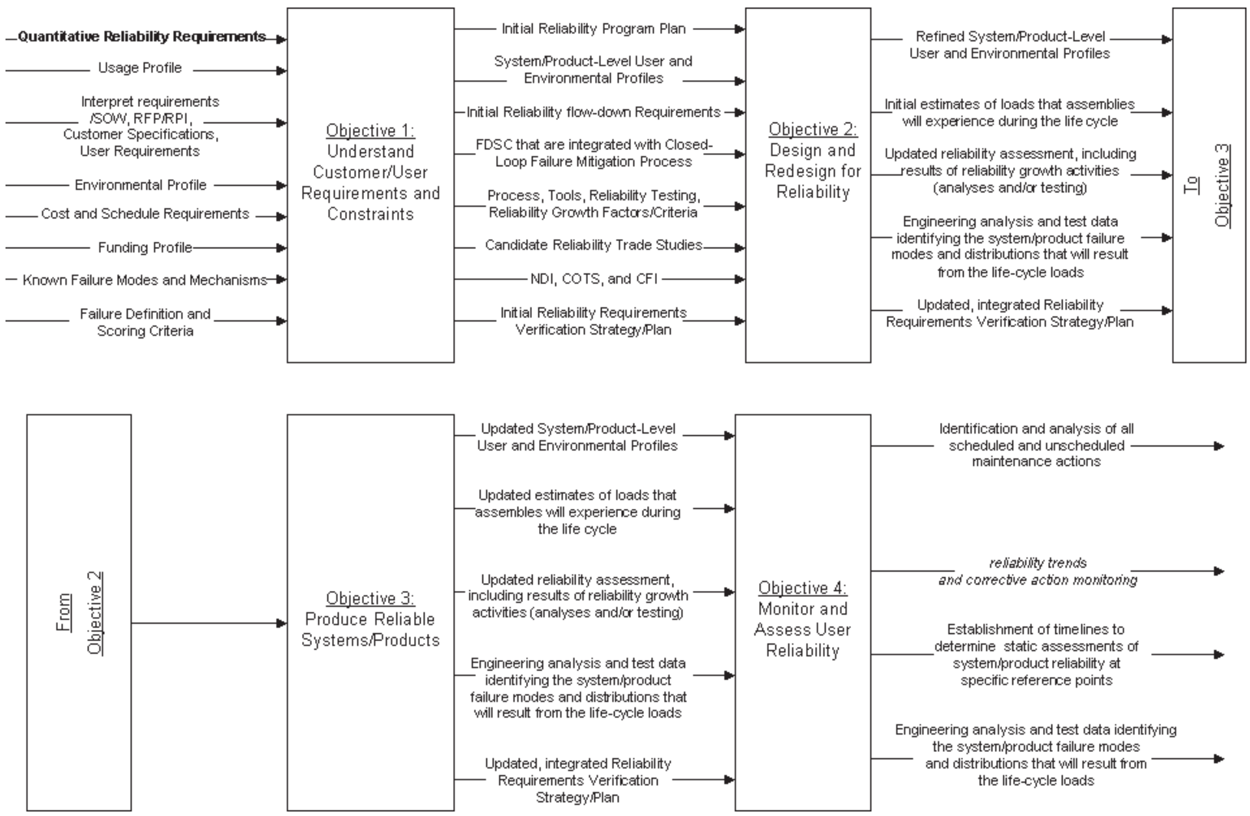


Figure 1. GEIA-STD-0009 objectives

the reliability activities. A reliability scorecard developed by AMSAA and the DoD Reliability Improvement Working Group can be used to guide the selection of reliability methods, tools, and best practices. The scorecard can be found at the Defense Acquisition University's website by selecting the Acquisition Community Connection, then the Reliability & Maintainability Special Interest Area, then tools. The current link is: https://acc.dau.mil/CommunityBrowser.aspx?id=210483&lang=en-US.

It is envisioned that the developer will be tasked up-front to draft a Reliability Program Plan, perhaps as part of the System Engineering Plan, so that the staffing and scheduling of the reliability program will be understood and budgeted from the beginning. Experience teaches that if the developer does not properly budget and plan for the reliability program before contract award, it is very difficult to fold it in afterwards.

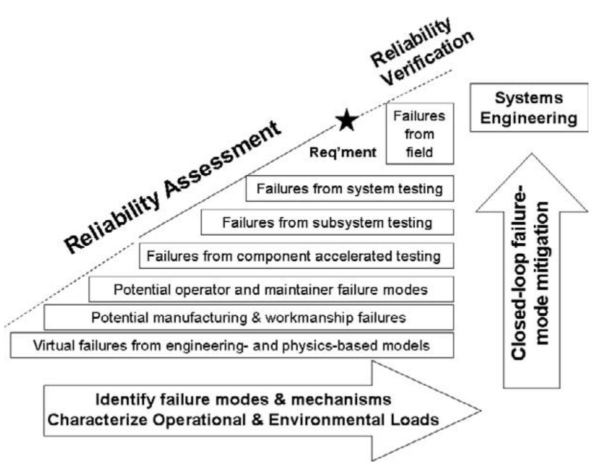


Figure 2. GEIA-STD-0009 integrated growth

Reliability engineering and growth integrated with systems engineering

As depicted in *Figure 2*, GEIA-STD-0009 embodies a systematic design-reliability-in process, not a process that focuses on identifying and improving a few reliability-critical components. There are three critical elements:

1. progressive understanding of the system-level operational and environmental loads and the resulting loads and stresses that occur throughout the structure of the system;
2. progressive identification of the resulting failure modes and mechanisms;
3. aggressive mitigation of surfaced failure modes.

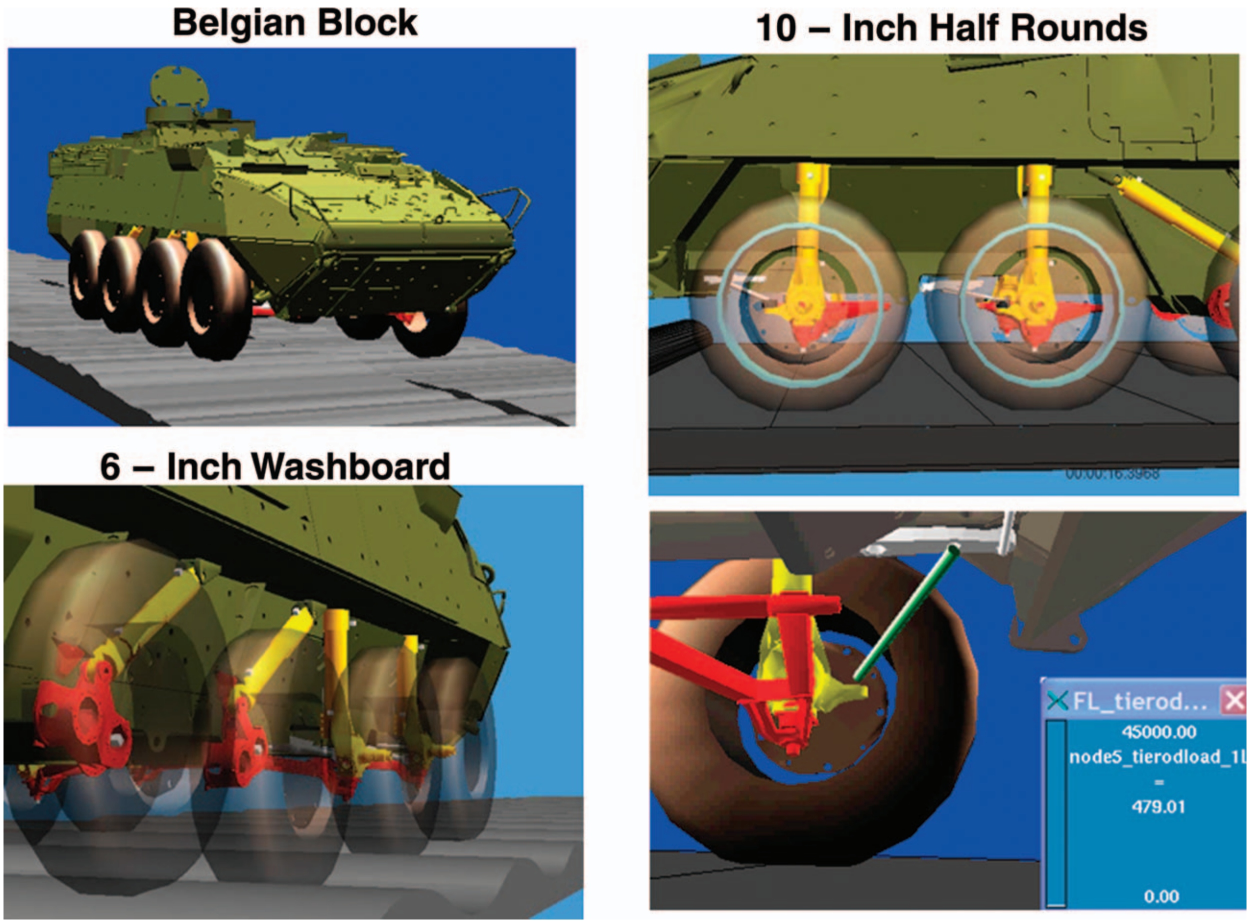


Figure 3. Dynamic simulation of tactical wheeled vehicle

Operational and environmental loads and stresses

The reliability that a system will demonstrate is, in part, a function of the life-cycle operational and environmental stresses that occur throughout the structure of the system. Operational loads result from user or maintainer actions as well as from external systems the system under development will interface with.

In GEIA-STD-0009, the operational and environmental loads to be imposed on the system are progressively characterized and designed for throughout development. This effort starts with information from the customer. For DoD customers, the system-level operational and environmental loads are typically defined by an Operational Mode Summary/Mission Profile (OMS/MP). GEIA-STD-0009 explicitly tasks the developer to study the OMS/MP and work with the customer in order to obtain added details if the OMS/MP is not specific enough for engineers to design to. If need be, the developer will seek access to customer assets (e.g., test courses or vehicles that the system will be integrated with) in order to obtain the needed specifics.

The developer progressively characterizes the resulting loads and stresses throughout the structure, down to components or assemblies being selected and integrated into the design, to include commercial off-the-shelf (COTS), nondevelopmental items (NDI), and government-furnished equipment (GFE). It is not possible to design reliable components, nor select and reliably integrate COTS, NDI, and GFE, without accurate estimates of the loads to be imposed on them. The operational and environmental load estimates must be verified to be operationally realistic with measurements using the production-representative system in time to be used for reliability verification.

The progressive characterization of loads and stresses is routinely done by the U.S. Army. *Figure 3* depicts the dynamic simulation of a tactical wheeled vehicle traversing some of the challenging road surfaces found at the Aberdeen Test Center. The simulation provided loading information on various suspension components including the A-arm. *Figure 4* depicts the

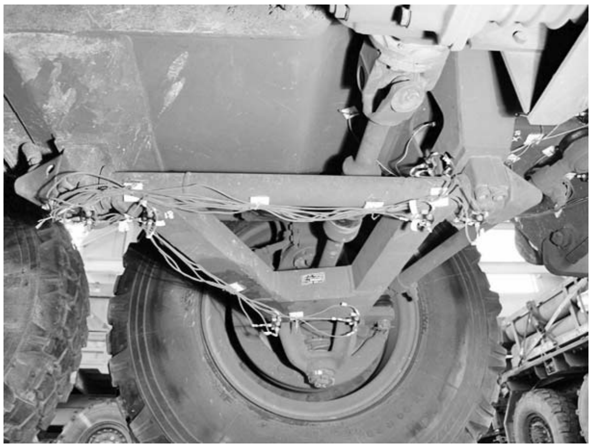


Figure 4. Instrumenting of A-arm

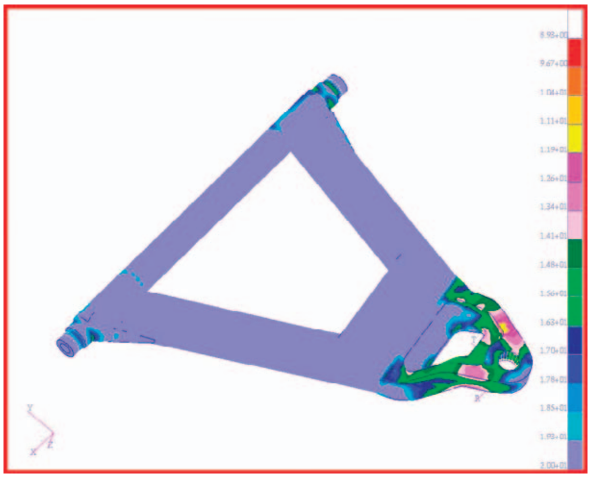


Figure 5. A-arm fatigue life calculations

instrumenting of an A-arm on an actual vehicle. The simulation and test data were compared in order to confirm the accuracy of the simulation model.

Identify and characterize failure modes and mechanisms

As depicted in *Figure 2*, GEIA-STD-0009 includes a robust effort to identify and characterize failure modes and mechanisms as soon as development begins. This is essential if the system is to enter subsystem test with a level of reliability that will lead to the successful achievement of reliability requirements.

Teams developing assemblies, subassemblies, and components for a system identify and confirm through analysis, test, or accelerated test the failure modes and distributions that will result when life-cycle operational and environmental loads are imposed on these assemblies, subassemblies, and components. Teams selecting and integrating items not specifically developed for this system (which may include COTS, NDI, and GFE, as well as other assemblies, subassemblies, and components) identify and confirm the failure modes and distributions that will result when these life-cycle loads are imposed on these items. Estimates of life-cycle operational and environmental loads on assemblies, subassemblies, and components are used as inputs to engineering-based and physics-based models in order to identify failure mechanisms and the resulting failure modes.

Figure 5 illustrates the A-arm from the tactical wheeled vehicle depicted in *Figures 3 and 4*. A likely failure mode for the A-arm is that a crack will develop and grow as a result of fatigue. Finite element analysis was used in order to estimate the stresses throughout the component that would result from the cyclic loads placed on it. *Figure 5* depicts the results, including when and where fatigue failure should first occur.

It is often the case that the discovery of failure modes that are typically charged to operators or maintainers does not occur until testing with actual operators and maintainers begins. GEIA-STD-0009 includes a proactive requirement that these failure modes are to be identified through analysis during system design. Failure modes and distributions that may be induced by manufacturing variation or errors are also to be identified during design rather than waiting until production. It is generally simpler and less expensive to mitigate failure modes the earlier they are discovered.

GEIA-STD-0009 requires that all failures that occur during accelerated, subsystem, or system testing are analyzed until the root-cause failure mechanism has been identified. Identification of the failure mechanism provides the insight essential to the identification and formulation of reliability improvements. The process of identifying and understanding failure modes and mechanisms continues as the design and manufacturing processes evolve.

Failure-mode mitigation

The developer aggressively mitigates failure modes to ensure the reliability requirements are successfully verified and do not degrade during production or in the field. Failure modes must be aggressively mitigated before subsystem testing begins in order to obtain a reliability level that will enable reliability growth to the requirement through subsystem and system testing. Failure modes are mitigated by one or more of the following approaches:

- eliminating the failure mode;
- reducing its occurrence probability or frequency;
- incorporation of redundancy; and/or
- mitigation of failure effects (e.g., fault recovery, degraded modes of operation, providing advance warning of failure).

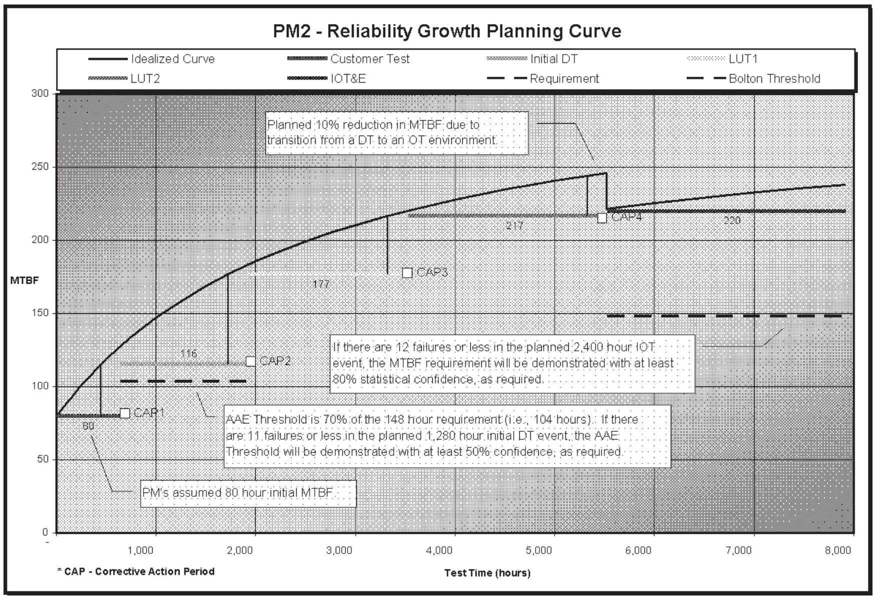


Figure 6. Reliability growth planning curve

The developer submits the potential reliability improvements identified during the execution of the Reliability Activities to the appropriate engineering organizations (e.g., Systems Engineering). The developer employs a mechanism that is accessible by the customer (e.g., a failure reporting, analysis, and corrective action system or a data collection, analysis, and corrective action system) for monitoring and communicating throughout the organization data regarding the identification and mitigation of failure modes. Failure modes that are expected to occur during the system life cycle are included in the system reliability model.

Reliability assessment

In GEIA-STD-0009 the term "*reliability assessment*" denotes the periodic assessment of reliability progress towards requirements and it is followed by "*reliability verification*" which denotes the formal verification that requirements have been met. The standard establishes seven general *reliability assessment* requirements:

1. The developer assesses the reliability of the system periodically throughout the life cycle using the system reliability model, the life-cycle operational and environmental load estimates generated from the OMS/MP, and the customer-supplied Failure Definitions and Scoring Criteria.
2. Reliability assessments are made based on data from analysis, modeling and simulation, test, and the field, and are tracked as a function of time and compared against reliability allocations and customer reliability requirements.
3. For complex systems, or when the customer requires this, the assessment strategy includes reliability values to be achieved at various points during development.
4. The developer monitors and evaluates the reliability impact of changes to the design or manufacture of the system.
5. The implementation of corrective actions is verified and effectiveness is tracked.
6. Formal reliability growth methodology is used where applicable (e.g., when failure modes are discovered and addressed with a test-analyze-and-fix process that is applied to complex assemblies) in order to plan, track, and project reliability improvement.
7. Predicted failure modes and mechanisms are compared with those from test and the field.

The third requirement in the list above is of particular interest for Army programs because new policy requires that at least one intermediate reliability-growth value be included in the request for proposals. Such an intermediate reliability-growth value will permit the early identification of a system that is not on-track towards meeting its reliability requirement, which will allow time to make program adjustments and intensify the reliability engineering and growth process. One approach to obtaining such an intermediate value is through the use of a reliability growth planning model. The customer can develop a reliability growth plan up-front based on the program schedule, test assets, and some assumptions concerning the intensity of the reliability growth effort.

Figure 6 depicts a notional Army reliability growth planning curve based on the PM2 model (Ellner and

Hall 2006). Using the PM2 model, the Army determines that the system mean time between failures (MTBF) must grow to 1,227 as a point estimate in order to have a reasonable chance (50 percent in this case) of demonstrating 690 hours with 80 percent statistical confidence (assuming a 10 percent drop from DT to operational test (OT)). This plan consists of four corrective action periods between five test events: a Customer Test, an Initial DT, a Limited User Test, a Low-Rate Initial Production DT, and an Initial OT. Since this is an Army program, the Initial DT is where the system MTBF must be demonstrated to be at least 70 percent of the requirement with 50 percent statistical confidence. The Army can incorporate this plan in the Request for Proposals so that the developer can design the reliability program accordingly.

The fourth reliability assessment requirement in the list above (i.e., monitoring and evaluating the reliability impact of changes to the design or manufacture of the system) is critically important to maintaining reliability during production and in the field, and it may require an intensive effort given the widespread use of complex global approaches to manufacturing. Several methods for implementing this requirement, such as parts control and supply chain management, are identified in Annex A of the standard.

The failure mode identification and mitigation activities discussed earlier lead to a two-part reliability growth program:

- Reliability growth driven by (a) engineering-based and physics-based models, (b) accelerated testing of low indenture-level items, (c) analyses that identify failure modes related to manufacturing variation and workmanship errors, and (d) analyses that identify failure modes that are typically charged to operators or maintainers.
- Reliability growth driven by a test-analyze-and-fix process applied under operationally-realistic conditions to complex assemblies such as subsystems and systems.

The first part of the growth program provides the high starting point for the traditional reliability growth program that is pivotal to success.

The Reliability Assessment process consists of two types of DT:

- testing, primarily accelerated testing, of low indenture-level items such as components and noncomplex assemblies, in order to surface and mitigate failure modes not readily identified with engineering-based and physics-based reliability modeling; and
- testing of complex assemblies such as subsystems and systems in order to surface and mitigate failure modes not readily identified otherwise.

Reliability assessment can be divided into three phases:

- assessment of requirements feasibility;
- assessment before subsystem testing begins; and
- assessment after subsystem testing begins.

Each will be addressed in more detail.

Requirements feasibility

During the execution of the first objective, the developer must acquire an understanding of the customer's reliability requirements. It is at this point that an assessment of the feasibility of the requirements is made. The system reliability model is used, in conjunction with expert judgment, to assess if the design (including COTS, NDI, and GFE) is capable of meeting reliability requirements in the user environment. If the assessment is that the customer's requirements are infeasible, the developer communicates this to the customer. Clearly this is not an analysis of a design but is rather an assessment of whether it is possible for a new design to meet reliability requirements given previous designs and projections of potential improvements.

Assessment before subsystem testing begins

In general, it is not possible to estimate the reliability that a system will demonstrate under operationally-realistic conditions until subsystem and system testing under these conditions begins. This is why this portion of the reliability growth curve is dotted in *Figure 2*. What can be done at this stage is an expert assessment of the quantity and quality of the failure modes identified, and the effectiveness of the associated mitigation. A key rule of thumb is that a high percentage of failure modes surfaced must be effectively mitigated in order to put the system on a successful reliability-growth path. Mitigation effectiveness can be evaluated in a variety of ways. If the failure mode was identified through the use of engineering-based or physics-based models, accelerated testing can be used to confirm that it occurs as expected and is well understood. It is also beneficial to compare the predicted to measured operational and environmental loads and stresses. This is beneficial because these loads are used to design reliability into new components as well as select and reliably integrate COTS, NDI, and GFE.

A major pitfall to be avoided concerns predicting system reliability under operationally-realistic conditions. This is generally not possible before system reliability testing begins which can be quite frustrating. Many programs perform handbook-based reliability predictions but such predictions are inaccurate because operational reliability is largely determined by stress and design specifics that handbook prediction models do not accept (Pecht and Nash 1994). Reliance on handbook predictions can lead a program to believe the system is ready for Reliability Verification when it is not.

Assessment after subsystem testing begins

Estimation of system reliability can begin once testing of subsystems or systems under operationally-realistic conditions begins. Testing of the first configuration establishes the initial reliability for reliability growth tracking. The implementation of corrective actions is verified and effectiveness is tracked. Predicted failure modes/mechanisms are compared with those from test and the field. Reliability growth methodology is used to plan, track, and project reliability based on failure data from complex assemblies tested under operational and environmental loads. Military Handbook 189, which is currently being revised, may be used as a guide. One may also consult the DoD Guide for Achieving Reliability, Availability, and Maintainability.

Reliability verification

As mentioned earlier, in GEIA-STD-0009 the term "*reliability verification*" denotes the formal verification that requirements have been met. The standard establishes six general reliability verification requirements:

1. The developer plans and conducts activities to ensure that the achievement of reliability requirements is verified during design.
2. The developer develops and periodically refines a Reliability Requirements Verification Strategy/Plan that is an integral part of the systems-engineering verification and is coordinated and integrated across all phases.
3. The strategy must further ensure that reliability does not degrade during production or in the field.
4. The verification is based on analysis, modeling & simulation, testing, or a mixture, and must be operationally realistic.
5. The verified system-level operational & environmental life-cycle loads, as well as the Failure Definitions and Scoring Criteria, must be used.
6. Additional customer requirements, if any (e.g., reliability qualification testing, testing in customer facilities, customer-controlled, customer-scored testing), must be included.

The latter portion of reliability assessment consists of testing activities that the DoD refers to as DT. DT is often followed by OT to assess how well the system will work when actual operators and maintainers use it under field conditions. The Standard facilitates the integration of DT and OT because the following are required:

- Operational loads (including from systems that interface with the system under development) and environmental loads are developed based on the OMS/MP, progressively refined, and eventually verified to be accurate and operationally realistic.
- Failure modes that are typically charged to operators or maintainers are identified earlier. These failure modes generally arise for the first time during OT and result in statistically-significant differences between the DT and OT reliability estimates.
- System reliability modeling is developed and refined as failure modes are identified, analyzed, mitigated, and incorporated in the modeling.

One item that needs to also be addressed to facilitate the estimation of reliability using both DT and OT data concerns balancing of the sample sizes so that a statistical comparison of the reliability estimates is credible. Even though the reliability estimates from DT and OT may appear to be quite different, it can be difficult to prove this statistically if either the DT or OT sample size is too small relative to the other. One must design the DT and OT sample sizes so they can be credibly compared before deciding whether to aggregate them.

MTBF-type reliability requirements are often verified using a fixed-configuration, fixed-length test plan from Military Handbook 781 (MH-781). One needs the following information in order to select such a test plan:

a) the MTBF to be demonstrated with statistical confidence;

b) the minimum level of statistical confidence that a) should be demonstrated with;

c) the best pretest estimate of the actual MTBF; and

d) the probability of passing the test if c) is accurate.

In MH-781, the MTBF to be demonstrated with statistical confidence is termed the "lower-test" MTBF and the minimum level of statistical confidence it should be demonstrated with equals one minus the "consumer risk." So in order to demonstrate an MTBF with at least 80 percent confidence, one should select a plan with a consumer risk of 20 percent. It is items c) and d) that are frequently misunderstood. The best pretest estimate of the actual MTBF is termed the "upper-test" MTBF in MH-781. This pretest MTBF estimate must be greater than the MTBF to be demonstrated with confidence. In order to use one of the standard plans the ratio of the pretest estimate to the MTBF to be demonstrated with confidence must

Figure 7. Army bridging system

be either 1.5, 2, or 3. It is unlikely that the pretest estimate and the MTBF value to be demonstrated with confidence will have this relationship so one should expect to design a custom plan. The probability of passing the test if the pretest estimate is accurate equals one minus the "producer risk." If one desires a test plan where the probability of passing is 80 percent, provided the pretest MTBF estimate is accurate, then a plan with a producer risk of 20 percent should be selected.

Many practitioners do not understand how to select a MH-781 test plan as just described which can result in the selection of a test plan that is unlikely to be passed. It is expected that MH-781, which is currently under revision, will be edited so that the logic described above is clearer. Regardless of the test-planning resource used, fixed-configuration, fixed-length test plans must be selected using items a) through d) so that the system will be highly likely to pass if it is as reliable as the developer believes and will be highly unlikely to pass if the MTBF is below the requirement.

In some cases it is impossible to rely exclusively on a reliability demonstration and a mix of modeling, analysis, and test may be needed. One example is a mobile Army bridging system, pictured in *Figure 7*, that can span gaps of up to 12 meters. Historically, the cost and time associated with conducting large scale bridge crossing tests precluded full testing of the requirement to levels of statistical confidence. To solve this problem, the Aberdeen Test Center developed the Bridge Crossing Simulator device, which physically simulates the loads imposed by a crossing vehicle on a bridge under test, allowing durability testing to be conducted quickly and economically. While the Army bridging system was under test on the Bridge Crossing Simulator a problem developed. The bridge center coupler connection failed before the bridge had reached its required durability life. Army Materiel Systems Analysis Activity (AMSAA) engineers used a physics-based computer modeling analysis technique—PoF—to identify the root causes of the failure and to recommend a design improvement.

The recommendation suggested adding structural angle sections to connect the center couplings of the bridge to the vertical webs, which would create a much stronger double-shear connection. The new design proposal eliminated the weak spot in the weld between the bridge bottom flange and vertical web where the previous failure had originated. U.S. Army Tank Automotive Research, Development and Engineering Center (TARDEC) engineers, along with the Product Manager (PM) Assured Mobility Systems, reviewed the results of an upgrade feasibility study performed by the bridging system prime contractor to address increased requirements. They determined that the suggested design improvement might not only fix the immediate problem, but would also provide the additional margin needed to upgrade the bridge's load capacity. The PM Bridging, located in the Program Management Office Force Projection, capitalized on the confluence of events and moved forward to upgrade the bridge. TARDEC and the bridging system contractor worked to implement the AMSAA recommendation and to add other enhancements to ensure that the system would meet the new, tougher requirements. After further testing the bridging system finished the durability testing with a few cracks but none that would impact the operational mission.

Summary

The DoD is working closely with the GEIA on the development of GEIA-STD-0009, *Reliability Program Standard for Systems Design, Development, and Manufacturing*, at the behest of the Defense Science Board. It is hoped that GEIA-STD-0009 will improve the odds that military systems will successfully demonstrate reliability requirements in both DT and OT. This article provides an overview of GEIA-STD-0009, along with initial guidance on its application with an emphasis on the assessment and verification of system reliability. ❑

DR. MICHAEL J. CUSHING is director (acting) of the U.S. Army Evaluation Center's Reliability & Maintainability Directorate. He earned a bachelor of science degree in electronics engineering and computer science from the Johns Hopkins University and both master of science and Ph.D. degrees in reliability engineering from the University of Maryland at College Park. During 25 years in U.S. military reliability, he has authored numerous publications, helped formulate and implement a variety of reliability policies, and contributed towards several reliability standards. Current activities include being a member of both the U.S. Department of Defense Reliability Improvement Working Group and the GEIA-Std-0009 Working Group. E-mail: michael.cushing@us.army.mil

MARGARET HOCKENBERRY is a senior operations research analyst in charge of the Reliability Analysis team for the U.S. Army Materiel Systems Analysis Activity (AMSAA). Her team at AMSAA is executive agent for Reliability and Maintainability (RAM) Standardization Reform and Standards Lead for the U.S. Army Research Development and Engineering Command and the Army. Her focus has been on updating reliability handbooks in order to provide the latest methodology. She is a member of the GEIA-STD-0009 Working Group and coauthor of section three of the standard. Hockenberry holds a bachelor of arts degree in mathematics and a master in business administration degree from the Florida Institute of Technology. E-mail: margaret.hockenberry@us.army.mil

E. ANDREW LONG has over two decades of experience in a broad range of systems analysis problems in performance and logistics reliability. He has participated in several efforts to help programs address supportability issues, and has performed both theoretical and applied studies of reliability and availability. Recent work includes logistics cost realism analysis for the Coast Guard's Integrated Deepwater System. He performed logistics modeling and footprint analysis for the Army's Future Combat Systems Program. Currently, he supports the OSD Director, Operational Test and Evaluation on reliability issues related to suitability. E-mail: andy.long.ctr@osd.mil

References

DoD Guide to Achieving Reliability, Availability, and Maintainability. August 3, 2005.

Ellner, P. M. and Hall, J. B. 2006. *Planning model based on projection methodology (PM2)*. Aberdeen Proving Ground (MD): U.S. Army Materiel Systems Analysis Activity. Technical Report No. TR-2006-9.

GEIA-STD-0009. 2008 (expected). *Reliability Program Standard for Systems Design, Development, and Manufacturing*.

Military Handbook 189. 1981. *Reliability Growth Management*.

Military Handbook 781. 1996. *Handbook for Reliability Test Methods, Plans, and Environments for Engineering, Development, Qualification, and Production*.

Pecht, M. G. and Nash, F. R. 1994. "Predicting the reliability of electronic equipment." *Proceedings of the IEEE*. pp. 992–1004.

ITEA Journal 2008; 29: 263–266
Copyright © 2008 by the International Test and Evaluation Association

Life Cycle Metrics and OSD Oversight: Discipline With Flexibility

Randy T. Fowler
Assistant Deputy Under Secretary of Defense (Materiel Readiness), The Pentagon, Washington, D.C.
Sean P. Killoy
Decisive Analytics Corporation, Arlington, Virginia

Recognizing that systems were not achieving adequate reliability during development, leading to reduced warfighting capability and tremendous cost growth, the Joint Requirements Oversight Council acted in August 2006 to require a mandatory sustainment key performance parameter, materiel availability, and two supporting key system attributes, materiel reliability and ownership cost, for all new major defense acquisition programs and other selected programs. This article summarizes the background and rationale for the metrics and discusses some of the additional actions being taken to implement a disciplined approach to reliability and availability.

