21st Annual National Defense Industrial Association
Systems and Mission Engineering Conference

Leveraging System Safety to Improve System Security

Melinda Reed
Michael McEvilley

Office of the Under Secretary of Defense for Research and Engineering

October 25, 2018

Background

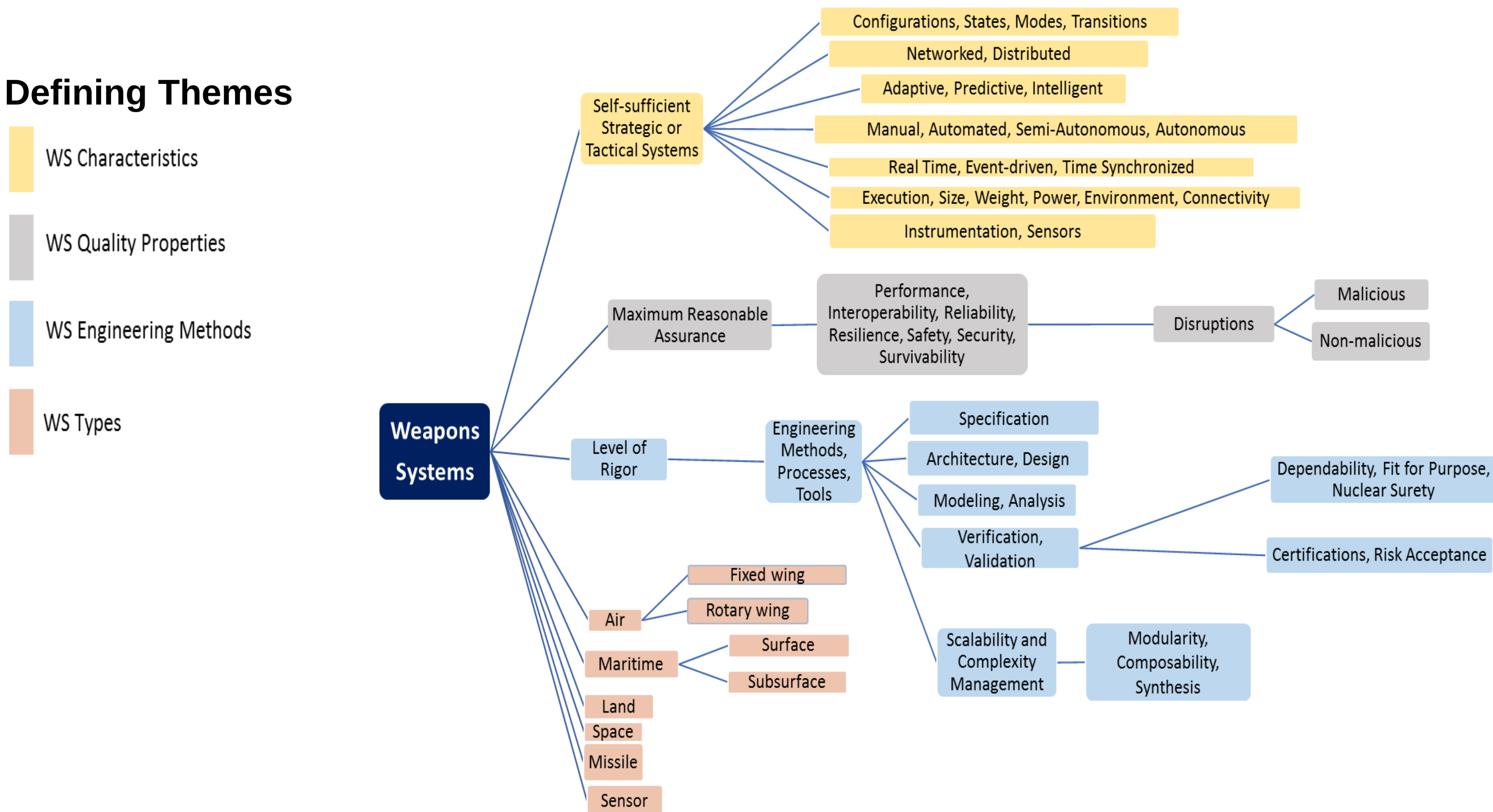
- Action item from the DASD-SE Cyber Resilient Weapons Systems (CRWS) Workshop Series
  - Leverage approaches and methods of system safety to improve systems engineering in response to DoDI 5000.02 Enclosure 14 Section 3.b "Design for Cyber Threat Environments"
- Provide recommendations for action
  - To achieve a strategic vision for increased synergy in safety and security engineering
  - To advance the practice of systems engineering for safe, secure, and resilient weapon systems

Why Safety

- System safety has the following characteristics:
  - Seeks to optimize safety performance within cost, schedule and technical performance constraints
  - Mature methodology in direct response to
    - Technology advances
    - Increased system complexity
    - Increased dependence on software
    - Lessons learned to correlate hazard analysis, mishap, and associated risk
  - Success integrating safety practices into systems engineering processes to enable more effective multidisciplinary collaboration and informed trade space decisions

