Integrated Test: Challenges & Solutions

Honorable Dr. Michael Gilmore
Director, Operational Test & Evaluation
Presentation to NDIA
March 15, 2011

Outline

- DOT&E Initiatives
- Integrated Test
- Challenges to Integrated Test
- Integrated Test Solutions
- Design of Experiments and Integrated Testing
- Conclusions

DOT&E Initiatives

Integrated Testing
Developmental
Operational
Live Fire
Use Scientific Test Design
(e.g. DOE)

Testers Engage Early
Testable, Mission-oriented Requirements

Field New Capabilities Rapidly
Accelerated Testing

Improve Suitability
Reliability Growth

Integrated Test

- What is Integrated Testing?
  - A cohesive test and evaluation plan that spans all stages of testing.
  - Integrated test is **NOT** simply combining data from different test events.
  - Integrated test is **NOT** a replacement for dedicated OT.

- Integrated Test methods:
  - Using data from CT, DT, and OT to inform the next stage of testing
  - When appropriate, combine CT, DT, and OT data
    - Reduce test time, increase statistical confidence and power
  - Integrate DT and OT test objectives
    - Enhance operational realism in DT to reduce OT requirements
  - Design of Experiments helps plan efficient, integrated testing
    - Plan testing as a sequence of tests

Integrated Test Can Be A Challenge

- Not business as usual
  - Unclear responsibilities. Who is in charge of the test?
- Contractual issues
  - Limited access to contractor test data and test procedures
- DT and OT test objectives conflict
  - Combining tests maybe impossible
- Combining data maybe irresponsible
  - How the test is executed affects results
  - How the system design evolves affects results
- Late involvement of OT testers
  - Affects all of the above

Integrated Testing Makes Sense!

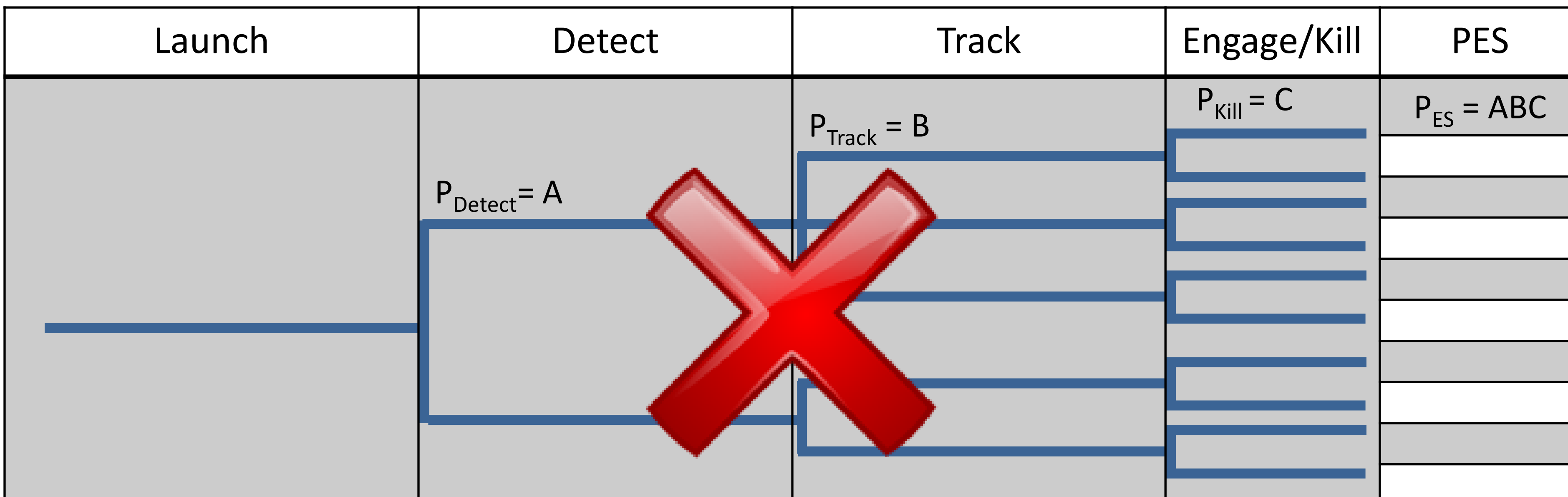
- Enables efficient testing
  - OT assessments can take advantage of CT and DT data
- Assessing system performance as the design matures requires consolidation of data
  - e.g., reliability growth
- System-of-systems requiring coordination of multiple test programs are increasingly common
- Discovery in OT is expensive
  - We need to find problems early in DT
- Design of Experiments facilitates efficient, integrated testing.

