INSTITUTE FOR DEFENSE ANALYSES

Reliability Best Practices and Lessons Learned in the Department Of Defense

Yevgeniya K. Pinelis, *Project Leader*

Jonathan L. Bell
Charles D. Carlson
Brent A. Crabtree
Rebecca M. Dickinson
Laura J. Freeman
Duane A. Goehring
Jason M. Gonzales
Allison L. Goodman
Dawn C. Loper
Max W. Roberts
Dean Thomas

January 2018

Cleared for Open Publication

IDA Non-Standard Document
NS D-8889

Log: H 2017-000706

INSTITUTE FOR DEFENSE ANALYSES
4850 Mark Center Drive
Alexandria, Virginia 22311-1882

About This Publication
This work was conducted by the Institute for Defense Analyses (IDA) under contract HQ0034-14-D-0001, Task 2299(90), “Test Science Applications,” for the Office of the Director, Operational Test and Evaluation. The views, opinions, and findings should not be construed as representing the official position of either the Department of Defense or the sponsoring organization.

Acknowledgments
Technical review was performed by Breena G. Anderson and Kelso W. Horst of the Operational Evaluation Division.

For more information:
Dean Thomas, Project Leader
dthomas@ida.org • 703-845-6986

Robert R. Soule, Director, Operational Evaluation Division
rsoule@ida.org • (703) 845-2482

INSTITUTE FOR DEFENSE ANALYSES

IDA Non-Standard Document NS D-8889

Reliability Best Practices and Lessons Learned in the Department Of Defense

Yevgeniya K. Pinelis, *Project Leader*

Jonathan L. Bell
Charles D. Carlson
Brent A. Crabtree
Rebecca M. Dickinson
Laura J. Freeman
Duane A. Goehring
Jason M. Gonzales
Allison L. Goodman
Dawn C. Loper
Max W. Roberts
Dean Thomas

Reliability Best Practices and Lessons Learned in the Department of Defense

Laura J. Freeman, PhD

Senior Technical Advisor to the Director,
Operational Test and Evaluation

64th Annual Reliability & Maintainability Symposium (RAMS) Reno, Nevada, January 2018

Outline

- Background
  - Director, Operational Test and Evaluation (DOT&E)
  - Reliability trends in the Department of Defense
- Reliability Best Practices and Lessons Learned in the Department of Defense
- Recommendations
- Future Research
- References

Background Section

DOT&E Background

- Director, Operational Test and Evaluation (DOT&E) was created by Congress in 1983

- Director is appointed by the President and confirmed by the Senate

- DOT&E reports go to:
  - Secretary of Defense
  - Defense Acquisition Executive
  - Vice Chairman Joint Chiefs of Staff
  - Senate/House Armed Services Committees
  - Senate/House Appropriations Committees-Defense
  - Army, Air Force, Navy, and Marine Corps

- Responsible for all operational test and evaluation and monitoring and reviewing of live fire test and evaluation within the Department of Defense

- Responsibilities outlined in Title 10, U.S. Code, Sections 139, 2366, 2399, and 2400

DOT&E Background: Congressional Objectives

- Independent oversight
- Coordination of Military Services planning and execution of Operational Test and Evaluation and Live Fire Test and Evaluation
- Thorough independent analysis and objective reporting of these results to decision-makers in the Department of Defense and Congress
- Fundamental concerns were that:
  - Military systems were not being tested thoroughly or realistically
  - Complete and accurate information was not being disseminated

DOT&E Background: DOT&E Focus

- Is the Operational Test and Evaluation and/or Live Fire Test and Evaluation **adequate**?
- Is the system **operationally effective**?
- Is the system **operationally suitable**?
- Is the system **survivable** and **lethal**?

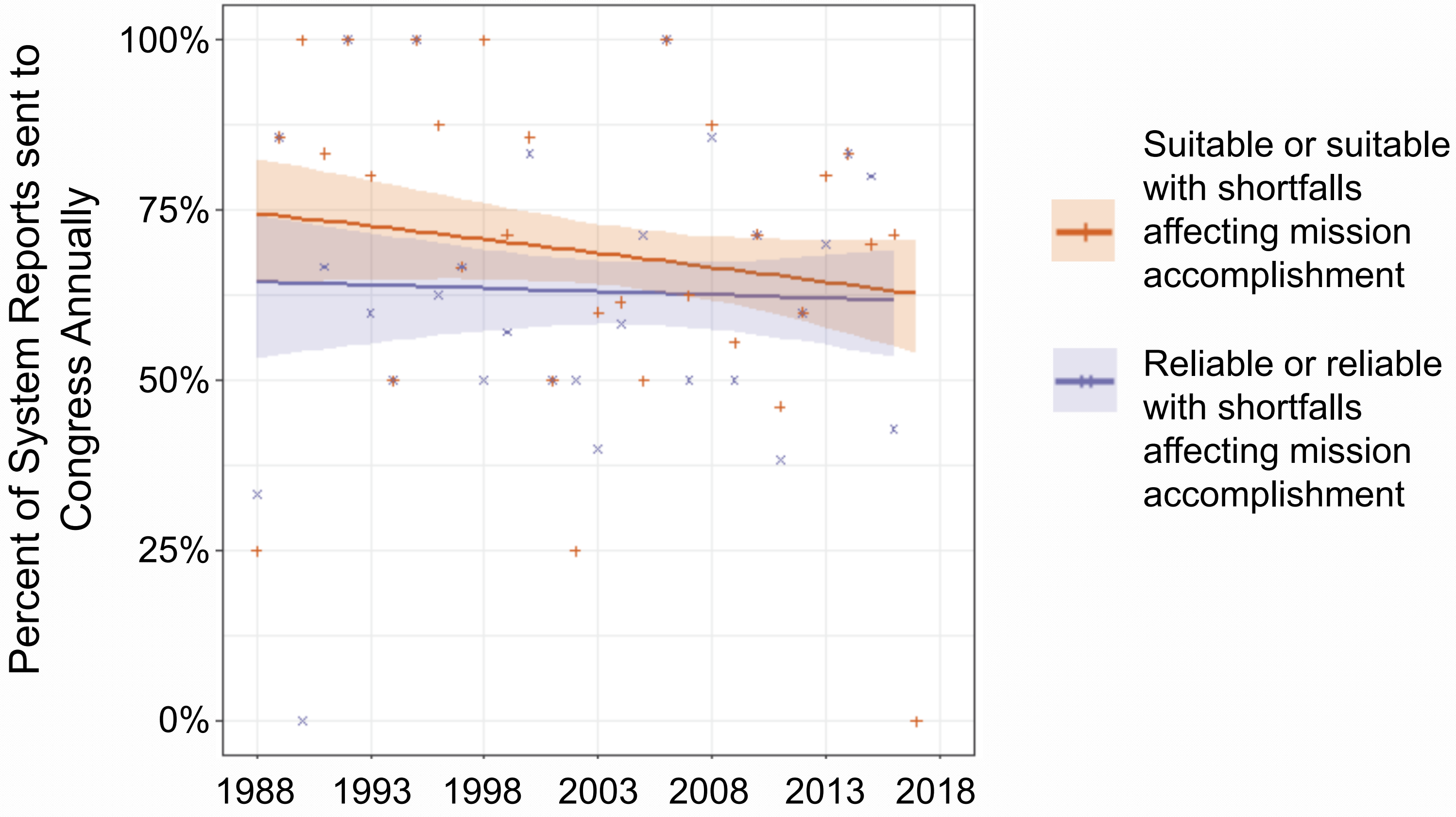
DOT&E Background: Operational Suitability

Operational suitability is the degree to which a system can be satisfactorily placed in field use, with consideration given to:

- Reliability
- Availability
- Compatibility
- Transportability
- Interoperability
- Wartime usage rates
- Maintainability
- Safety
- Human factors
- Manpower supportability
- Logistics supportability
- Documentation
- Environmental effects
- Training requirements

Background: Suitability Trends in the DoD

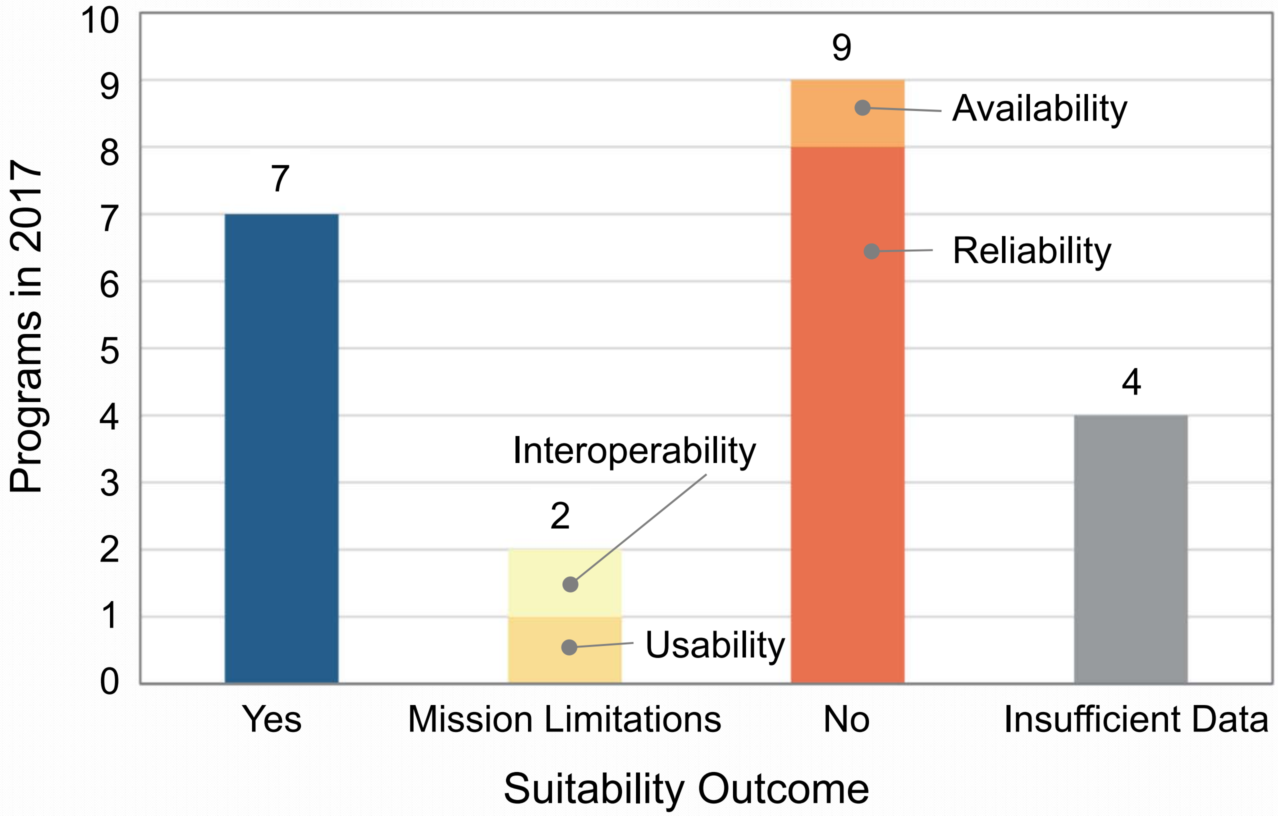
Many systems fail to meet reliability requirements, which affects the overall suitability evaluation of a system