Key words: Availability; cost of ownership; maintainability; operations; reliability; support.

Whether the symptoms are presented as quality problems, or shortcomings in system reliability as cited in recent Government Accountability Office reports, or as an alarming trend in systems being found unsuitable during operational test and evaluation, it has become clear that the discipline necessary to field reliable, high quality systems needs to be strengthened throughout the Department of Defense (DOD) acquisition system. DOD and Military Service policies continue to stress the importance of reliability, availability, maintainability, and quality, and there is no shortage of reference materials, guides, training, and tools available to support the practitioner, yet there is a systemic shortfall in implementing the best practices necessary to ensure that reliability and quality are cost effectively engineered into our systems. Somehow, the engineering discipline of reliability has been reduced to a minor factor associated with total ownership cost, instead of a fundamental characteristic of our systems. Availability considerations have faded into the background and many of our experienced acquisition professionals are no longer in the workforce. Recognizing the need to act at the enterprise level, the direction taken by the Office of the Secretary of Defense (OSD) Acquisition, Technology, and Logistics (AT&L) leadership and the Joint Requirements Oversight Council (JROC) was based on understanding the integrated acquisition and sustainment environment as it exists today, and was intended to use the primary management structures, tools, and processes in place.

Today's program managers are provided very challenging cost and schedule goals and a small number of key performance parameters (KPPs) and key system attributes (KSAs) that must be met; and everything else may be traded off if necessary to meet those minimum requirements. In reality, that means that everything else MUST be traded when necessary to meet performance, cost, and schedule. In order to get reliable, available, affordable systems, then reliability, availability, and ownership cost had to become firm requirements instead of vague objectives. With the signing of the Joint Requirements Oversight Council Memorandum 161-06 in August 2006, the JROC established the requirements, endorsing the importance of achieving reliability and planning for sustainment as key to future warfighting capabilities. The approach taken in developing the KPP and KSAs was to establish firm requirements for what was to be measured, but provide maximum flexibility to the services and program sponsors to establish the specifics applicable to each system.

Materiel availability—The capstone sustainment KPP

> Materiel Availability *is a measure of the percentage of the total inventory of a system operationally capable (ready for tasking) of performing an assigned mission at a given time, based on materiel condition. This can be expressed mathematically as (number of operational end items/total population).*

> Materiel Availability *also indicates the percentage of time that a system is operationally capable of performing an assigned mission and can be expressed as (uptime/(uptime + downtime)). (CJCSM 3170.01C)*

The introduction of a new term, *Materiel Availability*, has naturally raised questions and concerns, some of which will be addressed by providing the background and context for establishing the new KPP. As a metric, materiel availability provides insight into the usage of systems, but more importantly, requires that all of the major sustainment elements be considered, planned, and measured. The term *Materiel Availability* was deliberately chosen to emphasize that this metric is intended to support the management of the acquisition and materiel readiness processes whose function is to deliver capable, ready systems to the operational forces, as an input to operational readiness. Operational readiness is the concern of the operational forces and must still be measured and reported, and operational availability (A_o) no doubt will remain an important metric. Establishing and managing the materiel availability metric requires the consideration of all of the sustaining support that the acquisition and logistics professional must provide to sustain the capability being acquired, in addition to the reliability and maintainability characteristics of the system itself. The use of the term *Materiel Availability* was intended to highlight its difference from *Operational Availability*, although it can be argued that the definition of materiel availability is basically A_o at the highest level of the system, across the entire population. Since A_o is widely used and has become associated with very specific definitions in each service, selecting *Materiel Availability* further highlights that it is, by design, different from A_o as implemented across the services.

The numerical values for materiel availability will often differ from the values typically experienced for A_o, and may be significantly lower. The number itself provides some insight into the planned utility or capacity of the system, but the number itself is not as important as the discipline introduced by the need to plan and manage to objective values throughout the system. Achieving an arbitrary value for materiel availability is not the goal; rather the numerical value for availability should reflect the actual plans for operating and sustaining the system. For materiel availability, the entire population of the system must be accounted for, as does all time during the planned service life. Nonoperational units and time are included to provide a complete picture of the investment and sustainment required across the entire program life cycle.

The definition of *materiel availability* recognizes that in practical, concrete terms, the most direct way to measure readiness or availability in the field in many cases is to count how many end items are "up" each day. This approach works with some systems, but not all, and in order to establish the metric, analysis of downtime will still need to be conducted.

Generally, while the capstone metric is materiel availability, from the operator's viewpoint the starting point for establishing the metrics will be reliability, established as the new **Materiel Reliability** KSA. The intent of the KSA is to establish the reliability performance that is needed to make the system useful in its intended military context. This metric should be established with significant input from the operational users based on the planned employment of the system. Only the combatant commanders can really answer the question, "How reliable does the system need to be for it be useful in combat?" or "What probability of success must be achieved?"

> Materiel Reliability *is a measure of the probability that the system will perform without failure over a specific interval. Reliability must be sufficient to support the warfighting capability needed. (CJCSM 3170.01C)*

While the definition for *Materiel Reliability* in Chairman, Joint Chiefs of Staff Memorandum (CJCSM) 3170.01C goes on to discuss the use of mean time between failure, it does so in very general terms and was intended to allow the use of specific reliability metrics most appropriate to each system. For complex, multimission systems, it may be appropriate to establish more than one reliability metric or to use probability of mission success as the top level materiel reliability metric. In general, some form of mission reliability is most appropriate, although there may be cases in which logistics reliability would be recommended as the KSA. The selection and definition of the most appropriate metric for each system is left to the sponsor to recommend and support. It is critical to define the operating environments and mission profiles in which the system is intended to operate.

The ITEA Journal of Test and Evaluation jite-29-03-10.3d 18/8/08 11:32:54 264

Left unbounded by cost, systems could achieve availability objectives by requiring excessive spares, maintenance, or other support elements and reliability gains could be pursued beyond the point of diminishing returns. Establishing **Ownership Cost** as a KSA is intended to add cost discipline beyond that provided by the current approach to total ownership cost and life cycle cost estimates.

> Ownership Cost *provides balance to the sustainment solution by ensuring that the operations and support (O&S) costs associated with materiel readiness are considered in making decisions. For consistency and to capitalize on existing efforts in this area, the Cost Analysis Improvement Group (CAIG) O&S Cost Estimating Structure will be used in support of this KSA. Fuel costs will be based on the fully burdened cost of fuel. Costs are to be included regardless of funding source. The KSA value should cover the planned lifecycle timeframe, consistent with the timeframe used in the Materiel Availability KPP. Sources of reference data, cost models, parametric cost estimating relationships, and other estimating techniques or tools must be identified in supporting analysis. Programs must plan for maintaining the traceability of costs incurred to estimates and must plan for testing and evaluation. The planned approach to monitoring, collecting, and validating operating and support cost data to supporting the KSA must be provided. (CJCSM 3170.01C)*

Since acquisition costs are intensely monitored already, the KSA is focused on O&S costs, and is intended to elevate management attention to O&S cost considerations. However, it was also recognized that the quality and completeness of O&S cost data available is less than that of acquisition cost data, and the connection between the O&S estimates and the eventual costs incurred is soft in some areas. Over the long term, this area will continue to mature, with the objective of eventually being able to rely on the O&S cost KSA values as the basis for planning and budgeting.

Only the cost elements most directly associated with materiel readiness are required, but program sponsors are free to add other cost elements if appropriate. For example, manpower costs are not required, but there are some systems for which manpower costs are the focus of significant program effort and should be included. The CAIG O&S Cost Estimating Structure is used so that the KSA does not create a new or different cost structure that would differ from that used for other program estimates.

Supporting analysis

Clearly all of the analysis required to establish the KPP and KSAs will not be included in the Capability Development Document (CDD) or Capability Production Document (CPD) with the established KPP and KSAs. To support the immediate requirements of programs submitting CDDs or CPDs for JROC approval, the Office of the Deputy Under Secretary of Defense for Logistics and Materiel Readiness and the Joint Staff J4 developed a "Guide to the Sustainment KPP" issued by J4 and available on the Joint Staff Knowledge Management and Decision Support system. The Guide defines the requirements for supporting analysis and the process by which recommendations for KPPs and KSAs are reviewed. While the Guide does not mandate specific formats or products, it does describe the criteria by which the analysis will be evaluated. A more definitive document is currently in development which will formalize and standardize the required analytical products. Currently titled the "DoD Reliability, Availability, Maintainability and Cost Rationale Report Handbook," it is intended to support three primary objectives:

1. Provide guidance in developing and documenting realistic sustainment KPP/KSA requirements and the related supporting rationale;
2. Provide guidance to the acquisition community to understand how the requirements must be measured and tested throughout the life cycle;
3. Describe the processes for OSD AT&L, the Joint Staff and other stakeholders to follow in interfacing with the Services and programs in developing sustainment requirements.

Using the processes established in the handbook will assist in assessing alternatives considered during the Analysis of Alternatives, and in articulating the requirements and the supporting rationale needed for the CDD and CPD. Subject matter experts from OSD AT&L, DOT&E and the Joint Staff have worked together to develop the handbook which was inspired by the "RAM Rationale Report" used in the past by the Army.

Test implications

The test community, particularly the OSD DOT&E leadership, has actively campaigned for renewed attention to setting and achieving performance goals during the product development process well in advance of the operational evaluation. It is certainly preferable for the user community to define the required availability and reliability that would constitute a useful (suitable) system than it is to leave that determination entirely subjective on the part of the evaluator. A concern frequently voiced is that probabilistic measures such as availability and reliability are difficult to demonstrate during the operational test timeframe, and cost is almost impossible to verify directly. If we continued to view operational testing as a one time, pass/fail event as in the past, these concerns would be significant. Given the renewed leadership emphasis on using an integrated approach to testing intended to build confidence throughout the entire process leading to the operational test event, it certainly should be feasible to incrementally build up confidence in all of the sustainment elements.

Future steps

Within OSD, the new metrics have been endorsed by the Deputy Under Secretary of Defense for Logistics and Materiel Readiness for use even where they are not mandated as KPP or KSA values. Defense Acquisition Executive Summary reporting now includes sustainment, based on the new metrics discussed. Gradually, programs are developing benchmarks and assessing their status and contributing to our collective experience. Policy and guidance documents will address the metrics in their next revisions, and the status of these metrics is being added to major reviews. While the establishment of the mandatory Sustainment KPP is the cornerstone of our efforts, there is certainly additional work to be done to improve our ability to build in reliability and sustainment up front during the development stages of programs. Results from initial pilot programs exploring the relationship between funding and materiel readiness must be reviewed and action taken. There are improvements needed in collecting and analyzing system performance data across the enterprise, as well as in improving our O&S cost data collection and analysis. Shortfalls in our workforce skills are being identified and solutions are being developed. Finally, realistic approaches to testing the sustainment metrics will need to be established.

There is nothing more basic in the development of a weapon system than ensuring that when it is employed in combat, it will work when someone's life hangs in the balance. The progress made in restoring this discipline to the development process is significant, and will pay off in enhanced readiness and reduced cost over the life cycle of our weapon systems.

❑

RANDY T. FOWLER became the Assistant Deputy Under Secretary of Defense (Materiel Readiness) in May 2008, returning to the Pentagon from the Defense Acquisition University (DAU), where he was the director of the Learning Capabilities Integration Center (LCIC). He is a graduate of the Industrial College of the Armed Forces and holds a bachelor of science degree in economics from the University of North Texas, a master of business administration degree with a concentration in operations research from Jacksonville State University, and a master of science degree in national resource strategy from the National Defense University. E-mail: randy.fowler@osd.mil

SEAN P. KILLOY joined Decisive Analytics Corporation, Arlington, Virginia, as a senior consultant in 2005, providing support to the Office of the Deputy Under Secretary of Defense (Logistics and Materiel Readiness). He spent the majority of his career in industry, both as a consultant and with one of the major defense firms. He holds a bachelor of science degree in civil and environmental engineering from Cornell University, Ithaca, New York, and a master of business administration degree with a concentration in entrepreneurship from the Robert H. Smith School, University of Maryland, College Park, Maryland. E-mail: sean.killoy.ctr@osd.mil

The ITEA Journal of Test and Evaluation jite-29-03-10.3d 18/8/08 11:32:54 266

ITEA Journal 2008; 29: 267–274
Copyright © 2008 by the International Test and Evaluation Association

Improving Program Success by Applying Systems Engineering and Reliability Growth Analysis

N. Jenny Chang
U.S. Army Tank Automotive-Armament Command,
Tank Automotive Research and Development Center, Warren, Michigan
LTC David J. Rohall
Product Manager,
Stryker Brigade Combat Team-Development,
Program Executive Office Ground Combat Systems, Warren, Michigan

The Stryker Mobile Gun System (MGS) is a major, complex weapon system that presented a challenge in meeting its reliability requirement due to new technology revolving around the system's automatic ammunition handling system (AHS). However, as a result of a successful reliability growth management program, the Stryker MGS program experienced an unprecedented growth rate during developmental testing that led the program to meet its requirement. The program employed an effective systems engineering process to identify and implement effective corrective actions and adopted the Reliability Growth Analysis methodology to accurately track the resulting reliability growth. These tools provided the product manager with the information necessary to allocate resources and maintain support for the program throughout its development. Other similar complex systems may benefit by applying these processes and tools.

Key words: Reliability; Reliability Growth Analysis; reliability growth test; systems engineering.

MIL-HDBK-189 states that *"the Government's materiel acquisition process for new military systems requiring development is invariably complex and difficult for many reasons. Generally, these systems require new technologies and represent a challenge to the state of the art. Moreover, the requirements for reliability, maintainability and other performance parameters are usually highly demanding. Reliability growth management procedures have been developed for addressing the above problem."*

Stryker Mobile Gun System (MGS) was one of those complex and difficult development systems that presented a challenge in meeting the reliability requirement due to the new technology revolving around the system's automatic ammunition handling system (AHS). The AHS represents a significant portion of the vehicle's unique mission equipment package. Although reliability improvements were made to other subsystems, the AHS redesign contributed the most to the system's reliability growth. As a result of a successful reliability growth management program, the Stryker MGS program experienced successful reliability growth with an unprecedented growth rate during test.

This article reports on lessons learned from initiating a successful reliability growth management program that was based upon an effective systems engineering process to identify and implement effective corrective actions. It discusses the adoption of the Reliability Growth Analysis (RGA) methodology on the MGS development that provided the program management office with a tool to track the reliability growth accurately, which led to a successful reliability growth test. Additionally, the development of an idealized reliability growth curve provided astandard to measure progress. Finally, the tests were conducted in accordance with the system's operational profile and the assessment of data groups, based upon a balanced ratio of operational parameters provided, helped ensure the conclusions were relevant to intended operational use. Early and accurate assessment of the system's reliability was essential to maintain support for the program as it progressed through development, production, and fielding to soldiers.

RGA versus engineering analysis

Based on the MIL-HDBK-189, Reliability growth is the improvement in a reliability parameter over a period of time due to changes in product design or the manufacturing process. It occurs by surfacing failure modes and implementing effective corrective actions. In reliability growth management procedures, MIL-HDBK-189 introduces two methodologies, RGA and engineering analysis, which can be used to estimate the demonstrated reliability of the system if the configuration of the system is changing as a result of corrective actions to problem failure modes during testing. It also states that RGA is a preferred method since it provides an objective mathematical assessment of the reliability of the system being tested, that is, unless the RGA procedures cannot be applied to the test data because of data anomalies. It should be noted that if there is no change of configuration of the system during testing, then reliability growth procedures would not be necessary and the demonstrated reliability value would be cumulative reliability which is determined simply by dividing the total test time/miles/rounds, etc., by the number of charged failures.

The conventional way of assessing a demonstrated reliability using engineering analysis through an assessment conference has been used throughout combat vehicle history. For combat systems at Tank Automotive-armament Command (TACOM) and Program Executive Office Ground Combat Systems (PEO GCS), the engineering analysis technique has been the standard approach for estimating demonstrated reliability. However, the engineering analysis has several weaknesses. It is subjective and will therefore tend to be less definitive than data analysis based on reliability growth procedures. Engineering analysis involves using engineering judgment to assess the effectiveness of fixes that have been incorporated during the test program to determine the demonstrated reliability value. This technique uses the cumulative reliability adjusted based on a Fix Effectiveness Factor (FEF) applied for all fixes implemented. The FEF essentially provides the system "credit" for fixes applied and ranges from 0.0 (not effective) to 1.0 (failure mode eliminated).

To assess the demonstrated reliability using this conventional methodology, one typically waits until the test is completed to gain enough validated mileage after the fixes. The estimation of FEF is usually based solely on the concrete evidence from test data that the failure rate has been reduced in the operational environment and that it does not create any new failure modes. This methodology was not suitable for the Stryker MGS Production Verification Test (PVT), which lasted almost two years as the system went through many configuration changes due to corrective actions being implemented throughout the test period. Using this approach, the effectiveness of the fixes could not be tracked during the test and therefore the reliability growth could not be reported to the stakeholders.

On the other hand, the RGA technique lets the data speak for itself. In the presence of reliability growth, the data from earlier configurations may not be representative of the current configuration of the system. On the other hand, the most recent test data, which would best represent the current system configuration, may be limited so that an estimate based upon the recent data would not, by itself, be sufficient for a valid determination of reliability. Because of this situation, RGA offers a viable method for combining test data from several configurations to obtain a demonstrated reliability estimate for the current system configuration. Therefore, RGA allows for the effects of even recently introduced fixes into the system as its calculation incorporates the trend of growth established over the history, to date, of the development program.

Stryker Mobile Gun System is one of the programs that used RGA technique effectively to assess a demonstrated reliability during the system reliability availability and maintainability (RAM) testing during 2006–2008. The adoption of both the RGA technique along with an effective system engineering process led the MGS program through a very successful reliability growth test. This was accomplished after low reliability was demonstrated during its Production Qualification Test (PQT). Adopting the RGA methodology did require acceptance from the evaluation and user stakeholders.

What is the Stryker Mobile Gun System (MGS)?

MGS *(Figure 1)* is one of 10 variants of the Stryker Family of Vehicles (FOV). MGS was one of the two developmental variants while the other eight variants were ready for production based on technology readiness, integration readiness, and manufacturing readiness. The Stryker FOV shares a common chassis

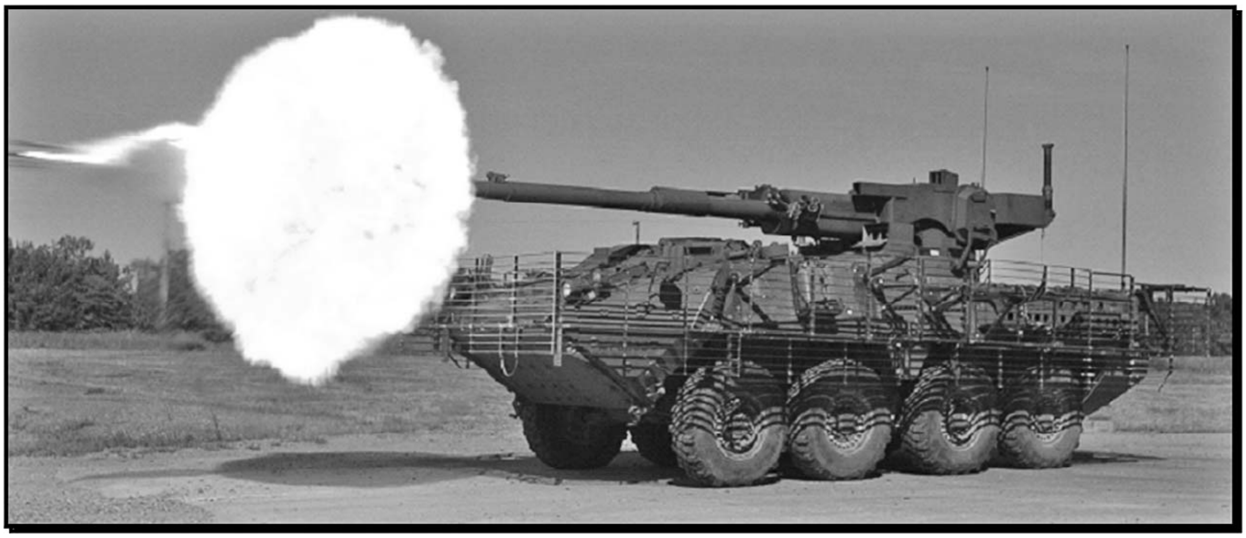


Figure 1. Stryker Mobile Gun System (MGS)—bunker buster

and many common components from the base vehicle—the Infantry Carrier Vehicle. Each variant is equipped with its unique mission equipment package. The MGS is equipped with a turreted, fully stabilized 105-mm main gun; a 7.62-mm coaxial mounted machinegun; a .50 caliber machinegun; and day and night optics. The 105-mm main gun ammunition is moved around the system and loaded in the breech by an automatic AHS. The AHS replaces some of the functions normally conducted by a loader in other weapon systems, e.g., Abrams Main Battle Tank.

The MGS went into PQT where the system demonstrated a small fraction of its system reliability requirement. The demonstrated reliability was too low and it was concluded that the MGS would require a redesign effort. The PQT was terminated about two thirds of the way through test. After reviewing and studying all the failure modes identified during PQT, it was concluded that the weakest link was the AHS replenisher. A redesign effort was launched for this and other AHS issues. Also, additional RAM test was conducted to prove the fixes that came in late in PQT, and the RAM community used the data from PQT and additional RAM testing to estimate where the reliability of the MGS would be at the start of the next test phase—PVT.

An accelerated reliability growth test was conducted upon the completion of the redesign of the replenisher and other AHS items. The test verified that the redesigned replenisher was robust. Before PVT was initiated, a short contractor's shake down test was conducted on all the redesigned AHS components including the AHS replenisher. The results of the test indicated that the system still had a challenge to meet the established reliability requirement of the mission equipment package. The Army community accepted the MGS PVT as a reliability growth test in place of a traditional verification test.

Reliability growth test

Ideally, the pure design process would be perfect with no testing required to improve reliability to meet the requirement. However, analytical tools, models, and engineering judgment are not perfect, so testing is always needed to fill in the gaps in knowledge and understanding. These tests have been specifically planned to stress the system components to predetermined realistic levels at which inadequate design features will surface as system failures. These failures are analyzed, design modifications incorporated, and then the modified system is tested to verify the validity of the design changes.

Most systems in the Army still rely heavily on the test-in rather than the design-in approach for reliability growth although a design-in approach is far more cost effective. This is due to ineffective design-in reliability practices. Making design-in reliability tools more effective will remain a challenge in the reliability domain. As a result of an ineffective design-in reliability practice, most systems have the initial reliability at the beginning of development test too low which leads to a lengthy test and often failure to meet the requirement at the end of test. The MGS contractor's brief shake down test after the redesign also showed the reliability growth of the MGS mission equipment package would have to rely heavily on the test-in approach during Government PVT.

The ITEA Journal of Test and Evaluation jite-29-03-02.3d 21/8/08 12:23:55 269

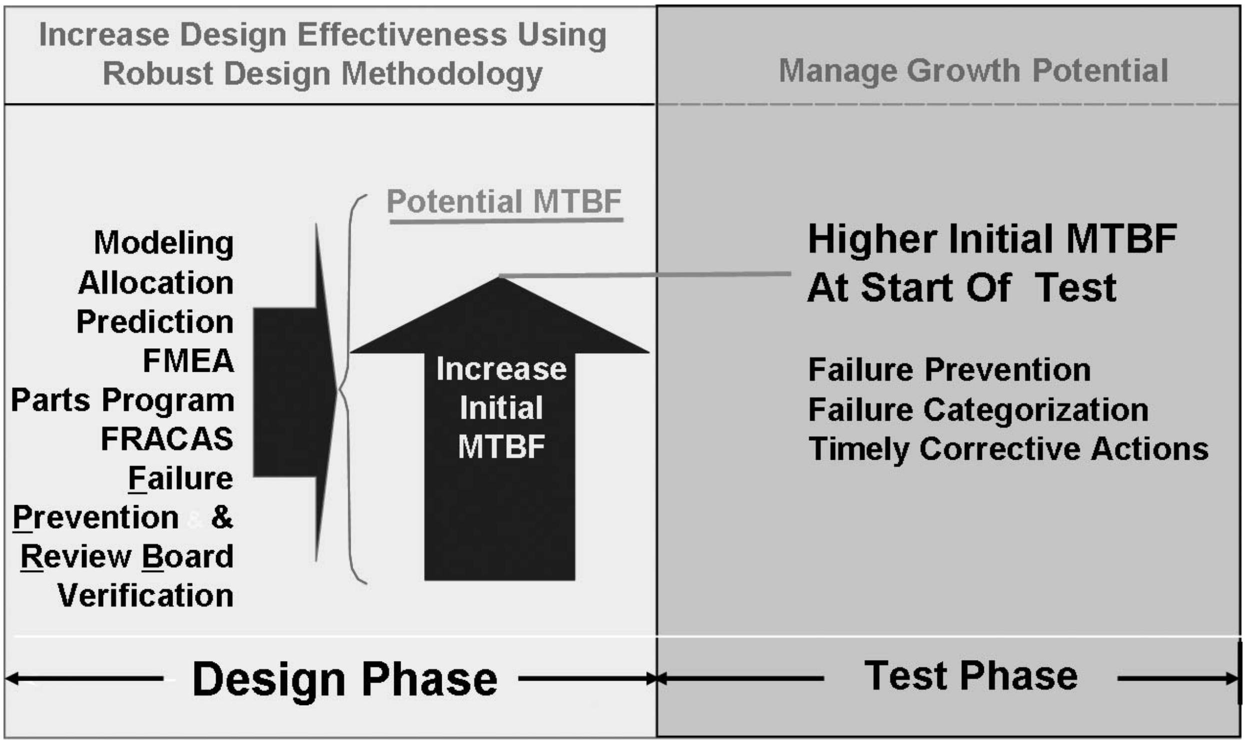


Figure 2. System engineering approach to reliability

This testing philosophy utilizes the Test-Analyze-Fix-Test (TAFT) procedure as the basic catalyst in achieving system reliability growth. The ultimate goal of a reliability growth program is to increase system reliability to the stated requirement levels by eliminating a sufficient number of inherent system failure modes.

Systems engineering

The growth rate experienced is a function of the design team's ability to identify and implement effective corrective actions and how quickly they are implemented. To achieve sufficient growth rate, a sufficient number of inherent system failure modes have to be eliminated. The U.S. Army Materiel Systems Analysis Activity (AMSAA) reports, on average, design changes are 70 percent effective in correcting a problem. The focus of MGS reliability growth management was to identify and close out failure modes from failure mode effects analysis and tests. The materiel developer's system engineering approach used during design phase and test phase for MGS reliability growth is depicted in *Figure 2*.

During the redesign phase after PQT, the system engineering process included performing Failure Mode Effects Analysis to identify, correct and close out issues found during design reviews and analyses as a preemptive action to potentially eliminate or greatly reduce the existing failure rate. This engineering process influenced the design to consider reliability so that the initial reliability of the system is high. However, initial prototype models of complex weapon systems will invariably have inherent reliability and performance deficiencies that generally could not have been foreseen and eliminated in early design states.

During the test phase, as performance deficiencies are observed and failures are uncovered, design engineers should properly analyze failures. Timely implementation of the corrective actions that can be taken to prevent recurrence or minimize the effects of failure are critical to any reliability growth program. The materiel developer implemented a very robust system engineering approach through a very effective and aggressive failure analysis and corrective action system with daily oversight activities by a Failure Prevention Review Board. The process included a closed-loop reporting system ensuring all test incidents were addressed. This systems engineering process during test phase was proven to be very effective with significant MGS mission equipment package reliability growth during PVT.

Constructing idealized growth for MGS mission equipment package

For a system under development, reliability generally increases rapidly early on and then at a much slower rate towards the end of development. It is useful at the

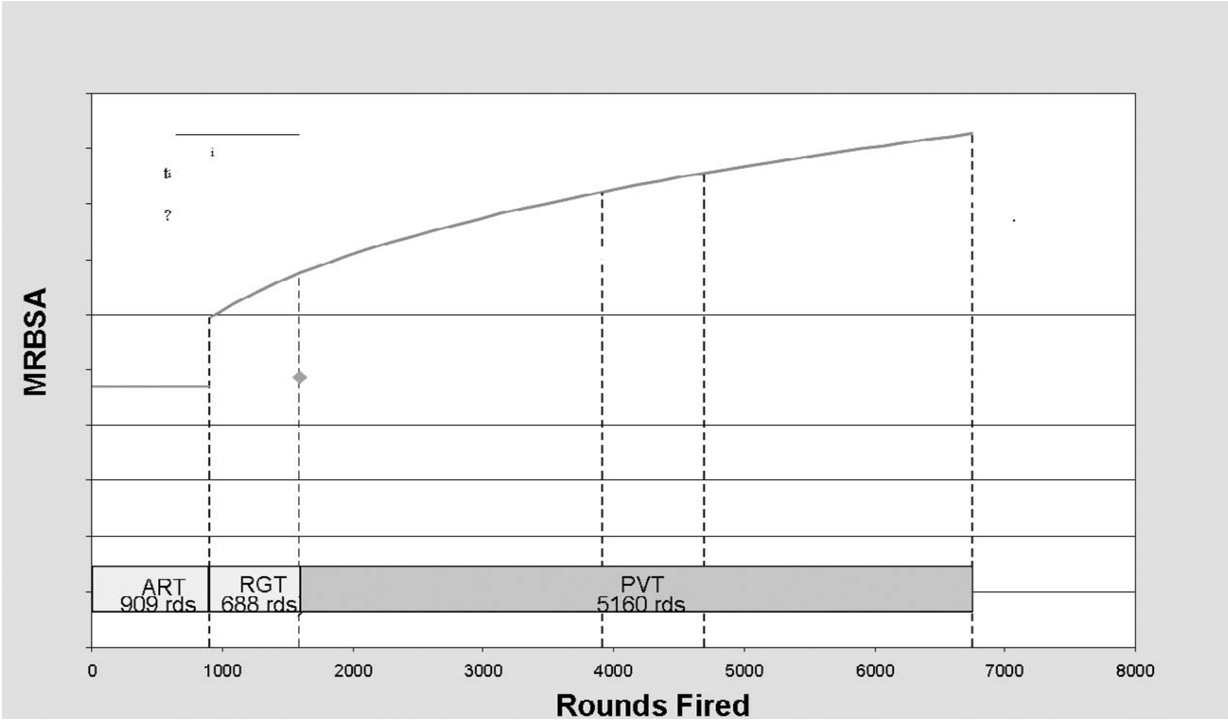


Figure 3. Idealized growth curve for Stryker Mobile Gun System (MGS) mission equipment package

beginning of a development program to depict the growth in reliability as a smooth curve which rises at slower and slower rates as time progresses. This curve does not necessarily convey precisely how the reliability will actually grow during development. Its purpose is to present a preliminary view as to how a program should be progressing in order for the final reliability requirements to be realized.

The RAM Integrated Production Team (IPT) led by the Army Evaluation Center (AEC) developed an idealized growth curve for the MGS mission equipment package using the previous test results as the initial reliability and the user's requirement as the target at the end of PVT. A conservative growth rate of 0.22 was used for planning purpose. The growth rate assumption was based on the historical combat system experiences. The idealized curve also showed that the system could reasonably be expected to meet its requirement. The idealized reliability growth curve developed for the MGS mission equipment package is depicted in *Figure 3*.

Growth tracking during MGS PVT

Reliability growth tracking is a process that allows management personnel the opportunity to gauge the progress of the reliability effort for a system by obtaining a demonstrated numerical measure of the system's reliability during a development program based on test data. Objectives for the reliability tracking include determining if system reliability is increasing with time and to what degree, and estimating the demonstrated reliability—an estimate based on test data for the system configuration under test at the end of each test phase.

The Stryker MGS PVT was conducted in cycles of 1,000 miles and 86 main gun rounds fired, approximating the operational mode summary/mission profile (OMS/MP). During PVT, three MGS vehicles were subject to run 20,000 miles and fired 1,720 main gun rounds each, for a total 60,000 miles and 5,160 rounds over the two-year period.

Test Incident Reports were prepared by the test centers (Aberdeen Proving Ground, MD and Yuma Proving Ground, AZ) and scored by the RAM scoring members (consisting of the evaluator, materiel developer and user representative) establishing the official Army database for estimating reliability. The data is further subdivided into chassis and mission equipment package failures. Since the chassis reliability was already proven as a common Stryker FOV subsystem, only the MGS mission equipment package reliability was tracked for growth. At the end of test, the MGS chassis reliability did prove to be reliable—just like the other variants in the Stryker FOV.