Integration of Available Data

Ballistic Missile Defense

Ballistic Missile Defense

- Motivation: Estimate system effectiveness with small sample sizes
- Probability of Success (PES) is the probability of successfully negating a ballistic missile threat using the Ballistic Missile Defense System (BMDS)
- Traditional probability based approaches are data intensive
  - Conditional probability model requires lots of data in each stage



PES for Ballistic Missile Defense

- DOT&E turned probability problem into sampling problem
  - PES = (# Kills)/(# Launches)
  - PES = (# Kills)/(# Detections) • (# Detections)/(# Launches)
  - PES = (# Kills)/(# Tracked) • (# Tracked)/(# Detections) • (# Detections)/(# Launches)
  - … repeat …

	Launch	Detect	Track	…	Intercept	Kill
Test 1	xx	xx				
Test 2			xx	xx	xx	
Test 3	xx	xx	xx	xx	xx	xx
Test 4	xx	xx	xx	xx	0	
Test 5	xx	xx	xx	xx	xx	xx

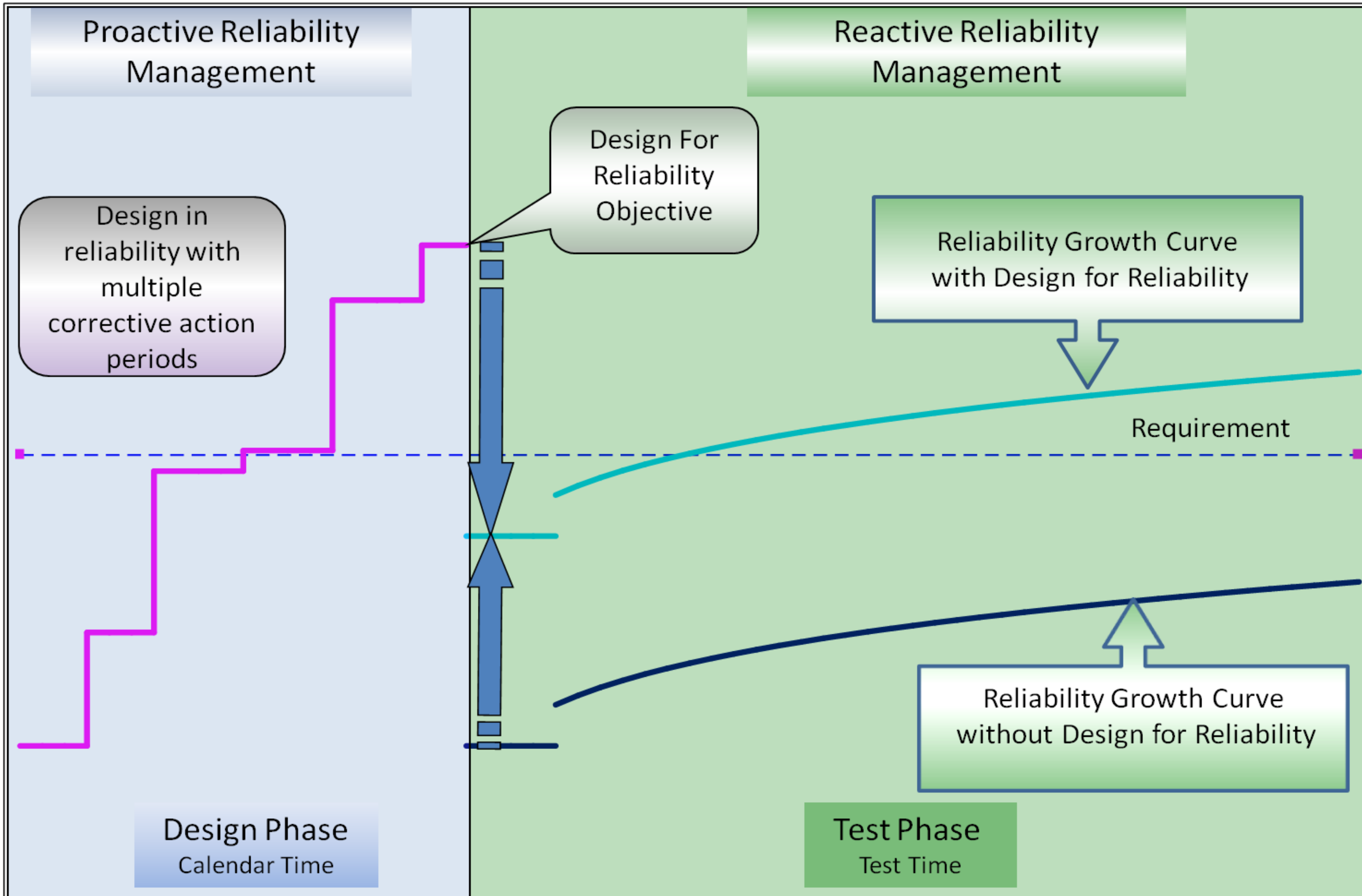
Partial Tests

Failure at Intercept Stage

- DOT&E PES methodology applied to Patriot data
  - Produces similar results to traditional analysis for large datasets (validates method)
  - Validation indicates that the similar results were achieved with less data
- DOT&E PES methodology applied to Aegis BMD (smaller dataset)
  - Refines the results from simple success/failure analysis to account for partial tests
  - Results included in DOT&E Report to Congress

Maximize use of data from relevant test events

Integrated Testing for Reliability



ANSI/GEIA-STD-0009

1. Understand user requirements and constraints
   - Reliability requirements include the anticipated use environment

2. Design for Reliability (DFR) and Re-design for Reliability
   - This means that user needs will be allocated through system model to reliability specifications at lowest component levels.
   - Lowest level reliability specifications include internal stresses and impacts of use environment
   - Redesign as needed to meet allocated reliability requirements

3. Produce reliable systems
   - During DT, all sub-assemblies, components, etc should demonstrate required reliability in anticipated use environments
   - Meeting reliability requirements will often require reliability growth programs for components utilizing repeated DT experiments