Background: Suitability Determination by Root Cause of Limitations for FY17 Program Reports

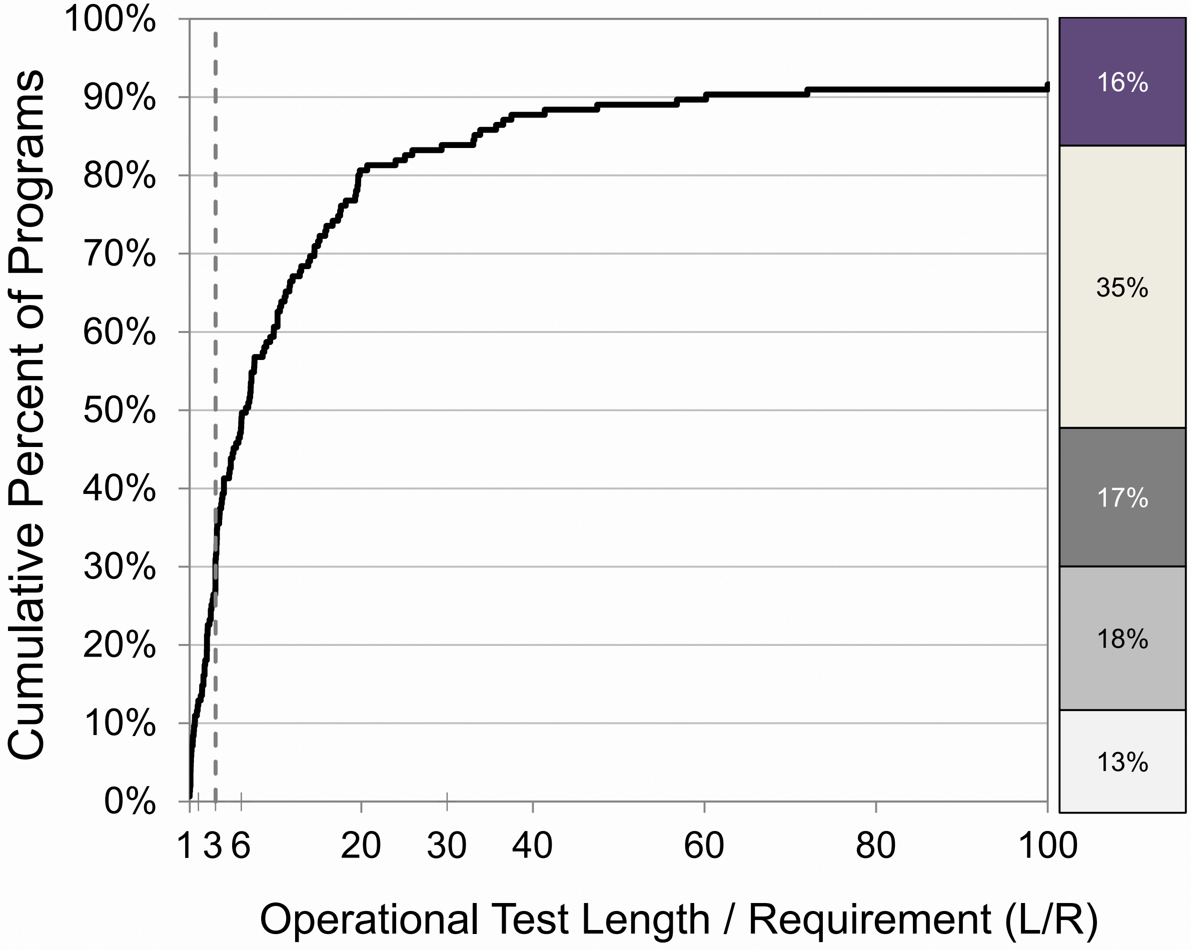
Poor reliability continues to drive suitability assessments

Primary Source of Limitations shown for “No” and “Mixed” Results



Background: Evaluation of Test Adequacy for Assessing Reliability

Operational test lengths can vary and affect reliability assessments



Can reasonably demonstrate reliability with 80% Power / 80% Confidence*
(L/R ≥ 30)

Can reasonably demonstrate reliability with 80% Power / 50% Confidence
(6 ≤ L/R < 30)

Satisfies previous rule of thumb
(3 ≤ L/R < 6)

Does not satisfy previous rule of thumb
(1 ≤ L/R < 3)

Test length shorter than requirement
(L/R < 1)

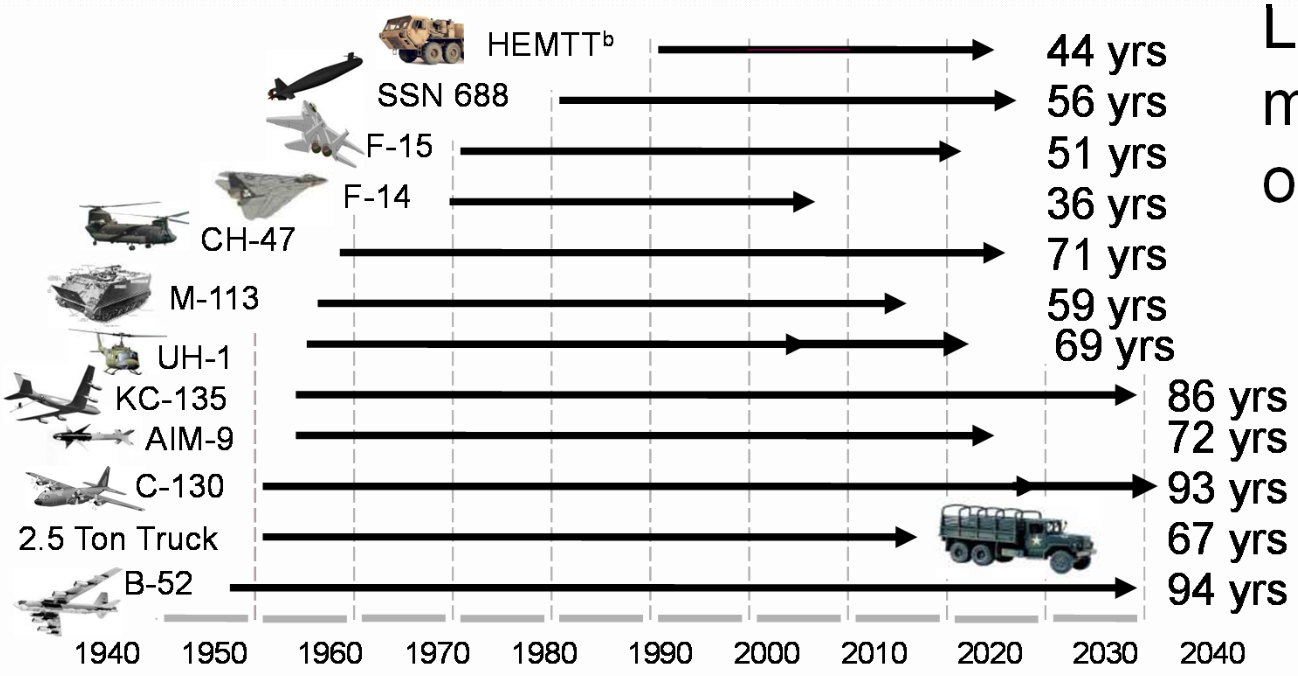
*A design margin of 40% was used to calculate confidence and power

Background: Motivation for Improving DoD System Reliability

System Type	Fraction of Total Cost		
	RDT&E	Procurement	O&S
Ground Combat	4%	28%	68%
Rotary Wing	4%	31%	65%
Surface Ships	1%	39%	60%
Fighter Aircraft	5%	29%	66%

Poor Reliability Drives Costs

a. RDT&E – Research Development Test & Evaluation
b. O&S – Operations and sustainment



Long service life magnifies problem of poor reliability

a. "Improving Reliability," Presentation to IDA by Dr. Ernest Seglie, 17 March 2009.
b. HEMTT – Heavy Expanded Mobility Tactical Truck

Background: DoD Steps Taken to Improve Reliability

2007	− Chairman of the Joint Chiefs of Staff 3170.O1C
2008	− USD(AT&L) RAM Memo (Young Memo) − Defense Science Board Task Force on Reliability − Department of Defense Instruction 5000.02
2009	− DOT&E Reliability Standard Operating Procedures − “Reliability Program Standard” (ANSI GEIA 009) − DOT&E Initiatives Memo
2010	− DOT&E State of Reliability Memo
2011	− USD(AT&L) Directive Type Memorandum 11-003
2012	
2013	− DOT&E Independent Assessment of Operational Testing Memo − Reliability Handbook (TAHB 009) − DOT&E TEMP Guidance
2014	− National Academy Study on Reliability
2015	− 5000.02 Incorporated Reliability Growth Guidance − DOT&E TEMP Guidance 3.0
2016	
2017	− Update to the “Reliability Program Standard” and affiliated Handbook (ANSI GEIA 009 and TAHB 009 Handbook)

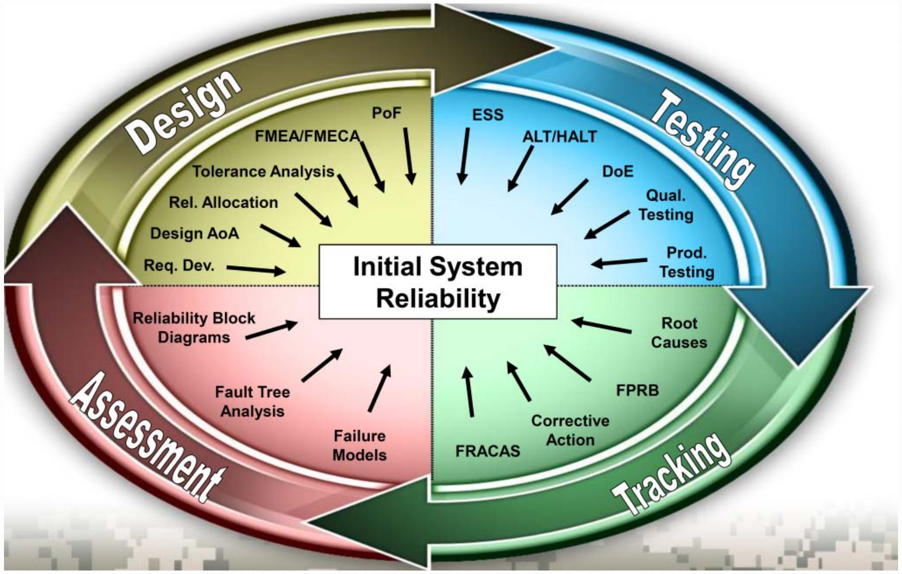
DOT&E – Director, Operational Test and Evaluation