To track the reliability growth during PVT, the RAM IPT led by AEC developed a data grouping methodology. Since PVT was Test-Analyze-Fix-Test and the fixes were being implemented as they were

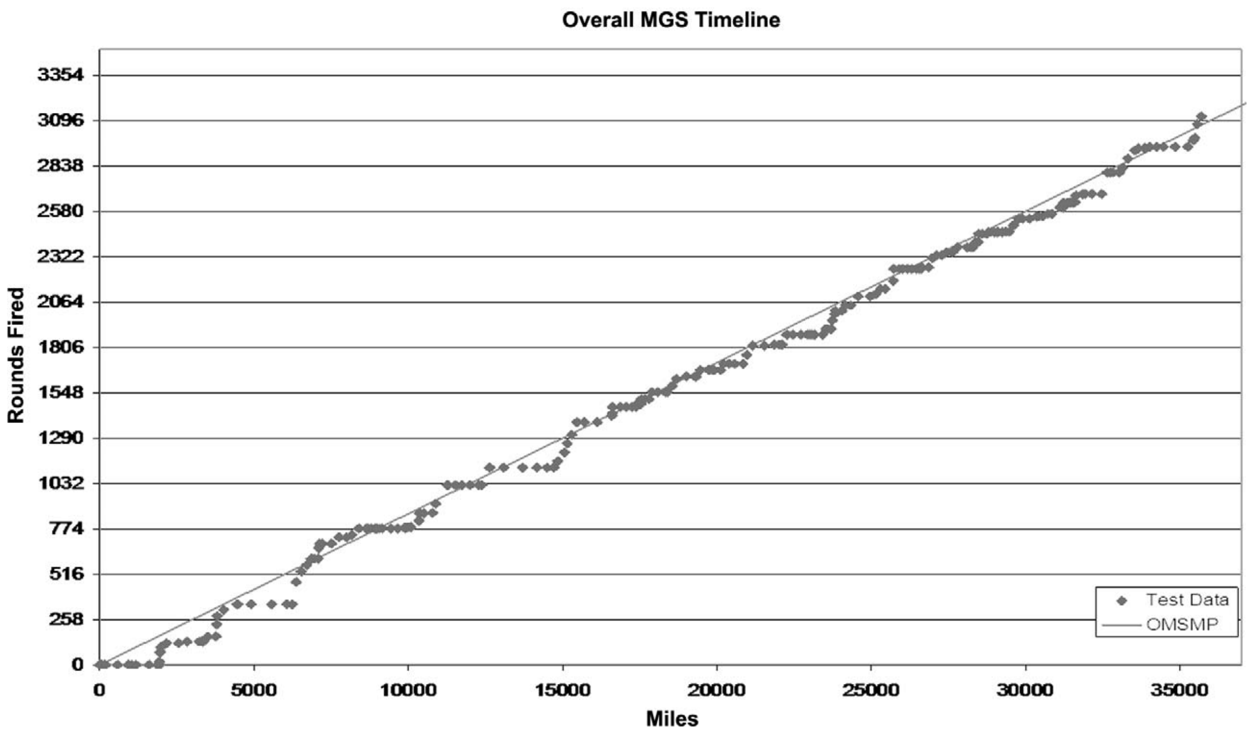


Figure 4. Cumulative rounds versus cumulative miles for all three vehicles in testing

available (in a rather random fashion) the RAM community had to establish a method to divide the PVT into distinct phases to track the growth of reliability. The approach for grouping the data first established a single timeline of events. This was done using daily updates from the test centers that documented the mileage accumulation and rounds fired for each of the test vehicles. The cumulative mileage and rounds fired were then summarized by date. The reliability failures Test Incident Reports were then aligned by date with these values to establish the timeline. *Figure 4* shows an example of the single timeline for the cumulative vehicle data as a function of the OMS/MP.

In order to plot the estimated expected mean time between failure (MTBF) versus the observed average MTBF, the Army Materiel System Analysis Activity (AMSAA) Reliability Growth Tracking Model for Continuous data (RGTM-C), which was selected to be used for MGS PVT, requires that each group contain at least one failure. Therefore, group selection was adjusted to accommodate this requirement. It became apparent as the MGS RAM test progressed that selection of the groups was becoming very subjective. Additionally, vehicle modifications, corrective actions, downtime for vehicle maintenance, and test conduct were impacting the analysis. The test firings were not evenly distributed across a test cycle but were conducted in groups when time permitted. Other contributing factors were firing range availability and weather conditions (e.g., main gun rounds could not be fired on test ranges when winds were high).

An alternative method was sought for establishing the groups used in the analysis. The approach for selecting the groups used three criteria:

- Have a ratio of rounds-to-miles closely approximating the OMS/MP (86:1,000) (minimize the delta between the number of rounds fired versus the expected value from the OMS/MP),
- For the vehicles within the groups, maximize the number of individual vehicles that are close to a multiple of 86 rounds,
- All test vehicles are represented in the group, i.e., at least one failure.

This method of grouping the data works for finding a mean mission equipment package reliability estimate given the large difference in mission equipment reliability of the individual vehicles. By requiring each vehicle to be represented in the group, impacts from extended downtime, configuration differences, and main gun firing were mitigated to the maximum extent possible. While this method is still subjective, the technique minimizes the variance in model output based on group selection.

Once the data was grouped, AMSAA RTGM-C was run. The chi-square goodness-of-fit statistic must be equal to or less than the critical value at the chosen

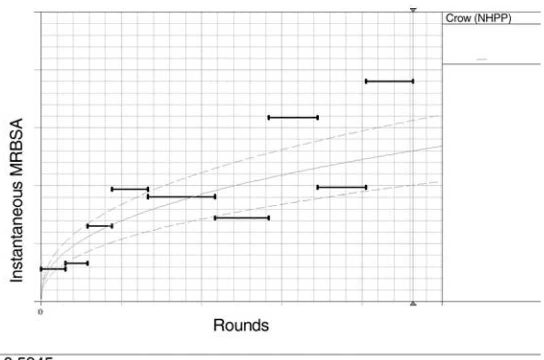


Figure 5. Estimated expected mean time between failure (MTBF) versus observed average MTBF

0.10 level of significance to accept the model. If this condition is met, then the model output is considered a viable estimate of the reliability of the MGS mission equipment package. The model provides estimates of the growth parameter β, the growth rate α (1-β), the scale parameter λ, and the MTBF of the last group. *Figure 5* shows a plot of the estimated expected MTBF versus the observed average MTBF.

Figure 6 shows a plot of the estimated expected and observed average MTBF superimposed on the idealized growth curve. The superimposed plot shows that with approximately two thirds of the PVT completed, the observed average MTBF was close to 0.4 growth rate and was exceeding the idealized growth curve. It also shows that initial reliability at the beginning of the test was much lower than expected which forced the growth rate to be much higher than planned to achieve the target at the end of the test.

The MGS mission equipment package system experienced a significant growth rate, 0.4 during PVT, with the system demonstrating above the target reliability at the end of the test. AMSAA reports the historical growth parameter to be in the range of 0.23–0.53 for time/mileage (continuous) systems. Typically for complex combat systems such as the Abrams Tank and Bradley Fighting Vehicle the growth rates were assessed to be approximately in the range of 0.2–0.25.

Conclusion

Understanding the status of a program at any given point is one of the challenges facing program

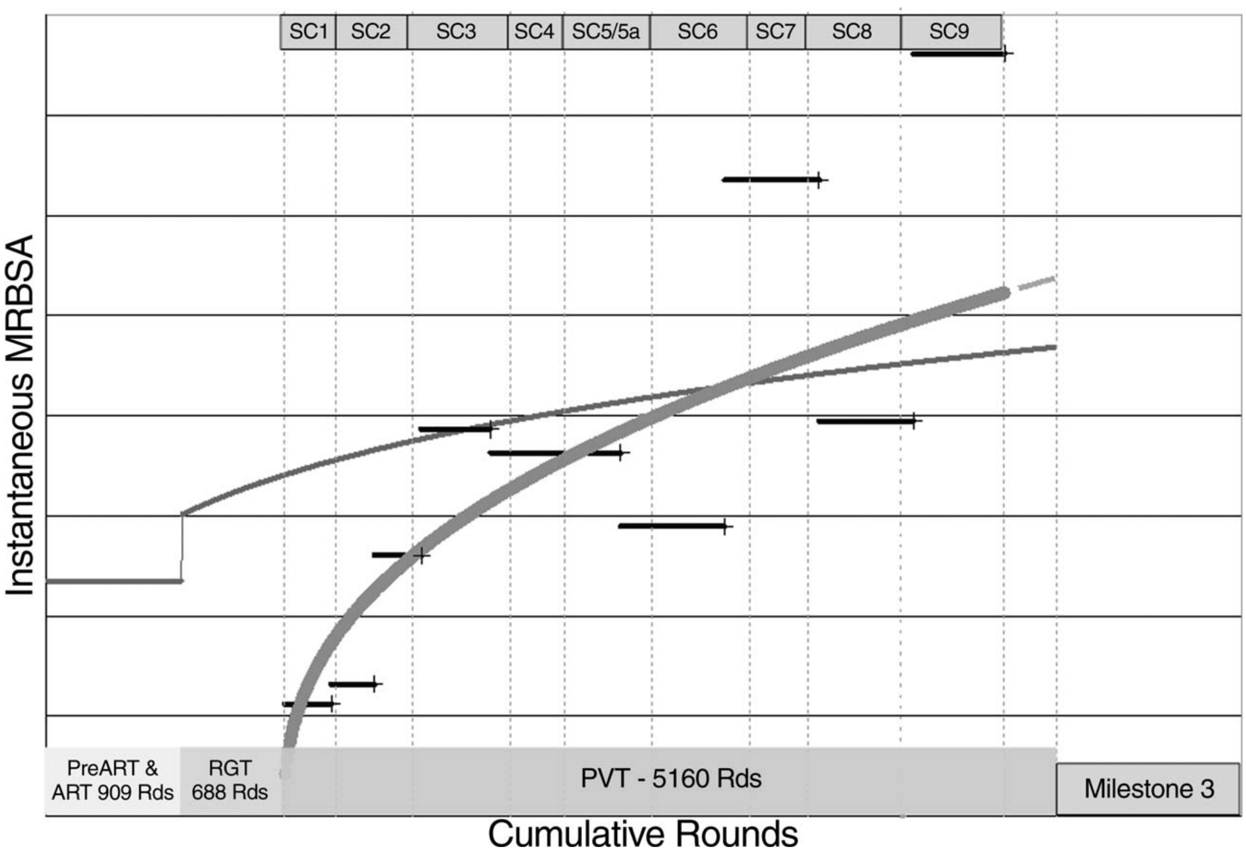


Figure 6. Mobile Gun System (MGS) mission equipment package observed average mean time between failure (MTBF) superimposed on the idealized growth curve

managers. Increasing system complexity coupled with increasing demand for more reliable systems causes members of the development team to rely on efficient and effective tools to report program health. RGA provides one such tool to understand reliability growth throughout the development without having to wait until the conclusion of development. When program management recognized the challenge lying with the reliability of the MGS mission equipment package after PQT, the Stryker Reliability Integrated Product Team was challenged to develop an effective reliability growth management program to meet the requirement. The Systems Engineering Team assembled reliability tools into disciplined processes and working organizations. When reliability assessment was reached through in-depth analysis coupled with a best fitting methodology, the result was the MGS mission equipment package experienced an unprecedented growth rate during PVT.

The successful mission equipment package system reliability growth program of MGS PVT can be attributed to the following factors:

1. The test program was planned to expose the system to test and stress levels adequate to uncover inherent failure modes.

2. The program office took into consideration the requirements of the test schedule and resources required to support the Test-Analyze-Fix-Test procedure.

3. The materiel developer conducted an effective system engineering process to identify and implement effective corrective actions.

4. The Stryker Reliability Integrated Product Team applied reliability growth analysis techniques and developed a methodology to track and assess the reliability growth at every test phase.

There is no simple way to ensure program success. However, program managers that encourage and demand a disciplined systems engineering process, enabled by fact-based recommendations, and implement tools to assess how the system reacts to changes will certainly increase the development success rates of challenging, complex systems. ❑

N. Jenny Chang is a senior engineer at the U.S. Army TACOM TARDEC, Warren, Michigan, and currently works as a lead reliability engineer for the Stryker system. Her previous assignments included working as a reliability engineer for Armored Gun System (AGS), Bradley Fighting Vehicle System, Light Armored Vehicle (LAV), and other systems. She is a 1983 graduate of the Georgia Institute of Technology with a bachelor's degree in industrial engineering. She also earned a certificate in quality and reliability engineering from the U.S. Army Red River Intern Program. E-mail: nanhi.j.chang@us.army.mil

LTC David J. Rohall is a U.S. Army acquisition officer. At the time this article was written, LTC Rohall was the product manager for Stryker BCT-Development, Warren, Michigan. He was responsible for the major development programs for the Stryker Family of Vehicles including the MGS, NBC Reconnaissance Vehicle, and the Stryker Family of Vehicles modernization program. He earned a bachelor of science in electrical engineering and a master of business administration, both from the Pennsylvania State University, University Park, Pennsylvania. E-mail: David.Rohall@us.army.mil

References

AMSAA Technical Report No. TR-652 September 2000. "AMSAA Reliability Growth Guide."

Ellner, P., October 18, 1993. "Reliability Growth Methodology: Fundamentals and New Directions." AMSAA Document.

General Dynamics Land Systems, Rev C, May 31, 2005. "Stryker MGS Reliability Growth Plan."

MIL-HDBK-189, February 13, 1981. "Reliability Growth Management."

Rose, A. J. and Mioduski, R. E., June 7, 2007. "Stryker Mobile Gun System (MGS) Reliability Growth Tracking Methodology for Mission Equipment Package (MEP)." Army Evaluation Center (AEC) Document.

ITEA Journal 2008; 29: 275–277

The 800-Pound (364 kg) Gorilla

W. D. Bell, D.Eng.
The MITRE Corporation, McLean, Virginia

Initial Operational Test and Evaluation (IOT&E)—the tests whose success is a necessary condition for entry into full rate production of a new system—has had increasingly poorer results over the last few years. These poor results are frequently the result of failures with respect to suitability issues in the areas of reliability and maintainability. This article addresses some of the reasons for that decline and suggests one very important change that could help turn this situation around.

This article is a reprint, with minor editing and without graphics, from an article originally published in Defense AT&L magazine.

Key words: Operational tests; program office; requirements; specifications; testers.

The purpose of this unusual article is to stimulate a healthy discussion, with no intent to be offensive to anyone. There is an 800-pound gorilla in the test and evaluation (T&E) and systems acquisition room! This gorilla is rarely acknowledged, sometimes fed and patted, but most often ignored. He has been on the prowl for decades and it is well past time to put him back into the zoo and permanent retirement. What is this so called gorilla? It will be revealed in a moment, but first a little background and perspective.

The situation and problem…

At the annual International Test and Evaluation Association (ITEA) Symposium in November 2007, numerous speakers alluded to the fact that about half of all systems undergoing Initial Operational Test and Evaluation (IOT&E)—the test whose success is a necessary condition for entry into full rate production of a new system—are at least partially not suitable or not effective or both. In prior years the IOT&E success rate certainly has varied but has hovered at more like the 20 percent level of either partially or totally not suitable or not effective. At the symposium, Dr. Castellano, Deputy Director, Assessments and Support, Office of the Deputy Under Secretary of Defense for Acquisition and Technology, U.S. Department of Defense (DoD) reported that over the past 10 years, DoD systems have experienced a 33 percent cost growth due to "RDT&E mistakes".

With all the modern day emphasis on systems engineering, why are we having the downturn in IOT&E success? Is it due to the testers, the acquisition program management offices (PMOs) or that darn gorilla? The gorilla, by the way, is the typically poor (traditionally strained) relationship between PMOs and testers.

Have you ever been a part of a systems acquisition program office? If so, you know it is high stress and fast tempo. Often a new capability is needed especially quickly—after all we are at war. Further exacerbating the situation, DoD acts as though it takes state-of-the-art technology to win a war. (Not always stated but usually assumed.) Add to this a program manager (PM), whose reward—like his/her next promotion—depends on some great event(s) occurring in the next 2 or 3 years. Then into the middle of this stress-soup comes a T&E professional (fondly called a tester) who tells the PM that his/her baby is ugly. Furthermore—the PM likely perceives—the tester is wasting precious time and money while performing tests to show that the baby is, or may be ugly!

Know any testers? They have for too long been the Rodney Dangerfields—the comedian known for his "I don't get no respect" phrase—of the engineering community! Why? There are at least two reasons. First, in the 1960s, 1970s, and 1980s, the test group was where engineers were sent to await their retirement. Now these folks were not usually the majority of the group, but the perception that some testers were biding their time tainted the image of testers. The second reason is bullying by program offices. Testers need system requirements and specifications from which to build their test plans. The PMs, who were and are the keepers of these documents and money, did not want the testers involved until the very last

The ITEA Journal of Test and Evaluation jite-29-03-01.3d 18/8/08 11:28:00 275

moment—if ever—thus diminishing the modicum of respect the testers might have otherwise received! Knowing that the PM did not think they were worth their cost, testers were—by golly—going to show that darn PM that they could find lots of problems that needed fixing. These are problems they would have found anyway but now—as a get-even ploy—the tester becomes an unhelpful pain until the problem is fixed.

So given this history—which is actually even more contentious, but you are being spared the gore—it is no surprise that the relationship between PMOs and testers is strained! Yet this strained relationship does not explain the recently degenerating IOT&E situation since these communities have been dueling for decades. There are many contributors to the recent decline in IOT&E results for which the PMOs and testers are uniquely qualified to turn this tide—but this can only happen if there is an end to the distrust and the beginning of trust and mutual respect.

In a nutshell, system complexity is the ultimate culprit—and **not** the gorilla—behind the decline in IOT&E results. Folks tout budget constraints, schedule pressures, incompetent engineers, slippery contractors, and so forth, but the root cause is the complexity of new and proposed systems; not just the systems themselves but the environment in which they must perform. This is especially true as we try to benefit by applying communication, computing, and internet technologies to our new systems. Net centricity is the answer to providing unparalleled capability to war-fighters just as it has provided this capability to the common man and businesses. It does this by exacting a terrible toll in added complexity. The complexity of new systems is beginning to challenge the cognitive capabilities of many—and maybe most—of our military operators. Such systems may be able to pass specification verification but when operators try to use them effectively in harsh environments, both physical and that induced by the "fog of war" or the "fog of competition," these systems are not effective or suitable. So while complexity of systems has increased, the relationship between PMOs and testers has not changed.

Here is why this matters. It is hard to imagine the specifics of the requirements in a complex system. It is difficult to get the requirements stated in a succinct and understandable way. It is impossible to develop appropriate specifications from poorly written requirements. The more complex a system, the greater are the opportunities for human error. From the definition of the need, to the requirements decomposition, to the building, coding, and integration, the difficulty skyrockets. This increased complexity causes more requirements, more applications, more environments, more failure modes, and increased sustainability challenges. In this environment there is no time for the Hatfield-McCoy behavior—the famous family feud from 1878–1891. PMs and testers have to pull together! No, this is not crazy!

Before the late 1990s when government-staffed PMOs had significant roles in the design and development of systems, PMOs could tell the end users that their requirements were unreasonable. Now contractors, who are the system developers or lead system integrators, would never do that, because there is another contractor right around the corner who will say that he can do it and thus win the contract. Therefore, almost every user requirement gets placed on a developing system, without a good reality check. Engineers, scientists and managers are literally "guessing" on feasibility, methods and resources until well into the development effort when it is often too late and too expensive for significant changes. It is human nature to underestimate (as one discovers with many home projects) so developments almost always overrun resources. However, the "guessing" can also cause designs and developed systems to be cumbersome, inadequate or even wrong for the requirements. Such systems will fail to be effective or suitable in IOT&E.

However...

Here is what can and often does happen when testers and PMOs become one group from the very beginning of the acquisition process:

- Requirements that are beyond the state-of-the-art for field deployable systems are questioned and eliminated. PMs often hear from their potential prime contractors that X can be done. The contractor is afraid of losing business if they say otherwise. Naturally PMs are hoping for the positive answer, but testers are accustomed to challenging and questioning things. Working together the PM and tester can sort out truth from fiction.
- The requirements are stated in a way that will ultimately be verifiable. The tester will make sure of this, because he has to provide the test to support this verification.
- A verifiable requirement is also one from which a verifiable specification can be written.
- Verifiable requirements and specifications are readily understood, i.e., not susceptible to misinterpretation during the requirements flow-down process.
- Appropriate testing-related schedule, budget, infrastructure, other resources, and personnel are planned early and become a part of such documents as the Test and Evaluation Master Plan, the Request for Proposal, the Initial

Capabilities Document, and the Systems Engineering Master Plan.

- Contractors can be prevented from under- or overbidding the T&E part of their proposal since testers as a respected part of the proposal evaluation team can assure test realism in the proposal or at least in the negotiations.
- Testers can help make the system integration lab a useful preparatory time and place for systems testing.
- Testers can encourage and double-check that proper reliability growth testing is planned and executed—a big contributor to successful IOT&E.
- PMO testers—when they exist—make the best interface to the independent operational testers in terms of communication between the IOT&E executors and the program office, that is English-to-English translation throughout the development cycle!
- Testers can plan and execute developmental test and evaluation (DT&E) thoroughly enough to virtually ensure success in IOT&E—especially valuable with respect to suitability and interoperability issues. The PMO testers learn what the independent IOT&E testers are planning and "dry-run" those tests to see how the system performs. Thus the PM has an excellent idea that his/her system will pass IOT&E before turning over the system for IOT&E. For example, if a developing system is technically tested without an end user as the operator, it is equivalent to testing that system with a critical component or subsystem missing or a surrogate in its place.
- PMOs—after a small initial investment—save huge sums of money and significant schedule reductions compared to the status quo.

These items explain why Dr. Charles McQueary, Director, Operational Test and Evaluation, and Mr. John Young, Under Secretary of Defense for Acquisition, Technology & Logistics signed a joint memo on December 22, 2007 where they state, "T&E expertise must be brought to bear at the beginning of the system life cycle to provide earlier learning about the strengths and weaknesses of the system under development".

The solution…

Put the 800-pound gorilla of **bad PMO and tester relations** into the zoo at the start of a program! He has been a problem way too long! How do we get rid of him? The solution starts with mutual respect. For example, testers must acknowledge the pressure and constraints under which PMOs work. Testers must be timely, helpful, and truthful. They must be the bearer of good news whenever it is appropriate! (Some testers mistakenly believe that if they cannot find a wart on the baby they were not successful.) Testers should be willing to suggest workable solutions to found problems.

On the other hand, PMs and their PMOs have to start taking the long-term or enterprise view. That is, it is **not** OK to delay the discovery of technical, schedule, or budget problems until a future PM has no choice but to acknowledge them. PMs need to be rewarded for solving problems, not for postponing them! Also PMs must recognize that testers are just as savvy and just as concerned about their program's success as they are and must treat them as partners in that success. Therefore, they must value their input and perspectives, i.e., do not shoot the messenger—after all, the messenger may have good or at least useful news! Testers provide knowledge, and a recent GAO report on weapon system acquisition cites lack of knowledge as a major problem with defense acquisition programs.

Some of you may recognize that the solution to the problem of unsuccessful IOT&E, in this age of complex systems, could be described as the implementation of an Integrated Product Team (IPT) as envisioned when IPTs were **first** invented and not as they exist today. This construct—again as originally created—would institutionalize (that is make it the norm) the respected and valued involvement of testers and many other "minimalized" engineering specialists (just a teaser for further discussion).

Finally…

We are at war, and the warfighters have immediate needs that include new and complex systems to accomplish their mission while staying safe. The aim of this light-hearted article (with a heavy-hearted message) is for folks to see a solution to one of the serious problems that hampers the successful creation of such systems. With an attitude change, from disdain to mutual respect, the talented combined teams of PMOs with testers can meet the challenge of defining and developing complex systems. ❑

Dr. Dave Bell is the senior technical advisor for the Joint Experimentation, Test, and Training Department of the MITRE Corporation. He has spent his career participating in all aspects of T&E and in many leadership aspects of program management as a U.S. Government employee—both Air Force and Navy—and with several large and small corporations. He holds a BS in physics from The Ohio State University, an MBA from National University and an Engineering Doctorate from Southern Methodist University. E-mail: wbell@MITRE.org

ITEA Journal 2008; 29: 278–284
Copyright © 2008 by the International Test and Evaluation Association

Leveraging SOA for Distributed Test and Evaluation: "To SOA or Not to SOA, That Is the Question"

Robert Heilman
Naval Air System Command (NAVAIR), Point Mugu, California
Gil Torres
Naval Air System Command (NAVAIR), Point Mugu, California
Ken Thomas
TSAE, Sierra Vista, Arizona
James Buscemi
GBL Systems Corporation, Camarillo, California

Can Service-Oriented Architecture (SOA) support distributed Test and Evaluation (T&E)? When will SOAs be suitable to support distributed testing data management requirements? What are the benefits of modernizing instrumentation to use an SOA for testing? Can SOAs improve reliability and composability of distributed T&E capabilities? These are just some of the questions that are being addressed in an ongoing Office of the Secretary of Defense distributed test infrastructure assessment that includes a study called "Applicability of Service-Oriented Architecture (SOA) to Distributed Testing Infrastructure." This article will give an overview of this quantitative/qualitative study and the Community of Interest being formed to support the study. In addition, the article will describe how the Netcentric Systems Test (NST) reference architecture developed under the T&E/Science & Technology NST focus area sponsored by the Test Resource Management Center is being used as a collaboration point to determine which T&E mission processes to consider in developing a use case for the study.

Key words: Community of interest; DoD architectural framework; Global Information Grid; netcentric systems test; netcentric web services; service-oriented architecture.

Testing of netcentric warfare systems requires bringing together a netcentric system under development with all of the interfacing systems in a scenario that represents the mission for the netcentric systems under test. Since the interfacing systems or their representations as hardware and software in the loop emulations are rarely located or available at a single location, a distributed network is required to link together all of the mission platforms at disparate locations together with the test management and evaluation tools. Further, newer netcentric systems are being developed using Community of Interest (COI)–defined warfare services in a service-oriented environment to take advantage of the agility and flexibility demonstrated in commercial Service-Oriented Architecture (SOA) environments. SOA introduces some new testing requirements and challenges that must be addressed.

It is into this environment that SOA might also be applied to facilitate development of common distributed T&E service applications for distributed test events.

Service-orientation describes an architecture that uses loosely coupled services to support the requirements of mission processes and users. Resources on a network in a SOA environment are made available as independent services that can be accessed without knowledge of their underlying platform implementation. These concepts can be applied to military missions, business processes, software, and other types of producer/consumer systems such as testing.

SOA applies to distributed applications and facilitates agility and flexibility by emphasizing *composability*—the ability to combine and recombine individual service applications in different configurations as long as service interfaces are satisfied. SOA uses coordina-

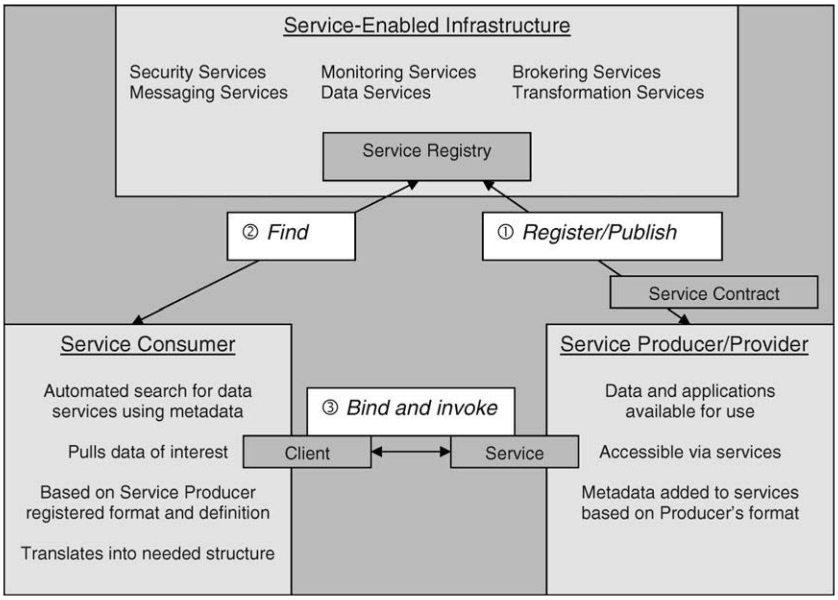


Figure 1. SOA service registration and binding

tion and orchestration services to combine fundamental services into mission activities, transactions, and processes.

The Department of Defense (DoD) Architectural Framework (DoDAF) v1.5 used to define the capability and structure of warfighting systems embraces the IEEE 1003.0 definition of service, *"a distinct part of the functionality that is provided by a system on one side of an interface to a system on the other side of an interface."* The DoDAF extends this definition to include those interfaces that allow execution of a business or mission process, or that exchange information among machines and humans using standard interfaces and specifications without regard for the underlying implementation. Note that while the netcentric guidance provided in DoDAF v1.5 focuses on Web-based services, much of the guidance is applicable to any form of electronic information processing or access service. Services (resources) may be registered by service providers within a registry of registers (itself a service) and made available to a COI with the right access privileges on a distributed network.

Armed with this interface information, clients can bind to service providers to utilize the resources. Across the SOA architecture, enterprise-wide services for registry, binding, access, instrumentation, messaging, security, and so forth can be specified by the architecture. These enterprise-wide services form the backbone upon which the services are built and accessed.

SOA is not a replacement for other software development architectures. Rather, its focus is on defining higher level mission or business process reusable and composable services that are platform and domain independent. Underlying code may be legacy applications or developed using usual methods as long as the SOA design principles and interfaces are met. *Figure 1* illustrates an early simplified SOA registry model demonstrating the potential interactions between a client (Service Consumer) accessing a particular service and the Service Provider offering that service.

In the diagram, step (1) shows the Service Provider registering or publishing the service/resource interface information and making it available for consumption with the Service Registry using a Registration Service. Already included in the registry are examples of Netcentric Core Enterprise Service offered by the Defense Information Systems Agency (DISA) Netcentric Enterprise Services (NCES). The Service Registry holds this information so that a Service Consumer may consult the Service Registry using defined interfaces (and protocols) to enumerate and obtain access to some service resource from the Service Provider. This is shown in the diagram as the (2) Discovery/Find Service. At this stage, it is possible the Service Consumer may not even know the specific Service Provider with which it will ultimately connect. After obtaining the necessary information describing the resource it is attempting to gain access to, the Service Consumer will then (3) bind and invoke the resource by contacting and negotiating access to the Service Interface offered by the Service Provider as specified by the Service Registry.

In practice, many commercial services and applications never used the registry. The registry provides

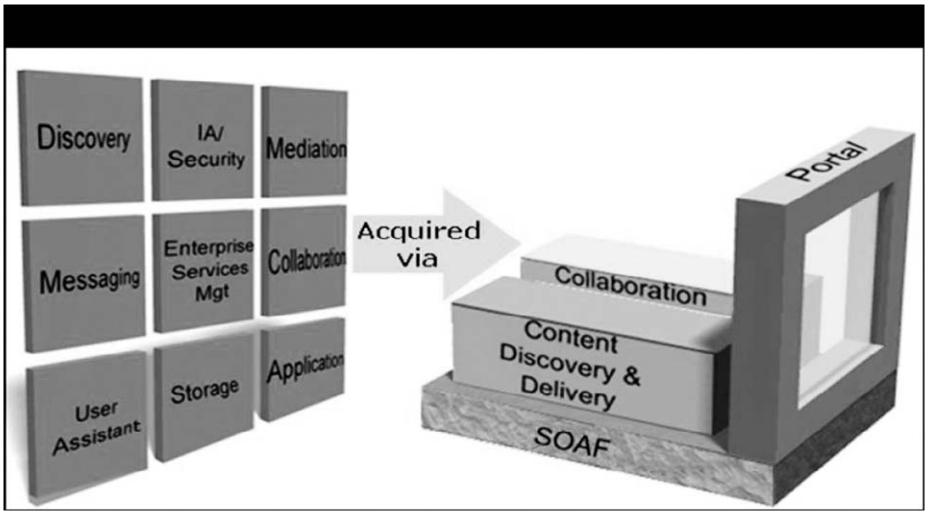


Figure 2. Netcentric Enterprise Services

additional agility and flexibility. A service requestor and/or service endpoint may be a human operator or a human-assisted service. A human reviewer or approval authority may also be an intermediate service. Coordination or Orchestration services may use the registers to compose complex activities using many services in serial and/or parallel processes.

A fundamental aspect of SOA operation is the ability of Service Providers to connect with Service Consumers in possibly unanticipated ways without coordination prior to the Service Consumer's binding and invocation. This is made possible by services designed to be stateless and composable with well-defined interfaces so that the Registration Service and Discovery Service can flexibly locate and bind consumer and provider services as needed to complete a mission activity. The discovery services may also be accessed through a human interface portal as well.