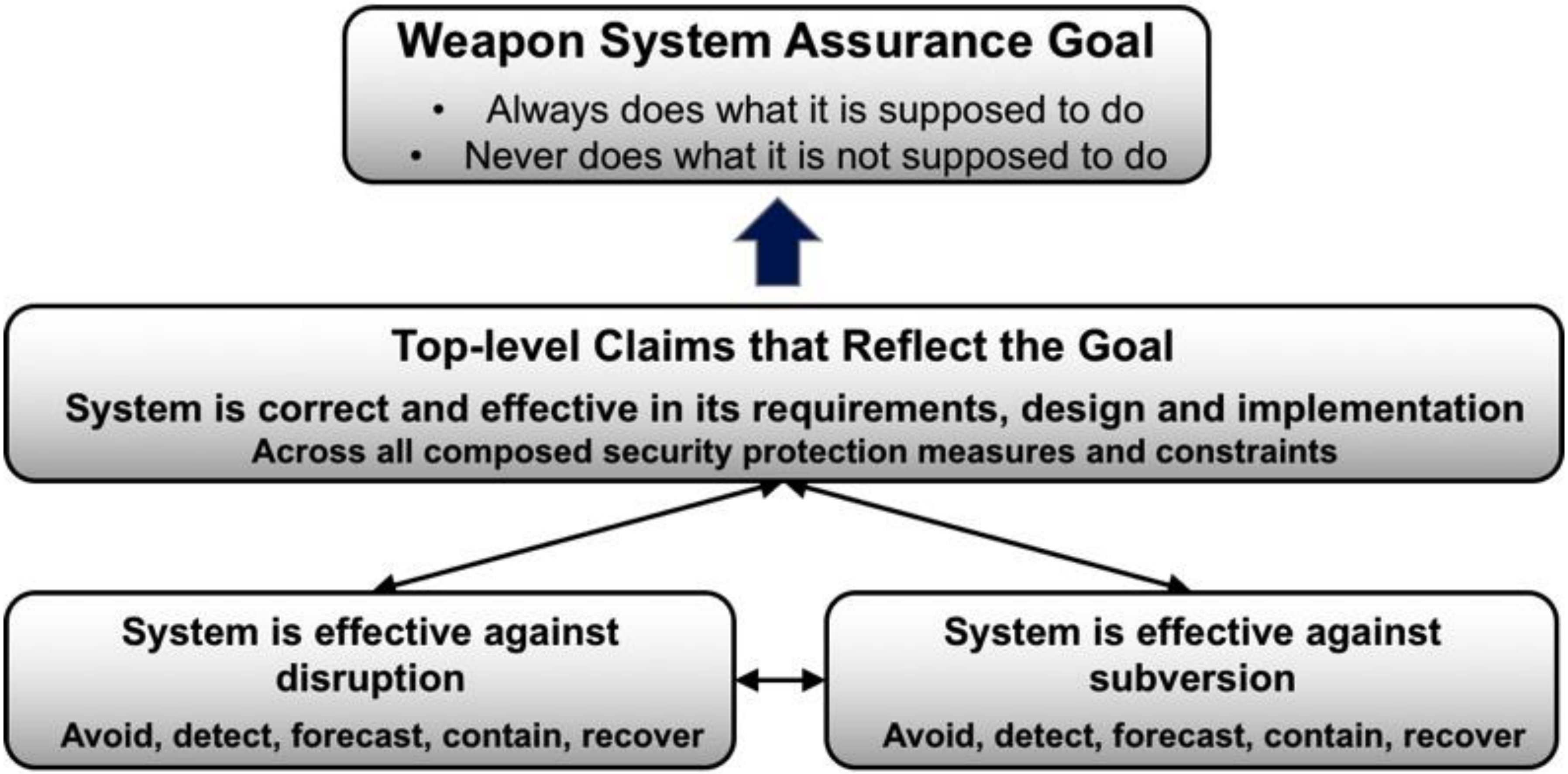
Safety and security share the common objective to prevent, control, and limit the extent of loss and associated loss effects

Leverage System Safety
Key Characteristics of Weapon Systems



Weapon systems deliver lethal force with the intent to cause harm

Weapon System Assurance



Ensure justified confidence in our approaches, decisions, and results

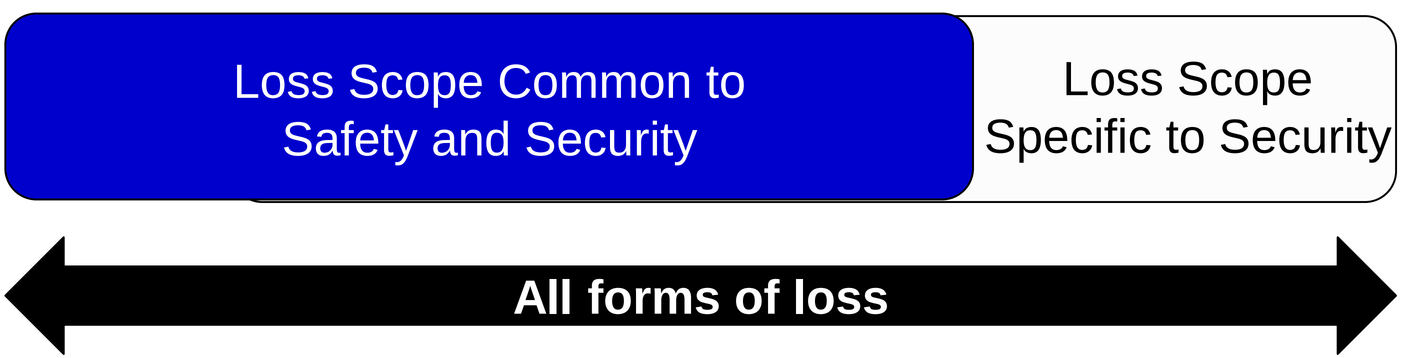
Strategic Vision for Safety and Security Engineering

- Seeks
  - To achieve stronger synergy in the engineering approaches and methods of system safety and system security
  - To improve systems engineering technical and risk and issue management practice
- Embodies
  - Foundational Concepts
  - Key activities
- Based on unacceptable loss effects with safety-relevance and/or security-relevance

Toward synergistic safety and security engineering

Safety and Security Synergistic Working Definitions

- Safety
  - Freedom from those conditions that can cause **death, injury, occupational illness; damage to or loss of equipment or property; or damage to the environment**. [DoD MIL-STD-882E, NASA System Safety Handbook]
- Security
  - Freedom from those conditions that can cause **death, injury, or occupational illness; damage to or loss of equipment or property, damage to the environment**; damage or loss of data or information; or damage to or loss of capability, function, or process. [adapted from DoD, NASA]