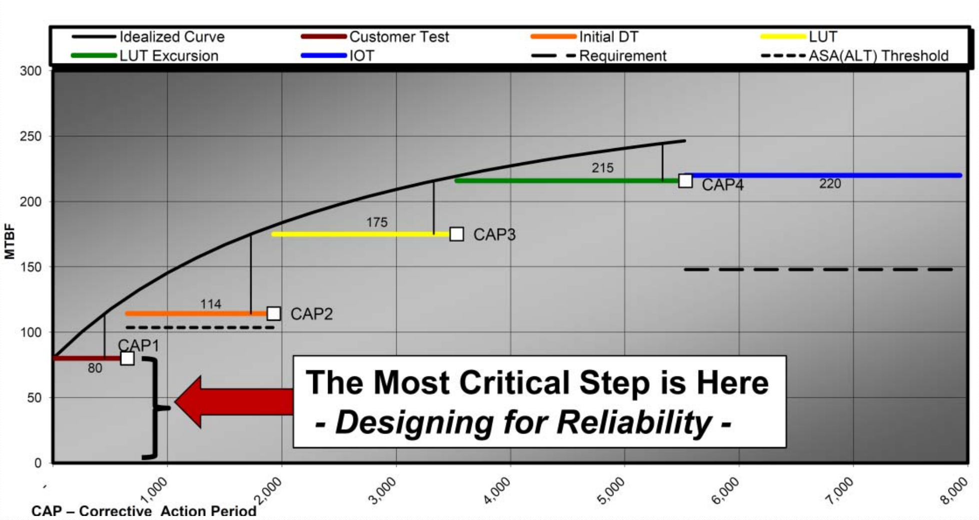
RAM – Reliability, Availability, Maintainability

TEMP – Test and Evaluation Master Plan

USD(AT&L) – Under Secretary of Defense, Acquisition, Technology, and Logistics

Design for Reliability (DfR)





The Most Critical Step is Here

- Designing for Reliability -

Reliability must be designed into the product from the beginning.

- Understand user requirements and constraints
- Design and redesign for reliability
- Produce reliable systems
- Monitor and assess user reliability

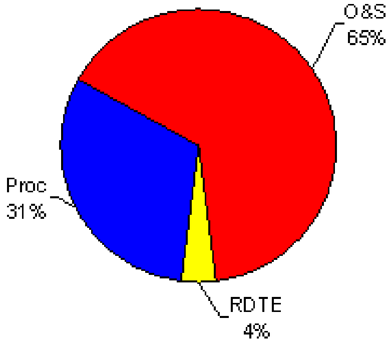
Program Example: Small Investment in Reliability Produced Dramatic Reduction in Life Cycle Cost

HH-60H

Component	MFHBR	APUC
CPU159/A AFCS Computer	582	$180
Auxiliary power systems	2,160	$86
Sections 2/3/4 drive shaft assembly	6,480	$4
CPI820/ASN150 navigational computer	434	$99
Stabiliator amplifier installation	549	$34
MLG drag beam/axle assembly	>10,000	$10
Floor assembly	>10,000	$10
T1360/ALQ144(V) transmitter	582	$52

2.4 Hrs. MTBF

$592.3 M



$6.6M Spent on Reliability

What Changed

Components (AUPC in $ thousand)
MTBF = Mean Time Before Failure
AUPC = Average Unit Production Cost

Reliability Change

(50 Percent Improvement)

Estimated 20-year LCC $M FY03

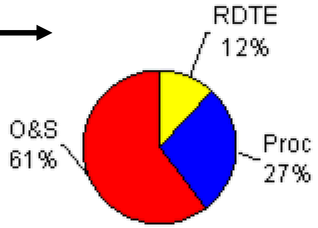
(LCC reduced by approximately 83 percent)

MH-60S

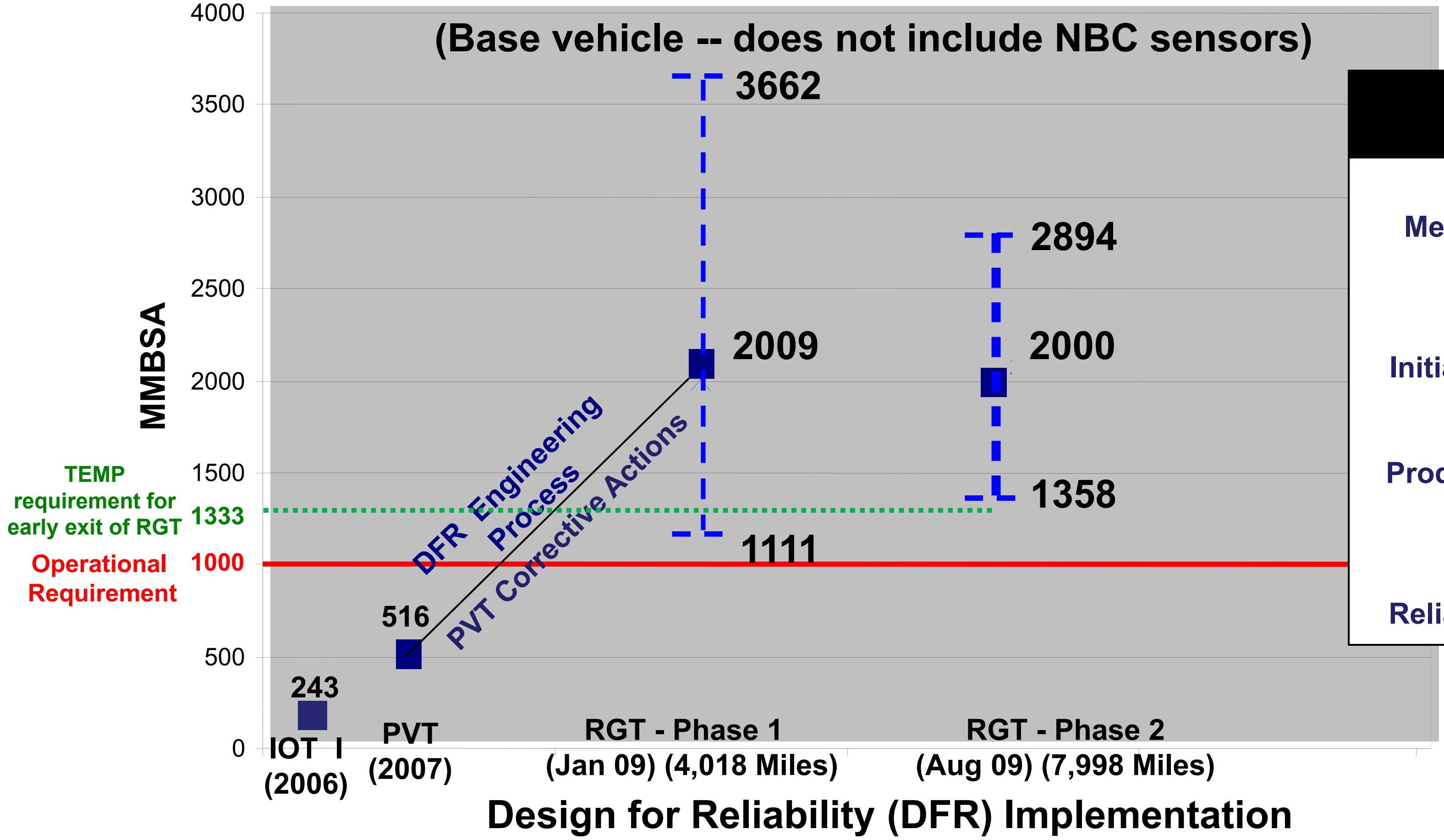
Component	MFHBR	APUC
CPU133/A Digital Computer	1944	$86
Aircraft power unit	10,000	$80
Sections 2/3/4 drive shaft assembly	10,000	$4
CP-2428/A digital data computer	2,236	$84
Stabiliator amplifier installation	1,351	$43
Beam-axle assembly	>10,000	$20
Aircraft floor	>10,000	$20
Light, infrared transmitter	>10,000	$5

3.6 Hrs. MTBF

$107.2 M



Stryker Nuclear, Biological, Chemical Reconnaissance Vehicle Reliability Growth



KEY

MMBSA
Mean Miles Between System Abort

IOT
Initial Operational Test

PVT
Production Verification Test

RGT
Reliability Growth Test

> ***"The amount saved from early discontinuation of Reliability Growth Testing, exceeded the spending on Design For Reliability phase almost 3 times."*** **– J. Ruma, VP Engineering, GD**

Examples of Reliability Best Practices and Lessons Learned in the Department of Defense

Warfighter Information Network-Tactical (WIN-T)

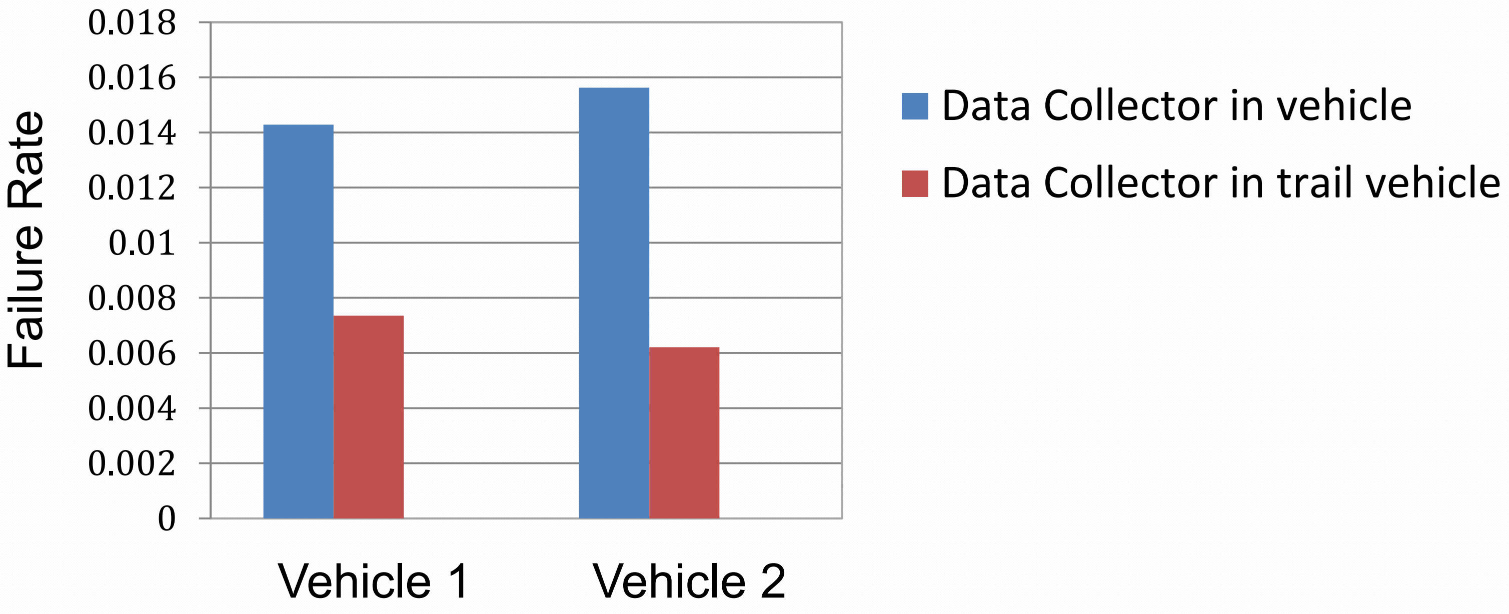
WIN-T is designed as a three-tiered communications system that uses space and terrestrial datalinks to allow soldiers to exchange information in tactical situations