The registry model may imply a simple request response Message Exchange Pattern; however contemporary SOAs support multiple Message Exchange Patterns, defined in evolving standards such as "fire and forget" and "publish and subscribe" that are also supported by the DISA NCES.

Clearly, though not always acknowledged in the literature, there is a potential performance penalty with registry access and with data conversions with loose coupling. SOA may be limited in hard real-time environments and may not be appropriate for every application.

SOA is the principle distributed architectural pattern used by the Global Information Grid (GIG) to support netcentric warfare and facilitates the secure and controlled sharing of data and services among warfighter applications over distributed networks. The DoD is relying on NCES to provision the GIG with SOA capabilities called Core Enterprise Services. NCES is composed of nine services grouped into four product lines (*Figure 2*).

In response to a Joint Capabilities Board Preliminary Review of Assessment (July 2007), the Joint Training Functional Capabilities Board, in coordination with service advisory members of the T&E community, recommended a Joint Distributed Test Infrastructure Capabilities Based Assessment (CBA) (Joint Requirements Oversight Council Memorandum 279-07, 10 Dec 2007) that focused on three potential gaps:

- Service Transition to Internet Protocol version 6;
- Applicability of Service-Oriented Architectures (SOA) to Distributed Testing Infrastructure;
- Transition to Distributed Testing using the Global Information Grid.

Responsibility for this CBA was then transferred to the Network-Centric Functional Capabilities Board (NC FCB), and the assessment is to be directed by the Office of the Secretary of Defense Test Resource Management Center (TRMC). This article will focus on the second component study of this CBA, the Applicability of Service-Oriented Architectures in Distributed Testing, which was initiated in March 2008.

Study objective

The primary objective of the "Applicability of Service-Oriented Architectures (SOA) to Distributed Testing Infrastructure" study is to determine what testing activities of netcentric systems test (NST), particularly distributed tests, can be beneficially and economically developed as reusable and composable test services and which activities would not be beneficially developed as SOA services.

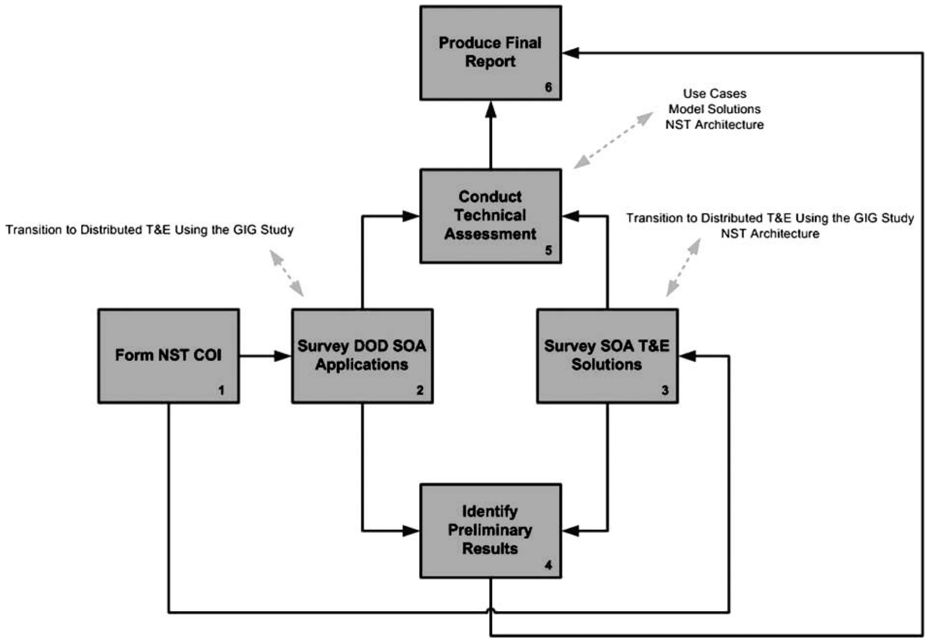


Figure 3. Study task relationships. Each box represents a large task described in this section, the small numbers in each box demonstrate rough sequencing of task execution, and the arrows demonstrate data relationships within internal (solid lines) tasks and external (dash lines) entities.

Study approach/methodology

In the context of the "testing and evaluation business process" for a distributed mission thread test, the study will identify potential SOA-based test tools in the areas of test control, synthetic battlespace environment, data analysis, and collection. It will survey commercial, joint services, and agency ongoing SOA activities; perform a technical assessment of these efforts; and then report on these findings. During the period of the study, preliminary and ongoing status will be briefed to relevant user groups (e.g., Joint Mission Environment Test Capability [JMETC] Users Group, Air Force SOA Symposium). *Figure 3* shows the work breakdown demonstrating task interrelationships and sequencing.

Form NST COI task

A COI representing distributed Netcentric System T&E will be formed and called the NST COI. This COI will collaborate across three portals: (a) Defense Architecture Repository System (DARS) https://dars1.army.mil, (b) Defense Knowledge Online (DKO) https://www.us.army.mil, and (c) TRMC www.trmc-test.org. The NST reference architecture developed as part of the NST Architecture and Technology Insertion Environment (NSTATIE) will be uploaded to both DARS and DKO for reference by the COI. The NSTATIE project will be described in more detail below. The COI will use the architecture to form research teams for evaluation of architecture operational activities (functions) for potential implementation as distributed T&E Services. They will also initially establish the potential benefits or problems of adopting or developing distributed services to implement each of the architecture functions.

The COI will leverage TRMC's new portal area called the Distributed Test Infrastructure Assessment collaborative environment to track status, act on items, and communicate information, including meeting information. Accounts may be established with all three portals to be formally part of the COI. However, one does not need to be a formal COI member to review documents, except on the DARS. For DARS, you must register for the netcentric system test area when requesting an account, located under the DoD portion of the directory. The netcentric system test COI already exists on the DARS. Once registered, you can request to join the COI in order to have access to the information published. For DKO, register at the site and then send an e-mail to Gil Torres (gilbert.a.torres@navy.mil) with your login ID to request access to the reference architecture. When registering for the TRMC portal, indicate that the project you support is the "Joint Distributed Test Infrastructure Capabilities Based Assessment project."

The COI is divided into the Core COI and extended COI membership. The Core COI and extended COI membership will be identified, and invitations for representatives from government and industry will be sent to form these teams. It is anticipated that periodic telecons will occur during the duration of the study, which is projected to end May 2009. These telecons, when required, may include the Joint Distributed Test Infrastructure CBA other two study tasks.

Throughout the study, the COI will brief key user communities on a regular basis, including each JMETC Users Group held during the duration of the study and the planned SOA/GIG Summit.

Survey current SOA applications task

The study will survey current ongoing SOA solutions sponsored by any of the DoD services or agencies that will need testing in a netcentric, distributed test environment. The survey questionnaire will be constructed to facilitate the collection of data from service and agency representatives of any SOA test tool efforts. The survey results will summarize any issues identified by DoD services and agencies regarding the use of SOA for T&E. Based on the survey and NST reference architecture, a use case will be defined for further investigation. This use case will identify potential SOA-based test services. The use case will also represent a joint mission thread test and identify potential areas where SOA-based tools might be applicable.

Survey current SOA T&E solutions task

The study will survey DoD test organizations to identify any service or agency T&E functions being developed as T&E services. Additionally, commercial vendors will be surveyed to identify any commercial T&E services that might be used.

These test services candidates will be initially identified using the NSTATIE architecture, and existing solutions will be compared to the candidate service requirements. The use case will be refined in this task, and qualitative measures will be determined to define testing in the NSTATIE Technology Insertion Environment laboratory.

Identify preliminary results task

This task is divided into three efforts: define initial evaluation criteria, conduct technical assessment, and produce draft report of preliminary results. Preliminary results for each area of research will be drafted into an agreed-upon format and presented to the COI. Publicly, these preliminary results may be presented to an interested external party such as the other GIG study. These preliminary results will be generated via ongoing technical research. The areas for research are initially identified with aid of the NST reference architecture and use case. The research will identify criteria for technical and performance evaluation of identified potential SOA services.

Conduct technical assessment task

After the preliminary survey results have been determined, a detailed technical assessment will be performed by each sub-working group defined in the study to identify and produce technical/performance measures for each area of research.

More specifically, this task is divided into five efforts: (a) model use case, (b) select candidate T&E SOA solutions, (c) define qualitative evaluation criteria and measures of effectiveness and measures of performance for prototype SOA solution experiments, (d) qualitatively evaluate the feasibility and effectiveness of the T&E services applied to the use case, and (e) conduct experiments of critical areas of the use case using available SOA services and prototyped applications. This activity is currently being finalized. The following are some of the metrics being considered:

- *Service time*: response time for synchronous services and delivery time for asynchronous services.
- *Scalability*: examples are user load and number of requests per second.
- *Availability*: includes planned maintenance and unplanned down time.
- *Reliability*: due to defects, rejected requests, message loss, etc (Lau, 2007).

To support this effort, the mission thread use case will be refined to identify candidate high-level T&E mission services to be evaluated for feasibility and benefit as SOA services that implement reference architecture operational activies. Using these inputs, candidate T&E services are identified from the survey for adoption or as candidates at a lower level for future implementation by the NST or Central Test and Evaluation Investment Program (CTEIP) projects. Some T&E services identified will be selected for prototype and experimentation in the Technology Insertion Environment laboratory during the study to collect quantitative results of the performance of these SOA-based services. These results will be compiled and presented to the COI in a draft report.

Produce final report task

Within two months of the preliminary report being presented to the COI, the final report will be generated. The final report will contain data and findings from the SOA experiments conducted. The final report and presentation will be vetted with the COI and service subject matter experts, then delivered to TRMC for further distribution across DoD via coordination with the senior advisory group.

NSTATIE project overview

The NSTATIE T&E/S&T NST Focus Area project depicted in *Figure 4* has two components that will be utilized in this study: NST reference architecture for use case development and qualitative parts of

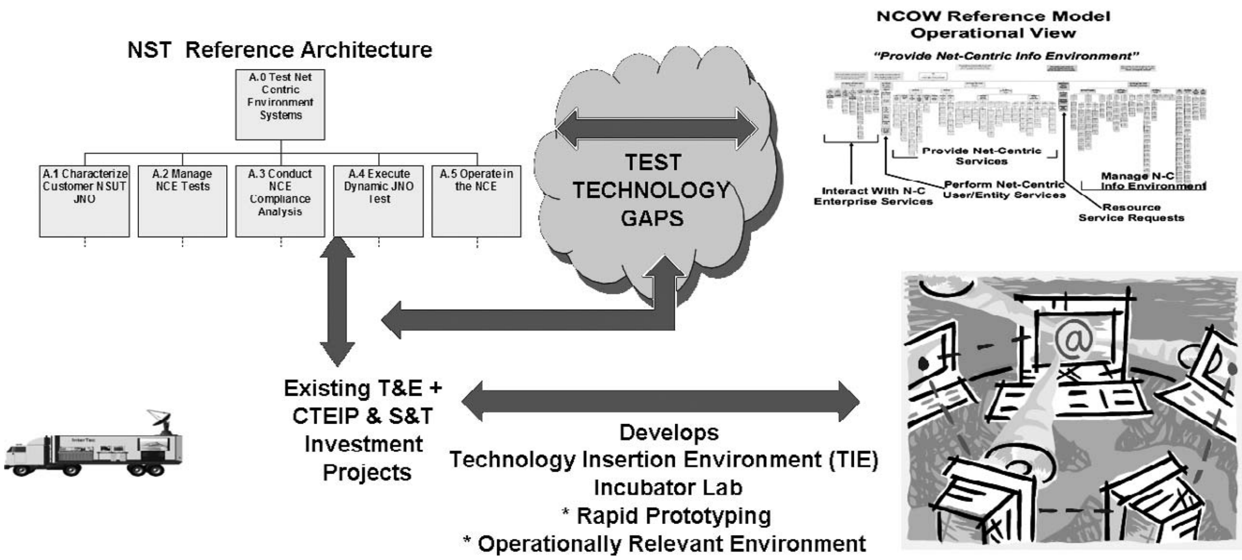


Figure 4. NST Architecture and Technology Insertion Environment (NSTATIE)

the study, and technology insertion environment for the quantitative parts of the study. The NSTATIE project is addressing technologies needed to define an NST architecture that stays lockstepped with the evolving Joint Netcentric Operations architecture. The prototype capability to accurately depict and organize NST technology gaps for current network-centric warfare systems and emerging Net-Centric Operations Warfare reference model compliant systems is being applied within the NST focus area for NST projects. The NSTATIE project is researching and developing the capability for NST S&T projects to perform R&D in a higher-fidelity and more relevant environment. The end product will be a prototype of a system available to all S&T projects to utilize as they mature through Technology Readiness Levels 5 and 6. In essence, this project becomes a technology sandbox and incubator for all T&E/S&T projects as they mature.

Current study accomplishments and status

The study team initiated the formation of the COI and solicited inputs on how to structure the study. Based on those inputs, a draft version of the Terms of Reference that describes how the study will be conducted was generated, released, and is currently under review. The Terms of Reference were presented at the first COI meeting held at the JMETC Users Group Conference held in May 2008. There was an entire track at the conference dedicated to SOA and test infrastructure. The JMETC track met in Charleston, South Carolina, with strong participation from industry, military service T&E, NASA, and DoD agencies. The presentation from this track can be found at www.trmc-test.org.

The study makes use of recent NST focus project outputs, in particular an NST reference architecture and Technology Insertion lab. In addition, the NST reference architecture was briefed at the SOA and Test Infrastructure JMETC Users Group track. Based on inputs from that meeting, the architecture was refined and then the first version was uploaded onto DARS. A survey was generated specifically to query all 18 of the T&E/S&T NST focus areas projects to determine in what ways, if any, they use SOA in the development of the technology for their project. The study team is currently compiling a report to summarize the results of the survey.

Summary

The "Applicability of Service-oriented Architecture (SOA) to Distributed Testing Infrastructure" study is under way and will involve both qualitative and quantitative measures. The COI is now requesting additional participation for the review and comment of the products produced by the core study group and COI. In general, acceptance by the larger netcentric test community is key to the success of the study, especially during the qualitative aspects of the study. The NST reference architecture will be used as a guide with the use case developed to identify candidate T&E services and as a collaboration point for discussions. The NST technology insertion environment will be used during the quantitative phase of the study to make actual measurements on current SOA technology as part of the technical assessment. Again, the goal of the study is to determine if SOA can improve the distributed testing infrastructure so that more thoroughly tested and timely capabilities are put in the hands of the warfighter. ❑

Acknowledgment

The authors would like to thank the Test Resource Management Center (TRMC) Test and Evaluation/Science and Technology (T&E/S&T) Program for their support. This work is funded by the NST T&E/S&T Program through the Naval Air Warfare Center Weapons Division Point Mugu.

ROBERT G. HEILMAN is employed with the Naval Air Systems Command (NAVAIR), Point Mugu, California. He is currently a project director in the NAVAIR Test and Evaluation (T&E) Directorate. He has more than 20 years of experience in senior engineering and project leadership positions in the field of test, training, and experimentation to include the weapons integration, flight testing, and systems T&E of numerous tactical aircraft; the development and operation of hypersonic missile targets; and, most recently, the development of next-generation testing and experimentation capabilities for Joint C4ISR system interoperability and netcentric systems. Heilman is currently performing technical and project leadership roles in investment projects developing netcentric warfare test, training, and experimentation capabilities for the Office of the Secretary of Defense, Joint Forces Command and Navy ranges. He received his bachelor of science degree in aeronautical engineering from California Polytechnic University, San Luis Obispo. He has attained the highest level of government professional certification in both systems engineering and program management. E-mail: Robert.heilman@navy.mil

GILBERT A. TORRES is employed with the Naval Air Systems Command (NAVAIR), Point Mugu, California. He is currently a project manager in the NAVAIR Test and Evaluation (T&E) Directorate. Torres is the deputy executing agent for the Netcentric Systems Test Focus Area, a Test and Evaluation/Science and Technology (T&E/S&T) advanced research project sponsored by the Office of the Secretary of Defense Test Resource Management Center. He received his bachelor of science degree in electrical engineering from California State University, Fresno, and his master of business administration with an emphasis in technology management from the University of Phoenix. He has attained the highest level of government professional certification in systems engineering, program management, and logistics.

JAMES S. BUSCEMI is the president and chief executive officer of GBL Systems Corporation, Camarillo, California. He is the principle investigator for "Applicability of Service-Oriented Architecture (SOA) to Distributed Testing Infrastructure" study. Buscemi also is the principal scientist for the Office of the Secretary of Defense Test and Evaluation/Science and Technology (T&E/S&T) Investment Program's Netcentric Systems Test (NST) Focus Area and a systems engineer for the JC4ISR/InterTEC program. He received a master's degree in electrical engineering from Polytechnic University and a bachelor's degree in electrical engineering from State University of New York at Stony Brook.

KENNETH THOMAS is the GBL lead system architect for the Science and Technology NST Focus Area and NST T&E Reference Architecture. Thomas retired from the Joint Interoperability Test Command (JITC) in June 2006 where he worked as the system architect and JITC subproject manager for the Interoperability Test and Evaluation Capability (InterTEC). He is a graduate of the University of Arizona College of System Engineering and of the George Mason University College of System Engineering. He has over 20 years of experience in interoperability testing of C4I systems and over 35 years of professional experience, including industry. Thomas is a member of ITEA and IEEE.

References

Lau, Y.-T. 2007. "Reference Metrics for Service-Oriented Architectures." *Cross Talk*, December 2007.

ITEA Journal 2008; 29: 285–288
Copyright © 2008 by the International Test and Evaluation Association

Army Realigns Headquarters' Test and Evaluation Office

Janet Garber
Director, Test and Evaluation Office,
Office of the Deputy Under Secretary of the Army, Washington, D.C.

On March 11, 2008, the U.S. Army Test and Evaluation Management Agency (TEMA) and the Office of the Army Test and Evaluation (T&E) Executive were realigned to form the U.S. Army Test and Evaluation Office (TEO). This merger consolidates and strengthens T&E oversight within the Office of the Secretary of the Army (SA) and provides a single focal point for Army T&E and, as executive agent, the Department of Defense (DoD) Chemical and Biological Defense Program (CBDP) T&E matters. The Director TEO serves as the senior advisor to the Secretary of the Army and the Chief of Staff on all test and evaluation matters and serves as the senior advisor to the Assistant to the Secretary of Defense for Nuclear, Chemical, and Biological programs (ASD [NCB]) for CBDP T&E matters.

Key words: Army materiel acquisition; Army Test and Evaluation Office (TEO, T&E) policy; overarching test procedures oversight; program oversight.

The Test and Evaluation Office (TEO) ensures that Test and Evaluation (T&E) associated with acquisition of Army materiel supports fielding the most operationally effective, suitable, and survivable warfighting equipment possible to current and future soldiers and commanders. This mission is accomplished through the establishment and enforcement of T&E policy, management of T&E resources, and continuous coordination with the T&E community, to include Army and Joint Program Managers, Office of the Assistant Secretary of the Army (Acquisition, Logistics, and Technology) and Director, Operational Test and Evaluation (DOT&E). The TEO has a director (SES Tier 2), deputy director, administrative staff, three divisions, and a special assistant. The TEO director reports to the Deputy Under Secretary of the Army (DUSA), *Figure 1*.

The Director TEO is responsible for T&E policy, program oversight, program, and budget analysis and serves as the Army and Chemical and Biological Defense Program (CBDP) T&E executive. The director advises key Army and Joint decision making panels (Army Systems Acquisition Review Council, Army Requirements Oversight Council, Army Acquisition Overarching Integrated Product Team, and Army-Marine Corps Board) on the testability of materiel requirements, the sufficiency of test plans and results, and the ability of tested systems to fill warfighter capability gaps. The Director serves as the Department of the Army (DA) staff interface with the Office of the Under Secretary of Defense, Director, Defense Research and Engineering, Office of the Under Secretary of Defense Test Resource Management Center and DOT&E for T&E related issues, policy, funding, and program coordination. The Director approves test-related documentation for HQDA and forwards, when required, to DOT&E and Under Secretary of Defense for Acquisition, Technology, and Logistics (USD AT&L) for final approval. The TEO director is the Acquisition Workforce T&E Functional Chief for the acquisition workforce Career Field, i.e., T&E. Additionally, the director interacts on a regular basis with the other Service's T&E directors to address and resolve current T&E issues. TEO's divisions provide guidance and subject matter expertise to Army and Joint test programs and promote early tester/evaluator involvement, thus ensuring T&E is integral to the entire acquisition cycle. TEO has a complementary mix of government civilians, military officers and contract employees.

Programs and Analysis Division

The Programs and Analysis (P&A) Division coordinates all T&E matters across program managers, the Army Test and Evaluation Command, HQDA staff, and Office of the Secretary of Defense (AT&L and DOT&E). This coordination is used to (a) ensure T&E programs support materiel requirements, (b)

The ITEA Journal of Test and Evaluation jite-29-03-08.3d 18/8/08 11:31:14 285

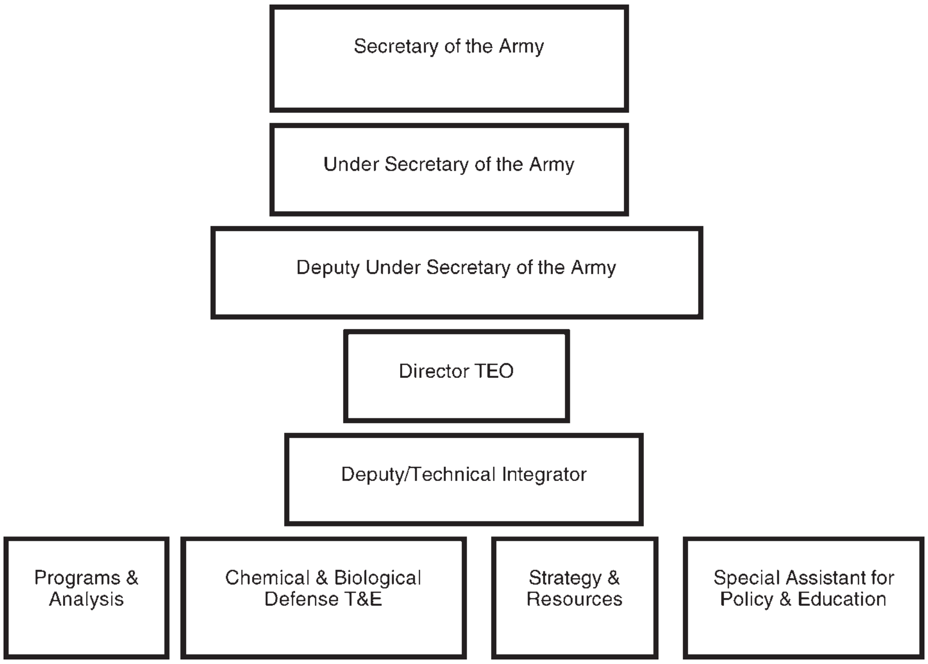


Figure 1. US Army Test & Evaluation Office organizational structure

assess the sufficiency of test plans and test results, (c) recommend modifications to the scope or focus of T&E activities, and (d) harmonize Army analysis with T&E activities in support of the acquisition process. Additionally, the division verifies T&E activities comply with Army and Office of the Secretary of Defense (OSD) T&E policy. *Figure 2* illustrates a program currently supported.

Chemical Biological Defense Test and Evaluation Division

The Chemical Biological Defense (CBD) T&E Division works closely with CBDP stakeholders to provide T&E input to the CBDP Program Objective Memorandum (POM), ensuring T&E infrastructure is maintained and modernized to support adequate developmental and operational testing. The division

Figure 2. Live fire testing verifies Mine Resistant Ambush Protected (MRAP) vehicle protects Army soldiers

develops the T&E Infrastructure Investment Strategy, test standards and processes, and provides T&E program oversight to ensure credible and robust testing and evaluation in support of programs of record. They coordinate actions throughout the CBDP community including DOT&E, Joint Program Executive Office for Chemical and Biological Defense, Defense Threat Reduction Agency, Joint Science and Technology Office, Joint Staff J-8, Joint Requirements Office-Chemical, Biological, Radiological, and Nuclear Defense, Service Operational Test Agencies, and various test facilities.

Strategy and Resources Division

The Strategy and Resources (S&R) Division serves as the proponent for Army T&E resources at HQDA by developing and defending the Army T&E funding, of about $900M per year, to the Army, OSD, and Congress, and serves as the HQDA staffing and approval agent for all T&E resource programming. The S&R division develops and monitors the Army test capabilities in the DoD's Major Range and Test Facility Base and provides HQDA oversight of the funding of Army instrumentation, targets, and threat simulator programs. This division also develops the Army T&E strategy for the test community, administers the Army portion of Department of Defense's (DoD's) Central Test and Evaluation Investment Program, and oversees the Army validation of threat representations used in testing.

Special Assistant for Policy and Education

The Special Assistant for Policy and Education develops and promulgates Army T&E policy and procedures. The Special Assistant authors Army Regulation and Department of Army pamphlet 73-1, manages the HQDA Test and Evaluation Master Plan (TEMP) approval process, supports the T&E education and training programs within the Army and OSD, and chairs the DA-wide chartered Test and Evaluation Managers Committee. These initiatives improve T&E processes in support of rapid acquisition, volatile acquisition schedules, and specialized multi-Service acquisition programs.

Value added

TEO provides senior Army leadership with expert advice on acquisition programs from an independent, overarching perspective, ranging from adequacy of requirements analysis, documentation, and testing to best use of information to support acquisition decisions.

By overseeing Army T&E activities in close collaboration with other key Army and OSD agencies, TEO ensures our soldiers receive the best possible warfighting systems and equipment within available cost and time constraints. This is accomplished by:

- Setting the strategic direction for the Army T&E community
- Facilitating and adjudicating T&E issues between the T&E community and the program executive offices
- Coordinating with DOT&E to resolve test documentation issues for OSD T&E oversight programs
- Overseeing the T&E infrastructure and minimizing duplication of capabilities
- Facilitating TEMP development and approval
- Improving T&E processes
- Standardizing test procedures and methodologies
- Ensuring Army and Joint doctrine and capabilities are represented in analyses supporting Army Requirements Oversight Council/Joint Requirements Oversight Council approvals of joint capability documents and related studies
- Fostering international agreements for use of test facilities and other cooperative T&E projects
- Overseeing Army responsibilities in JT&E, Foreign Comparative Testing, and multi-service and multinational T&E acquisition programs
- Reviewing study plans, scenarios, and results of Analysis of Alternatives
- Establishing and maintaining Army T&E policy and procedures to ensure efficient and effective application of T&E in support of the defense acquisition process
- Establishing CBD T&E collaborations with Australia, Canada, and the United Kingdom to share T&E information and fill gaps in US capabilities where appropriate
- Developing and defending the Army T&E and DoD CBDP POM to ensure T&E infrastructure, manpower, analysis and instrumentation requirements are sufficiently funded.

The Army T&E Strategic Plan expresses four overarching goals that TEO uses to set the future direction for the T&E community.

- Goal 1: Cultivate a highly skilled, multi-disciplinary professional workforce with skills and competencies to provide current and future T&E capabilities.
- Goal 2: Tailor the T&E infrastructure to meet current and future needs.
- Goal 3: Improve the quality, rigor, and responsiveness of Army testing and evaluation to better meet decision-making needs of senior leaders.

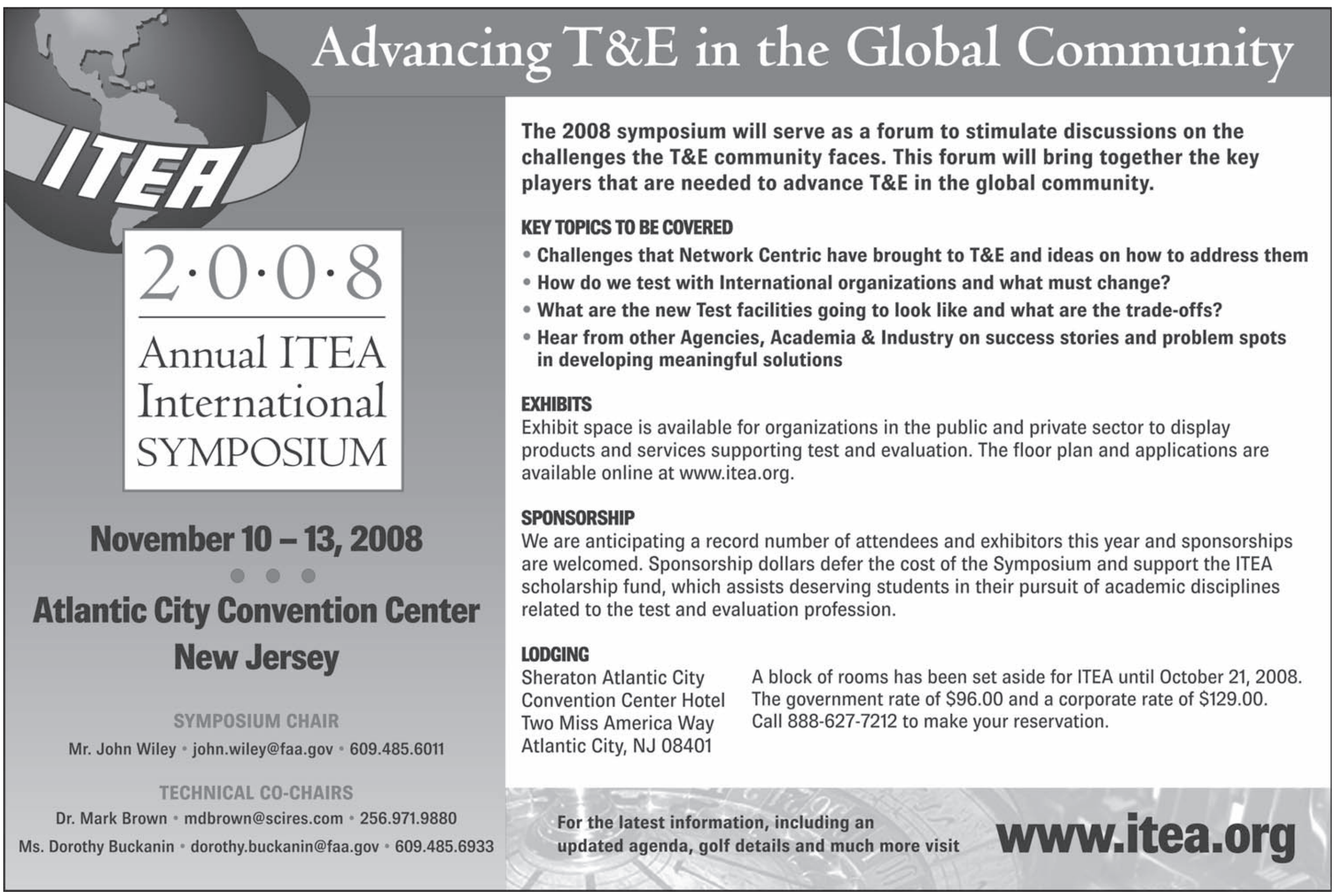
- Goal 4: Continuously improve the T&E enterprise through the development of better corporate processes.

The plan also outlines strategies and initiatives to achieve the above goals. TEO will work with the T&E community to move these initiatives forward.

Conclusion

Combining TEMA and T&E Executive offices has increased efficiency, streamlined functionality, improved value-added to the community while saving time and reducing expenses. The establishment of TEO not only provides HQDA with a single focal point for all Army and CBDP T&E matters, but also centralizes T&E subject matter expertise and analysis to ensure timely, accurate, and relevant information is provided to decision makers to support the warfighters. More information about TEO and the T&E strategic plan is available on Army Knowledge Online—search for TEO in the search box. ❑

JANET GARBER entered the Senior Executive Service in November 2003 as the deputy director, Office of Program Analysis and Evaluation (PAE) at the Department of Energy (DOE). In July 2004, Garber became the director of the DOE PAE office. In May 2005, she transferred to the Army to serve as director, Test and Evaluation Management Agency, Office of the Chief of Staff. She became the director, Test and Evaluation Office, Office of the Deputy Under Secretary of the Army in March 2008. Garber is a graduate of Michigan State University, class of 1974, earning a bachelor of science degree in multidisciplinary social science and a master of science degree in business administration from Boston University in 1986.