4. Monitor and assess user's experienced reliability

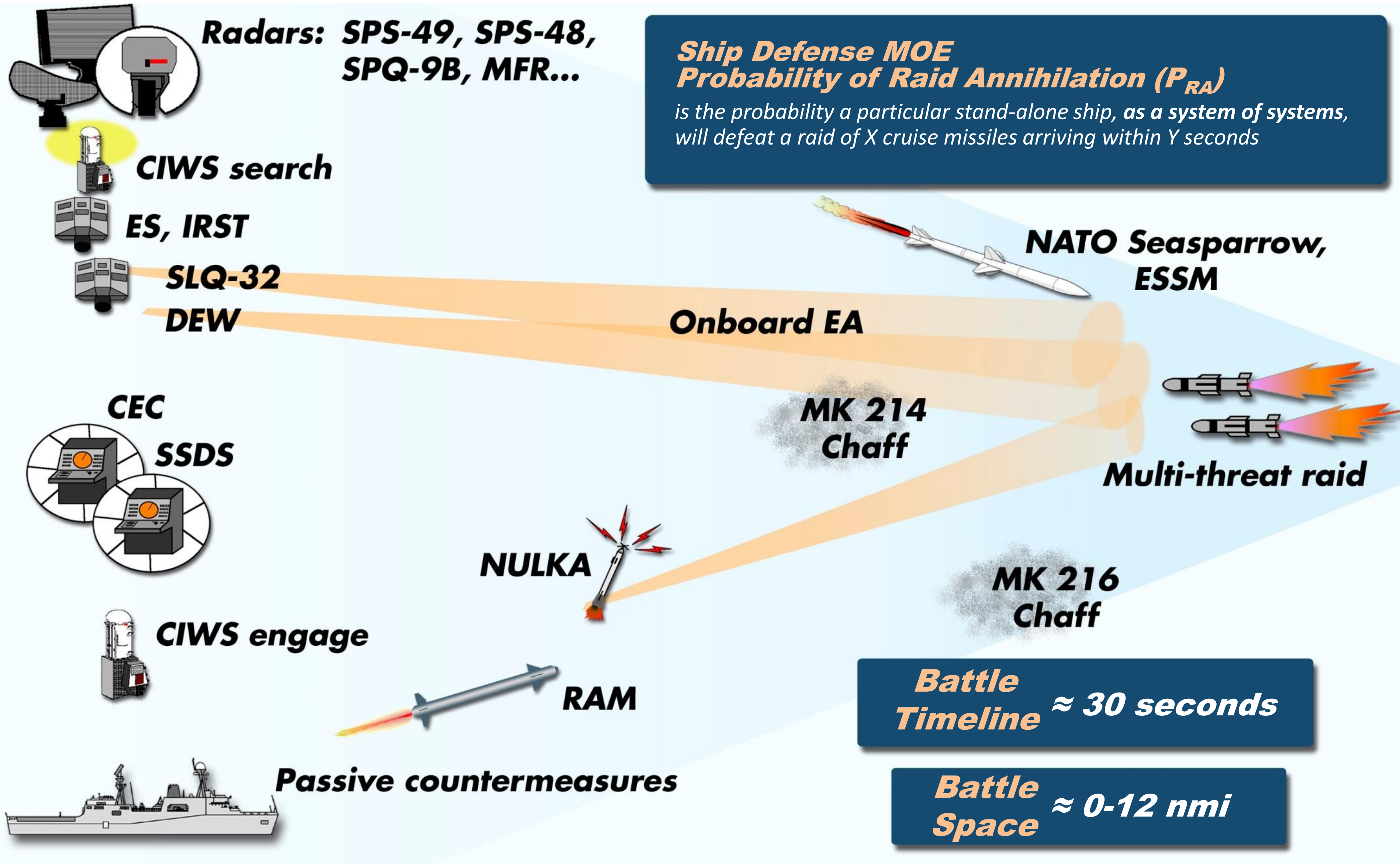
Integrated Reliability: Each stage informs the next

Stryker NBCRV Design For Reliability

1. Production Verification Testing (PVT) was halted prematurely due a large number of System Aborts
   - Did not meet the user requirement of 1000 Mean Miles Between System Aborts (MMBSA) for the base vehicle
   - No reliability requirement for NBC sensors
2. System contractor implemented Design For Reliability to improve base vehicle reliability (2007-2008)
3. NBCRV underwent 8000 mile Reliability Growth Test (RGT) in 2009 to determine whether reliability had improved.
   - Base vehicle reliability dramatically improved over PVT (2000 MMBSA).
   - Little change in NBC sensor reliability.
4. Dramatic improvement in reliability between PVT and RGT but no reliability growth seen during RGT itself.
5. Requirements drove the focus of DFR, but requirements addressed only the base vehicle and not the NBC sensors
6. DFR is a powerful tool to improve reliability, but must address entire system to be effective