Tactical Communications Node

Point of Presence

Soldier Network Extension

Warfighter Information Network-Tactical (WIN-T)



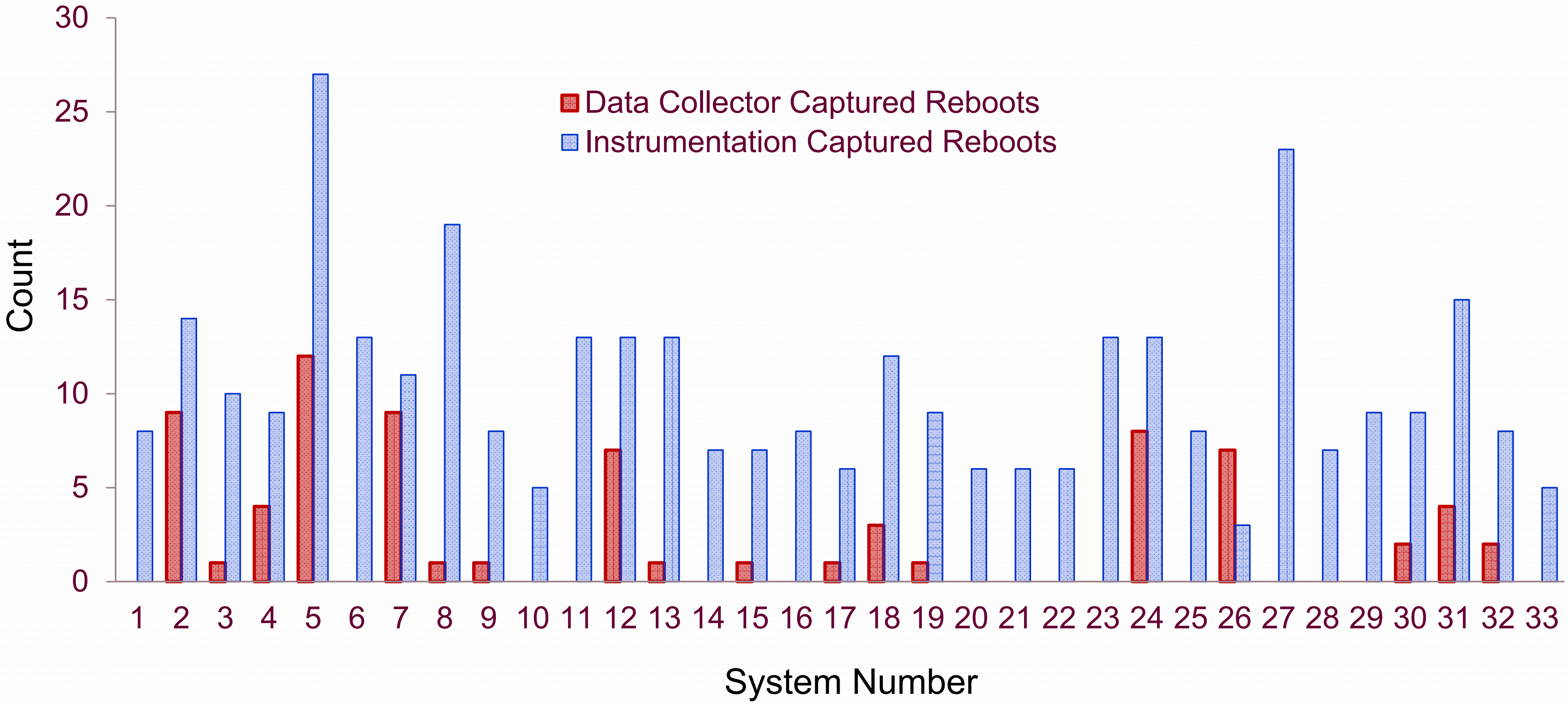
- Data collectors captured more failures when they were in the vehicles during testing, resulting in a higher failure rate

- Data collectors in trailing vehicles did not record all failures

Good data collection procedures are essential for providing the most accurate assessment of system reliability.

Warfighter Information Network-Tactical (WIN-T)

Instrumentation can capture failures that data collectors miss



Good data collection procedures are essential for providing the most accurate assessment of system reliability.

Terminal High Altitude Area Defense System (THAAD)

THAAD is a ballistic missile defense system consisting of mobile launchers with interceptors, a radar, and a command and control unit designed to defend against short-to-intermediate-range ballistic missiles

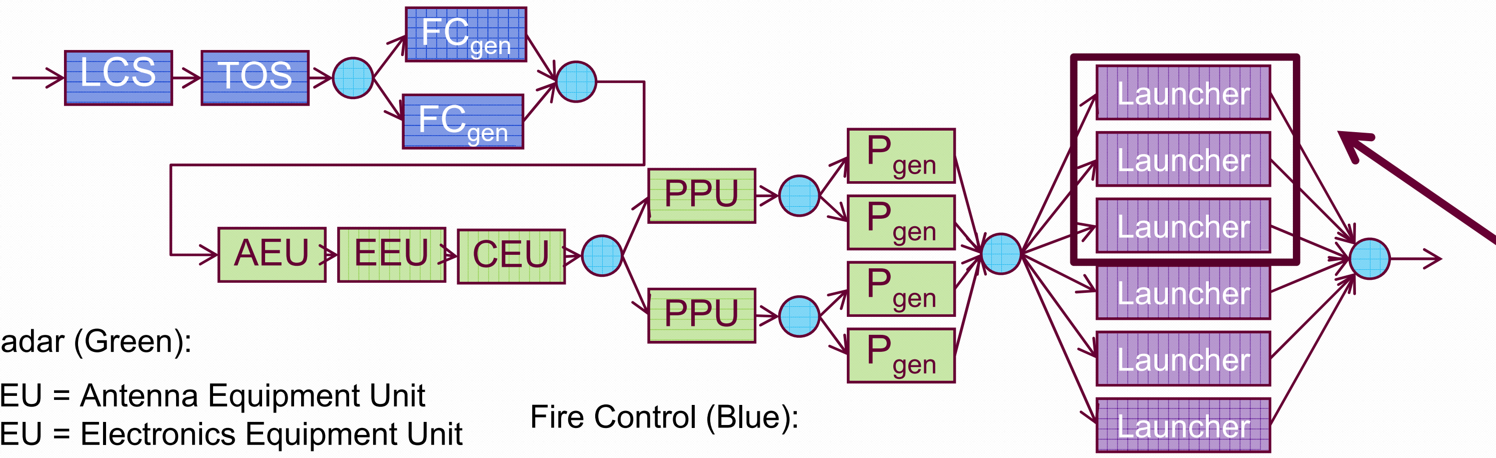
Launcher

Radar

Command and Control Unit

Terminal High Altitude Area Defense System (THAAD)

THAAD Component Configuration



Minimal configurations have three launchers. During flight tests, often only one or two launchers participate (non-operationally realistic)

Radar (Green):

AEU = Antenna Equipment Unit
EEU = Electronics Equipment Unit
CEU = Cooling Equipment Unit
PPU = Prime Power Unit
P_{gen}= PPU Generators

Fire Control (Blue):

LCS = Launch Control Station
TOS = Tactical Operation Station
FC_{gen}= Fire Control Generators

- Early testing considered a scaled-down version of a THAAD fire unit consisting of one to three launchers
- Later testing considered the full THAAD configuration with six launchers

Terminal High Altitude Area Defense System (THAAD)

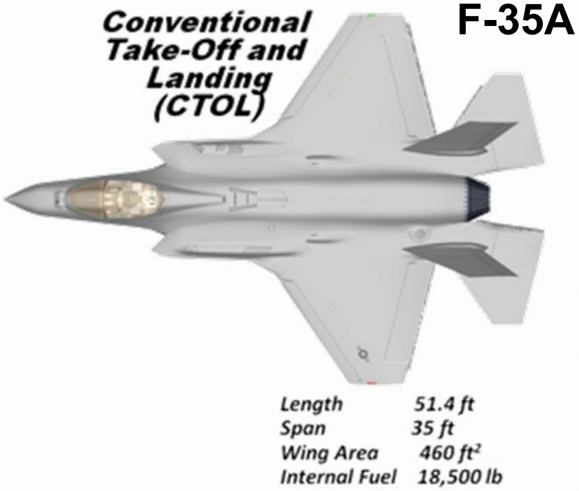
- The full THAAD configuration revealed information that the reduced configurations did not
- Testing found statistically significant reliability differences among the individual launchers, and the crew made decisions based on having to service all six launchers simultaneously
  - Three of the launchers had significantly higher failure rates than the other three, indicating variability between the units
- The resulting reliability estimate for the six configuration launcher was lower that block diagrams and smaller system testing predicted.

Testing the full system leads to greater insights into equipment variability and crew actions.