Key Concepts of the Strategic Vision



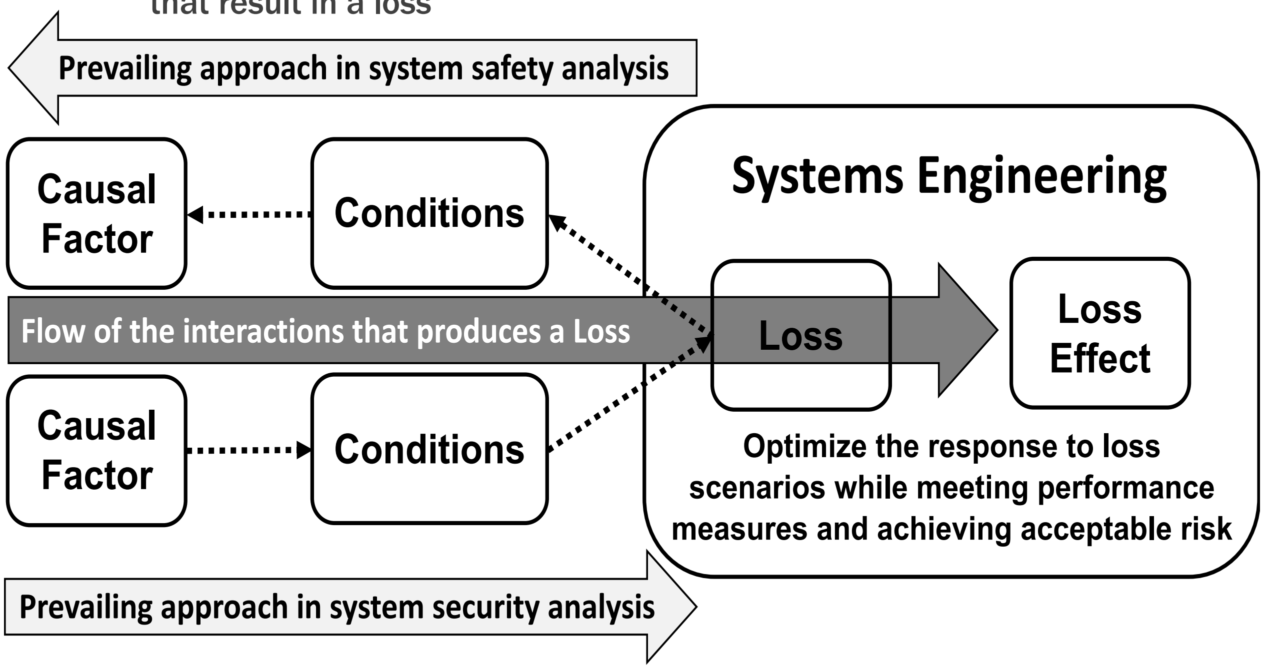


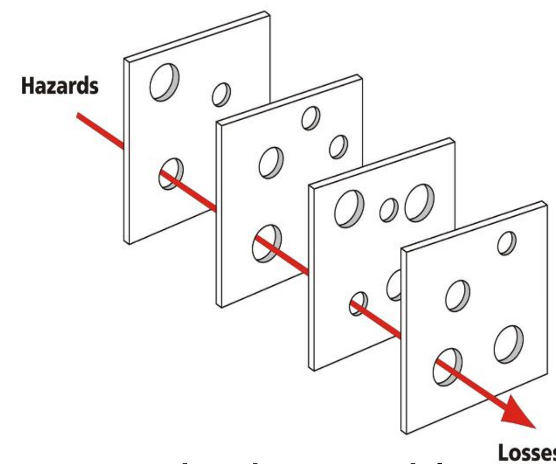
Loss, Loss Effect, Loss Scenario

- **Loss**
  - Degradation, removal, or destruction of an asset (tangible and intangible)
  - Loss drives all safety and security activity
- **Loss effect**
  - Undesirable or unacceptable outcomes associated with a loss
- **Loss scenario**
  - Interaction amongst causal factors and conditions within a specific system and environmental context that result in a loss

Loss scenarios

- Describe the constituent elements and relationships that result in a loss
- Informs analysis to determine response action and to assess the effectiveness of response action
- Informs risk and issue management activity





Swiss cheese model of accident causation

Confidence, Assurance, Risk

- Assurance is grounds for justified confidence that a claim has been or will be achieved [IEEE 15026]
- Confidence is directly related to risk
  - Insufficient confidence about the system may translate to risk that must be identified, assessed, accepted, or mitigated
- Assurance is a trade space
  - Selection of assurance approach must consider the assurance need and assurance ROI
    - Level of confidence sought
    - Limits of the confidence that can be obtained for a given method
    - Cost expended to acquire that confidence
- Rigor is a means to achieve assurance
  - Rigor identifies the formality, thoroughness, accuracy, and precision to be applied to achieve the required level of confidence

Key Activities of the Strategic Vision

Key Activities:
Planning, Analysis, Design

- Planning, Assessment, and Control
  - Plan and execute to optimize system capability – including security protection capability – in meeting design intent and achieving technical performance measures with acceptable risk
- Security Requirements Analysis
  - Stakeholder and system loss-driven security protection needs
    - Prevent loss effects from occurring
    - Minimize extent and/or duration of loss effects
    - Recover from loss effects
- Design for Assurance
  - Builds confidence through proper application of security design principles, concepts, and patterns
  - Design selection and design alteration removes security-related exposure, hazards, and vulnerabilities
  - Accounts for known, unknown, and underappreciated security loss scenarios