Integrated Testing for System of Systems

Air Warfare Ship Self-Defense Enterprise



Air Warfare Ship Self-Defense Enterprise

- Combat systems for aircraft carriers and amphibious ships composed of systems from various program offices
  - Previously, each program office developed its own test program
  - Each test program focused on an individual system, not on the integrated combat system or the overall air defense mission
- Ship Self-Defense Enterprise coordinated these various test programs
  - Provides significantly better end-to-end testing of the ***integrated*** combat system, focusing on the air self-defense mission
  - Used principles of Design of Experiments to develop test plan
- For air self-defense, the Navy estimates:
  - Before Enterprise, testing cost about $1.1 Billion FY05 through FY15
  - Enterprise saved $240 Million out of $1.1 Billion

Better testing for less money

Integrated Testing to Avoid Late Problem Discovery

Integrated Defensive Electronic Countermeasures (IDECM)
&
Miniature Air Launched Decoy (MALD)

Late Discovery of Problems
IDECM and MALD

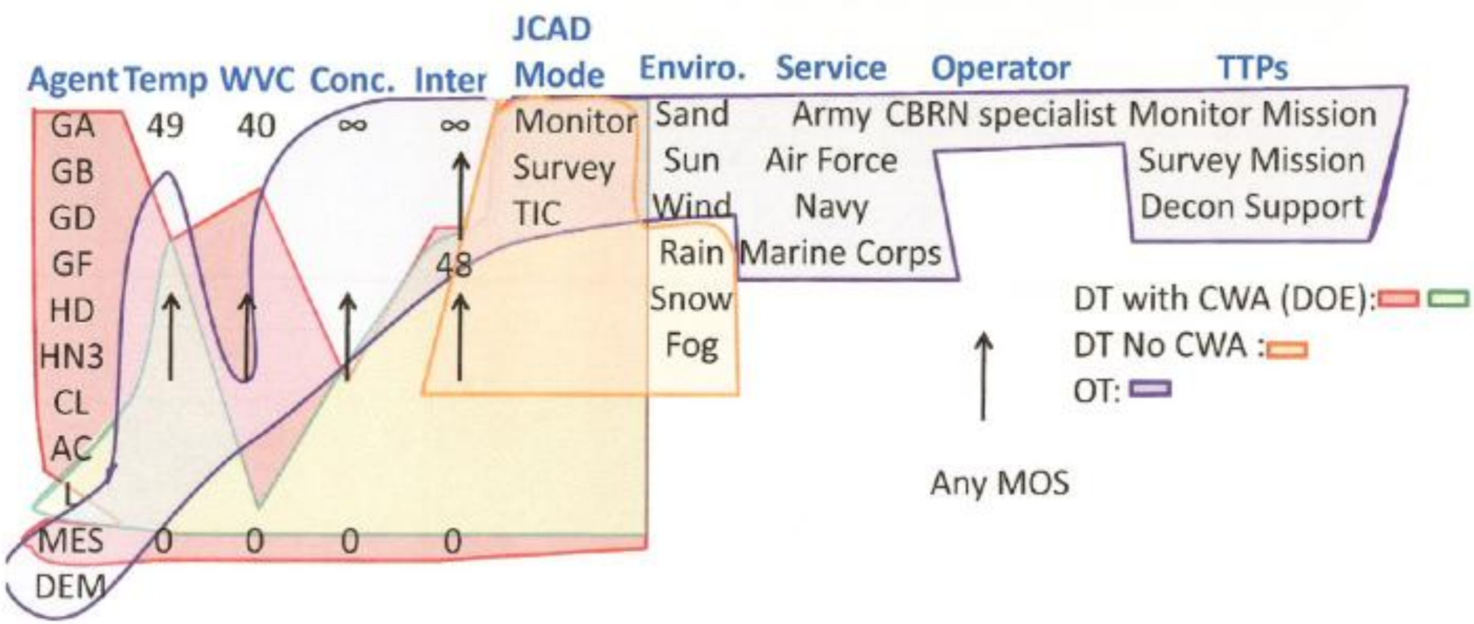
- Limited operational realism in early testing
  - IDECM – use of special DT equipment to reduce test costs
  - MALD – no long-duration carriage of decoys
- Significant problems discovered in IOT&E
  - IDECM
    - Uncommanded deployments and problems severing decoys created safety problem for ground crew
    - Intermittent failures resulted in decoys being prematurely discarded and in poor reliability
  - MALD
    - Long-duration flight caused premature failures when decoys were launched.