F-35 Joint Strike Fighter (JSF)

The F-35 JSF is a multi-role fighter aircraft being produced in three variants for the United States Air Force, Navy, Marine Corps, and partner nation services

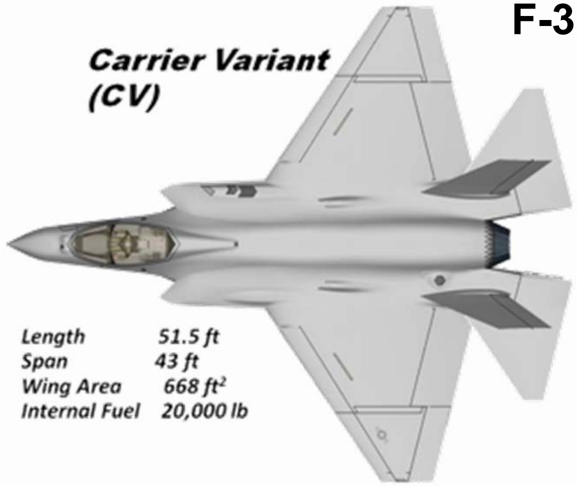
F-35A



Conventional Take-Off and Landing (CTOL)

Length	51.4 ft
Span	35 ft
Wing Area	460 ft²
Internal Fuel	18,500 lb

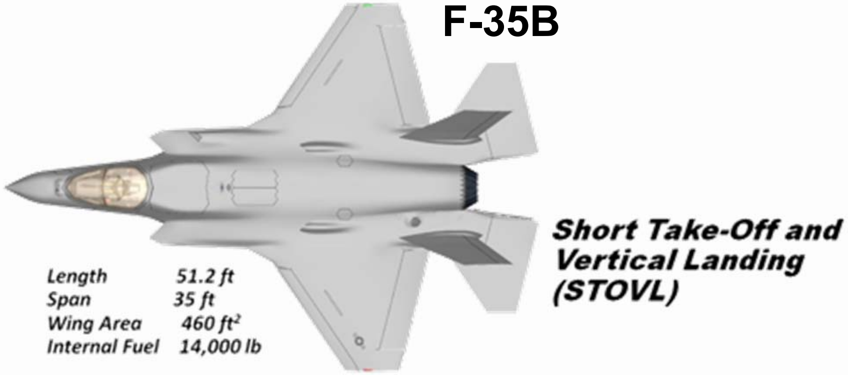
F-35C



Carrier Variant (CV)

Length	51.5 ft
Span	43 ft
Wing Area	668 ft²
Internal Fuel	20,000 lb

F-35B



Short Take-Off and Vertical Landing (STOVL)

Length	51.2 ft
Span	35 ft
Wing Area	460 ft²
Internal Fuel	14,000 lb

F-35 Joint Strike Fighter (JSF)

- With three variants and multiple software block configurations, the JSF program is undertaking simultaneous system design, testing, and production of a modern weapon system to an unprecedented degree
- JSF program releases developmental “blocks” to the field incrementally throughout system development
- As of June 2017, the program had delivered 218 production aircraft to the Services, and 376 aircraft are scheduled to be delivered by the end of 2018; up to 639 aircraft may be fielded by the time operational testing completes in 2020
- Early reliability growth analyses showed mixed results depending on metric and variant, but mostly insufficient growth

F-35 Joint Strike Fighter (JSF)

- Developmental testing and fielded flight operations have surfaced new failure modes, or exacerbated existing ones, as production has continued to increase
- These failure modes are driven not only by the expanding flight envelope, but also by the use of new mission systems capabilities delivered with each block or increment of new capability and flight envelope
- Design changes to fix reliability problems discovered later in concurrent programs are harder for the supply system to support

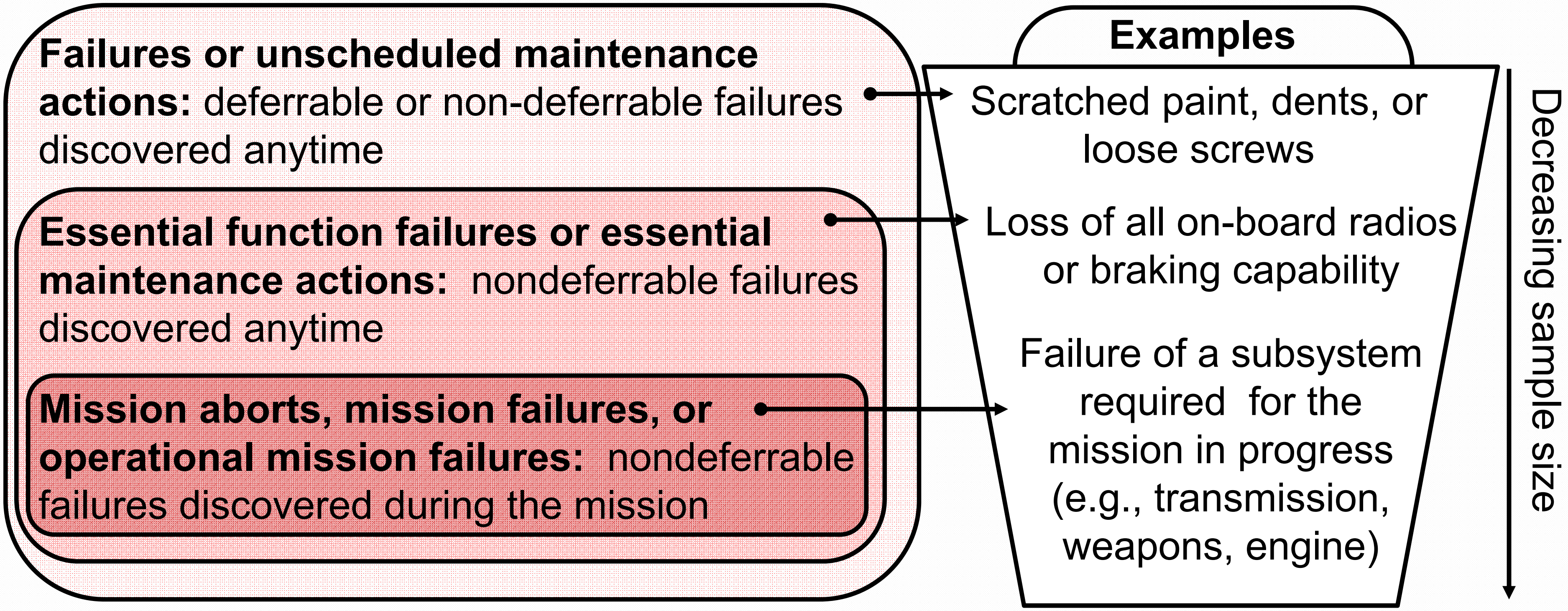
Highly concurrent programs face unique challenges in meeting reliability growth objectives, but such issues can be mitigated by addressing reliability early in the design phase.

AH-64E Apache Helicopter

The Army's AH-64E Apache is modernized version of the AH-64D Attack Helicopter that is being fielded with added capability increments (lots) over time

AH-64E Apache Helicopter

Reliability Requirements in the DoD



Failures or unscheduled maintenance actions: deferrable or non-deferrable failures discovered anytime

Essential function failures or essential maintenance actions: nondeferrable failures discovered anytime

Mission aborts, mission failures, or operational mission failures: nondeferrable failures discovered during the mission

Examples

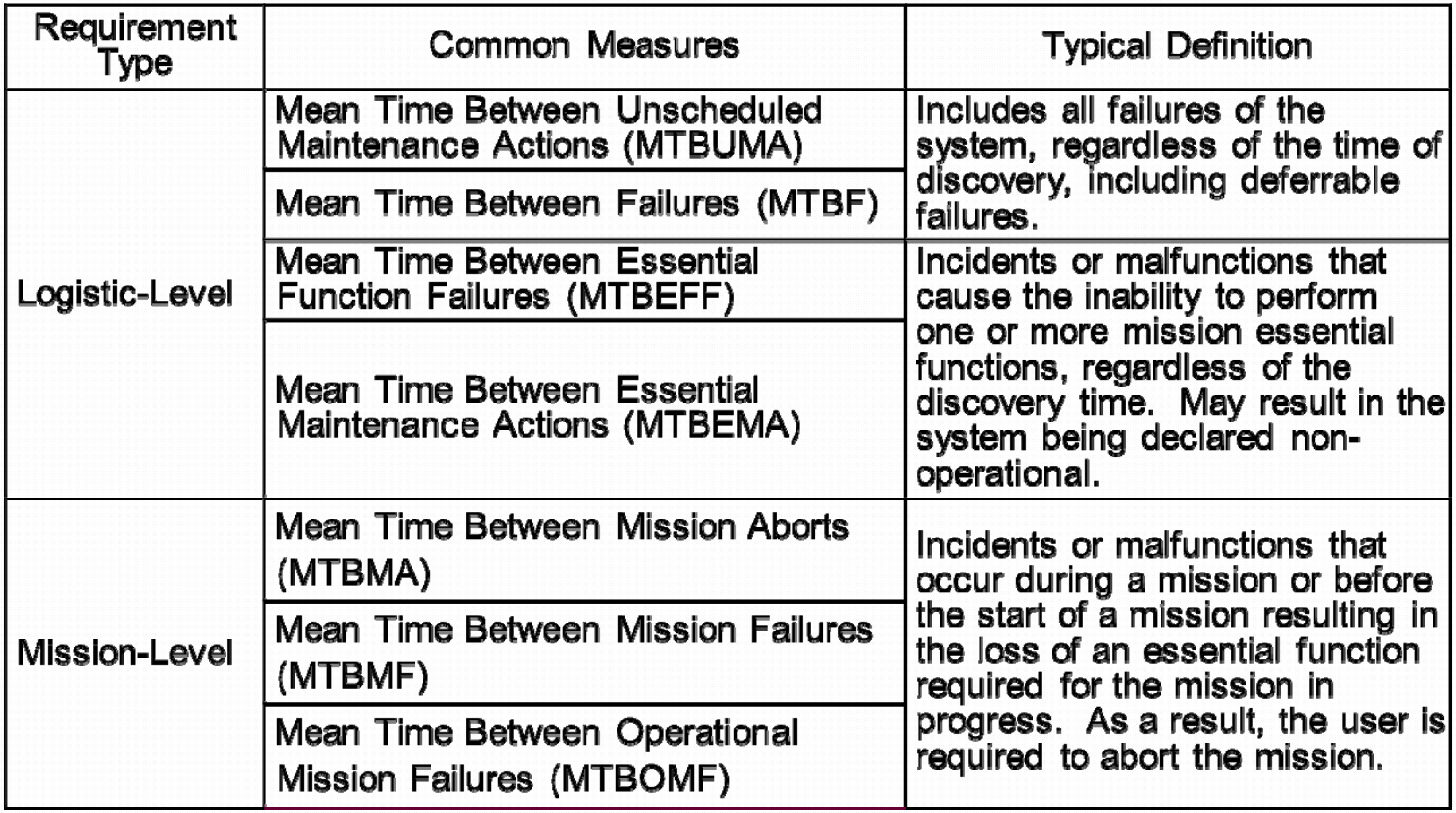
Scratched paint, dents, or loose screws

Loss of all on-board radios or braking capability

Failure of a subsystem required for the mission in progress (e.g., transmission, weapons, engine)

AH-64E Apache Helicopter

During development, the Apache program established reliability requirements for: (1) failures that cause mission failures or aborts, and (2) failures that would not cause aborts but would require essential maintenance actions (EMAs)



Requirement Type	Common Measures	Typical Definition
Logistic-Level	Mean Time Between Unscheduled Maintenance Actions (MTBUMA)	Includes all failures of the system, regardless of the time of discovery, including deferrable failures.
	Mean Time Between Failures (MTBF)	
	Mean Time Between Essential Function Failures (MTBEFF)	Incidents or malfunctions that cause the inability to perform one or more mission essential functions, regardless of the discovery time. May result in the system being declared non-operational.
	Mean Time Between Essential Maintenance Actions (MTBEMA)	
Mission-Level	Mean Time Between Mission Aborts (MTBMA)	Incidents or malfunctions that occur during a mission or before the start of a mission resulting in the loss of an essential function required for the mission in progress. As a result, the user is required to abort the mission.
	Mean Time Between Mission Failures (MTBMF)	
	Mean Time Between Operational Mission Failures (MTBOMF)	