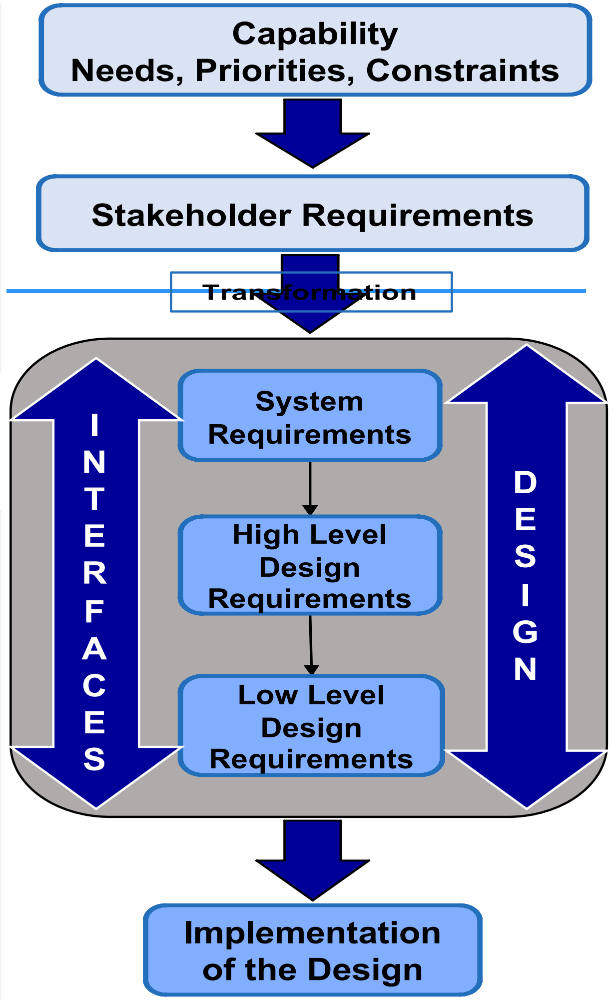
Key Activities:
Synergistic System Security Analysis

NEED

Capability needs, loss concerns, acceptance
- Mission
- System
- Regulatory, statutory, certification, policy
- Assurance

STRUCTURE

System architecture, design, interfaces, interconnections
- Exposure, hazards, vulnerabilities
- Critical functions
  - Mission
  - System
  - Security
  - Safety



ADVERSITY

Loss scenarios
- Causal factors
  - Attack, subversion
  - Error, fault, failure
  - Abuse, misuse
- Conditions
  - Exposure, hazard, vulnerability
- Adversarial threat informed
  - Threat data-dependent
  - Threat data-independent

BEHAVIOR

System function, interfaces, data, interconnections
- Functional, data, control flow interactions
- Interactions not anticipated by the system requirements
- Exposure, hazards, vulnerabilities

Applied with rigor necessary to achieve the targeted level of confidence

Key Activities:
Risk, Issue, Opportunity Management

- Differentiate
  - Risk and issue
  - Known, unknown and underappreciated loss scenarios
- Recognize
  - Limitations of probabilistic risk methods
  - Insufficient confidence is risk

Known Loss Scenarios

- Leverage broad and deep knowledge and experience base to optimally apply solutions with high confidence to nullify causal factors and conditions of exposure, hazard, and vulnerability that lead to loss and the associated effects

Unknown and Underappreciated Loss Scenarios

- Creatively apply broad and deep knowledge and experience to design-in "margins" to reduce the likelihood, duration, or severity of loss and associated effects despite the inherent uncertainty in threat data, attack methods, exposure, hazards and vulnerability

Safety Concepts and Methods

List of Candidates

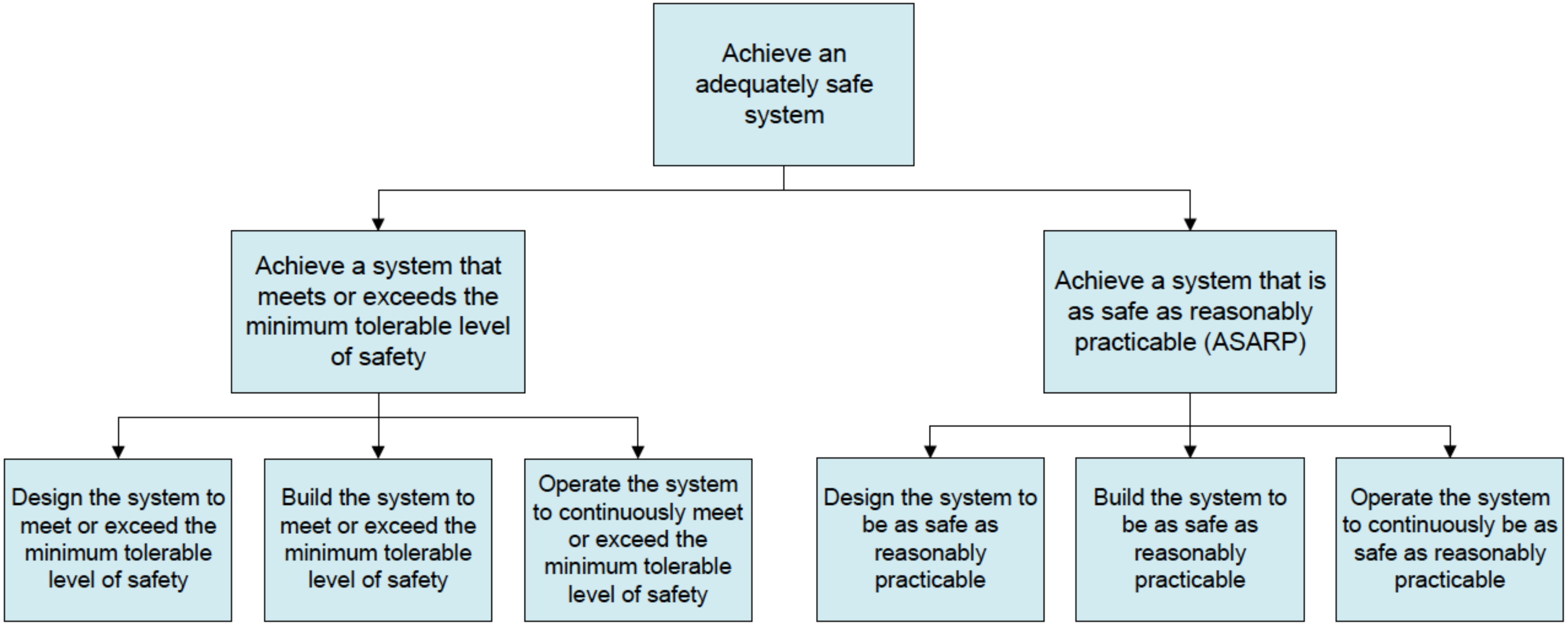
- Adequate safety
  - Minimum threshold level of safety
  - As Safe as Reasonably Practicable (ASARP)
- Design order of precedence
- Structured evidence basis for engineering reviews
  - Risk-Informed Safety Case (RISC)
- Level or Rigor (LoR)
- Risk
  - Actual risk
  - Aggregate risk
  - Assurance deficit
- Risk-Informed Decision Making (RIDM)