ITEA Journal 2008; 29: 289–293
Copyright © 2008 by the International Test and Evaluation Association

Test Strategy for Net-Centric C4ISR System

Anil N. Joglekar, Ph.D.
Institute for Defense Analyses, Alexandria, Virginia

Today, almost all command, control, communications, computer, intelligence, surveillance, and reconnaissance (C4ISR) systems are acquired via spiral development. In order for these systems to support warfighter missions, they must be networked together with various weapon systems. Current operational testing is generally dedicated but limited in scope and is conducted by stimulating the system under test with master scenario event lists. Such testing is not fully operational and does not focus on effectiveness in accomplishing the warfighter mission.

This article proposes a strategy to improve the operational realism of testing net-centric C4ISR systems. The strategy is based on the United States Department of Defense's lessons learned from the Year 2000 operational assessments conducted in 1999 in all combatant commands. Developmental C4ISR systems can be deployed in the field but isolated from real world operations using different crypto for the networks. Thus, the systems can ride on operational networks and use real-world combatant command databases, but using different crypto and hence not mixing in real-world operations. In this strategy, the systems being tested with supporting networks and databases can be assessed to support the warfighter mission in a field training exercise. The feasibility of this strategy is now possible because of the improved net-centric environment in the field today. This strategy can also deal with interfacing coalition C4ISR systems, which are generally excluded during testing today.

Key words: C4ISR; crypto; IA; IV&V; net-centric; radiant mercury.

Command, control, communications, computer, intelligence, surveillance, and reconnaissance (C4ISR) systems are increasing in number, yet their procurement is fragmented. These systems are procured by the four Services, the Defense Information System Agency (DISA), and by combatant commands. Further, these systems are constantly upgraded through spiral developments and capability packages/modules. Over the past 10 years, the ability to network these C4ISR systems together has significantly enhanced the C4ISR system operations. The present status of networking is achieved using different data links and some unclassified and classified networks.

In procuring net-centric C4ISR systems, the major emphasis is primarily on integration rather than development of new hardware and software. The military is relying heavily on commercial off-the-shelf hardware and software. Future command and control systems might rely on service-oriented architectures. Many future command and control systems will simply ride on information networks and use generic services provided to them rather than develop networks and develop specific applications software. Services provided are: situational awareness, discovery, collaboration, security, availability, and authorized user pull of information at all levels of war (strategic, operational, and tactical) and all ranges of military operations. For example, net-centric collaboration service enables decision makers and staff to change the operational paradigm from sequential planning and decision-making to synchronous planning and decision-making through simultaneous information sharing.

Figure 1 depicts the current environment versus how the military environment will change in the future with net-centricity. Currently, functions such as logistics, intelligence, ground fires, and air defense are stove-piped. A significant level of system-of-systems integration effort is required for these systems to inter-operate together. Many times such integration is by "sneakernet," e-mail, or other means, and the user is responsible for integration. To complicate matters, modifications to these systems are made by software releases every few months.

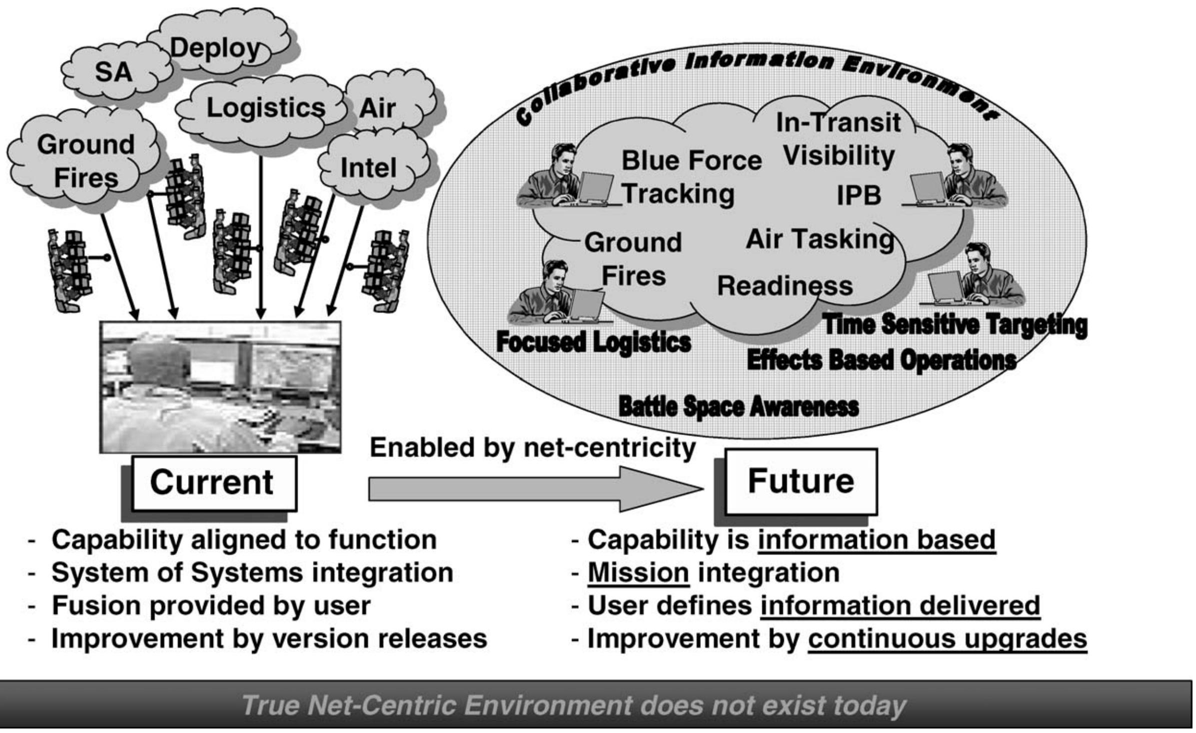


Figure 1. Expected evolution of net-centric command and control

In the truly net-centric vision of the future, all capabilities and functions will be information based. Similarly, all missions will be integrated rather than stovepiped. Users will define what information to push and what information to pull. Upgrades and changes to these systems will be made continuously in a plug-and-play environment.

Uniqueness of C4ISR programs

C4ISR programs are commonly technology driven. Users generally do not know all the advances in technology and contractors generally demonstrate new technologies to the users before an acquisition program is initiated. The pace of both hardware and software technologies is evolving very rapidly in the commercial field. Hardware is undergoing miniaturization, and advances have been made in reduction of the heat generated by hardware. Display technologies are improving at a rapid pace. Software desirability is gravitating towards Windows rather than UNIX or other specialized software. Software advances are allowing improved collaboration, security, access lists, discovery, and situational awareness. Improvements in the network technologies involve higher bandwidth, secure mobile networks, and multi-level security networks. Currently, networks are very much stove-piped, and network domains do not permit crossing network boundaries. Roaming capabilities and mobility for users is significantly lacking in military operations. When a military user goes from one area of responsibility to another area of responsibility, he/she cannot seamlessly continue to operate. These problems are unique to C4ISR systems as compared to weapons systems because the acquisition of C4ISR systems is fragmented, and most contractors use whatever hardware/software they are familiar with rather than use standard protocols and formats. Sometimes by using unique software, contractors can receive lifetime maintenance contracts for a system. Similarly, contractors also benefit by building interface systems for two disparate systems to operate together.

Thus, in general, C4ISR programs do not have well articulated requirements or capability definitions. Most C4ISR acquisition programs are technology driven rather than requirements driven. Many times the requirements are very technical rather than operational. Systems' operational missions and mission-level measures are not well articulated.

Increasing net-centric capability achieved through links and networks has improved C4ISR systems to interoperate more effectively together. Very few C4ISR systems are based on Internet protocols.

For disparate C4ISR systems to interoperate, they must be able to share and contribute to a common database. This challenge is not yet easily solved. Databases are in different formats and protocols, and transferring data generated by one system to another many times requires manually entering the data because of lack of standardization in formats and protocols.

In the effort to achieving the net-centric environment, the military has developed four networks in addition to the open Internet, which is not secure. One network is unclassified but for official use only and can be accessed only with some type of user authentication; the second is classified secret; the third is secret releasable to allies; and the fourth is top secret. This arrangement lets many C4ISR systems operating at one of these classifications to be networked together, but problems arise when systems with different classifications need to interoperate. Radiant Mercury is a software application that automatically sanitizes and downgrades multilevel classified formatted information. Radiant Mercury guards are used between systems with different security classifications, but it is still a cumbersome solution. Many times the alternative is to have all systems operate at the highest security classification level.

Information assurance issues

Another key area for net-centric C4ISR systems is information assurance. Increased networking and improved interoperability make the networks more vulnerable for jamming, hacking, and exploitation, so they must be well protected. Information assurance is achieved by the following actions, which are not comprehensive:

- Use of hardware tools such as firewalls and intrusion detection devices
- Closing unneeded computer ports
- Enforcing strong password policies
- Keeping vulnerability patches up to date
- Remote access authentication
- Permit booting systems from internal hard drives only
- Monitor security logs.

Information assurance also requires education and awareness among users and systems administrators. They are key to protecting the networks and information. Further, periodic vulnerability assessments and penetration tests should be conducted to identify vulnerabilities, and systems administrators should have the ability to respond to detection of penetrations.

Challenges for the Net-Centric C4ISR system testing

Because C4ISR systems are technology driven, they become obsolete quickly. Hence, the desire is to keep development and testing time to a minimum and with limited resources. This desire, however, has to be matched with test rigor so that users are not receiving immature products. Discipline is necessary during the development process to ensure that strict software configuration control is maintained and that an independent verification and validation (IV&V) team is overseeing software development. Also, development contractors should be encouraged to have a hardware-in-the-loop laboratory to periodically check software functionality. Also, reused software should be used without modification. Otherwise, modified software needs to be retested. Many times the tendency is to use prior developed software and make modifications without retest, and assume it will work fine.

Another significant need is to have operational users participate in the design process and examine the new C4ISR systems early. Operational users also need to develop concepts of operations for new systems. Many times a designer develops the concepts of operations using sound technical capabilities, but without considering operational consequences. Another pitfall for C4ISR systems is that when operational users are exposed to these systems early, they like to add capabilities and requirements. Contractors do not mind changes to the requirements, as they can receive additional funds. However, changes generally delay the program, which should be avoided. New capabilities or requirements should be considered in subsequent block upgrades.

Y2K operational evaluations (OPEVALS)

In 1999, Congress mandated that all combatant commands should verify through operational testing that all of their C4ISR systems would be operational when the clocks rolled over to the Year 2000 (Y2K). Many old and obsolete systems were phased out because they could not be made Y2K-compliant. For each mission area, a "thin line" of C4ISR systems was identified—only the minimum of systems that could execute that mission area. *Figure 2* shows an example of the United States Forces in Korea (USFK) targeting mission system thin line.

The thin line of systems was operationally tested for the clock rollovers using the actual hardware, software, and networks. Y2K assessments could not use simulated systems or networks; all assessments had to be conducted using operational systems and networks. A clever approach to accomplishing this feat was to use combatant commands' operational systems and networks, but use different crypto to isolate test networks from day-to-day, real-world operational networks at each combatant command. Using the Y2K test systems and networks, all essential mission actions were conducted to verify that the clock rollover did not cause any operational problems. For some combatant commands, such as Pacific Command and Central

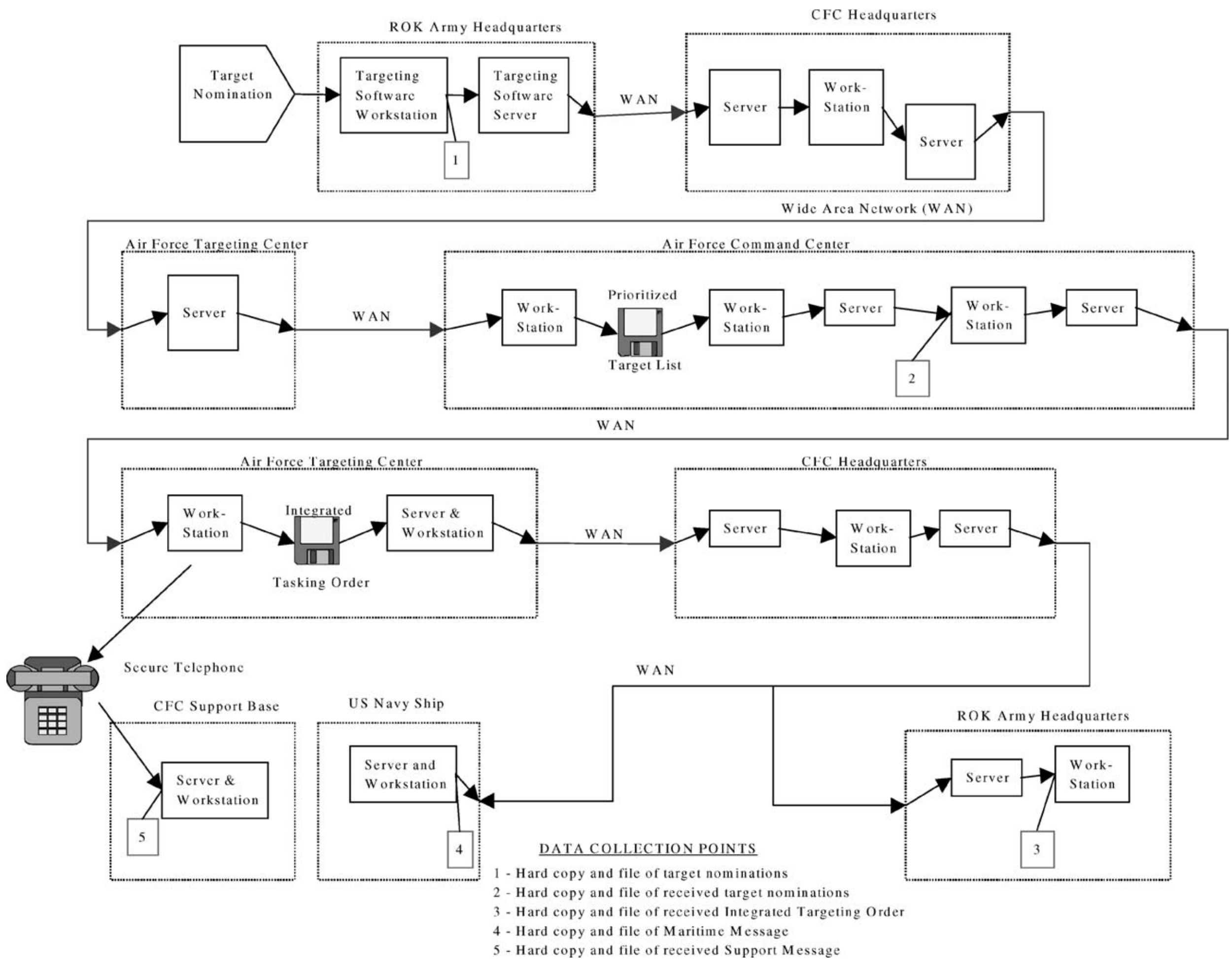


Figure 2. Flow diagram for United States Forces in Korea (USFK) targeting mission thin line (Joglekar et al. 2000)

Command, C4ISR systems were connected in 13 different time zones. The testing was very realistic and valuable, and because of that rigor, no operational problems occurred when actual clocks rolled to the year 2000.

After the completion and success of the Y2K testing, this technique was somehow lost. This article recommends resurrecting this testing technique to conduct testing in parallel with real-world operations for future net-centric C4ISR testing.

Combatant command and Service training exercises provide an operationally representative environment in which to assess C4ISR system operational performance. *Figure 3* shows the hierarchy of how systems, functions, and missions can be tested in progression.

By piggybacking on training exercises but isolating the test systems with different crypto, test realism could be achieved without interfering with real-world operations. This technique may not work adequately if the network bandwidths are not adequate. However, most combatant commands not actively involved in combat should have sufficient bandwidth to use this technique.

Suggested improvements for net-centric C4ISR testing

Future net-centric C4ISR testing should be flexible and responsive to the acquisition process while ensuring that warfighters get robust, mature, and operationally effective capabilities. To accomplish this, the following improvements are suggested:

- Conduct operational test and evaluation in warfighters' mission contexts.
- Maintain early and continuous involvement by combined/integrated (developmental and operational) test teams
  - Maintain independence of testers
  - Integrate development and operational testers
- Use real-world networks and C4ISR systems, but isolate test networks using different crypto
- Piggyback testing on combatant commands' exercises

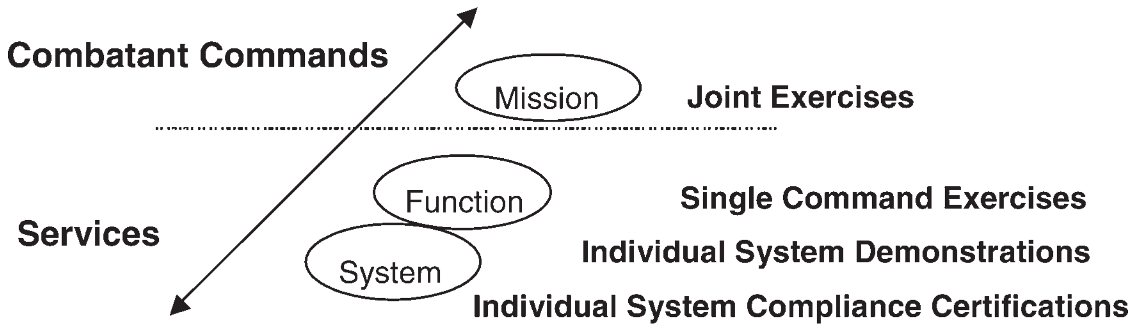


Figure 3. Example of how command, control, communications, computer, intelligence, surveillance, and reconnaissance (C4ISR) systems, functions, and missions can be tested

- Conduct risk-based testing for intermediate capabilities, packages, blocks, modules, etc.
  - Determine level of test by risk analysis
  - Release capability packages only when ready
  - Conduct rigorous operational testing when a Block or Increment is complete.

Summary

Net-centric C4ISR operational testing should focus on operational mission or capability rather than technical assessment. The test strategy should be flexible while maintaining test discipline. Testing should be risk based and focus on high-risk areas for intermediate capabilities. Further, the Y2K testing methodology suggested earlier should be employed. This means exploring the fielding of intermediate capabilities for trial periods in parallel with real systems and networks, but separated by crypto. Combined developmental and operational testing teams should be used and involved early and continuously. The hope is that with combined test teams, rigorous development will be accomplished. ❑

DR. ANIL JOGLEKAR is a senior research staff scientist with the Institute for Defense Analyses (IDA) in Alexandria, Virginia. He oversees the operational test and evaluation of net-centric command, control, and communications systems, the Joint Test and Evaluation (JT&E) program, and was responsible for an effort to develop the National Personnel Recovery Architecture that covers all government agencies including the Department of Defense. Dr. Joglekar received the National Defense Industrial Association's 2000 Tester of the Year Award. He has been with IDA since 1979 and has led studies related to NATO airspace management, combat identification systems, Patriot fielding in NATO, and the Joint Air Defense Operations/Joint Engagement Zone Joint Test. Dr. Joglekar led the DOT&E's team for the CINC year-2000 (Y2K) operational evaluations. He was instrumental in the development of operational test and evaluation policies for interoperability and information assurance (IA). Dr. Joglekar received his Ph.D. degree from Stanford University in 1973 in aeronautics and astronautics. E-mail: ajogleka@ida.org

References

Joglekar, A., et al. April 12, 2000. IDA Paper P-3517, "Year 2000 (Y2K) Operational Evaluations by Warfighting Commanders."

ITEA Journal 2008; 29: 294–300
Copyright © 2008 by the International Test and Evaluation Association

Tests and Experiments: Similarities and Differences

Richard A. Kass, Ph.D.
GaN Corporation,
U.S. Army Operational Test Command, Fort Hood, Texas

Test and experimentation are integral to the capability development process. This is the second of a two-part discussion on experimentation. This article considers the similarities and differences between experimentation and testing. While the two endeavors address different questions and exhibit some differences in the planning and execution process, overall similarities outweigh differences especially in event resources suggesting potential gains from sharing resources.

Key words: Demonstration; interdependence; shared resources; terminology; tests experiments; training.

Test and experimentation are two primary information venues in the Department of Defense (DoD) research and development. Testing is associated with system acquisition. Developmental and operational testing assess system progress toward acquisition milestones. Warfighting experiments[1] on the other hand, are associated with concept development. This is the second of two articles on experimentation. The previous article (Kass 2008) illustrated the uses, components, and validity requirements for warfighting experiments. Building on that description, this article discusses the similarities and differences between a test and an experiment.

This experiment versus test presentation[2] is intended to start the discussion. As experimentation and testing continue to evolve, the characteristics contrasted here will certainly change. The main thesis of this comparison is that while notable differences are evident, overall similarities are more significant than differences. Given the similarities, this article suggests that the resources employed in both endeavors can be shared to the mutual benefit of both. In the process of comparing and contrasting experimentation and testing, associated aspects of training are discussed. This will further illustrate the interconnectedness of DoD activities.

Terminology confusion between tests and experiments

Our language promotes confusion between tests and experiments.

We conduct experiments to test hypotheses.
We employ an experiment design to test systems.
Experimental systems undergo testing.

This confusion is exacerbated by common practices. Some test-like activities are renamed as "assessments" or "demonstrations" in order to reserve "testing" to specific agencies with identified acquisition requirements or to avoid consequences of negative results. Likewise, the "experiment" title can be attached to a number of activities that others would call "wargame" or "demonstration."

Terminology confusion suggests a close connection between test and experiment. The following definitions are provided:

Test: to assess the presence, quality, or genuineness of anything (Random House 1982);

Experiment: to explore the effects of manipulating a variable (Kass 2008).

Tests are one way to assess the quality of something. Other means include reliance on logical and mathematical relationships, authority, historical precedent, and natural observations. Assessments derived from testing imply empirical measurements under specified conditions. An example will illustrate the different but complimentary focus of experiment and test.

A math *test* is given to confirm whether students have attained certain levels of math proficiency using familiar letter-grade scale of A through F. A math *experiment* has a different purpose. Math experiments are designed to explore something new, for example, to determine the best way to teach math. The primary purpose of a math *experiment* is not to assess participants' level of math ability; but rather to examine the effect of various teaching methods on participants' math ability. During the *experiment*, each participant's math ability will be assessed by a math *test* to determine higher math ability from lower. The purpose of this *test* is not to pass or fail the participants; but to quantify the effect of the *experiment* treatment. The *experiment* hypothesis might be: If teaching methods (a) are used, then math scores and (b) will increase. The way to determine whether math scores increased is to give the students a math *test* before and after the treatment.

The example signifies that testing is a method for assessing trial outcome. An experiment can be viewed as a sequence of tests. Each experiment trial is a test of one experimental treatment condition. An experiment is a systematic sequence of individual tests to examine a causal relationship, while a test is conducted to quantify an attribute.

Misperceptions of test and experiment distinctions

Given the interconnectivity of experiment and test, it is inevitable that misperceptions arise. One often hears experimenters caution their visitors: "Remember, this is an experiment, not a test." Why this admonition? Acquisition systems that do poorly in tests are in jeopardy of being cancelled. Tests include the idea of pass or fail. Experiments do not. Failure to produce a hypothesized experimental effect is more forgiving: "Let's try this and see what happens."

Experimenters sometimes push the forgiving nature of experiments too far in the phrase—"Tests can fail, experiments never fail." If this statement is interpreted to indicate that experiments rarely impact system acquisition decisions, the statement is understandable. If however, the statement is interpreted to mean, "there are no useless experiments" the statement is wrong. As discussed in the previous article, experiments can fail to provide sufficient information to resolve the experiment hypothesis.

Another misperception is that "Experimenting is messy, but testing is precise." This perception may reflect difficulties in representing the complexity of warfighting in experiments. It is difficult to conduct precise experiments in the operational environment. However, it is equally difficult to conduct precise operational tests in a realistic representative environment for the same reasons. This then cannot be the basis for distinguishing warfighting experiments from operational tests. Both depend on the expertise and experience of the experimenter and tester to balance the requirement for realistic operations against the needs to detect a change and to understand why the change occurred.

A third misperception is that "testing requires detailed data, while experiments use only high-level data." This distinction would not apply to warfighting experiments conducted in constructive or human-in-the-loop simulations because simulation outputs in these experiments are very precise and the experimenter is often inundated with detailed second-by-second interaction data on every entity in the simulation.

*Table 1. Four different perspectives of hypothesis elements**

Hypothesis: If capability A (new sensor), then effect B (increased detections).	
Demonstration	Show how A works to produce B
Training	Practice using A to produce B
Experiment	Determine better way to produce B
Test	Determine if A works to produce B

* Adapted from Figure 42 in Kass, R. A. 2006. *The Logic of Warfighting Experiments.* Published in 2006 by the Command and Control Research Program (CCRP) of the ASD/NII. Used with the permission of the CCRP.

This "data" distinction is derived from the circumstances in which tests and experiments are conducted in the field environment. Test agencies have accumulated sophisticated data-collection instrumentation for use in field tests. When the acquisition community needs to make a decision on a multibillion dollar program, it can justify the development of sophisticated instrumentation to provide maximum information to the acquisition decision. Conversely, experiments designed to examine the potential of a new technology do not have the same incentive today to invest large resources in a detailed answer.[3]

Conceptual difference between tests and experiments

The difference between an experiment and test cannot be based on precision or level of data alone. So is there a difference? One way to formulate an answer is to compare how various disciplines approach a new capability exemplified in the experiment hypothesis paradigm:

If capability A (new sensor), then effect B (increased detections).

Table 1 depicts how the elements of this hypothesis are viewed from the perspective of a demonstration, training, experiment, and test.

A ***demonstration*** is an event orchestrated to show how a process or product works. Demonstrations exhibit how a capability can produce an effect. In the military arena, demonstrations are commonly used as the initial step in training. An instructor demonstrates the correct procedures to follow with A to produce B. In the commercial world, product demonstrations are useful to convince others to buy the product or to illustrate the correct way to use it. While tests and experiments examine the effectiveness of capabilities, demonstrations assume the product works.

Training can be characterized as practice with A in order to accomplish B. This is easy to see when B is defined as a task with conditions and standards (more on this later). If the general task is to detect targets, the task conditions would specify the environment in which detections need to occur. The task standard would indicate the percent of the targets to be detected to meet the training objective.

Most ***experiments*** begin with a "capability gap." In our example, detections need to be increased as indicated on the right-hand side of the hypothesis. An experiment then is a trial-and-error process in search of a good solution for the left side of the hypothesis paradigm to fill the capability gap. Typical experiment questions are expressed in broad terms such as—"Does this approach produce a favorable outcome?" and "Can this problem be solved with X?"

In contrast, ***tests*** can be viewed as examining the goodness of a particular solution with respect to producing its intended effect. Tests are not searches for solutions, but rather a search for the strength of a solution's relationship to its effect. Typical questions for testing are expressed as, "How well does this item work?" and similarly, "How well does this item meet its requirements."

Thus far we have discussed some useful, and some not so useful, ways to think about the differences between experiments and tests. The remainder of this article will address the practical similarities and differences when it comes to planning, executing, and resourcing each event.[4]

Planning process

Planning coordination

Planning processes for experiments and tests employ different terminology but are quite similar functionally. Large tests are collaboratively designed and resourced using test and evaluation working-level integrated product teams (T&E WIPTs). This group meets periodically and is chaired by the capability Program Manager or operational test agency (OTA) depending on whether it is planning a developmental or operational test. Subgroups devoted to M&S, scenario, instrumentation, training, and so forth meet more frequently and less formally. A series of test readiness reviews (TRRs) brings together senior stakeholders to assess progress in the development of the system-under-test (SUT), test planning, and test-resource commitments.

Similar planning processes occur for major experiments. A concept development conference is followed by three planning conferences—initial planning conference (IPC), mid planning conference (MPC), and final planning conference (FPC). These serve the same purpose that T&E WIPTs and TRRs serve in testing. Again, smaller experiment planning IPTs can be formed to focus on M&S, scenario, analysis and data collection, training, and initiative development. An "initiative development" IPT is the experiment corollary to the capability Program Manager. Often capability initiatives for experimentation begin as "good-ideas" that need to be fleshed out so a concrete instantiation can be brought to the experiment. Capability instantiation can include adjustments to simulation, creation of new procedures, early prototypes when available, or implementation of low-fidelity surrogates when prototypes are not available.

Event rigor

The previous article identified four validity requirements for rigorous experiments:

1. Ability to employ the new capability;
2. Ability to detect a change;
3. Ability to isolate the reason for change;
4. Ability to relate results to real operations.

These experimentation requirements are applicable to testing. If the test unit is not able to employ the new system, or if the tester cannot detect a change in performance when the new system is employed, or cannot isolate the reason for any observed performance, or cannot relate the test environment and test results to actual operations; then the test has validity deficiencies. Experimenters and testers consult the same "design of experiment" textbooks to design their events.

While tests and experiments have similar validity requirements, they have different review processes. Test agencies and ranges conducting developmental testing have detailed test protocols and test plans that have increased in rigor through refinement over many years. Deviations from these protocols often require prior approval from both the tester and program manager. Operational testing includes an external review. Operational test plans of major acquisition systems are formally reviewed by the Director, Operational Test and Evaluation (DOT&E) and operational testing cannot begin until DOT&E approves the test plan.

Experiment agencies typically do not have the historical heritage found in the major test ranges and do not have detailed experiment protocols. Experiment plans are often reviewed internally and these reviews tend to focus on scenario realism, adequacy of the experiment initiative instantiation, and availability of experiment resources.

Results utilization

Testing has a major advantage over experimenting in results utilization. The results of developmental test (DT) or operational test (OT) support decisions about capability development programs. Test results assist

*Table 2. Three different portrayals of measures and goals**

	Measure	Goal
Training standard	Minutes to complete attack after target identification	Criterion (provided by commander)
Test criterion	Time to complete task after target identification (MOE/MOP)	Threshold (X minutes)
Experiment measure	Time to complete task after target identification (MOE/MOP)	(not usually available)

* Adapted from Figure 43 in Kass, R. A. 2006. *The Logic of Warfighting Experiments.* Published in 2006 by the Command and Control Research Program (CCRP) of the ASD/NII. Used with the permission of the CCRP.

program managers in assessing whether system performance is on schedule and where to focus system corrections.

In contrast, it not so easy for experimentation programs to show examples where their experiment results have changed the military environment. One reason for this is that most experiments are conducted on future prototypes or concepts outside the programmed acquisition realm. Any good ideas from experiments are initially unfunded and will struggle to find a "funded home."

Interestingly, the impact of many experiment programs may be more indirect than direct. One of the most visible legacies of the Millennium Challenge Field Experiment conducted in 2002 (MC02) was the follow-on creation of the Joint National Training Capability (JNTC) in Joint Forces Command (JFCOM). JNTC was built on the distributed live, virtual, and constructive (LVC) simulation architecture designed for experiment execution. While not the focus of the experiment, most experiment agencies can point to technologies developed to support their experiments that have found use (reuse!) in the operational forces as enhancements to the training environment or operations themselves.

Execution process

Type event

Experiments have an advantage over tests in flexibility—design space—to explore new ideas. Acquisition tests are restricted to testing something concrete—in hand, a component or prototype—even if it is only software algorithms. Experiments, in contrast, have few reality constraints. Experiments can be conducted on future weapons that exist only as concepts. These experiments can be executed entirely in simulation as constructive experiments or as analytic wargames. The focus of these experiments is not "does it work;" but on the potential impact of these ideas on future warfighting operations.

Unit tasks and measures

Tests and experiments are both concerned with realistic scenarios based on defense planning scenarios (DPS). Both look to the Joint and Service description of strategic, operational, and tactical tasks with their associated conditions and standards to provide the basis for unit activity during the event trial. Joint tasks and standards are identified in the Universal Joint Task List (CJCSM 2002)[5] (UJTL). Test and experiment use of the standardized tasks, conditions, and standards originally developed by the training community has been a positive development. The training, testing, and experimentation community can now speak a common language.

The UJTL conditions can provide the basis for the test or experiment trial conditions and UJTL standards can provide a starting point for developing the measures of effectiveness (MOE) and measures of performance (MOP) for tests and experiments. A closer look at the terminology for standards and measures will show that differences in terminology might blur similarities across the three communities. The UJTL task "Provide firepower in support of operations" includes the standard provided in the first row in *Table 2*.

The UJTL notes that training standards have two parts: a *measure* and *criterion*. While numerous quantifiable *measures* are provided in the UJTL, the *criterion* component is not included. The UJTL document asserts that the criterion, the specific time (in this example) in which the task is to be completed, is to be provided by the commander of the unit undergoing training. The commander might select 6 minutes as the task criterion and the unit would continue to re-execute the task until they accomplish it within the allotted time. It is a common misperception that the UJTL includes training standards—it only includes the measure portion of the standard. Training measures without criteria are still quite useful to testers and experimenters.