AH-64E Apache Helicopter

- Most DoD programs focus their reliability strategy and requirements exclusively on failures that cause mission aborts
- By focusing also on Essential Maintenance Action (EMA) failures, the Apache program was able to:
  - Identify a larger share of the failure modes that negatively affect availability and maintainability with a minimum amount of testing
  - Improve the program evaluators' ability to assess and track reliability growth during developmental and operational tests
  - More objectively score reliability incidents during developmental testing
- Ultimately, the AH-64E demonstrated high reliability during operational testing

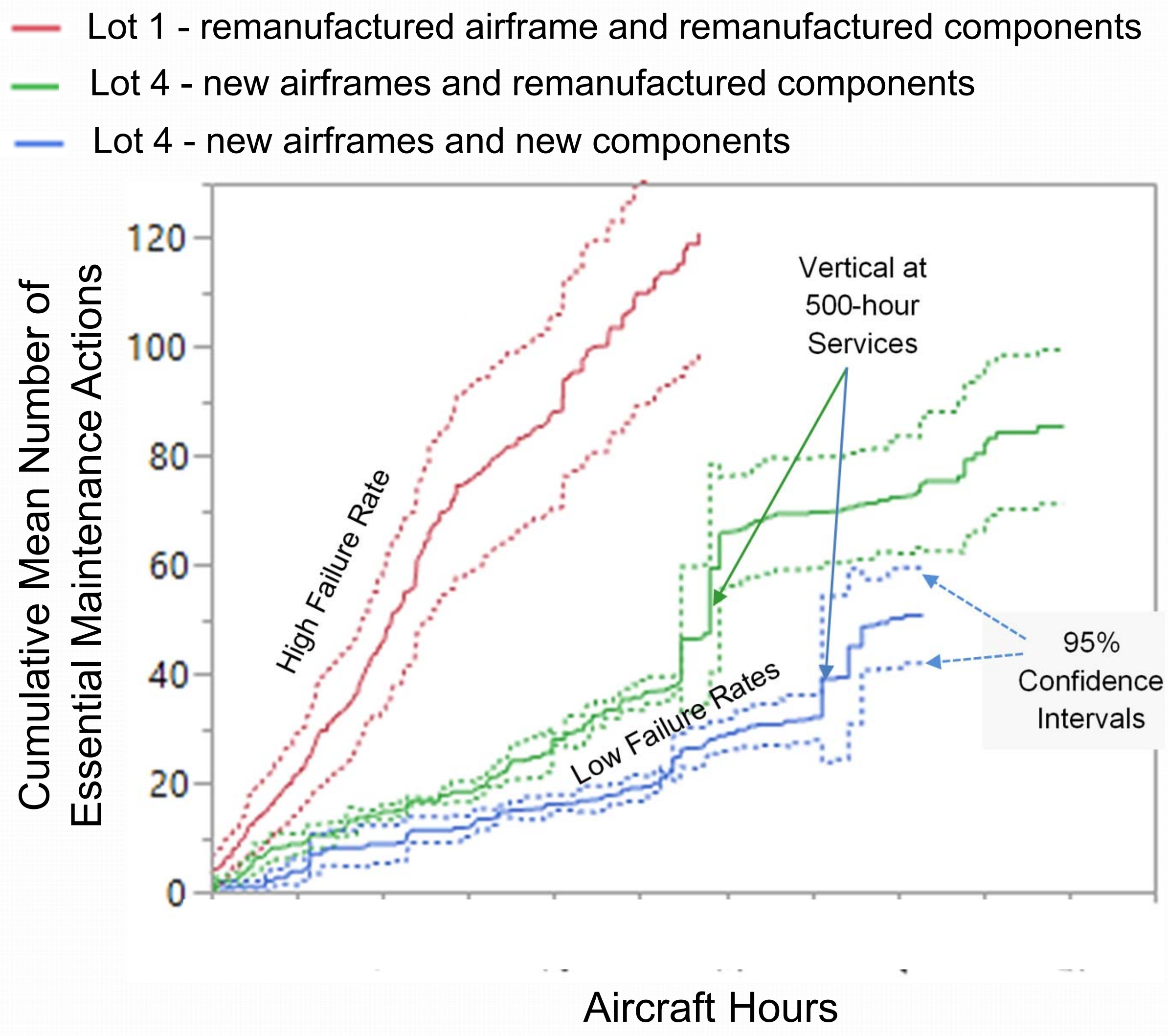
Focusing on both mission and non-mission terminating failures can lead to a more robust reliability growth strategy.

AH-64E Apache Helicopter

- Throughout DoD, a common acquisition approach is to upgrade systems with new hardware and software, and reuse and refurbish old parts to the maximum extent possible
- During production of the of AH-64E aircraft, Boeing and the Army discovered that it would be cheaper to use new airframes than to refurbish and reuse existing airframes
- The first two Lots of AH-64E used refurbished or remanufactured airframes adding new AH-64E fuselages, and re-using AH-64D common hardware
- Beginning with the third Lot, the Army fielded all AH-64E aircraft with new airframes and new AH-64E fuselages, overhauled AH-64D components when possible, added new parts to replace certain life-limited parts, and reused common hardware from the legacy AH-64D

AH-64E Apache Helicopter

Although the change to new airframes was not intended to improve reliability, AH-64E aircraft with new airframes demonstrated a significantly lower failure rate compared to remanufactured aircraft



Acquisition strategies based on reusing old parts may not be optimal from a cost or reliability perspective.

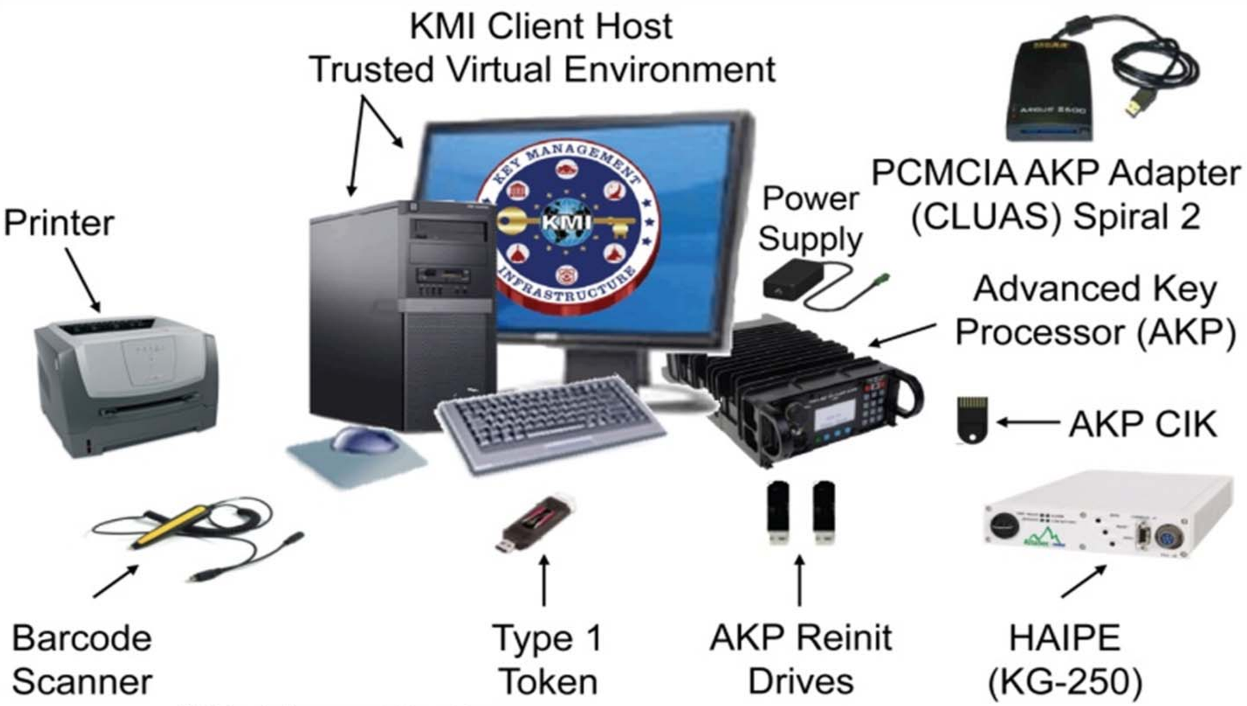
Key Management Infrastructure (KMI)

- KMI is being developed to provide an integrated, consolidated, and automated capability for requesting, producing, delivering, monitoring, and auditing the status of all cryptographic products
- KMI consists of core nodes that provide web operations at sites operated by the NSA, as well as individual client nodes distributed globally, to enable secure key and software provisioning services for the DOD, the Intelligence Community, and other Federal agencies

Key Management Infrastructure (KMI)

The Client Node consists of:

- Commercial computer (Windows Desktop)
- Government High-Assurance Internet Protocol Encryptor
- Advanced Key Processor (AKP) – Custom Hardware
- Type 1 Hardware Token
- Commercial Printer
- Commercial Barcode Scanner



AKP - Advanced Key Processor
CIK -- Crypto Ignition Key
CLUAS -- Card Loader User Application Software
HAIPE -- High Assurance Internet Protocol Encryptor
PCMCIA - Personal Computer Memory Card International Association
Reinit -- Reinitialization

Key Management Infrastructure (KMI)

- During early testing and in its Initial Operational Test and Evaluation (IOT&E), KMI experienced several high-priority software defects and poor token reliability, which resulted in schedule slips and failures during live mission operations

Software Defects per IEEE Standard 12207.2, Annex J

Priority	Number
1	1
2	36
3	48
4	39
5	3
TOTAL	**127**

- Prior to IOT&E, KMI did not have a robust developmental or regression test program

Key Management Infrastructure (KMI)

- The KMI program was ultimately able to address client workstation reliability problems after its Initial Operational Test and Evaluation by establishing a rigorous configuration management and control process
- KMI program conducted developmental and regression testing using mission-based scenarios that were developed by the Operational Test Agencies and vetted by users during operational testing
- In the 2017 Operational Assessment (OA) there were only five system reboots in 6,300 hours of testing, compared to the 149 reboots in 17,919 hours of testing observed in the 2011 OA

Mission-oriented developmental testing and early regression testing is important for discovering high-priority software defects in software systems.