Security-Relevant Constructs

- Risk that software contributes to achievement of stakeholder capability objectives
- Limits of probabilistic methods for assessing risk of inherently anomalous, unpredictable, unknown, and underappreciated loss scenarios
- Limits of threat-dependent security analysis due to insufficient volume or quality of threat data
- Inherent uncertainty about the adversary's methods, timing, and objectives for an attack

Adequate Safety

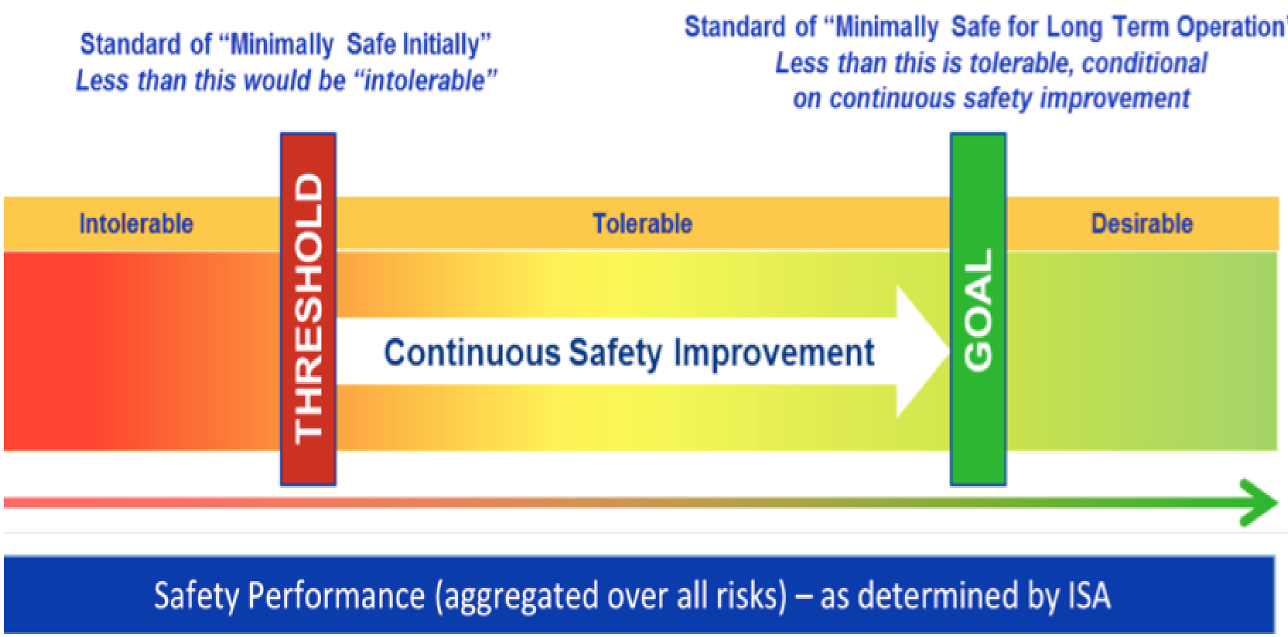
- NASA defines an adequately safe system as one that achieves the following two principles:
  - Meets the minimum threshold level of safety
  - Is As Safe as Reasonably Practicable (ASARP)



Source: NASA

Minimum Threshold Level of Safety

- The minimum threshold has an associated safety goal for expectations about safety growth in the long term. Below the threshold level, the system is considered unsafe.
- Achievement or exceeding the minimum threshold of safety is determined by analysis, operating experience, or a combination of both.
- The minimum acceptable level of safety may shift as the system is operated and information is gained as to its strengths and weaknesses
- Design and operational modifications can be made to improve safety performance toward the goal, while ensuring the minimum threshold is always met or exceeded



Source: NASA

As Safe as Reasonably Practicable (ASARP)

- **Entails weighing the safety performance of the system against the sacrifice needed to further improve it**
  – Any further incremental improvement in safety would require a disproportionate deterioration of system performance in other areas and/or unacceptable or intolerable commitments