Starting points for measuring success in the test community are requirements identified in the initial capability document (ICD) and deployment capability document (DCD). While translating acquisition requirements directly into test criteria can be challenging,[6] some are relatively straightforward. Requirements for mean-time-between-failures, message completion rates, and detection ranges would be associated with measures like *time-between-system-failures*, *percent of messages successfully completed*, and *range of detection* with associated *"thresholds."* Notice the shift in terminology between the test and training community. For trainers, the "criterion" represents only the threshold component of the training standard. For the testers, "criterion" includes the measure and the threshold.

Testers and experimenters use identical terms for measures. The primary difference is that experimenters avoid the term "criterion" because they rarely have available thresholds to evaluate success. Consequently, experiments rely on comparative analysis based on alternate treatment conditions—different proposed capabilities or a single capability under different scenario conditions.

Table 2 above highlights the ease in bridging terminology differences. Use of common measures among trainer, tester, and experimenter would increase mutual synergies among the three communities. Operational forces continuously undergo training that could yield realistic, operational thresholds of baseline-force capability based on a heterogeneous mixture of units under a wide variety of operational conditions. This training data, if systematically collected and used by experimenters, would greatly enhance the relevancy of experimentation in answering "so what" questions. Even if an experimental capability performs better under some conditions than others, how much better it is than what is available today?

The test community could also benefit when quantifiable thresholds on current mission performance are available from the training community. Test criteria are based on *system* performance rather than *unit* mission accomplishment. In operational testing, there is increasing emphasis on assessing system capabilities and limitations with respect to overall unit mission accomplishment—especially on the Joint battlefield.[7] While *system performance thresholds* are readily available from requirement documents, there is yet no agreement on how to arrive at a *unit mission success threshold*. A systematic data collection effort of unit mission successes during training exercises might provide the operational baseline for testing system contribution to mission success.

*Table 3. Comparison of resource requirements**

Requirement	Experiment	Test	Training
System Realism			
Simulated (*constructive*)	X	x	X
Simulator (*virtual*)	X	x	X
Prototype (*live*)	X	XX	-
Operational system (*live*)	X	x	X
Trained operators	X	X	x
Instrumentation			
System-level diagnostic collection	-	XX	-
System-range interactions collection	X	X	X
Feedback	X	X	XX
Networks/communications	X	X	X
Exercise Management			
Controllers	X	X	X
Observers	X	X	XX
Trainers	x	x	X
OPFOR unit/equipment	X	XX	XX
Analysts	XX	XX	-

* Adapted from Figure 44 in Kass, R. A. 2006. *The Logic of Warfighting Experiments.* published in 2006 by the Command and Control Research Program (CCRP) of the ASD/NII. Used with the permission of the CCRP.
"x," "X", and "XX" indicate increasing emphasis.

Event resources

If you fell into the middle of a warfighting field experiment, operational test, or training exercise it would be difficult to know which one you had fallen into. In any one, you would observe military operators performing tasks to accomplish a mission. In the extreme case, one might detect operators employing novel procedures or equipment. This could indicate an experiment or test on advanced technology. With this exception, almost nothing else during actual execution from the player perspective would indicate experiment, test, or training. It is only when the purpose of the event is known as discussed above that subtle differences between experiment, test, and training may be evident.

Given this similarity in execution, it is not surprising that the resources to execute an experiment, test, or training exercise are quite similar with only a few notable exceptions. *Table 3* provides a comparison of resource requirements.

System realism. Earlier it was noted that experimentation is the most flexible enterprise. It can be executed with live, virtual, or constructive systems as the primary system of interest. Testing has the most stringent system requirements. Developmental and operational testing are conducted on live prototypes and preproduction systems. While testing does employ some constructive and virtual representations, these are used to save resources in populating a realistic test environment, not to represent the primary SUT. While most operator training is conducted on operational systems, use of air, ground, and sea virtual simulators is continuing to expand to save training resources. Most operational staff training is also accomplished in an LVC environment.

Instrumentation. Range instrumentation and communication networks for event control, execution, distribution, data collection, and player feedback is becoming more similar for test and training as a result of several ongoing initiatives to further common support architectures.[8] The primary exception to eventual complete commonality is that testing requires diagnostic performance data from the SUT, as discussed earlier.

Exercise management. The lead-in to this section suggested that event execution was quite similar among the three communities; this is also true for event management resources. All three communities require similar event controllers, event observers or data collectors, and analysts for interpreting results. Similarly all three require a representative threat force. Operational test approval often requires that threat representativeness be certified by an external agency. While the training community emphasizes training, test and experiment communities employ trainers to ensure operators can use the new capability. The table highlights that experiments and tests place a higher premium on statistical analysis of the event data.

Summary

Experimentation and testing are both important to capability development. They provide empirical data for different questions. However, they have more similarities than differences. They have similar planning processes; similar validity requirements; use similar language for tasks and measures; and for the most part, employ the same resources to design, execute, and report events.

There are some differences to keep in mind. Experiments have greater flexibility to explore a wider variety of warfighting questions and alternatives using virtual and constructive simulations, since experiments do not have to wait for actual prototypes. Conversely, experimentation has far less formal methodological oversight and it is not always easy to link experiment outcomes to implemented operational changes. Tests, on the other hand, have more explicit measures of success (e.g., system procurement requirements) and provide far greater system-diagnostics data collection to know what to fix. Moreover, test results always impact capability development.

Confluence of test and training architecture is assisted by the fact that they have clearly delineated sponsors in the DoD—Under Secretary of Defense (Personnel and Readiness) manages training while the Under Secretary of Defense (Acquisition, Technology, and Logistics) and Director, Operational Test and Evaluation provide test management. There is no corresponding high-level sponsor and policies for experimentation. Consequently, experimentation policy is decentralized making it more difficult to build a coalition with the test and training communities from the top down. However, the sharing of expertise, data, and resources can only benefit all three.

A realization from these comparisons is that predominantly the same resources can be used for experimentation and testing, as well as training. This suggests that efficiencies can be gained if experimentation, testing, and training continue to progress towards shared resources. Increased emphasis is being directed at finding interdependent investment strategies for overlapping infrastructure to support testing and training.[9] A similar interdependency case can be, and should be, made for testing and experimenting. ❑

RICK KASS has 25 years in designing, analyzing, and reporting on operational field tests and military experiments. He held multiple positions as test officer, analyst, and test director for 18 years with the U.S. Army Test and Evaluation Command (USATEC) and was chief of analysis for 7 years with the U.S. Joint Forces Command (USJFCOM) joint experimentation program. Currently Rick works for GaN Corporation supporting the Army's Operational Test Command at Fort Hood, Texas. He has authored over twenty-five journal articles on methods for research, experimentation, and testing and was the primary architect establishing the permanent Warfighting Experimentation Working Group in the Military Operations Research Society (MORS). Rick is a graduate of the National War College and holds a Ph.D. in psychology from Southern Illinois University. E–mail: rick.kass@us.army.mil

Endnotes

[1]Warfighting experiments are distinguished from experiments used in medical research and early technology research.

[2]This article expands on ideas previously printed in Kass R.A., 2006. The Logic of Warfighting Experiments, published by the Command and Control Research Program (CCRP) of the ASD/NII, which has graciously granted permission to include the material in this work.

[3]In the 1970s and 1980s the U.S. Army sustained a Combat Development Experimentation Center (CDEC) with dedicated operational forces, advanced range instrumentation, and scientific methodology for experiments. This center no longer exists. There are costly experiments today. These costs are mostly associated with force operations and M&S development and execution; not the cost of collecting detailed system performance data.

[4]The following comparison of test and experimentation is more applicable to developmental and operational testing following early prototype development that are assessed against military tasks as opposed to engineering thresholds.

[5]CJCSM 3500.04C, July 2002. The Services have augmented the Joint list with their respective tactical tasks: Army Universal Tactical List (AUTL), Universal Navy Task List (UNTL), and Air Force Task List (AFTL).

[6]See Kass, R. A. "Writing measures of performance to get the right data." *The ITEA Journal of Test and Evaluation*, June/July 1995, vol. 16

(2)) for a discussion of pitfalls when translating requirement statements into performance measures for test plans.

[7]In 2007 DOTE chartered a JT&E to develop methodology to conduct and assess system performance within a system-of-system approach to accomplishing military missions in a Joint environment. This JT&E is called Joint Test and Evaluation Methodology (JTEM).

[8]Some examples are the Central Test and Evaluation Investment Program (CTEIP) Common Range Integrated Instrumentation System (CRIIS), Army's One-Tactical Engagement Simulation System (ONE-TESS), the Joint Mission Environment Test Capability (JMETC), and Test and Training ENablingArchitecture (TENA).

[9]Joint memorandum "Test and Training Interdependency Initiative" September 7, 2006 signed by the Under Secretary of Defense (Acquisition, Technology, and Logistics), the Under Secretary of Defense (Personnel and Readiness) and the Director, Operational Test and Evaluation provided a common vision for interdependent test and training solutions to achieve a single, more realistic operational environment. See the Test Resource Management Center *FY2007 Annual Report* (January 2008) for implications of this memorandum.

References

Kass, R. A. 1995. "Writing measures of performance to get the right data." *The ITEA Journal of Test and Evaluation*. June/July 1995, Volume 16-2, pp. 2–6.

Kass, R. A. 2006. *The Logic of Warfighting Experiments*. Washington, D.C.: Command and Control Research Program (CCRP) of the Assistant Secretary of Defense and Information Integration (ASD/NII), Available from the CCRP website at http://www.dodccrp.org.

Kass, R. A. 2008. "Framework for understanding experiments." *The ITEA Journal of Test and Evaluation*. Volume 29-2, pp. 167–174.

Random House. 1984. Random House College Dictionary New York, NY: Random House, Inc.

The ITEA White Sands Chapter introduces...

LIVE-VIRTUAL-CONSTRUCTIVE CONFERENCE

Formerly known as the ITEA M&S Conference

JANUARY 12 - 15, 2009

El Paso, Texas

*Our conference committee, chaired by **Mr. Thomas Berard**, White Sands Missile Range Executive Director, has decided to take the highly successful ITEA Modeling and Simulation Conference that has been held in the month of December in Las Cruces, New Mexico for the past thirteen years and move it to El Paso, Texas in January. This conference will still have the traditional components of the M&S conference with new topics for you to consider. We hope that you continue to support this event by providing outstanding abstracts, exhibiting your products and services, and contributing your sponsorship dollars for the scholarship program.*

CONFERENCE PLANNING COMMITTEE

PROGRAM CHAIR
Gilbert Harding, Director of Systems Engineering Directorate

TECHNICAL CO-CHAIRS
Doug Messer • 575-430-2951 or 1825 • dmesser99@aol.com
Hank Newton • 575-385-7270 • hnewton@aoc.nrao.edu

TOPICS

- Distributive Live-Virtual-Constructive and T&E
- High Performance Computing in Test and Experimentation
- Civil-Military Operations
- Collaborative Simulation & Testing
- Testing in a Networked Environment
- Requirements for M&S in Test & Evaluation
- M&S Tools
- M&S Technologies
- Autonomous and Cognitive Systems
- Netcentric Testing

CALL FOR PAPERS

All sessions will be at The Wyndham Hotel and will be unclassified. Abstracts will be reviewed for presentation during a conference session or as a poster paper. Papers selected for presentation will be published in proceedings on the World Wide Web. Visit **www.itea-wsmr.org** for previous Modeling & Simulation proceedings.

EXHIBITS

Your company or government organization will want to take advantage of the premium and yet limited space that is available for you to display and demonstrate products and services for the test and evaluation community. To obtain an application to exhibit or to see the floor plan, visit the ITEA website.

SPONSORSHIP

Fours levels of sponsorship are available for your company to participate in: Platinum $2500, Gold $1000, Silver $500 and Bronze $250. Your sponsorship dollars will defray the cost of this event and support the ITEA scholarship fund, which assists deserving students in their pursuit of academic disciplines related to the test and evaluation profession. For more information on the benefits of sponsorship, or to obtain a pledge form, please visit www.itea.org.

For the latest information, including an updated agenda, applications to exhibit and sponsor, the floor plan, lodging information, register on line, and much more visit **www.itea.org**

ITEA Journal 2008; 29: 301–309
Copyright © 2008 by the International Test and Evaluation Association

Multipath RDMA Instrumentation

Brett McMillian, Dennis Ferguson, and Gary McMillian, Ph.D.
Crossfield Technology, Austin, Texas
Donald R. Snyder, III
Air Force Research Laboratory, Eglin AFB, Florida

Multipath remote direct memory access (RDMA) is a new networking technology that provides an order-of-magnitude increase in bandwidth through a full-mesh backplane and through a standard Ethernet switch fabric. This article first describes full-mesh backplane technology, RDMA access over internet protocol, and Multipath RDMA, then introduces two novel Multipath RDMA instrumentation systems: an instrumentation module and an instrumentation blade. The instrumentation module supports Giga-sample per second (GSPS) instrumentation and operates at the end of the wire using Power-over-Ethernet Plus. The instrumentation blade goes directly into a Multipath RDMA-enabled blade chassis, supports multi-GSPS instrumentation, and gives rise to real-time or near–real-time parallel processing of data produced by multi-GSPS instrumentation.

Key words: Data acquisition; Ethernet; instrumentation; hardware-in-the-loop; Multipath RDMA; power over Ethernet; remote direct memory access (RDMA).

Novel networking, instrumentation, and computation hardware is introduced for multiple Giga-sample per second (GSPS) data acquisition, hardware-in-the loop simulation, signal processing, and spectrum analysis. The system builds on 10 Gbits/s remote direct memory access (RDMA) over internet protocol (IP) and uses full-mesh backplanes to create a new networking technology called Multipath RDMA. A Multipath RDMA full-mesh bridge chip provides over an order-of-magnitude increase in bandwidth between computing nodes and supports real-time or near–real-time parallel processing in multi-GSPS instrumentation systems.

Full-mesh networking

A full-mesh bridge chip enables Multipath RDMA through a full-mesh passive backplane. The backplane incorporates a full-mesh network topology into a passive backplane, or midplane, which provides every processor blade a direct connection to every other processor blade. *Figure 1a* illustrates this full-mesh topology while *Figure 1b* shows how the same blades are also mapped to a dual-redundant star topology, the network topology used in a typical Ethernet switch fabric. Each switch provides a cut through switch connection between any two blades in the system. The full-mesh fabric on the other hand allows for a blade to have as many connections as there are other blades in the backplane. Hence, the full-mesh fabric provides much higher aggregate bandwidth into and out of any one node.

Both the AdvancedTCA (PICMG® 2003) and VITA 46/48 standards (VITA 2008) currently define full-mesh backplanes. These backplanes support up to 16 slots with 15 serial communications channels per slot. Each channel supports full-duplex 10G Ethernet with a 10 Gigabit physical layer comprised of eight differential pairs, each operating at a 3.125 GHz signaling rate. Hence each slot has a maximum 150 Gbits/s of bandwidth in both directions through the backplane. Unfortunately, little exists on the market today that can use this bandwidth outside of the backplane. Current switch chips support the full bandwidth through the backplane but bottleneck down to one or two 10G Ethernet connections to the processors on the blades. A full-mesh bridge chip with a native host processor interface and Multipath RDMA capability is designed to alleviate this bottleneck and provide over an order-of-magnitude greater bandwidth between processors.

The full-mesh bridge chip, shown in *Figure 2*, provides a bridge between a processor's high-speed host interface and a full-mesh backplane interface. The host interface can take the form of any high-speed,

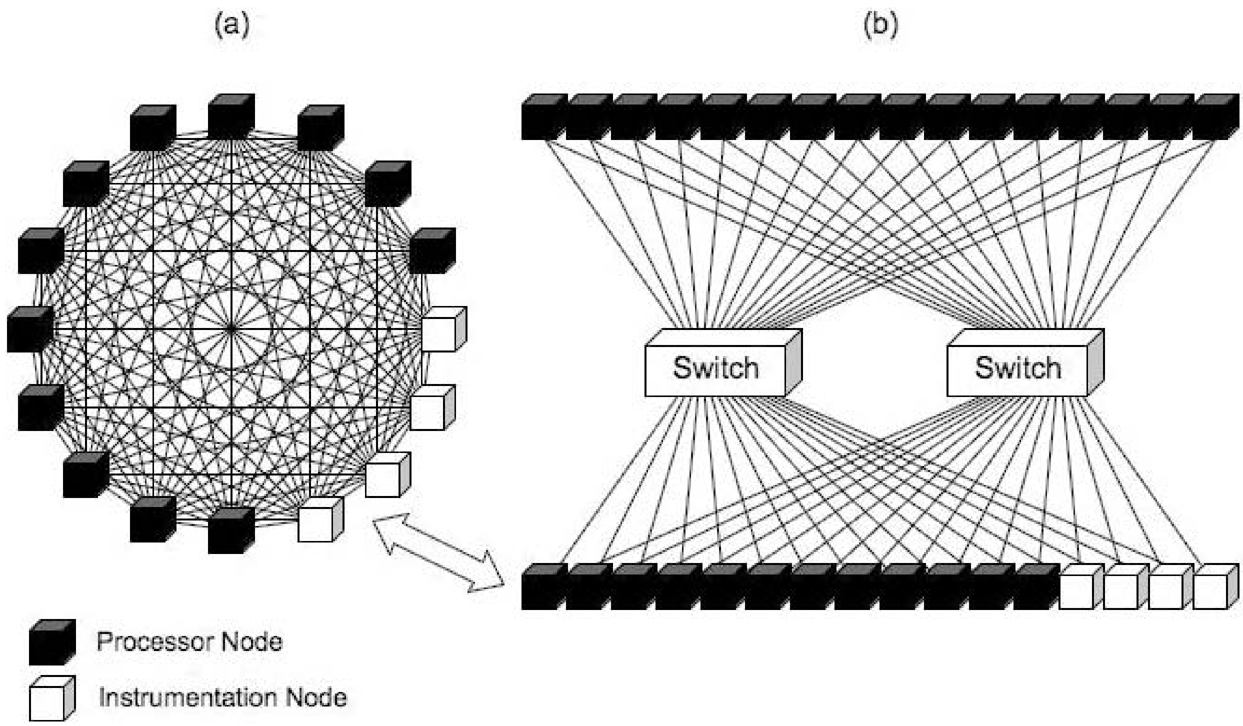


Figure 1. (a) Full-mesh network topology, (b) dual-redundant star network topology

chip-to-chip interconnect technology like Advanced Micro Device's (AMD's) HyperTransport™ (HT) (HyperTransport Consortium 2001), Intel's Quick-Path Interconnect (QPI) (Intel 2008), and Rambus' FlexIO™ (Rambus 2008), which is currently used in the IBM, Sony, and Toshiba Cell Broadband Engine (Cell), or other future interconnect standard. Dual host interfaces support dual-processor blades, with the full-mesh bridge providing a peak bandwidth of around 20 GBytes/s (160 Gbits/s) to each processor. The backplane interface supports the maximum fifteen channels, which means that two processors connected via host interfaces will be able to share the full 150 Gbits/s through the backplane.

Remote direct memory access over IP

In addition to the fifteen serial channels to the backplane, the full-mesh bridge chip provides two interfaces for use as external 10G Ethernet ports into a dual redundant switch fabric as shown in *Figure 3*. In other words, the full-mesh bridge chip has the equivalent of 17 10G serial connections: two external Ethernet ports and 15 backplane channels. Each processor can use these 17 10G connections to transfer data directly from its memory to another processor's memory on any other blade in the system via RDMA.

RDMA over IP (RDMA Consortium 2003) differs from traditional Ethernet transmission in that RDMA eliminates unnecessary buffering in the operating system when transmitting and receiving packets. Instead of copying packets into a buffer in the operating system before sending, an RDMA Network Interface Controller (RNIC) takes data directly from application user-space memory, applies the appropriate network layer protocols and Ethernet link layer frame and sends the packet across the network. On the receiving end, another RNIC receives the packet and places the payload directly into application user-space memory. By removing the unnecessary data copying in the kernel and off-loading network protocol processing at both ends of the link, RNICs alleviate the latency issues normally associated with Ethernet and make Ethernet a viable solution for high-speed, low-latency instrumentation systems.

To get into more detail, the RDMA Consortium (RDMA Consortium 2003) defined a suite of protocols at the transport layer that enables cooperating Direct Memory Access (DMA) engines at each end of a network connection to move data between user-space memory buffers with minimal support from the kernel and with "zero copy" to intermediate buffers. The Remote Direct Memory Access Protocol Verbs specification (Recio 2007) describes the behavior of the protocol off-load hardware and software, defines the semantics of the RDMA services and how they appear to the host software, and defines both the user and kernel application programming interface. The Verbs specification defines RDMA *READ*, *WRITE*,

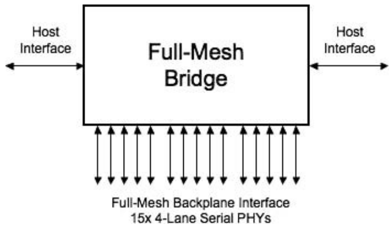


Figure 2. Full-mesh bridge chip

and *SEND* operations that transport data between user space memories or into a receive queue, respectively, and defines send and receive queues and queue pairs to control data transport and completion queues to signal when an operation is complete. Work requests are converted into work queue elements which are processed in turn by the off-load engine. An asynchronous event (interrupt) signals when work is complete. Also, data need not be in contiguous locations in either the source or destination memories as scatter-gather lists define the physical memory locations of data segments.

The full-mesh bridge includes built-in support for RDMA over IP, and the processors control the device through their host interface. However, since the RNIC accesses processor main memory directly, the RNIC must have a high bandwidth path to main memory in order to not bottleneck. XDR and DDR3 are the only memory interfaces that currently match the 20 GBps bandwidth of advanced host interfaces. The Cell Broadband Engine, which uses XDR memory, is the only processor that currently provides both a host and memory interface at this bandwidth, but Intel and AMD should reach this level of performance in the near term.

Multipath RDMA

While full-mesh connectivity provides high aggregate bandwidth between multiple processors distributed across the backplane, the bandwidth between any two processors is the same as if the backplane were a simple switch because only one direct connection exists between any two slots in the backplane. The same is true for connections between processors connected through an external dual-redundant Ethernet switch fabric. Processors connected across the dual-redundant switch fabric have at most 20 Gbits/s of bandwidth for RDMA data transfer (using both paths through the redundant fabrics). However, a new technology called Multipath RDMA, incorporated into the bridge chip, allows two individual processors located anywhere in the network to use the full 150 Gbits/s bandwidth through the backplane plus the 20 Gbits/s bandwidth through the external switch fabric for RDMA data transfer via multiple parallel paths.

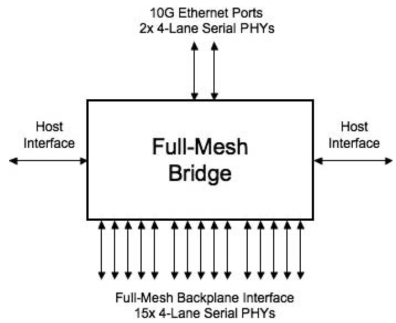


Figure 3. Full-mesh bridge chip with Ethernet ports

Multipath RDMA gets its name from its similarity to multipath in wireless communication. In wireless communications, multipath refers to the multiple paths traveled by electromagnetic waves between antennas due to reflections in the environment. New Multi-Input, Multi-Output (MIMO) technologies like those incorporated into IEEE 802.11n (IEEE 2008) exploit multipath transmission to increase the overall data rate. Similarly, Multipath RDMA uses every possible route through both the full-mesh backplane and the dual-redundant switch fabric to route packets between any two processors. The most significant improvement Multipath RDMA offers over current networking technology is that it provides over an order-of-magnitude increase in bandwidth between processors in the same backplane and between blades connected through a dual-redundant switch fabric.

First, when processor blades share a common full-mesh backplane, they have 17 available channels to transmit data to any other processor. Any two processors in the network have one direct channel through the full-mesh backplane plus two switched connections. However, with Multipath RDMA the two processors also gain 14 one-hop connections through the remaining full-mesh channels via the other 14 blades' full-mesh bridge chips. The full-mesh bridge chips act as cut-through switches between the sending and receiving full-mesh bridges on the source and destination blades. *Figure 4* shows an example of Multipath RDMA between nodes in the same backplane.

Second, when two blade chassis are connected through a dual-redundant switch fabric, Multipath RDMA exploits the two full-mesh backplanes to increase the bandwidth through the switch fabric. On the sending side, the full-mesh bridge chip segments the payload and transmits the resulting packets through all 17 available channels. The two switch fabric ports send their packets to the Ethernet ports on the receive side in the standard way. The 15 connections through the backplane interface, however, route through the other 15 full-mesh bridge chips in

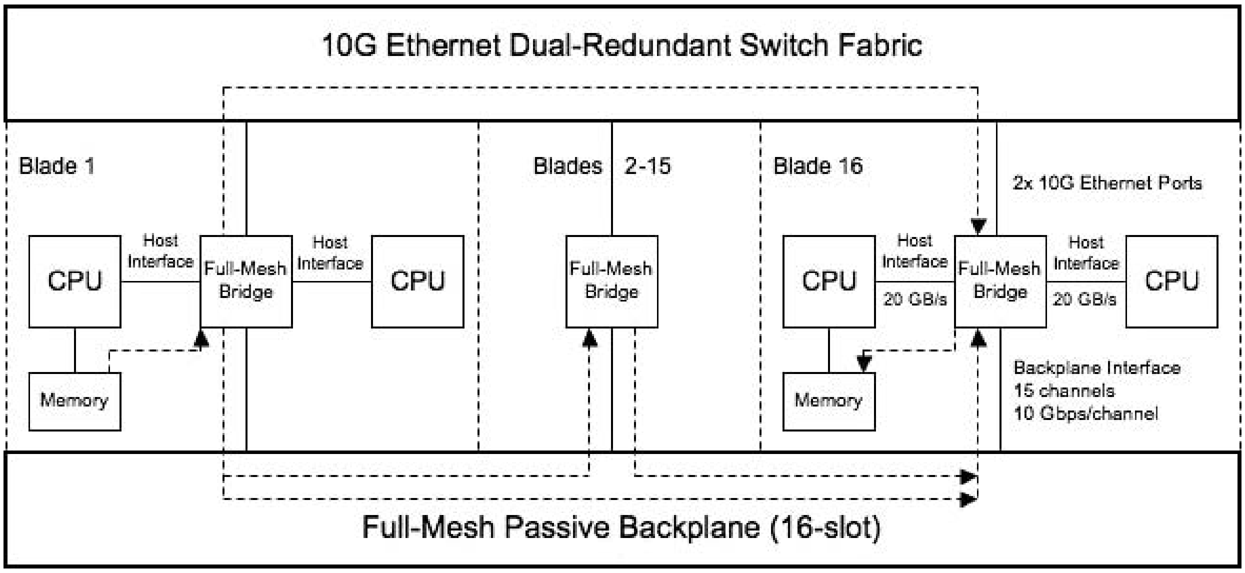


Figure 4. Multipath RDMA through a full-mesh backplane

the backplane, which act as cut-through switches, to the switch fabric ports on each blade. Then the packets are sent across the switch fabric simultaneously to the 15 blades connected to the receiving blade through the other backplane. The full-mesh bridge chips on those blades then act as cut-through switches to the channels connected to the receiving blade's full-mesh bridge chip. The full-mesh bridge on the receiving blade then combines the payloads from each packet and places the data directly into the receiving processor's memory. Hence for a Multipath RDMA transmission through the dual-redundant switch fabric, a processor has two one-hop switched connections and 15 three-hop connections through full-mesh bridges. *Figure 5* shows an example of Multipath RDMA between nodes in separate full-mesh backplanes.

In both cases, the bandwidth between any two full-mesh bridge chips is a multiple of 10 Gbits/s up to the maximum 170 Gbits/s. The number of channels utilized by Multipath RDMA can change on the fly such that the number of communication paths between any two processors is dynamic. As a result, Multipath RDMA can support a large number of connections at dynamically changing bandwidths in order to optimize connections for a wide range of algorithms. When compared to current commercial off-the-shelf systems, Multipath RDMA provides more than 16 times the maximum bandwidth between processors connected through a full-mesh backplane and more than eight times the maximum bandwidth between processors connected through, and utilizing, a dual-redundant switch fabric. However, real-time or near–real-time systems can only take advantage of this bandwidth if the instruments are able to obtain the same amount of bandwidth into and out of the network fabric.

Multipath RDMA instrumentation

In order to achieve high bandwidth between instrumentation and a Multipath RDMA-enabled supercomputing cluster, a new instrumentation system had to be architected from the ground up that could incorporate the concept of using multiple connections to achieve higher bandwidth. First, the instrumentation systems must accommodate analog-to-digital converters (ADCs), digital-to-analog converters (DACs), or other sensors or actuators that source or sink high bandwidth data streams. With up to 170 Gbits/s bandwidth through a blade chassis, Multipath RDMA supports ADCs and DACs with multi-GSPS data rates. Second, the instrumentation systems must either have a single high bandwidth connection into the computing cluster or several lower bandwidth parallel connections.

To this end, designs for two different instrumentation systems developed. The first design is a 10/20 Gbits/s RDMA Power-over-Ethernet Plus (PoEP) Instrumentation Module that supports ADCs and DACs up to 2 GSPS (at 8 bits of resolution). The second design approach puts the ADCs and DACs directly onto the blades in the supercomputing cluster and supports sampling rates up to 20 GSPS (at 8 bits of resolution). The first approach has the advantage of placing the ADCs and DACs close to the sensors and actuators, which eliminates noise on the analog side and decreases the number of cables. The second

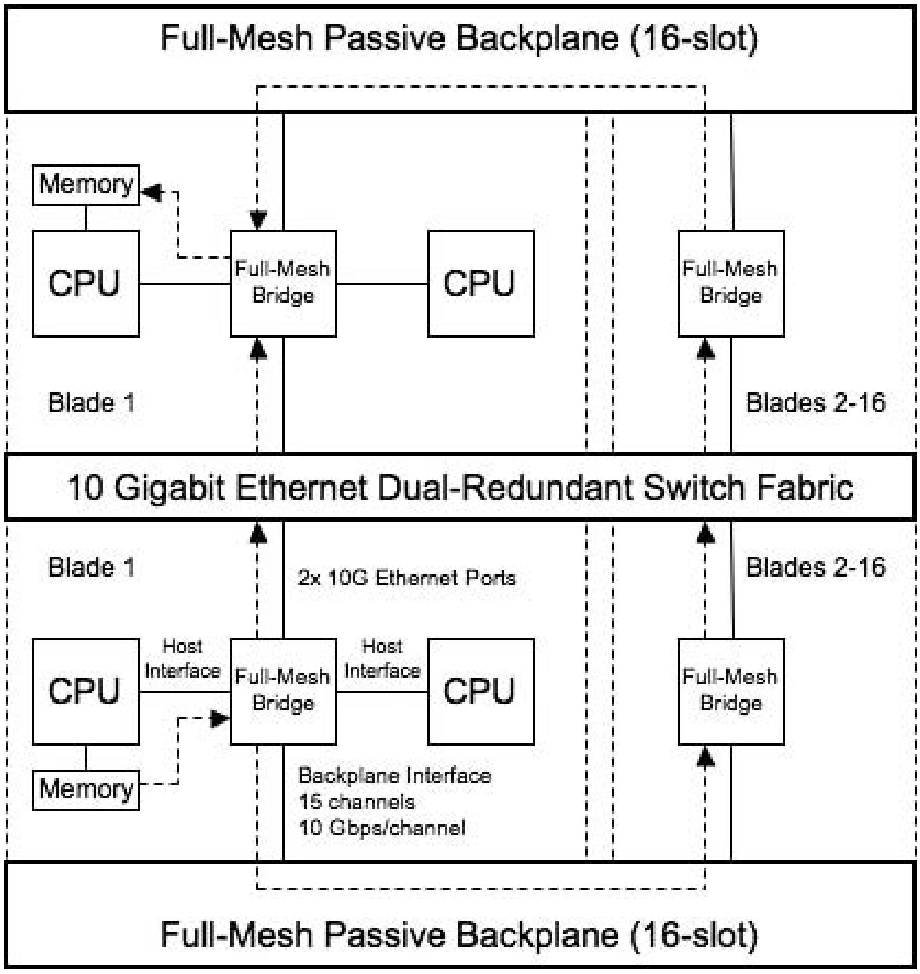


Figure 5. Multipath remote direct memory access (RDMA) across a dual-redundant switch fabric

approach has the advantage of supporting the fastest ADCs and DACs currently on the market and provides enough bandwidth into a full-mesh backplane to distribute full-rate sensor data across parallel signal processors using Multipath RDMA.