Design of Experiments (DOE)

- A method for planning efficient integrated testing.
- For integrated testing, DOE can inform:
  - Plan testing as a sequence of tests
  - Screen out insignificant factors in DT to focus OT
  - Control factors in DT that are difficult to control in OT
  - Split factors across test periods
  - Ensure that operational envelope is covered
- DOE is an Industry Best Practice
  - DOE traditionally applied in DT context, but we are seeing great gains using the methodology in integrated testing and operational testing
- Example of DOE in DT: wind tunnel testing
  - Characterize the aerodynamic behavior of the X-31 Enhanced Fighter
  - Traditional techniques would require 1000 + test points
  - DOE applied & testers were able to characterize aerodynamic performance in 104 test points.

Example of Integrated Testing Employing DOE Joint Chemical Agent Detector

- Problem: Agents are unable to be tested in an OT.
  - Agent, temperature, water vapor content, operating mode and agent concentration were systematically varied in DT using a Response Surface Design.
  - Allowing for operational factors affecting performance to be assessed in OT (Service, environment, and mission tactics)



Conclusions

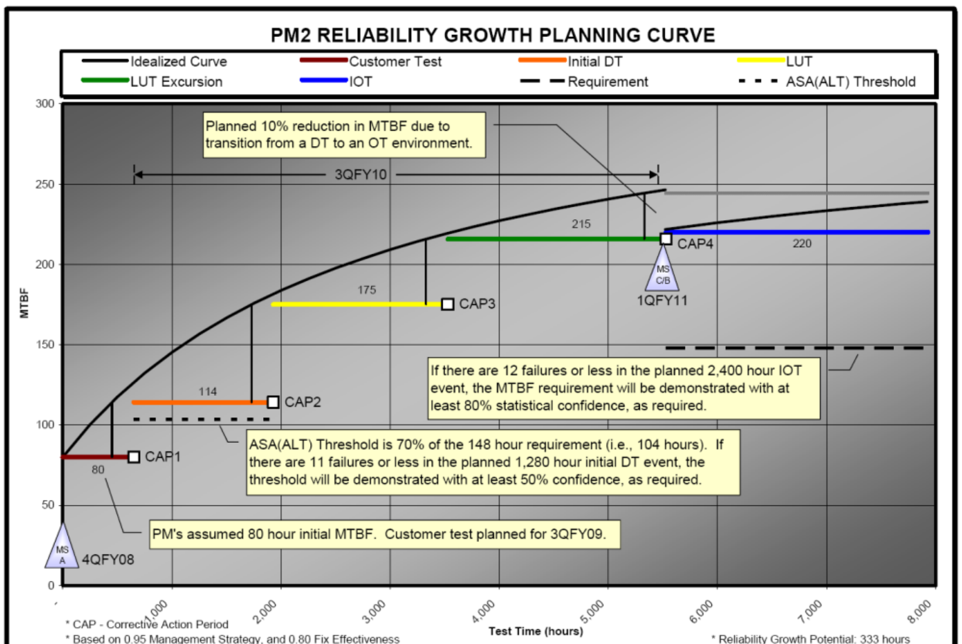
- Efficient integrated testing is a must.
- Integrate Test solutions are as unique as the challenges
  - Plan CT and DT tests to enable OT use of the data.
  - Assessing system reliability requires integrated test.
  - System-of-systems requires integration of multiple test programs.
  - Operational realism in DT allows problems to be discovered early
- Key Ingredients for Integrated Testing
  - Early engagement of Operational Testers
  - Robust data collection and documentation
  - Experimental Design
    - Can help ensure integrated testing is comprehensive
    - Provide confidence and power across the operational envelope