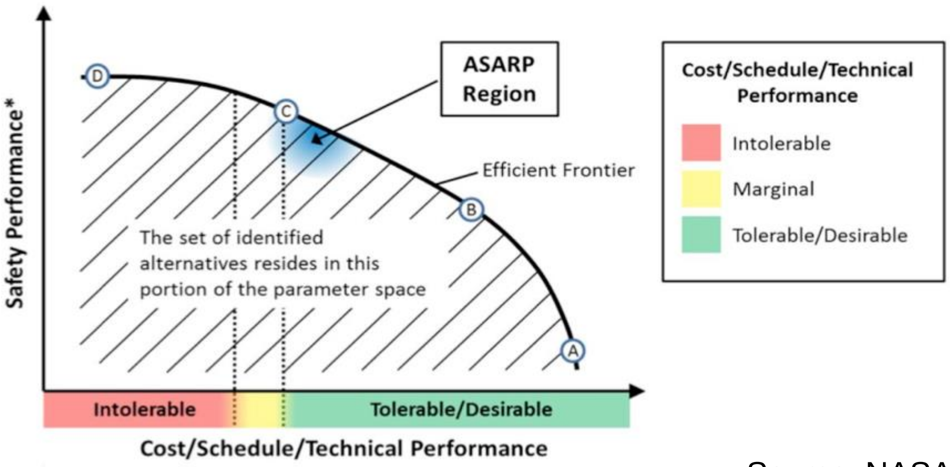
- **Elements of ASARP**
  - **A comprehensive range of alternative means for achieving operational objectives has been identified**
  - **Performance of each alternative has been characterized in sufficient detail to support an assessment of the relative gains and losses in performance that would result from selecting one alternative over another**

(A) Large increases in safety can be achieved by addressing low hanging fruit. Little cost/schedule/technical impact for doing so.

(B) Low hanging fruit has been addressed. but significant increases in safety can still be "bought" without failing to meet cost/schedule/technical performance requirements

(C) Limit of ASARP regime has been reached. Increased safety cannot be "bought" without exceeding tolerable limits of cost/schedule/technical performance

(D) Limit of achievable safety has been reached. Increased safety cannot be "bought" at any cost.

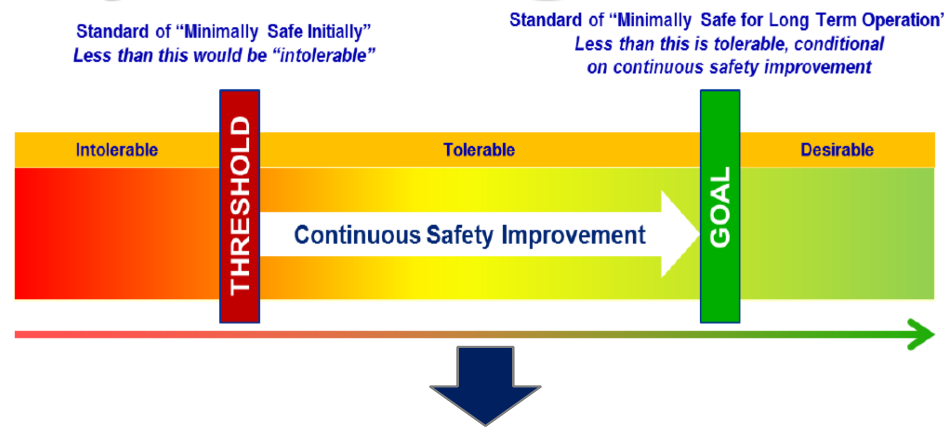


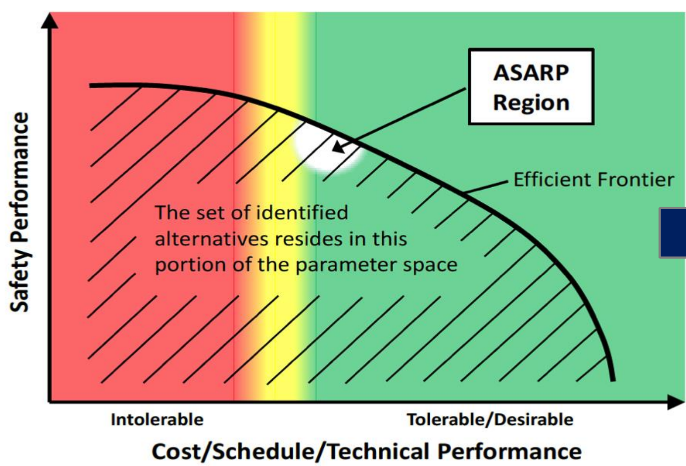
Source: NASA

How the Elements of Adequate Safety Come Together

Source: NASA

The concept of adequate safety provides a basis to optimize the engineering return on investment in achieving system performance objectives inclusive of safety and security performance





System meets minimum tolerable level of safety?

System is ASARP?	Yes	No
Yes	System is adequately safe	System is inherently unsafe
No	System is sub-optimally safe as designed	System is unsafe as designed

ASARP provides a basis for engineering trade space optimization

Design Order of Precedence
(MIL-STD-882E)