Lessons Learned

DOT&E Case Studies Highlight Best Practices

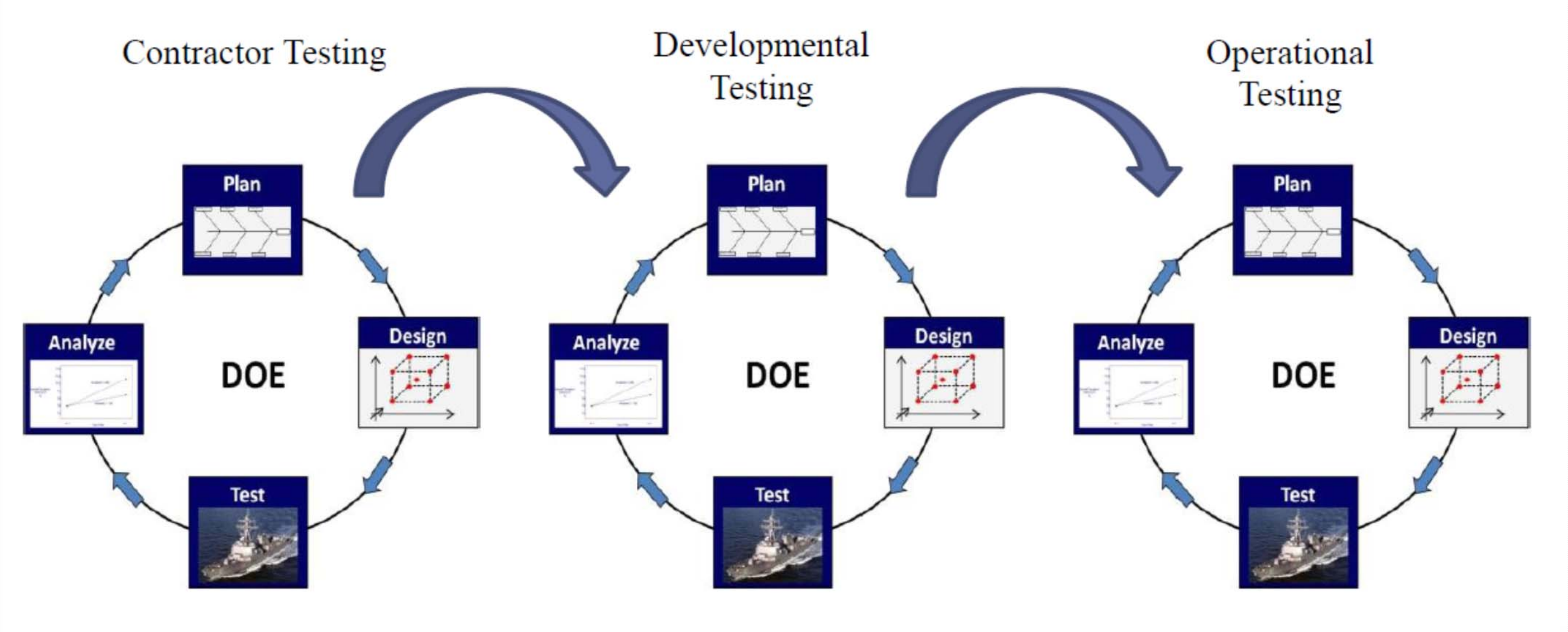
- Investigate and select optimal methods to support collection of failure data, acknowledging that some systems might require dedicated instrumentation
- Focus reliability testing on the fully configured and operationally representative system, as this can offer greater offer insights into equipment variability and crew actions
- Place greater emphasis on addressing reliability in the design phase for programs such as the Joint Strike Fighter with highly concurrent acquisition strategies
- For complex systems, ensure there is adequate time and funding for prototype development and early reliability testing
- Focus reliability efforts on both mission and non-mission terminating failures, as this can lead to a more robust reliability growth strategy
- Consider whether reusing old parts is really optimal, from a cost and reliability perspective, in acquisition strategy analyses
- Conduct mission-oriented testing and robust regression testing during development of software-intensive systems

FUTURE RESEARCH

Future Research: Designing Tests for Reliability

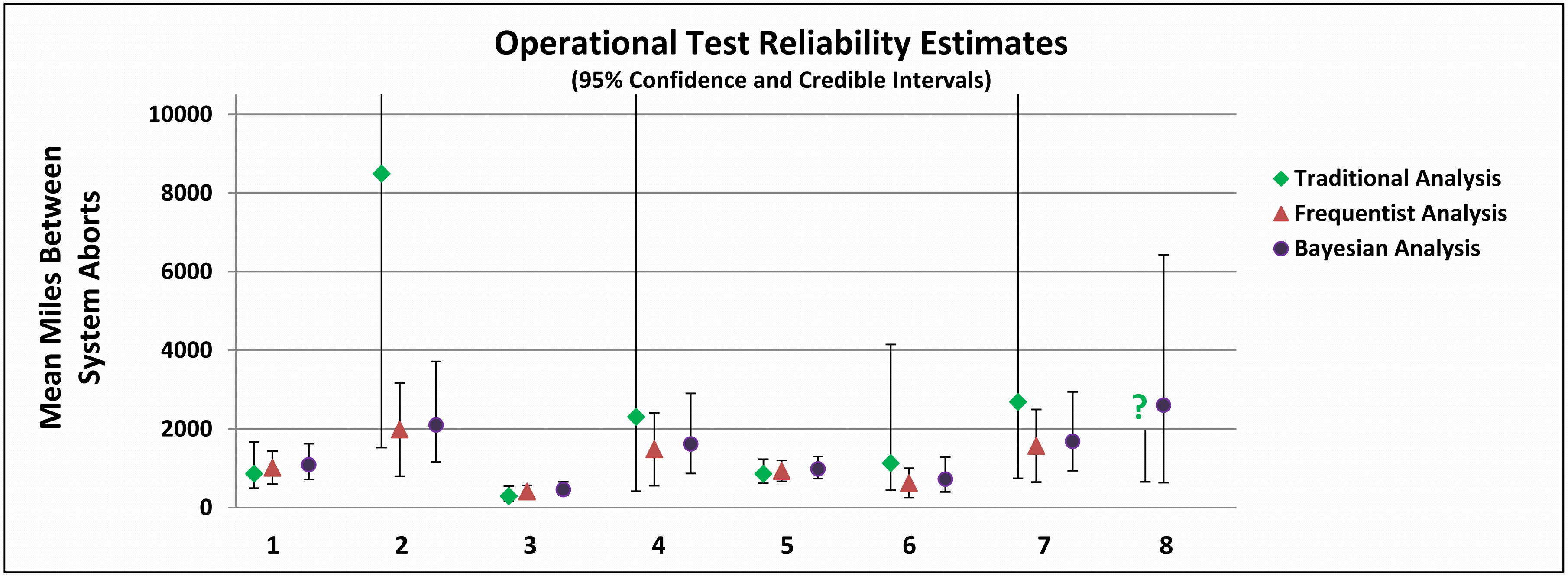
- Standard approach for assessing reliability in DoD is to scope testing to assess whether the program meets its system-level reliability requirement considering all operational test data
- In many instances, it might be desirable to discern differences in reliability between system variants or as a function of test environment (e.g., desert, littoral)
- Design of Experiments (DOE) can be used to systematically investigate whether the system reliability varies significantly based on the variables (factors) of interest
- DOE can also be employed in the design phase to optimize product reliability and save costs
- Incorporate Design for Reliability Efforts in Prototyping and Experimentation

Improved Test Strategies that Account for Use Conditions can Improve Reliability Assessments



Future Research: Methods for Leveraging all Test Data

Bayesian approach can be used to combine data across different variants to produce more realistic estimates and narrower confidence bounds



(Numbers represent different vehicle variants)

Future Research: Methods for Leveraging all Test Data

A Bayesian assurance testing approach to test planning may be used to leverage all test data to reduce test duration and control both consumer and producer risk

Failures Allowed	Bayesian Assurance Test Miles Required 10% Consumer Risk 5% Producer Risk	Classic Operating Characteristic Curve Miles Required 10% Consumer Risk Producer Risk Varies
1	2,940	7,780 58% Producer Risk
2	4,280	10,645 50% Producer Risk
3	5,680	13,362 43% Producer Risk
4	7,120	15,988 37% Producer Risk
5	8,580	18,550 32% Producer Risk

It is an Exciting Time to Work in Reliability and Maintainability!