Every Program and every challenge has a unique solution to Integrated Testing

Backups/Extras

Integrated Testing for Reliability

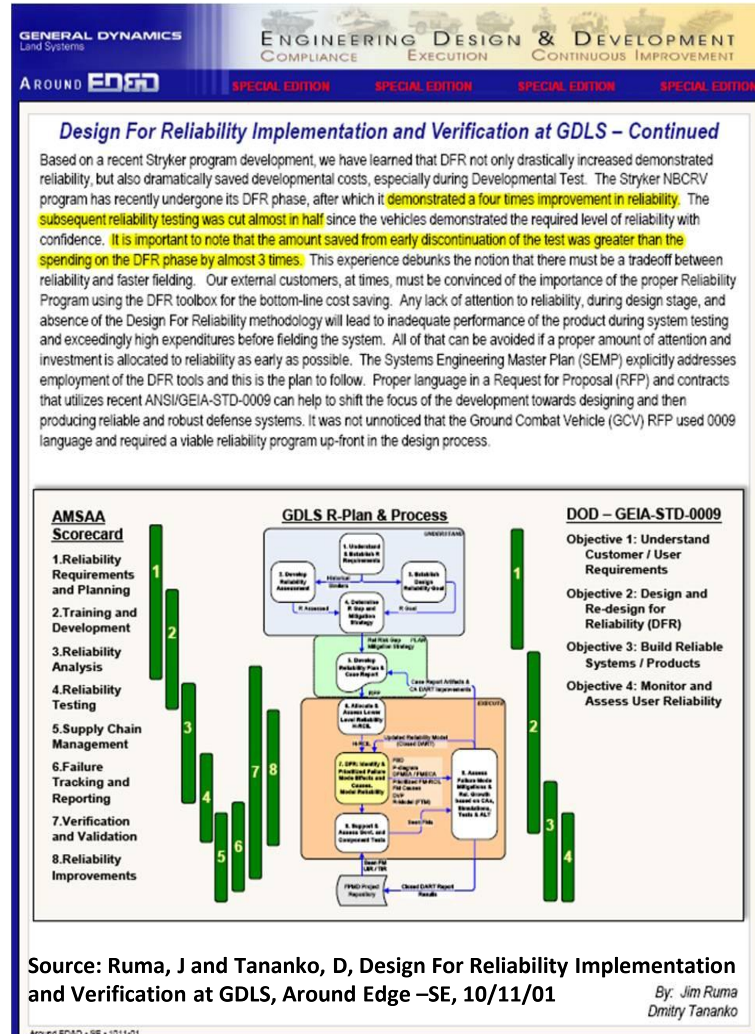
Design for Reliability
&
Reliability Growth



Stryker NBCRV
Design For Reliability Case Study

- **Reduced Risk:** "... program has recently undergone its DFR phase, after which it demonstrated a four times improvement in reliability."
- **Reduced Acquisition Time:** "... subsequent reliability testing was cut almost in half since the vehicles demonstrated the required level of reliability."
- **Reduced Cost:** "...the amount saved from early discontinuation of the test was greater than the spending on the DFR phase by almost 3 times."