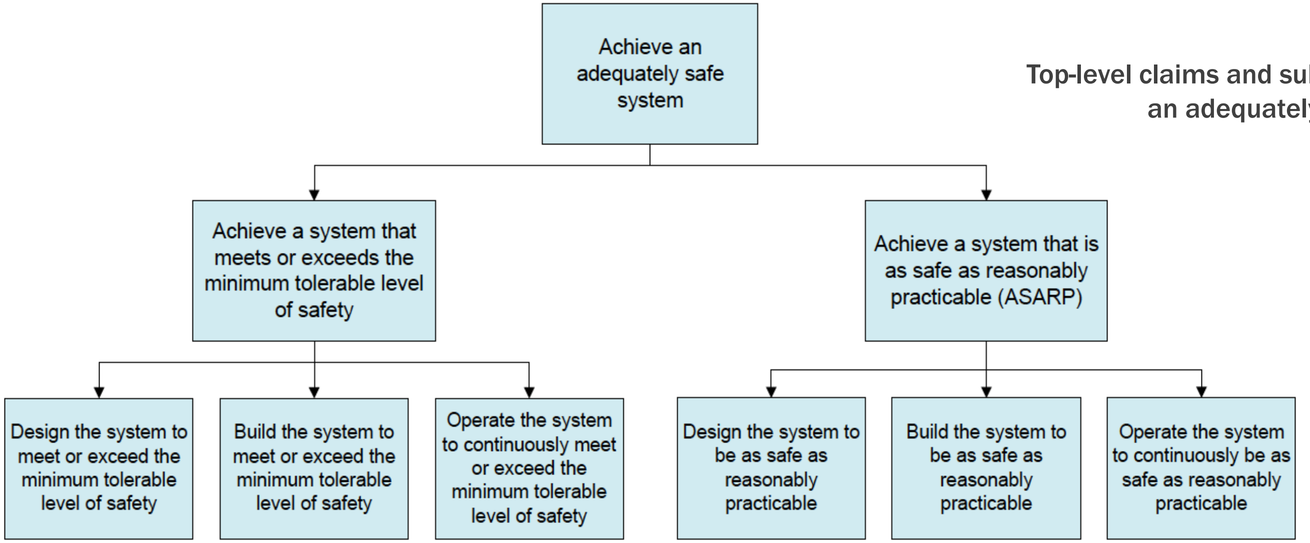
1. Eliminate hazards through design selection
   - Ideally, the hazard should be eliminated by selecting a design or material alternative that removes the hazard altogether
2. Reduce risk through design alteration
   - Design changes that reduce the severity and/or the probability of the mishap potential caused by the hazard(s)
3. Incorporate engineered features or devices
   - Reduce the severity or the probability of the mishap potential caused by the hazard(s) using engineered features or devices
   - In general, engineered features actively interrupt the mishap sequence and devices reduce the risk of a mishap
4. Provide warning devices
   - Detection and warning systems to alert personnel to the presence of a hazardous condition or occurrence of a hazardous event.
5. Incorporate signage, procedures, training, and PPE
   - Incorporate signage, procedures, training, and personal protective equipment (PPE), along with appropriate warnings and cautions

"When a hazard cannot be eliminated, the associated risk should be reduced to the lowest acceptable level within the constraints of cost, schedule, and performance"

MIL-STD 882E System Safety

Structured Evidence Basis for Engineering Reviews

- NASA Risk-Informed Safety Case (RISC) is presented and defended at major milestone reviews as an aspect of milestone decision making
  - Emphasizes that judgments of adequate safety result from a deliberative risk-informed decision-making process
  - Evidentiary in nature
  - Evolves over the course of the system life cycle
  - Continuously evaluated to assess the veracity of the safety claims made therein
- RISC and all associated judgments are based on the top-level claim and principles of adequate safety



Top-level claims and sub-claims of the RISC for an adequately safe system

Source: NASA

Level of Rigor (LoR)

MIL-STD-882E

- A specification of the depth and breadth of software analysis, development, and verification activities necessary to provide a sufficient level of confidence that a safety-critical or safety-related software function will perform as required

- LoR is employed in response to the anomalous and unpredictable nature of software behavior
  - LoR applied is determined by the safety criticality of the software
  - LoR achieves confidence about how the software can be expected to behave under varying conditions and stresses
  - Insufficient confidence about the software drives effort to identify and assess the associated system-level risk, and due to the safety-relevance of the software that risk is managed as safety risk

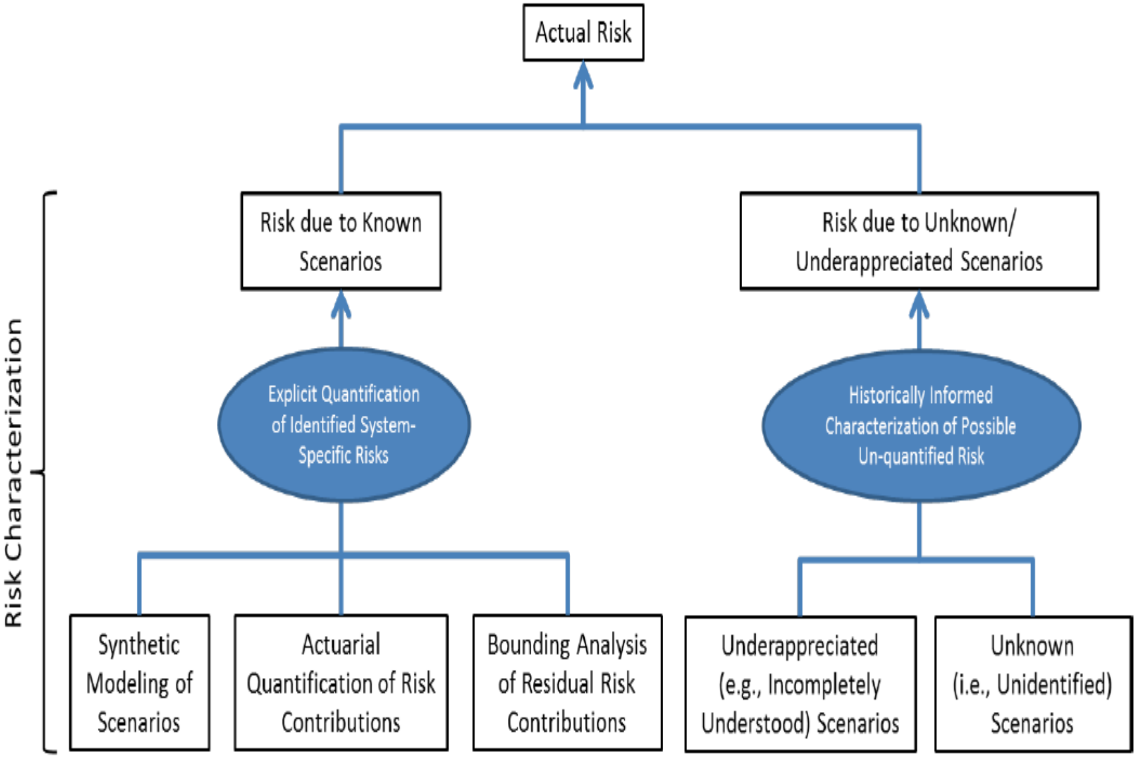
Risk associated with incomplete LoR activities must be identified, assessed, and accepted if not mitigated through other means