Big data

Software

Autonomy

Smart censors

Real time monitoring

Preventative maintenance prediction

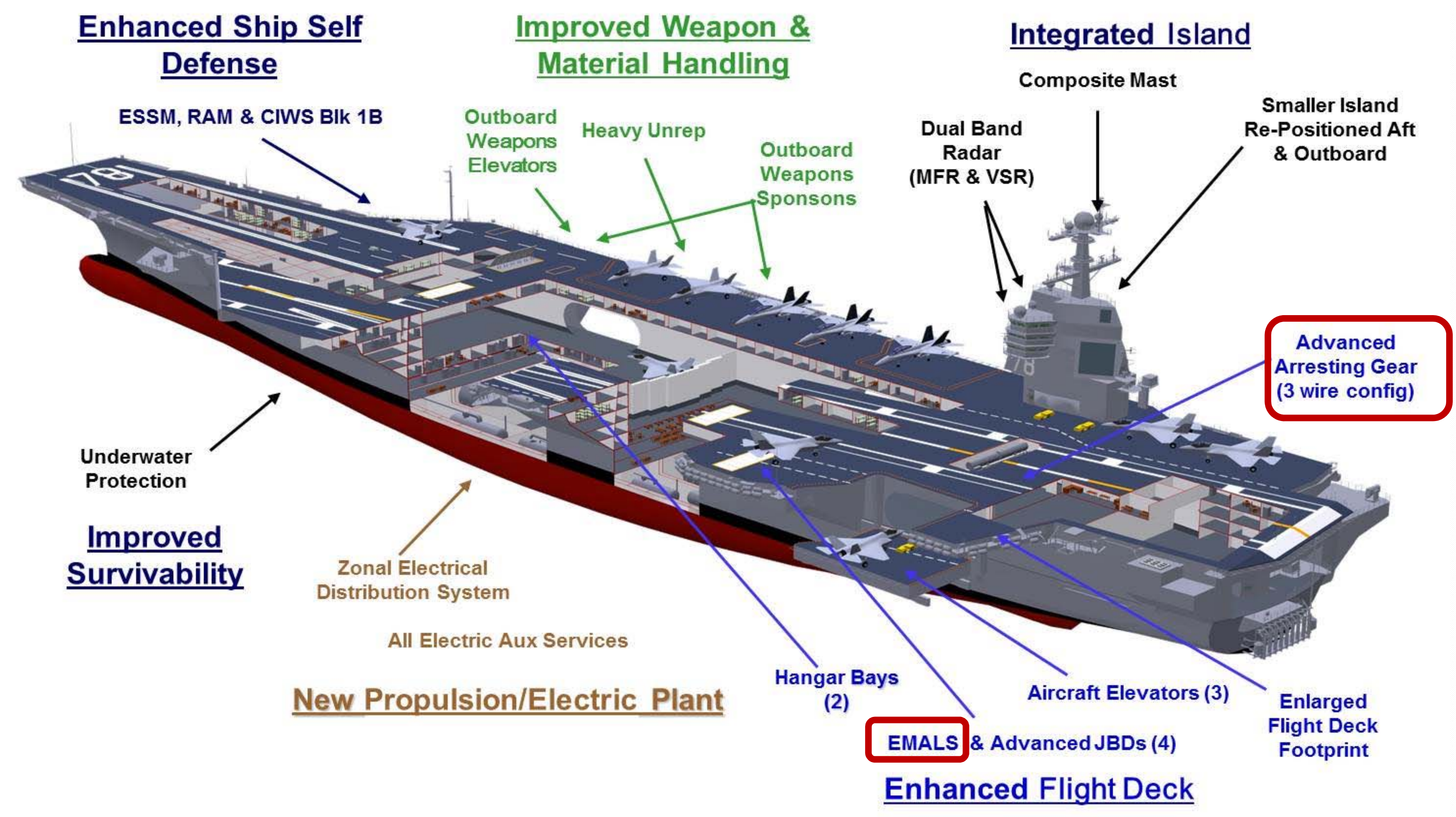
Hardware and software interactions

Reliable hardware enables software

BACKUP

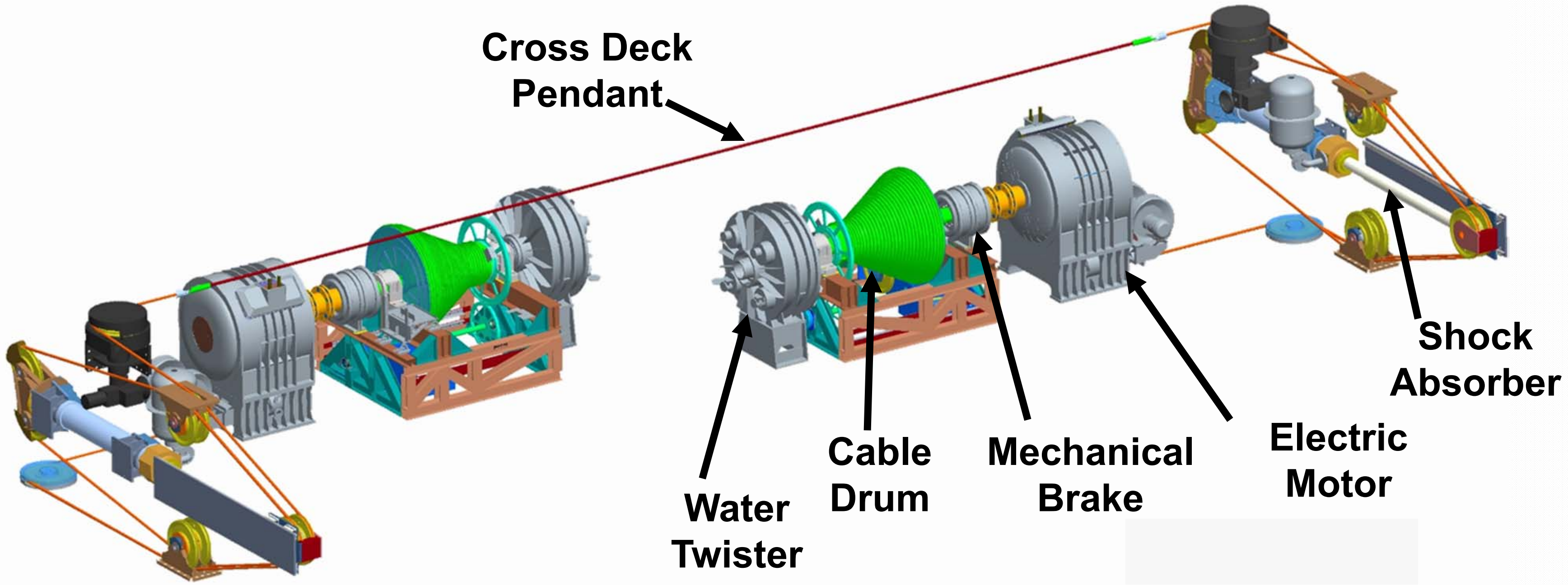
Gerald R. Ford Aircraft Carrier (CVN 78)

The CVN 78 is the Navy's newest aircraft carrier, and it incorporates many new systems such as the Advanced Arresting Gear (AAG) and the Electromagnetic Aircraft Launch System (EMALS) that are required to recover and launch fixed-wing aircraft


Enhanced Ship Self Defense
ESSM, RAM & CIWS Blk 1B
Improved Weapon & Material Handling
Outboard Weapons Elevators
Heavy Unrep
Outboard Weapons Sponsons
Integrated Island
Composite Mast
Dual Band Radar (MFR & VSR)
Smaller Island Re-Positioned Aft & Outboard
Advanced Arresting Gear (3 wire config)
Underwater Protection
Improved Survivability
Zonal Electrical Distribution System
All Electric Aux Services
New Propulsion/Electric Plant
Hangar Bays (2)
Aircraft Elevators (3)
Enlarged Flight Deck Footprint
EMALS & Advanced JBDs (4)
Enhanced Flight Deck

Gerald R. Ford Aircraft Carrier (CVN 78)

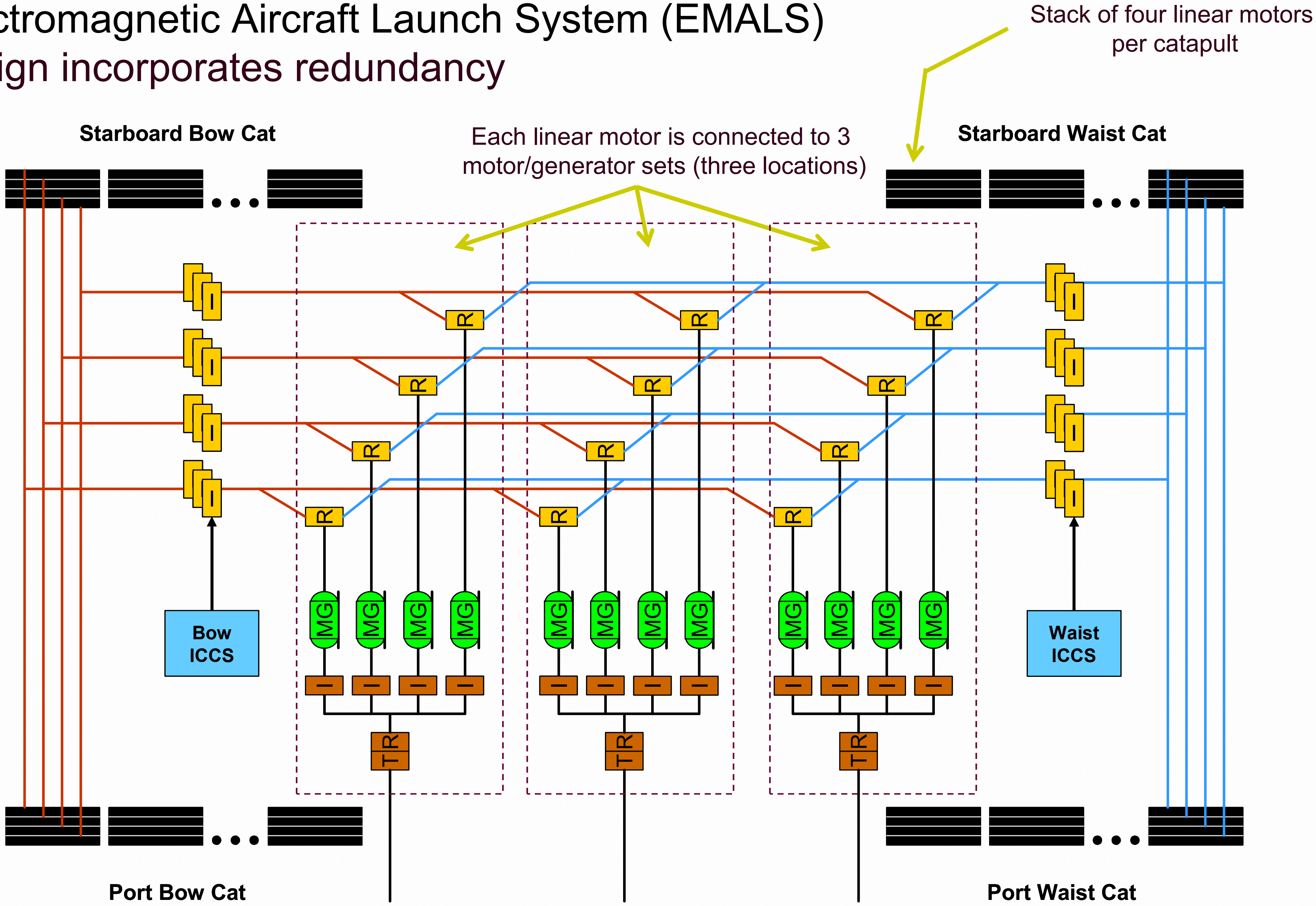
CVN 78 has three Advanced Arresting Gear engines, each engine has one wire (*aka* cross deck pendant)



- Water Twister passively removes bulk of energy
- Electric motor supplements twister, providing nearly constant deceleration
- Mechanical brake backs up twister or motor
- Any two out of the three will provide a safe arrestment
- Cable shock absorber damps “kink waves” from initial landing shock

Gerald R. Ford Aircraft Carrier (CVN 78)

Electromagnetic Aircraft Launch System (EMALS) design incorporates redundancy


Stack of four linear motors per catapult
Starboard Bow Cat
Each linear motor is connected to 3 motor/generator sets (three locations)
Starboard Waist Cat
I
R
MG
TR
Bow ICCS
Waist ICCS
Port Bow Cat
Port Waist Cat

Acronyms on this slide: Invertor (I), Integrated Catapult Control Station (ICCS), Motor Generator (MG), Rectifier (R), Transformer/Rectifier (T/R)

Gerald R. Ford Aircraft Carrier (CVN 78)

DoD review noted that development of the Advanced Arresting Gear (AAG) did not follow reliability engineering best practices in multiple areas

- AAG schedule and funding did not allow for a prototype development and risk reduction effort as was done during Electromagnetic Aircraft Launch System (EMALS) development
- Initial AAG design assumed that the system was merely an integration of existing technologies
- Prototype testing would have revealed a low technology readiness level and uncovered design deficiencies prior to the start of the AAG development phase
- AAG reliability testing was limited to approximately 5,500 arrestments due to schedule and cost constraints

Gerald R. Ford Aircraft Carrier (CVN 78)

- In contrast, the Electromagnetic Aircraft Launch System (EMALS) development effort led to fewer test failures and less redesign than the Advanced Arresting Gear (AAG) development program
- The EMALS program conducted highly accelerated life tests (HALT) and high cycle tests (HCT) that simulated a much larger number of aircraft cycles, providing more rigorous testing and earlier technical discovery
- EMALS also executed a more rigorous reliability growth program compared to AAG

Following reliability engineering best practices improves system reliability.

Terminal High Altitude Area Defense System (THAAD)

- For many DoD systems, the reliability evaluation focuses on the overall system-level reliability
- THAAD is comprised of many complex components that can be organized into various configurations, and reliability block diagrams are used to determine the system-level reliability for a particular configuration
- Therefore, it is important to have credible estimates of component reliabilities to support reliability block diagram modeling. THAAD components required to have high levels of reliability necessitate longer test periods to assess component failure distributions with statistical confidence.

Test periods must be long enough to support statistically credible reliability assessments, especially when evaluating components in complex, configurable systems.