The Cost AND Schedule Optimal Solution is to Design for Reliability


GENERAL DYNAMICS Land Systems

ENGINEERING DESIGN & DEVELOPMENT
COMPLIANCE EXECUTION CONTINUOUS IMPROVEMENT

AROUND ED&D SPECIAL EDITION SPECIAL EDITION SPECIAL EDITION SPECIAL EDITION

Design For Reliability Implementation and Verification at GDLS – Continued

Based on a recent Stryker program development, we have learned that DFR not only drastically increased demonstrated reliability, but also dramatically saved developmental costs, especially during Developmental Test. The Stryker NBCRV program has recently undergone its DFR phase, after which it demonstrated a four times improvement in reliability. The subsequent reliability testing was cut almost in half since the vehicles demonstrated the required level of reliability with confidence. It is important to note that the amount saved from early discontinuation of the test was greater than the spending on the DFR phase by almost 3 times. This experience debunks the notion that there must be a tradeoff between reliability and faster fielding. Our external customers, at times, must be convinced of the importance of the proper Reliability Program using the DFR toolbox for the bottom-line cost saving. Any lack of attention to reliability, during design stage, and absence of the Design For Reliability methodology will lead to inadequate performance of the product during system testing and exceedingly high expenditures before fielding the system. All of that can be avoided if a proper amount of attention and investment is allocated to reliability as early as possible. The Systems Engineering Master Plan (SEMP) explicitly addresses employment of the DFR tools and this is the plan to follow. Proper language in a Request for Proposal (RFP) and contracts that utilizes recent ANSI/GEIA-STD-0009 can help to shift the focus of the development towards designing and then producing reliable and robust defense systems. It was not unnoticed that the Ground Combat Vehicle (GCV) RFP used 0009 language and required a viable reliability program up-front in the design process.

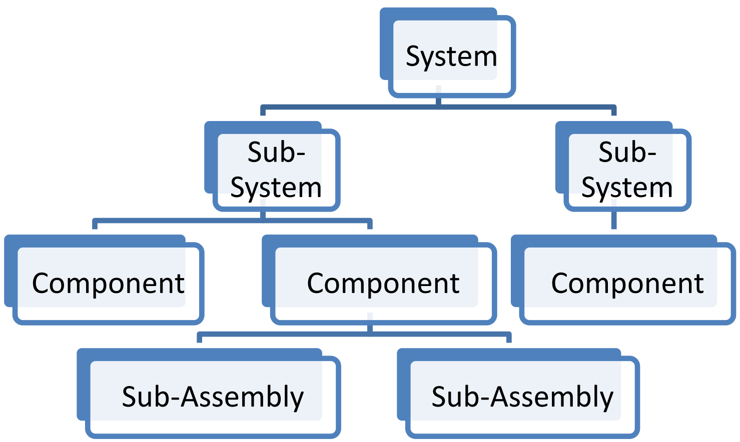
Source: Ruma, J and Tananko, D, Design For Reliability Implementation and Verification at GDLS, Around Edge –SE, 10/11/01

By: Jim Ruma
Dmitry Tananko

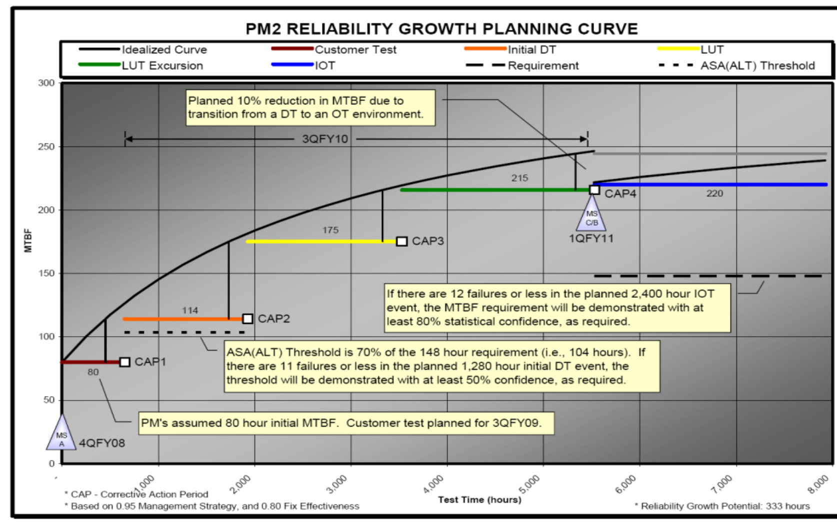
Around ED&D - SE - 1011-01

Reliability growth depends on two distinct reliability models

- Full system model guides integrated testing.
  - Provides an initial guess at system reliability
  - The goal is NOT to create a complete model of what will fail when and why



Full system model used to allocate system reliability down to required reliability at lowest levels.



Growth model used to track and predict reliability of individual pieces(sub-systems, etc) in DT/IT and of full system in IT/OT

As design matures…Reliability Growth