10/20 Gbits/s RDMA PoEP instrumentation modules

The 10/20 Gbits/s RDMA PoEP Instrumentation Module is a remote high-speed sensor data acquisition and actuator controller system that both transmits data through and gets its power from a single CAT6 or CAT7 Ethernet cable. The module consists of a motherboard and an XMC card, the first of which is responsible for handling the Ethernet communication, IEEE 1588 time synchronization, and main power supplies. The XMC card contains high-speed ADCs or DACs and analog instrumentation. With pluggable XMC cards, the module is extremely versatile and can be used in almost any data acquisition, signal processing, hardware-in-the-loop, or other type of instrumentation system.

XMC is the VITA 42 standard (VITA 2008) that extends the PCI Mezzanine Card (PMC) to several different high-speed buses including PCI Express (VITA 42.3) and HyperTransport (VITA 42.4). The XMC standard supports up to 16 data lanes per slot. With 8-lane PCI Express V2.0 the XMC slot provides up to 32 Gbits/s of bandwidth, which means that a plug-in card can support a single ADC sampling at 4 GSPS with 8-bits of resolution or a 2 GSPS DAC with 16-bits of resolution or multiple ADCs or DACs with lower sampling rates or resolutions. However, the module itself only supports up to 20 Gbits/s into a switch fabric, which means the XMC card is limited to a single ADC sampling at 2 GSPS with 8 bits of resolution or a 1-GSPS DAC with 16 bits of resolution.

XMC cards with various ADC and DAC configurations are currently available as commercial off-the-shelf boards. The real innovation in the instrumenta-

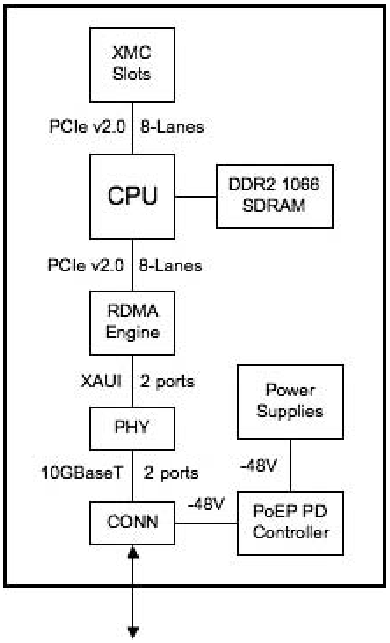


Figure 6. 10G PoEP motherboard

tion module is the motherboard. The motherboard supports up to two 10G Ethernet ports, IEEE 1588, and PoEP, and has an extremely small form factor about the same size as a double-wide XMC card (15 cm × 15 cm), and as a result, can be placed relatively close to the sensors or actuators, which reduces the lengths of the analog sensor and actuator wires. *Figure 6* shows a simplified block diagram of the motherboard.

A dual-port RDMA ASIC, like NetEffect's 10GbE ASIC (NetEffect Inc., Austin, Texas) for example, on the motherboard provides two 10 Gbits/s RDMA over IP Ethernet ports. With RDMA support, the instrumentation module can transfer data directly between its own memory and tens, hundreds, or even thousands of processor memories in a supercomputing cluster through an Ethernet switch fabric. While only two 10G Ethernet connections may not warrant using Multipath RDMA, it can be used on the server end to send data to multiple instrumentation modules from a single processor or read data from multiple instrumentation modules into a single processor.

IEEE 1588 Precision Time Protocol (PTP) (IEEE 2002) is a time synchronization protocol that allows the motherboard to synchronize a local clock with a master clock located somewhere in the network. Once the motherboard clock is synchronized with the IEEE 1588 PTP master clock, it can then be used as a reference for the ADC and DAC clocks on the XMC card. By using IEEE 1588 PTP time synchronization, every instrumentation and processing node in the network can have clocks synchronized to within 100 ns of each other or better.

PoEP is the IEEE 802.3at (IEEE 2005) extension to the IEEE 802.3af (IEEE 2003) Power-over-Ethernet standard that allows a powered device to use up to 70 W of power supplied by power sourcing equipment. The IEEE 802.3at draft standard defines a midspan power sourcing equipment device that connects in series to insert power into an Ethernet link. Since the motherboard accepts PoEP, the motherboard and XMC card can use up to 70 W of power, which is more than enough power for most applications. PoEP products are currently available for Gigabit Ethernet, but extending PoEP technology to 10G Ethernet is straightforward since 10GBase-T uses the same 4 pairs as 1000Base-T.

As mentioned above, the instrumentation module connects to the Multipath RDMA enabled cluster and receives its power via a PoEP midspan device or switch. With a midspan configuration and CAT7 cable, each instrumentation module can be located up to 100 m from a switch. Hence, the instrumentation modules can be located large distances from the processing cluster, which is often convenient when the cluster cannot be physically located nearby. *Figure 7* shows an illustration of the module network.

The only disadvantage of the instrumentation module approach is the limited bandwidth between the ADCs and DACs and the cluster. With only 20 Gbits/s maximum bandwidth into the switch fabric, the modules cannot support the sampling rate and resolution of the fastest ADCs and DACs currently available. For this reason, a second, faster instrumentation system approach was developed that has sufficient bandwidth to support state-of-the-art ADCs and DACs and can use Multipath RDMA to distribute data throughout the cluster.

Multipath RDMA instrumentation blades

In order to achieve the maximum bandwidth into a Multipath RDMA system, the ADCs and DACs must have a high-speed interface into a full-mesh backplane. A straightforward way to gain access to the full-mesh backplane is if the ADCs and DACs are on the blades themselves. The ADCs and DACs can be interfaced to the full-mesh bridge using an FPGA, which buffers data in memory and provide a host interface to the full-mesh bridge.

FPGAs can communicate with the fastest ADCs and DACs on the market, through a high speed interface. The latest FPGAs provide a DDR3 interface supporting up to several Gigabytes of memory, which, for example, can be used to buffer data from an ADC before it is distributed to processors in the cluster.

The ITEA Journal of Test and Evaluation jite-29-03-04.3d 18/8/08 15:54:07 306

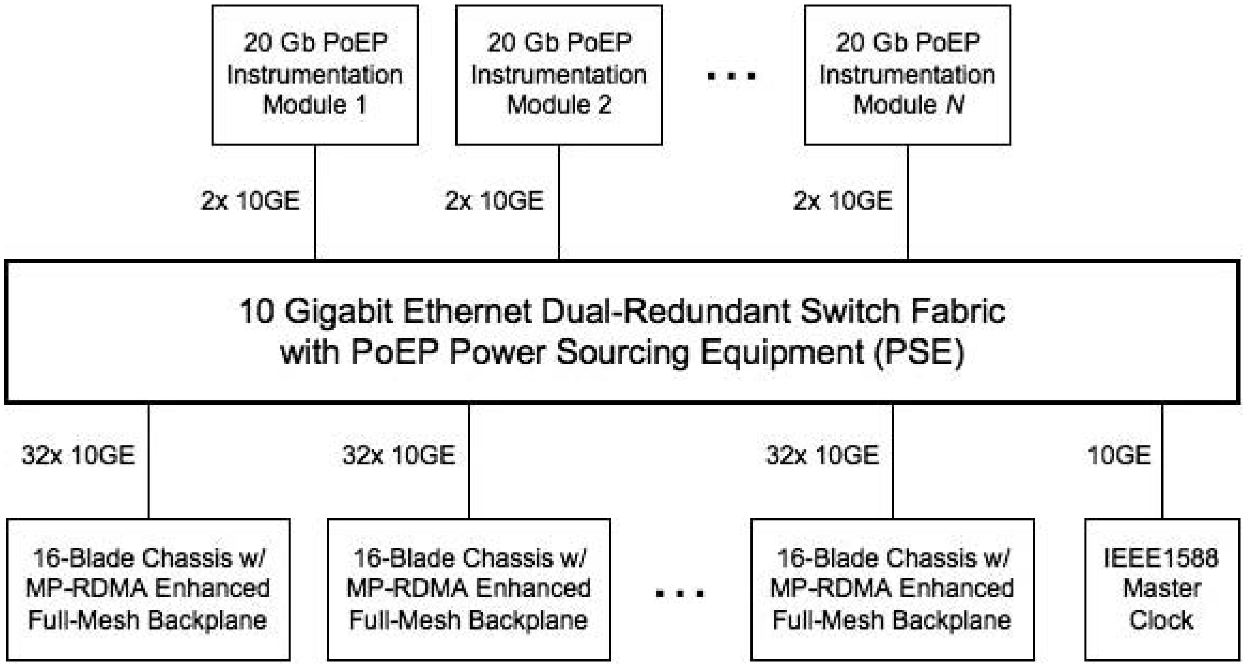


Figure 7. Instrumentation module network with the IEEE 1588 master clock

Finally, the FPGA's host interface bridges the instrumentation into the full-mesh bridge. The full-mesh bridge is then able to use Multipath RDMA to transfer data between the FPGA's memory and the memory of any processor in the cluster. Furthermore, with IEEE 1588 PTP support built into the full-mesh bridge, instrumentation blades and processing blades will all support clock time synchronization in the same fashion as the instrumentation modules. *Figure 8* shows a block diagram of the instrumentation blade architecture.

Application example

As a demonstration of how the Multipath RDMA instrumentation system works, the following section shows how the system could compute a billion-point complex fast Fourier transform (CFFT) of acquired sensor data in near–real-time using the fastest commercial off-the-shelf ADCs and multiple Cell Broadband Engines from IBM. First, assume two 8-bit 5-GSPS ADCs capture an arbitrary waveform starting at time $t = 0$ seconds on an instrumentation blade connected to a full-mesh backplane. The two ADCs sample the waveform in-phase and quadrature-phase (I & Q). A billion complex data points are sampled every 214.7 ms (a billion in this case is actually 2^{30}). This data rate corresponds to a bandwidth of 10 GBytes/s from the ADCs. Since the ADC interface and the DDR3 interface on the FPGA have more than 10 GBytes/s bandwidth, the data points are stored in the FPGA memory as fast as they are sampled.

As the data is stored in memory, the FPGA simultaneously tasks the full-mesh bridge to transport the data to a Cell processor somewhere in the cluster using Multipath RDMA. Since the host interface and the network bandwidth to a Cell processor through the full-mesh bridge using Multipath RDMA are greater than the sampling rate of the ADC, the data is streamed into a Cell processor's memory as fast as the ADCs sample the waveform.

The IBM Cell Broadband Engine is a multicore heterogeneous processor with one PowerPC Processor Element (PPE) and eight Synergistic Processor Elements (SPEs). The most recent update to the Cell, the PowerXCell 8i (IBM, Los Alamos, New Mexico), used in the Los Alamos Roadrunner QS22 blade cluster (Koch 2007), has a high speed FlexIO interface that supports up to 20 GBytes/s I/O bandwidth and a DDR2 memory controller that supports up to 64 Gigabytes of memory. This latest update also boasts 230 GFLOPS single precision floating-point performance and over 100 GFLOPS double-precision floating point performance.

In order to perform a near–real-time billion-point CFFT on a Cell, the processor must have both enough memory to support the large data set and a significant amount of computational power to perform the calculations in the required time. A billion-point single-precision floating-point CFFT requires 8 Bytes

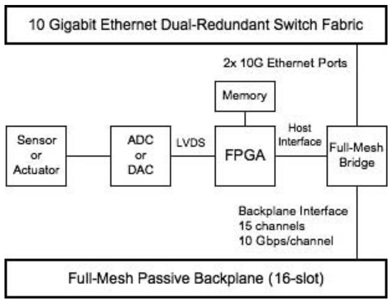


Figure 8. Multipath remote direct memory access (RDMA) instrumentation blade

of storage per complex data point and thus needs 8 GB of memory for storage. While the first-generation Cell's XDR interface did not support this much memory, the PowerXCell 8i's DDR2 controller interface can easily meet this requirement.

The Cell provides unmatched computational power in a single chip. With 230 GFLOPS single-precision performance, the Cell can perform a 16M-point CFFT in 0.043 seconds (Chow, Fossum, and Brokenshire 2005). The time required to compute a 16M-point CFFT on a Cell processor can be scaled to estimate the time required to calculate a billion-point CFFT. The first part of the calculation involves determining the relative complexity factor of a billion-point CFFT to a 16M-point CFFT. The complexity of a CFFT is determined by $N \log_2 (N)$, where N is the number of points. Hence the relative complexity of a billion-point CFFT when compared to a 16M-point CFFT is $[1B \log_2(1B)]/[16M \log_2(16M)]$ or 80. Assuming the billion-point CFFT is as equally parallelizable as the 16M-point CFFT the time a Cell takes to calculate a billion-point CFFT is 80* 0.043 seconds or 3.44 seconds.

Since the ADC can sample one billion complex data points in 214.7 ms, computing the billion-point CFFT in real time is not currently possible on a single Cell. However, by distributing the computational load across multiple Cell processors, the CFFT can be performed in real-time or near-real-time. A Multipath RDMA Instrumentation system with one instrumentation blade and nine Cell blades containing 18 Cells is able to compute a billion-point CFFT on a 5 GSPS complex signal in real-time.

Conclusion

In conclusion, Multipath RDMA is a new networking technology that exploits multiple routes through a fabric to achieve an order-of-magnitude increase in bandwidth between processors in instrumentation systems and supercomputing clusters. It also provides a novel means to interface ultra-high-speed instrumentation to a computing cluster and makes real-time or near–real-time data acquisition and parallel processing possible. Multipath RDMA is a quantum leap in the battle for more bandwidth. While several organizations are currently working on 100G Ethernet standards, Multipath RDMA offers even greater bandwidth using current 10G Ethernet technology. As Ethernet speeds increase, Multipath RDMA will maintain this advantage. Multipath RDMA is an innovative way to get to the next level of bandwidth capacity by exploiting an underutilized resource in a blade system, the passive backplane. ❑

BRETT MCMILLIAN is an engineer at Crossfield Technology in Austin, Texas. He has worked on several programs in the areas of circuit design, instrumentation, high-speed networking, Linux driver development, analog and mixed-signal design, and wireless systems. He codeveloped the concept of Multipath RDMA and also recently developed a PoEP instrumentation module similar to the module described in this article for strain gauge, pressure, and temperature sensors as part of a program with the Arnold Engineering Development Center. His areas of expertise include printed circuit board design and computer networking. He holds a B.A. in physics from Grinnell College, Grinnell, Iowa. E-mail: brett.mcmillian@crossfieldtech.com

DONALD R. SNYDER is a senior scientist with the Air Force Research Laboratory Munitions Directorate and is the lead scientist for the Advanced Guidance Division Kinetic Kill Vehicle Hardware in the Loop Facility (KHILS). He currently directs a research portfolio of over 30 research efforts, including hardware-in-the-loop test technologies. He is a fellow of SPIE and a member of IEEE/Laser and Electro Optics Society, Society for Information Display, NDIA, National Training System Association, and SPIE High Speed Instrumentation Working Group. Snyder is the recipient of the ITEA Modeling and Simulation Award for MDA Modeling and Simulation Scene Generation. E-mail: snyder@eglin.af.mil

DENNIS FERGUSON has held technology management and engineering positions at several companies over his 32-year career. At Crossfield Technology, Ferguson has worked on software radios and harsh environment wireless instruments. At Honeywell, he led many technology developments involving integrated circuits and software radios for commercial aviation and communications. Ferguson holds a BSEE and an MBA. He has been a member of IEEE for 36 years and is currently vice chair of the Central Texas Solid State Circuits society. Ferguson has been awarded six patents and has published or presented over 37 articles. E-mail: dennis.ferguson@crossfieldtech.com

GARY MCMILLIAN is a scientist and partner at Crossfield Technology in Austin, Texas. He has served as principal *investigator for numerous R&D programs in the areas of instrumentation, high-speed networking and signal processing. His technical areas of expertise include sensors, instrumentation, systems engineering, signal processing, wireless systems, and network data communications. He is a member of the IEEE and previously served as the secretary of the Central Texas Section of the IEEE Communications/Signal Processing Society. He holds a B.S. in physics from Sam Houston State University, Huntsville, Texas and a M. A. and Ph.D. in physics from Rice University, Houston, Texas. E-mail: gary.mcmillian@crossfieldtech.com*

References

Chow, A. C., Fossum, G. C. and Brokenshire, D. A. 2005. "A Programming Example: Large FFT on the Cell Broadband Engine." White Paper available at http://www-01.ibm.com/chips/techlib/techlib.nsf/techdocs/0AA2394A505EF0FB8725.

HyperTransport™ Consortium. 2001. Hypertransport Link Specification. Sunnyvale, California: Hypertransport Consortium. Available at http://www.hypertransport.org. Accessed June 6, 2008.

IEEE (Institute of Electrical and Electronics Engineers). 2002. Standard for Precise Time Synchronization as the Basis for Real Time Applications in Automation. Piscataway, New Jersey, IEEE available at http://ieee1588.nist.gov/ or http://www.ieee1588.com. Accessed June 6, 2008.

IEEE (Institute of Electrical and Electronics Engineers). 2003. IEEE Std. 802.3af—IEEE Standard for Information technology-Telecommunications and information exchange between systems—Local and metropolitan area networks. Specific requirements. IEEE Computer Society.

IEEE (Institute of Electrical and Electronics Engineers). 2005. IEEE P802.3at DTE Power Enhancements Task Force authorization. http://www.ieee802.org/3/at/

IEEE (Institute of Electrical and Electronics Engineers). 2008. Wireless Networking Standard 802.11n, proposed amendment n, Multi-Input, Multi-Output (MIMO) technologies. New York: IEEE.

Intel. 2008. White Paper Intel QuickPath Architecture. Santa Clara, California: Intel. Available at http://www.intel.com/pressroom/archive/reference/whitepaper_QuickPath.pdf Accessed June 6, 2008.

Koch, K. 2007. Roadrunner System Overview. Powerpoint briefing. Available at http://www.lanl.gov/orgs/hpc/roadrunner/rrinfo/RR%20webPDFs/Roadrunner%20Syste m%20Overview%20Oct%2017,%202007%20(LAUR).pdf. October 17, 2007.

PICMG® (PCI Industrial Computers Manufacturers Group). 2003. Advanced TCA Specifications for Next Generation Telecommunications Equipment. Wakefield, Massachusetts: PICMG. Available at http://www.picmg.org/v2internal/newinitiative.htm. Accessed June 6, 2008.

Rambus. 2008. FlexIO™ Processor Product Brief. Available at http://www.rambus.com/assets/documents/products/FlexIO_ProcessorBus_ProductBrief.pdf. Accessed June 6, 2008.

RDMA Consortium. 2008. Architectural Specifications for RDMA over TCP/IP. available at http://www.rdmaconsortium.org.

Recio, R. 2007. Remote Direct Memory Access Protocol Verbs specification. RDMA Consortium. available at http://www.rdmaconsortium.org.

VITA. 2008. VPX and REDI Specifications. Fountain Hills, Arizona: VITA. Available at http://www.vita.com. Accessed June 6, 2008.

Acknowledgments

This research was sponsored by the U.S. Department of Defense Missile Defense Agency under U.S. Army contract number W9113M-06-C-0163.

Chapter News

The Roadrunner Chapter Albuquerque, NM

June Luncheon Highlights: The Roadrunner Chapter held a luncheon meeting on June 10, 2008, featuring guest speaker, Brenda S. Lee, Senior Analyst with Parallax, Inc., an *Energy*Solutions company. Photo 1. In Albuquerque, Parallax provides technical writing, requirements analysis, training, and system test support for the Transportation Safety and Surety Program (TSSP) at Sandia National Laboratories (SNL). Ms. Lee provided a presentation entitled In troduction to Session-Based Exploratory Testing.

Photo Caption: Patrick Dunn, President of the ITEA Roadrunner Chapter, presents Ms. Brenda S. Lee a token of appreciation during the June 10, 2008 chapter luncheon meeting.

Session-Based Exploratory Testing is a technique that allows for simultaneous test design and test execution. It is typically used when requirements have not been documented or the available time to test is short. Key elements of Session-Based Exploratory Testing include charters, sessions, session notes, and debriefings. Charters are predetermined areas of functionality or interest to test. Sessions are dedicated times to execute the charters and session notes are the charters, metrics captured, notes, test cases, resources used, and test artifacts collected during the session. Debriefings are meetings held between the test manager and the tester(s) to discuss the test activities and identify potential new charters to execute.

Ms. Lee described a scenario at the SNL that initially led her group to adopt this test technique: a project with signed system-level requirements, a highly dynamic development environment, a non-negotiable delivery date, and a two-week window to conduct the system test. She described the risks involved, and how her group was able to modify the technique to include associated requirements and sample test cases as part of the charters. The test results were seven business days to execute the test, seventeen charters executed, over 200 test cases executed, forty-one out of forty-five expected requirements validated, and fifteen defects identified.

As part of her summary, Ms. Lee provided a comparison between Session-Based Exploratory Testing and more traditional testing techniques. She recommends that any group planning to implement this test technique should receive formal training. Software Quality Engineering (SQE) offers two courses that cover the method; Just-In-Time Testing and Exploratory Testing. These courses are offered in various cities throughout the year. Ms. Lee s presentation was followed by a robust question and answer period.

February Luncheon Meeting Highlight:

←

Maj. Gen. Stephen T. Sargeant, Commander, Air Force Operational Test and Evaluation Center addressing the February 4, 2008 Roadrunner Chapter Luncheon Meeting.

Emerald Coast Chapter Eglin Air Force Base, Florida

The Emerald Coast Chapter is proud to announce the 2008 1st and 2nd Quarter, "ITEA Tester of the Quarter Award."

1st Quarter ITEA Tester of the Quarter Award

This quarter's award goes to Mr. Jonathan Reeves and Mr. Terren Niedrauer from the 782nd Test Squadron at Eglin Air Force Base, Florida.

Jon is an electronics engineer and Terren is an optics engineer working at the Guided Weapons Evaluation Facility (GWEF) at Eglin. Responding to a quick reaction requirement for the Navy CH-53 Infrared Countermeasures (IRCM) system, these men designed a $328K laser radiometer calibration capability within the facility with new control software code and a new automated log feature. Within a few days of the notice, a fully functional radiometer calibration station was operational completing 60 radiometer calibrations in four days (doubling the throughput). Most importantly, they provided very rapid, high quantity, high quality customer support to the OSD Center for Countermeasures, Joint Mobile IRCM Test System (JMITS) program.

The 782nd Test Squadron and the Air Armament Center at Eglin AFB are proud to demonstrate the exceptional talents of Jonathan Reeves and Terren Niedrauer.

■ ■ ■

The 1st Quarter 2008 ITEA Testers of the Quarter Award went to two people, Mr. Jonathan Reeves and Mr. Terren Niedrauer, both from the 782nd Test Squadron at Eglin Air Force Base , Florida.

2nd Quarter ITEA Tester of the Quarter Award

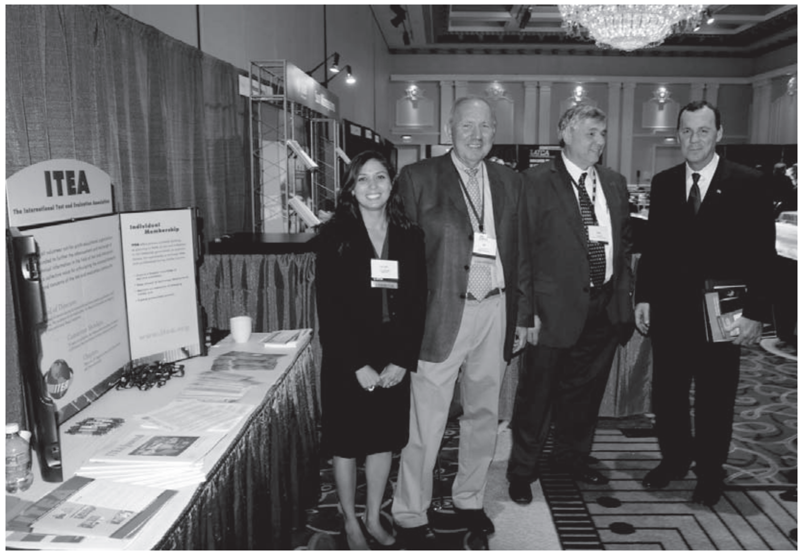
This quarter's award goes to Maj. Matt Hayden, Maj. Kyle Kolsti, Maj. Mark Massaro, Capt. Catherine Porcher, 1st Lt. Sidney Usry, 2d Lt. Ryan Moss, Mr. John Akers, Mr. Al Berard, Mr. Steve Chandler, Mr. James Dubben, Mr. Frank Gallegos, Mr. Will Graf, Mr. Charlie Holman, Mr. Steve Musteric, Ms. Julie Saladin Mr. Dennis Manning, and Mr. Chris Stewart representing military, civilian, and contractor expertise from the entire spectrum of DT&E at the 46th Test Wing at Eglin Air Force Base, Florida. This group of engineers established the first wireless instrumentation network to be implemented on an Air Force flight test aircraft. Advanced Subminiature Telemetry (ASMT) is a concept demonstration to reduce the impact of aircraft instrumentation. The first test used "stick-on" external, battery-powered sensors to gather data on an F-16 flutter flight test program. ASMT used these wireless accelerometers to send data to an instrumentation pod wirelessly via a Bluetooth (TM) frequency. These low-powered Bluetooth (TM) transmitters eliminated the need for a hard-wired data bus for these sensors and elimination of instrumentation bundles in the airframe. Under the present system initial modification of a flutter-instrumented aircraft can take 6 months and $600,000. This decreases aircraft availability which introduces lengthy delays in fulfilling test objectives and program milestones. Bottom line, this translates into unsatisfactory delays in delivering critical combat capabilities to GWOT warfighter. The team members driven by a common vision cut across traditional organizational lines capitalizing on the specialized skills of members and an open climate to encourage innovation. The 46th Test Wing and the Air Armament Center at Eglin AFB are proud to demonstrate the exceptional talents of ASMT test team.

■ ■ ■

U.S. Air Force photo by SSgt. Kevin Davidson

South New Jersey Chapter

South New Jersey Chapter held a contest to find a new logo for their chapter. Kristin Snyder, Hitec Systems has won one year ITEA membership for submitting the winning logo. Congrats to Kristen for a job well done.



FAA Administrator Visits ITEA Booth at the Air Traffic Control Association (ATCA) Technical Symposium ↑

Photo caption: FAA Administrator Bobby Sturgell visits the International Test and Evaluation Association (ITEA) Booth at the recent ATCA Technical Symposium in Atlantic City and chats with Don Marple, Chapter President, Al Mancini, Treasurer and Ly-Lan McCarthy, ITEA member.

■ ■ ■

Mark Your Calendar

TEST INSTRUMENTATION WORKSHOP

May 12-14, 2009
Ridgecrest, California

Check the ITEA website for more details, including the Call for Papers.

For additional information, contact:
Mr. Kurt Rockwell
kurt.rockwell@navy.mil

This event is hosted by the ITEA China Lake and Antelope Valley Chapters

www.itea.org

MARK YOUR CALENDARS

TEST WEEK 2009

Changing our T&E Processes to be More Responsive to the Needs of Our Customer:
Right Size, Right Price, Right Stuff, Best Value!!

June 1-4, 2009
Huntsville, Alabama • Von Braun Center

Exhibit Floor Plan is Open!
Contact Jeanette Wilson
jwilson@trideum.com
256-704-6121

Call for Papers coming soon!
Check www.testweek.org for all the details

Tutorials hosted by ITEA.
Contact Eileen Redd at eredd@itea.org, if you have a subject matter for instruction to the T&E community

For additional information, contact:
HOST
Mike McFalls
mmcfalls@trideum.com
256-704-6148

EVENT COORDINATOR
Sherry Hilley
sherry.hilley@us.army.mil
256-842-6715

This event is sponsored by OUSD (AT&L), Test Resource Management Center

South New Jersey Chapter
Continued

Advanced Applications of Metrics

On July 22, 2008, Mike Paglione of the Simulation and Analysis Group at the William J. Hughes Technical Center presented "Advanced Application of Metrics to Test and Evaluate New Systems for the National Airspace System".

The presentation discussed a proposed process to apply metrics studies once the key performance parameters (KPP) are identified among the requirements to help ensure test sufficiency, as well as used as presented several examples of the test metrics applied to the En Route Automation Modernization (ERAM) Testing Program.

Paglione discussed the testing of a complex computer system and emphasized that testing is an iterative process that is an integral part of all phases of software development,
from initial design through deployment.

Defining test sufficiency for complex computer systems can be a daunting task. If test and evaluation methods and measurements are not adequately scoped, testing can never be considered to be complete.

The V&V Protocol of Operations Project has identified a need to define testing metrics at the requirements phase of development.

After the presentation, Don Marple South New Jersey President presented a Certificate of Appreciation to Mike Paglione for his outstanding presentation.

Government News

WRC 2007 REPORT Big Gains for Flight Test Users

Tim Chalfant, Chief
Airborne Instrumentation,
Air Force Flight Test Center
Edwards AFB, California

As part of the Electromagnetic Spectrum Encroachment issue the USA Department of Defense (DoD) sought additional spectrum to augment our existing flight test telemetry frequency bands. We lost significant spectrum to commercial reallocation in 1993, 1997, and most recently in 2005. Future commercial encroachment would continue and future flight test would be impacted. After a long campaign, with both industry and government participation (including significant Air Force (AF) involvement), we have secured access to additional spectrum to augment and protect flight test.

As a result of the recent adoption of resolutions associated with Agenda Item 1.5, the 2007 World Radiocommunications Conference (WRC) granted Internationally- and Regionally-protected spectrum for flight test telemetry. In the USA this includes over 1300 MHz of new spectrum for wideband flight test telemetry. We must still negotiate and secure our rights to this spectrum in the USA, upgrade our ranges and test vehicles, and protect these new allocations.

Background: Flight Test Telemetry Encroachment The Flight Test Telemetry community lost 35% of its manned vehicle telemetry spectrum in 1993 due to U. S. A. Congressional legislation. The band 2310-2360 MHz was transferred from the DoD to the FCC to auction in support of Digital Audio Broadcasting Services (Sirus, XM Radio). This loss negatively impacted military and commercial flight test facilities that had utilized this band. In response, several DoD organizations engaged to protect our flight test spectrum. The Office of the Secretary of Defense (OSD) quickly formed the Research, Development, Test and Evaluation (RDT&E) Spectrum Requirement Working Group which brought together the National Aeronautics and Space Administration (NASA), DoD, and industry. The Telemetry, Frequency Management, and Sustainability Groups of the Range Commanders Council (RCC) quickly responded with tasks and studies to address the encroachment issue. The prime objectives were to 1) defend our electromagnetic spectrum against further encroachment, 2) increase our efficient use of the existing bands, and 3) pursue additional bands to augment existing allocations.

Plan of Action: The DoD successfully led the way on studies and efforts to articulate the impacts of spectrum loss to Test & Evaluation (T&E) via the RCC and improve efficiencies

Figure 1. WRC Plenary Session- 192 countries, each with a single vote, involved in making decisions

through Central Test and Evaluation Investment Program ((CTEIP) research and development programs). Today we are far better equipped to identify the impacts of spectrum encroachment and better defend our spectrum. We are also able to increase the efficiency of our use by many times due to CTEIP efforts like advanced modulation and the Integrated Frequency De-conflicting System developed under the Advanced Range Telemetry project.

We needed to secure access to additional spectrum for flight test and determine the means to protect it. An effort was started to address this issue via the international forum: The UN-chartered International Telecommunication Union (ITU). Success in this forum would be protected by international treaty.

The USA and France led the charge to address this need in the ITU sponsored WRC (figure 1), which meets every three or four years and includes a long preparatory process that typically requires multiple regional conferences to make a decision. It also includes the involvement of 189 administrations spread across the three ITU Regions. To make this work we would need to get support across the globe.

With USA DoD and French Ministry of Defense (MOD) endorsement, The International Foundation for Telemetry (IFT) chartered the International Consortium for Telemetry Spectrum (ICTS). With charter members from the US, France, Germany, the United Kingdom, and Australia, the ICTS sought to enlighten international administrations on the issue and provide a forum to exchange information on this topic. The ICTS formed an international grassroots network among telemetry practitioners in commercial and Government flight test areas in 2000 to garner the support of the many nations.

USA members of the ICTS include Darryl Holtmeyer (Boeing, ICTS Secretary), Mike Ryan (Navy, ITU Region II Coordinator), Chuck Irving (AF, past Chairman), Darryl Ernst (OSD/Mitre, advisor), and Tim Chalfant (AF, past Chairman, past Vice Chairman). The current ICTS Chairman is Dr. Gerhard Mayer (Germany) and the Vice-Chairman is Dr. Michael Harris (Australia).

Efforts Provide Fruit

In 2000, supporters of this effort successfully prevented a NASA-introduced topic, previously tabled at the WRC in 1997, from being deleted for consideration and was kept alive for discussion at the next WRC. With key administrations engaging (mostly the five Charter Members of the ICTS), and much ground work by the ICTS to share information and inform the gentry, the need was internationally surveyed and identified. At the 2003 WRC, we successfully passed a resolution (#230) recommending this topic be included in the agenda for the next WRC in 2007.

Since 2003, the ICTS initialized an extensive public relations campaign. This included presenting papers at international conferences, conducting workshops, and meeting with industry and Government officials, both nationally and internationally. Key support for the ICTS was provided by OSD's CTEIP Office, the US Navy, the US Air Force, the International Foundation for Telemetering, and the Aerospace & Flight Test Radio Coordinating Council (commercial aircraft manufacturers).

■ ■ ■

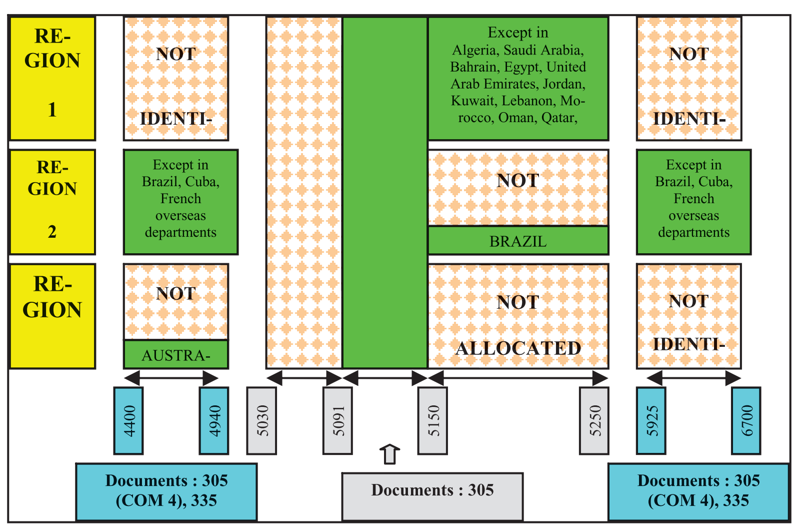


Figure 2. As a result of the recent adoption of resolutions associated with Agenda Item 1.5, the 2007 World Radio-communications Conference (WRC) granted Internationally- and Regionally-protected spectrum for flight test telemetry.

Assessment of the V&V Pilot Program

Jason McGlynn

The Test Standards Board (TSB) conducted a Verification and Validation (V&V) Pilot Program at the FAA William J. Hughes Technical Center (WJHTC) from August 27, 2007 through May 15, 2008. The Pilot Program focused on Test and Evaluation (T&E) processes for the En Route Automation Modernization (ERAM), Automatic Dependent Surveillance - Broadcast (ADS-B), Alaska Flight Service Modernization (AFSM), Data Communications (DataComm), System Wide Information Management (SWIM), and Business Continuity Plan (BCP) programs. Since each program was at a different phase of T&E, the pilot activities were tailored to appropriately assess program needs and to not adversely impact the programs.

The Pilot Program started by familiarizing the test community with the processes contained in the ATO-P T&E Handbook and V&V Operations Guide, which are based on T&E best practice studies. Twenty-three T&E employees attended the familiarization. Participant evaluation forms, discussion points, observations, and feedback were gathered and recorded for use in improving future sessions and as input for updating the T&E Handbook and V&V Operations Guide. Overall, the sessions were well-received, and there were no concerns with the proposed T&E process changes and enhancements.

On February 21-27, 2008, the TSB conducted Mid-Point Additional training is needed on Reviews to gather feedback from the ERAM, ADS-B, and AFSM programs and to assess the usefulness and effectiveness of the T&E Handbook and supporting processes. Before the review sessions were held, the program Test Directors completed a questionnaire which was provided to the TSB at the sessions. The discussions held during the review sessions, feedback received throughout the pilot activities, and the data from the questionnaires were assessed to ensure that the T&E teams were fully engaged in the processes delineated in the T&E Handbook.

After completing the Pilot Program, Pilot Program Debriefings were conducted by the TSB from May 8-15, 2008 where final feedback was obtained from all six technical programs on the Pilot Program activities. Eight T&E employees participated in the debriefing sessions. Similar to the Mid-Point Reviews, debriefing questionnaire forms were completed by the program Test Directors and provided to the TSB at the debriefings. The debriefing sessions and the data from the questionnaires helped the TSB to determine the usefulness and effectiveness of the T&E processes, the T&E Handbook, and the oversight of the TSB in monitoring T&E activities. As a result of the Pilot Program activities, the participating T&E employees identified the following recommended improvement areas to the TSB:

Additional T&E guidance is needed (i.e., on T&E process tailoring, T&E roles and relationships with System Engineering and other T&E cross-matrix organizations, and on initiating and integrating test programs)

Additional training is needed on the T&E Handbook, Critical Operational Issue (COI) decomposition, the Next Generation Air Transportation System (NextGen), and personnel certification on the T&E process.

Overall, the V&V Pilot Program was extremely useful and beneficial in engaging the ATO-P improvement from cross-cutting groups of test programs. Specific recommended improvement areas, such as those mentioned above, technical programs in the T&E processes and assessing where T&E improvements are still needed. The Pilot Program surfaced opportunities for were highlighted. This additional input by the technical programs will assist the TSB's efforts in refining the T&E practices and providing standardized quality T&E methods and practices across the ATO-P. The completion of the Pilot Program marks a major milestone for implementing the Technical Center's standards for T&E by September 2008.

■ ■ ■

Corporate News

- **L-3 BT Fuze Products** announced today that it has been awarded a $13.4 million second option release by the U.S. Army Sustainment Command at Rock Island Arsenal to produce M67 fragmentation hand grenades. This award brings the total hand grenades contract value to over $30 million. The M67 fragmentation grenade is a common close combat weapon for soldiers and Marines.

This is a major award for our division, sai d Bill Mohler, vice president and general manager of L-3 BT Fuze Products. In addition to the M67 fragmentation hand grenade production effort, we are actively pursuing additional product enhancements in an effort to provide a better product for the warfighter.

■ ■ ■

ITEA News

News from ITEA HQ

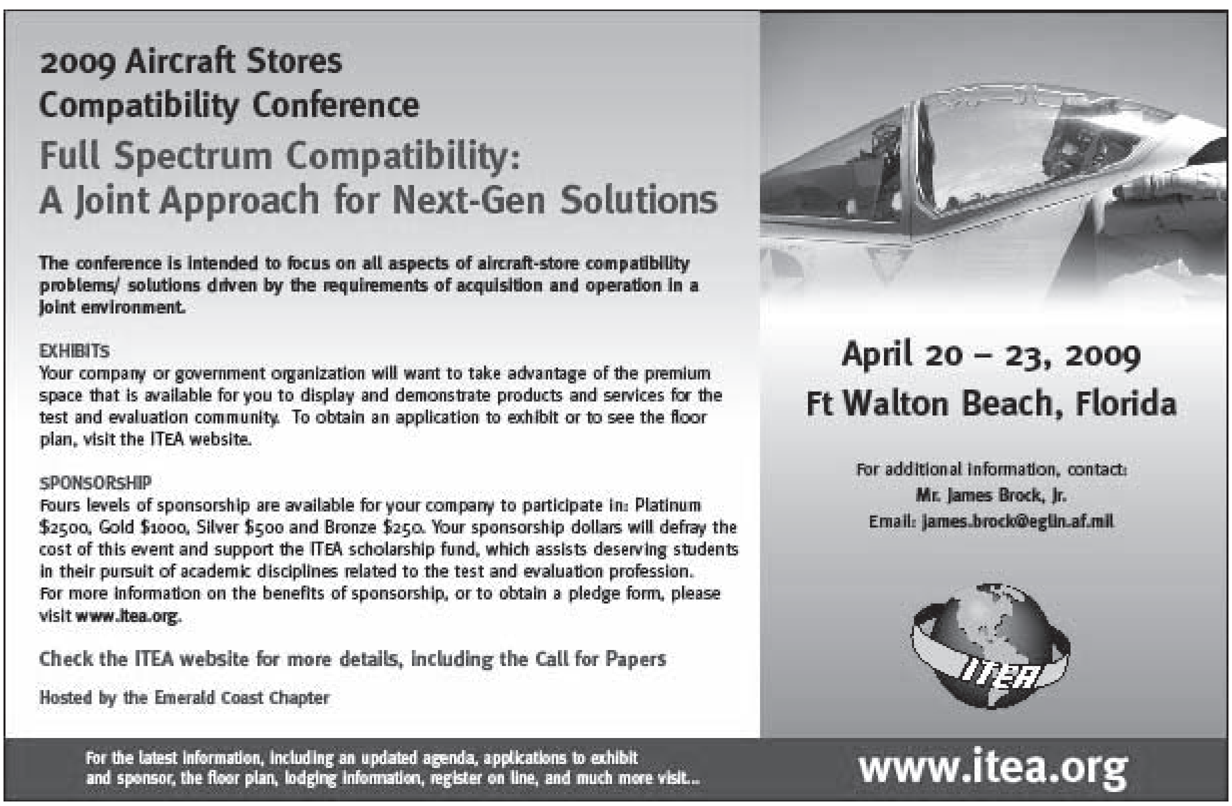
Journal Improvements : Earlier this year, *The ITEA Journal* editorial functions were officially outsourced to Allen Press in Lawrence, Kansas. ITEA has always had an outstanding, well-reputed quarterly journal, and our new relationship with Allen Press affords us even greater opportunities to improve on our successes, automate the journal style as much as possible, and raise the quality of the publication further.

We have made several significant changes to the journal over the past two issues. *The ITEA Journal* now includes consecutive numbering from one issue to another within a volume year. The spine of the journal is now printed with the journal name, volume, and issue so it can be identified standing on a shelf. A style sheet has been developed to provide consistency in areas such as table sizing, referencing, and fonts. Running headers, footers, a copyright line, and key words in the Technical Articles have been introduced and their appearance standardized. Inside and outside covers are now prime advertising spaces available for our corporate members and partners. Finally, internal ITEA advertising has been minimized to half-page advertising to allow the technical content to take precedence.

In 2009, The *ITEA Journal* will introduce a redesigned front cover that features a full color, full bleed photograph on each issue. These changes may not seem significant at first glance; however, in tandem they have improved the professionalism, richness, and look of our flagship publication.

Lori Tremmel Freeman,
Executive Director

■ ■ ■



ITEA Corporate Members –September 2008 *Last Update: July 31, 2008*

A2LA - American Association for Laboratory Accreditation
ACRA CONTROL, Inc.
Advanced Test Equipment Rentals
AEgis Technologies Group Inc.
Agency for Defense Development
AIM-USA
Alion Science and Technology
American Systems
Ampex Data Systems
Applied Resources, Inc.
Arcata Associates Inc.
ARES Systems Group
Argon ST, Network Systems
ARINC Engineering Svcs, LLC
AVW Technologies
BAE SYSTEMS Technical Services
Battelle Memorial Institute
BMC Software, Inc.
Booz Allen Hamilton
CALCULEX, Inc.
Calspan Corporation
COLSA Corporation
Computer Sciences Corporation
Cubic Defense Applications Group
Data Systems & Technology Inc.
Dawson Group, Inc.
Defense Acquisition University
Dewetron, Inc.
DigiFlight Incorporated
DRS Training & Control Systems (TCS)
Eagan McAllister Associates Inc.
EC III, LLC
EDO Technical Services Operations
Electronic Warfare Associates, Inc. GSI
EMC Corporation
EMRTC New Mexico Tech.
Epsilon Systems Solutions, Inc.
ERC Inc.
ESEI
EWA Services, Inc.
FLIR Systems
General Dynamics C4 Systems
General Dynamics InformationTechnology
Georgia Tech Research Institute
Glacier Technologies, LLC
Herley Industries, Inc.
InDyne, Inc.
ITCN, Inc.
Jacobs Technology Inc.
JHU Applied Physics Laboratory
JT3 LLC
L-3 Linkabit
L-3 Communications Telemetry & RF Products
L-3 Communications Titan Group
Log.Sec Corporation
MacAulay-Brown, Inc.
ManTech Systems Engineering Corporation
MIL Corporation
Modern Technology Solutions Inc.
NCI Information Systems Inc.
NewTec
NewTec LLC
NMSU/PSL, 21st Century Aerospace
Northrop Grumman IT
Photo-Sonics, Inc./IMC
PURVIS Systems Inc.
QinetiQ Managed Services
QUADELTA, Inc.
RAN Test Evaluation & Analysis Authority
Raytheon Missile Systems
Rotating Precision Mechanisms Inc.
Sabre Systems Inc.
Sandia National Labs/Interactive Systems Simulation & Assessments
Science Applications International Corporation
Scientific Research Corporation
Shape Technologies LLC
Spiral Technology, Inc.
SRS Technologies
Summit Instruments Inc.
SURVICE Engineering Company
SYMVIONICS Inc.
Sypris Data Systems
System Development Center - CSIST
Syzygy Technologies Inc.
Tactical Information Exchange Integration Office
TRAX International Corporation
Trideum Corporation
TYBRIN Corporation
U.S. Army Developmental Test Command (DTC)
U.S. Army Evaluation Center (AEC)
Weibel Scientific A/S
Weir Canada/NETE
Westech International Inc.
Windmill International, Inc.
Wyle

Chapter Directory

NORTHEAST REGION
Donald R. Marple
FAA/ACB-700, WJHTC
Atlantic City Airport, NJ 08405
609-485-6732 (tel)
609-485-6696 (fax)
donald.marple@faa.gov

CONNECTICUT & RHODE ISLAND
Narragansett Bay Chapter
Robert A. Connerney/NUWC
Division Newport, Code 72
1176 Howell St.
Newport, RI 02841
401-832-2151 (tel)
401-832-1954 (fax)
connerneyra@npt.nuwc.navy.mil

MASSACHUSETTS
New England Chapter
Michael E. Keller, Sr.
Gemini Industries
11 Barksdale St.
Bldg. 1614
Rm. 1A0776
Hanscom AFB, MA 01731-1625
781-266-9227 (tel)
781-266-9736 (fax)
michael.keller.ctr@hanscom.af.mil

NEW JERSEY
South Jersey Chapter
Donald R. Marple
FAA/ACB-700, WJHTC
Atlantic City Airport, NJ 08405
609-485-6732 (tel)
609-485-6696 (fax)
donald.marple@faa.gov

EAST REGION
Robert A. Vargo (Acting)
NAWC-AD, Unit 5/Bldg. 2109
48150 Shaw Rd.
Patuxent River, MD 20670
301-342-6166 (tel)
301-342-6023 (fax)
robert.vargo@navy.mil

UNITED KINGDOM
Vacant

FRANCE
Lafayette Chapter
Vacant

ISRAEL
Israeli Chapter
Aaron Leshem
Rafael 3M, P.O. Box 2082
Haifa, 31021, Israel
972-3-572-3204 (tel)
972-3-572-3314 (fax)
aleshem@netvision.net.il

OHIO
Miami Valley Chapter
Bruce D. Guindon
Northrop Grumman
IT Defense Group
Midwest Department
2555 University Blvd., Ste 100
Faiborn, OH 45324
937-431-5882 (tel)
937-205-7500 (cell)
Bruce.Guindon@ngc.com

MARYLAND
Francis Scott Key Chapter
https://www.sksi.net/fskitea/index.html
Cathy N. Pritts
US Army Developmental Test Command
314 Longs Corner Rd.
Aberdeen Proving Ground, MD 21005-5055
410-278-1435 (tel)
410-278-3710 (fax)
cathy.pritts@us.army.mil

Southern Maryland Chapter
Bill Darden
Naval Air Warfare Center AD
Atlantic Test Range
23013 Cedar Point Rd. Bldg. 2118
Patuxent River, MD 20670
301-342-3818 (tel)
bill.darden@navy.mil

DC/NORTHERN VIRGINIA
George Washington Chapter
http://gw-itea.org
Cathy Cordova
7100 Defense Pentagon
Washington, DC 20301
703-614-2227 (tel)
Catherine.Cordova@mda.mil

VIRGINIA
Tidewater Chapter
John P. Bunn
Commander, Operational Test & Evaluation Force
7970 Diven St.
Norfolk, VA 23505-1498
757-282-5546 ext. 3402 (tel)
757-282-5520 (fax)
bunnj@cotf.navy.mil

SOUTHEAST REGION
Daniel J. Pierre
Jacobs Sverdrup
600 William Northern Blvd.
Tullahoma, TN 37388
931-393-6551
931-393-6389 (fax)
pierredj@sverdrup.com

ALABAMA
Rocket City Chapter
Dr. Mark Brown
Scientific Research Corporation
103 Quality Circle, Ste 220
Huntsville, AL 35806
256-428-9285 (tel)
256-971-9890 (fax)
mdbrown@scires.com

FLORIDA
Central Florida Chapter
Keith H. Gardner
PEO STRI
12350 Research Pkwy.
Orlando, FL 32826-3276
407-384-5257 (tel)
407-384-5255 (fax)
keith.gardner@us.army.mil

Emerald Coast Chapter
http://itea-ecc.org
Chris Nixon
46 TW/TSDE
303 North 7th St., Ste 205
Eglin AFB, FL 32542
850-855-4003 (tel)
david.nixon@eglin.af.mil

GEORGIA
Atlanta Chapter
Russell S. McCrory
Georgia Tech Research Institute
400 Tenth St. NW
CRB 425
Atlanta, GA 30318
404-407-7096 (tel)
404-407-9096 (fax)
russell.mccrory@gtri.gatech.edu

TENNESSEE
Volunteer Chapter
Joe Babilon
AEDC/XPR
100 Kindel Dr. Ste A214
Arnold Engineering Development Ctr.
Arnold AFB TN 37389-1214
931-454-5346 (tel)
joe.babilon@arnold.af.mil

SOUTHWEST REGION
Gregory D. Lamberth
5540 Equestrian Ave.
Sierra Vista, AZ 85650
520-538-4940 (tel)
520-538-4853 (fax)
greg.lamberth@epg.army.mil

COLORADO
Rocky Mountain Chapter
http://www.itea-rmc.org
Patricia A. Ratterree
P.O. Box 182
Colorado Springs, CO 80901-0892
719-387-6426 (tel)
719-237-4394 (fax)
email@itea-rmc.org

ARIZONA
Huachuca Chapter
Denny Beaugureau
Booz Allen Hamilton
3103 Avenida De Suenos
Sierra Vista, AZ 85650
520-456-4644 (tel)
beaugureau_denny@bah.com

NEW MEXICO
Roadrunner Chapter
Patrick T. Dunn
AFOTEC (SAIC)
8500 Gibson Blvd. SE
Kirkland AFB, NM 87117
505-846-8272 (tel)
505-830-6765 (fax)
patrick.t.dunn@saic.com

White Sands Chapter
Douglas D. Messer
P.O. Box 398
White Sands, NM 88002
505-430-2951 (tel)
505-678-5162 (fax)
doug.messer@us.army.mil

TEXAS
Lone Star Chapter
Vacant

UTAH
Great Salt Lake Chapter
Jefferey D. Peterson
U.S. Army-Dugway Proving Ground
Dugway, UT 84022
435-831-5346 (tel)
petersonj@dpg.army.mil

WEST REGION
Keith Sutton
Syzygy Technologies Inc.
1450 Frazee Road, Ste 701
San Diego, CA 92108
619-297-0970 x232 (tel)
619-297-0975 (fax)
ksutton@syzygy-tech.com

CALIFORNIA
Antelope Valley Chapter
http://www.iteaavchapter.org
Reynaldo Garza
JT3 LLC
190 S. Wolfe Ave., Bldg. 1260
Edwards AFB, CA 93524-1033
661-277-9246 (tel)
661-277-9279 (fax)
reynaldo.garza.ctr@edwards.af.mil

Channel Islands Chapter
Christopher J. Weal
NAVAIR, NAWC-WD
Code 5C3000E
575 I Ave., Suite 1
Point Mugu, CA 93042
805-989-7947 (tel)
805-989-5224 (fax)
christopher.weal@navy.mil

China Lake Chapter
Bettye R. Moody
NAVAIR, Code 473000D
Bldg. 11010, Rm. 21
2400 E. Pilot Plant Rd., MS-5300
China Lake, CA 93555
760-939-7252 (tel)
760-939-7479 (fax)
bettye.moody@navy.mil

Greater San Diego Chapter
Jack Sears
Epsilon Systems Solutions, Inc.
1565 Hotel Circle South
Ste 200
San Diego, CA 92108
619-702-1700 (tel)
jsears@epsilonsystems.com

WASHINGTON
Pacific Northwest Chapter
Jon Hanan
The Boeing Company
M/S 4W43
9725
East Marginal Way
Seattle, WA 98101
206-940-1162 (tel)
Jonathan.B.Hanan@Boeing.com

PACIFIC REGION
Stewart V. Burley
Strategic Theories Unlimited
Lawai, HI 96765
808-651-1111 (tel)
808-338-2409 (fax)
stu@stukavai.com

AUSTRALIA
Southern Cross Chapter
Dr. Vivian H. Crouch
University of South Australia
SEEC, Mawson Lakes Campus
Room 2.F.2.9
Mawson Lakes, Australia 5095
61-08-8302-3989 (tel)
61-08-8302-5344 (fax)
vivian.crouch@unisa.edu.au

HAWAII
Mid-Pacific Chapter
Stewart V. Burley
Strategic Theories Unlimited
Lawai, HI 96765
808-651-1111 (tel)
808-338-2409 (fax)
stu@stukauai.com

INTERNATIONAL TEST AND EVALUATION ASSOCIATION
APPLICATION FOR MEMBERSHIP (INDIVIDUAL)
FOR ON-LINE REGISTRATION/INFORMATION, VISIT WWW.ITEA.ORG

Date ____________________________

Name____
Last First Initial

____Mr. ____Mrs. ____Ms. ____Miss ____Dr. ____Other

Preferred Mailing Address ____Work ____Home

Company/Organization ____

Title ____

Street____

City _________________State __________Zip_____________________Country _________________

Office Tel (Country Code + Area Code) ______________________________________

FAX Number (Country Code + Area Code) ____________________________________

E-mail ____

Company/Organization World Wide Web URL (if applicable) __________________________

Home Address: ____
Street

____
City/State/Zip/Country

Home Telephone (Country Code + Area Code) ____________________________________

____Industry ____Military ____Government ____Academia

ITEA membership dues are $45 for individuals and $25 for full-time students.
Overseas addresses require additional $20 for postage.

ITEA Lifetime membership dues are as follows:

Age 60+	$375
Age 50-59	$575
Age 40-49	$690
Below age 40	$790

Annual dues include a one-year subscription to *The ITEA Journal of Test and Evaluation*.

____Check payable in U.S. dollars to: International Test and Evaluation Association

____Credit Card (Circle One) AMERICAN EXPRESS VISA MasterCard

Card Number __ __ __ __ __ __ __ __ __ __ __ __ __ __ __ __ Expiration Date _____________________

Signature____ Date ______________________

RETURN PAYMENT TO
International Test & Evaluation Association
4400 Fair Lakes Court
Fairfax, Virginia 22033-3899
Tel: (703) 631-6220 Fax: (703) 631-6221
E-mail: *itea@itea.org* Web: *http://www.itea.org*

The
ITEA Journal

ARTICLE SUBMISSION GUIDELINES

Thank you for considering the *ITEA Journal* for your work. Use these guidelines in preparing your submission.

Text Format: Please submit papers in MS Word format. Avoid complicated MS Word formatting/design, because articles are re-designed in other software.

Graphics: In order to achieve the best photographic resolution, all electronic artwork (pictures, graphs, tables, mathematical equations) must be submitted separately, each as its own separate file in one of the following formats: jpg, tiff, eps, bmp, xls, labeled as “Figure 1,” “Figure 2” etc. to correspond with figure call outs and captions. Graphics must be clear, good quality, and saved at no less than 300 dpi resolution, with actual dimensions of no less than 2 inches by 4 inches. We need to receive the native or original file the graphic was created in because as soon as graphics are placed into Word and PowerPoint their resolution is reduced to 72 dpi.

Tables: Submit tables in Excel (xls) or in Word (doc).

Content Order: Papers must contain the following elements:

1) **Clearance:** All submissions must be cleared for publication BEFORE being submitted to the *ITEA Journal*. An *ITEA Journal* Copyright Agreement will be provided upon request.
2) **Title** of paper.
3) **Byline** containing full names of each author; each author’s place of employment, city and state.
4) **Abstract** paragraph (150 words or less).
5) **Key words.** 4 to 7 words or phrases that convey the most important concepts in the article and assist in cross indexing the article. Key words are not required for editorials, commentaries, historical perspectives, or technotes.
6) **Figure call outs** in text: If article contains artwork, the body of the article must contain Figure call outs (e.g., Figure 1, Figure 2). Please call figures out in the order in which they appear in your article or they will be renumbered.
7) **Figure and Table Captions:** Figure and table captions must be provided for each figure and/or table.
8) **Biography** for each author (separate paragraph for each, and not to exceed 100 words each) containing author’s full name, current job title, place of employment, city & state; brief description of any pertinent professional experience; and any degrees held, listing the university/institution with city & state.
9) **E-mail address and phone number**: Complete telephone number, e-mail address, and address for each author. Please indicate which author is the corresponding author.

Word Count/Graphics Limit: Required word limit: 4,800 words. This includes abstract, article text, captions, biographies, references & endnotes. Please limit the number of graphics to 12 figures or less.

Submit Material electronically to: Rita Janssen, Managing Editor, ITEA Journal, at rjanssen@allenpress.com; or by mail/courier with a cd to ITEA, Attention: Rita Janssen, 810 East 10th Street, Lawrence, KS 66044. (785) 865-9137.

Questions: Contact Rita Janssen, Managing Editor, ITEA Journal, at rjanssen@allenpress.com or 785-865-9137.

The ITEA Journal of Test and Evaluation

Effective January 1, 2008

Rates

Black & White	Base Rate	4X
Full Page	$840	$788
1/2 Page	$541	$504
1/4 Page	$441	$410
2 Page Spread	$2,011	$1,806

4 Color Rates	Base Rate	4X
Full Page	$1,680	$1,575
1/2 Page	$1,208	$1,123
1/4 Page	N/A	N/A
2 Page Spread	$3,455	$3,150

Premium Positions

Cover 2 (ISFC)	$1920
Cover 3 (ISBC)	$1920
Cover 4 (Back Cover)	$2300
Page 3	$1840

2008 Production Schedule

Issue	Space Reservation	Ad Material Deadline
March	1/28/2008	2/11/2008
June	4/25/2008	5/9/2008
September	7/28/2008	8/1/2008
December	10/24/2008	11/17/2008

Mechanical Specifications

Ad Sizes	Non Bleed W X H	Bleed W X H
Full Page	6 1/2 X 9	8 3/4 X 11 1/4
Half Page (H)	6 1/2 X 4 1/2	
Half Page (V)	3 X 9	
QTR Page	3 X 4 1/2	
2 Page Spread	15 X 9	17 1/4X 11 1/4

Trim Size: Journal trims 1/8 " off top, bottom and outside edge. Final trim size of publication is 8 1/2" x 11". Live matter should be a minimum of 1/2" inside trimmed edges, and a minimum of 1/2" should be allowed for the bind.

Binding: Perfect

Paper Stock: Inside Pages: 60 lb gloss Covers: 100 lb gloss

Insert Requirements: All inserts must be furnished printed, Folded, ready for binding. Inserts must be folded to 8 3/4 " by 11 1/8". Folded inserts will be placed between signatures.

Electronic Files: If you are supplying the images on disk, please include a directory of files contained on the disk and a hard copy printout. Supplied hard copy must match the file. Digital color files must be submitted on CMYK mode. Digital color proofs supplied with the file must be output in CMYK.

Halftone Screen: 133—175 line screens are acceptable

Resolution: Graphics should be a minimum of 350 dots per Inch (dpi) or higher.

Formats: We accept TIFF or EPS format for both Macintosh and PC platforms. We also accept files in the following Native Application File Formats:
Adobe Acrobat(.pdf) (use Press settings under Job Options)

Adobe Photoshop (.psd)	Illustrator (.ai)
Macromedia FreeHand (.fh)	Corel Draw (.cdr)
Canvas (.cvs)	PowerPoint (.ppt)
InDesign (.id)	Pagemaker (.pmd)
QuarkXPress (.qxd)	

Fonts: Include the screen and printer font files for any text. PC or Macintosh versions of Adobe PostScript fonts should be used. Do NOT use True Type fonts. Do NOT use system "bitmap" fonts.

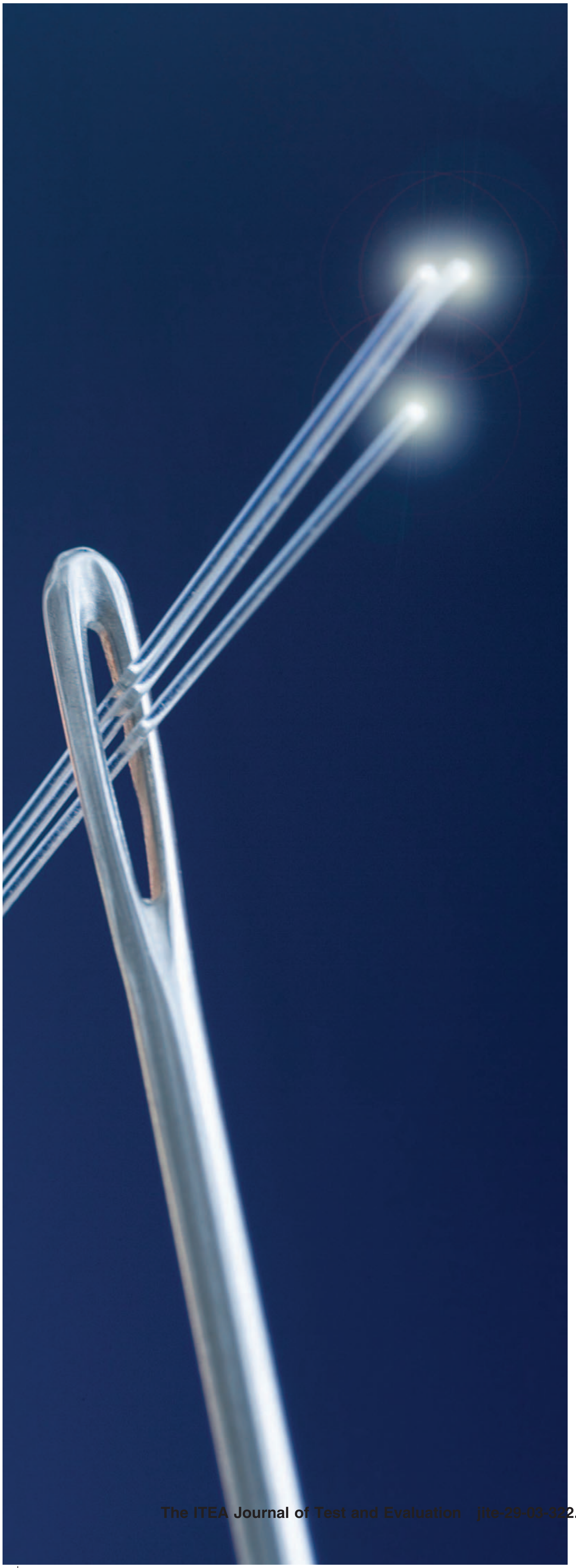
Data Transfer: We can accept the following media:

Macintosh/PC Platforms	3.5 Floppy Disk
Zip Disks	Jazz Disks
CD_ROM	DVD

We can also accept or retrieve files via File Transfer Protocol (FTP). If you will be posting to the FTP site please follow the following directions:
1. The user should FTP to: ftp.allenpress.com
2. Login with the username of anonymous
3. Use your email address as the password
4. Place files in the inbox

Compression: Large files should be compressed with Stuffit or Win-Zip if possible.

Disclaimer: All claims for errors in advertisements must be made in writing and received within ten days of publication and will be considered only for the first insertion of the advertisement.



www.boozallen.com

complex challenges tailored solutions

Program managers today face more complex challenges with fewer resources. Cost, schedule, and performance assurance require a dedicated, trained, and professional T&E workforce that supports the acquisition lifecycle by focusing on aligning resources, reinforcing systems engineering principles, and improving management processes.

Booz Allen Hamilton, a global strategy and technology consulting firm, delivers tailored solutions to help clients horizontally integrate T&E across the acquisition lifecycle enterprise. Our proprietary and strategically-designed *T&E Bootcamp* builds a common core of basic T&E competencies while providing the baseline knowledge required to initiate a comprehensive training program. Developing industry relationships and fostering a collaborative environment between program management, test and evaluation, and systems engineering teams is our mission.

Booz Allen is the one firm that helps clients solve their toughest problems, working by their side to help them achieve their missions. Booz Allen is committed to delivering results that endure.

For more information, contact Tim Cerniglia at cerniglia_tim@bah.com or 719/387-3771.

Booz | Allen | Hamilton

delivering results that endure