NASA Risk Concepts

- Actual risk
  - Reflects the combination of the risk associated with known loss scenarios and risk associated with unknown and underappreciated loss scenarios.
- Aggregate risk
  - System-level risk that results from the accumulation of a set of complex and dynamic interactions, typically the result of multiple factors aligning in ways that may or may not be predicted
  - Recognizes that any argument that system safety has been optimized is more plausible when the risk argument explicitly accounts for risk in its aggregated form and the associated consequences
- Assurance Deficit
  - Identify, assess, and continuously manage the risk associated with the failure to acquire sufficient confidence

Actual Risk

- **Known Loss Scenario**
  - Correctly identified and accurately assessed with respect to its likelihood of occurrence and potential severity of harm or loss
- **Underappreciated Loss Scenario**
  - Correctly identified but for which the likelihood of occurrence and/or potential severity of harm or loss are underestimated
- **Unknown Loss Scenario**
  - Has not been identified and is therefore unknown at the time of analysis.



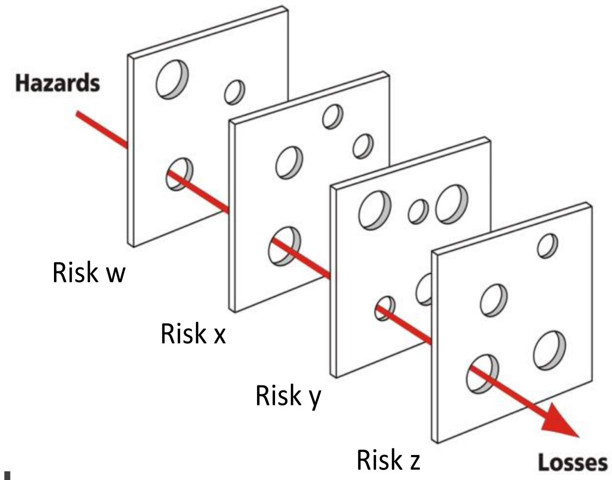
Source: NASA

Loss scenarios affect safety performance

Aggregate Risk

The accumulation of risks from individual loss scenarios that lead to a shortfall in system level safety performance

- An argument that system safety has been optimized is more plausible when it accounts for risk in its aggregated form and the associated loss effects
- Without measures employed in response to aggregated risk concerns, it is not reasonable to expect that safety has been optimized with respect to other technical and programmatic objectives
- System analysis methods that account for aggregate risk yield confidence that the design and employed engineering features and devices
  - Handle risks that have been identified and properly characterized
  - Provide a general, more holistic means for protecting against unidentified or uncharacterized risks



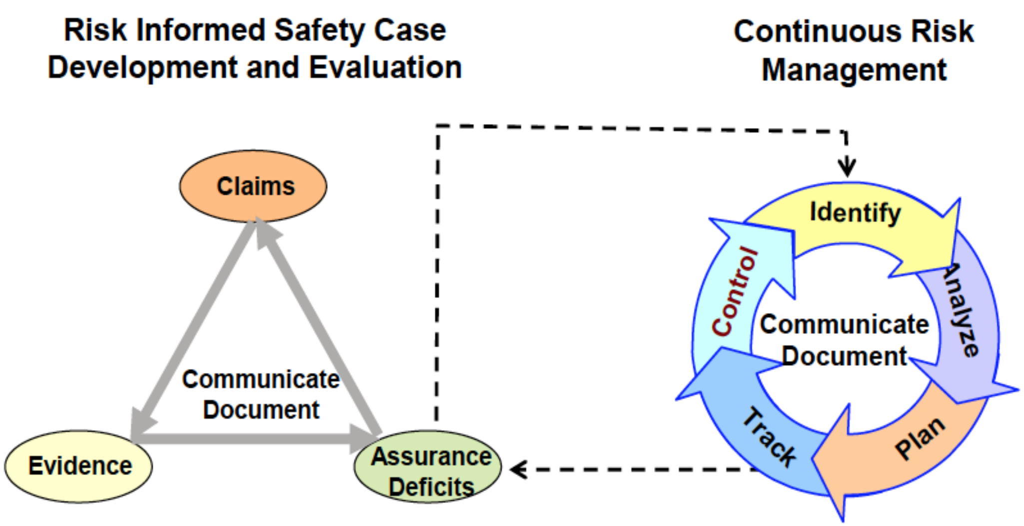
Assurance Deficit

Any knowledge gap that prohibits perfect (total) confidence

- Assurance deficits are caused by
  - Variability or lack of knowledge concerning the data, analysis, or models used to produce the evidence
  - Parameter inputs to models and methods
  - Interpretation of model and methods outputs
- Three cases describe judgments about the confidence achieved
  - Sufficient confidence that the objective is achieved
  - Insufficient confidence that the objective is achieved
  - Sufficient confidence that the objective has not been achieved

Reflected in NASA and DoD safety concepts

- DoD Level of Rigor – any risk associated with the failure to complete LoR activities must be managed
- NASA Risk Tolerance - the lack of confidence that can be accepted in the argument that the system meets an aggregate performance requirement



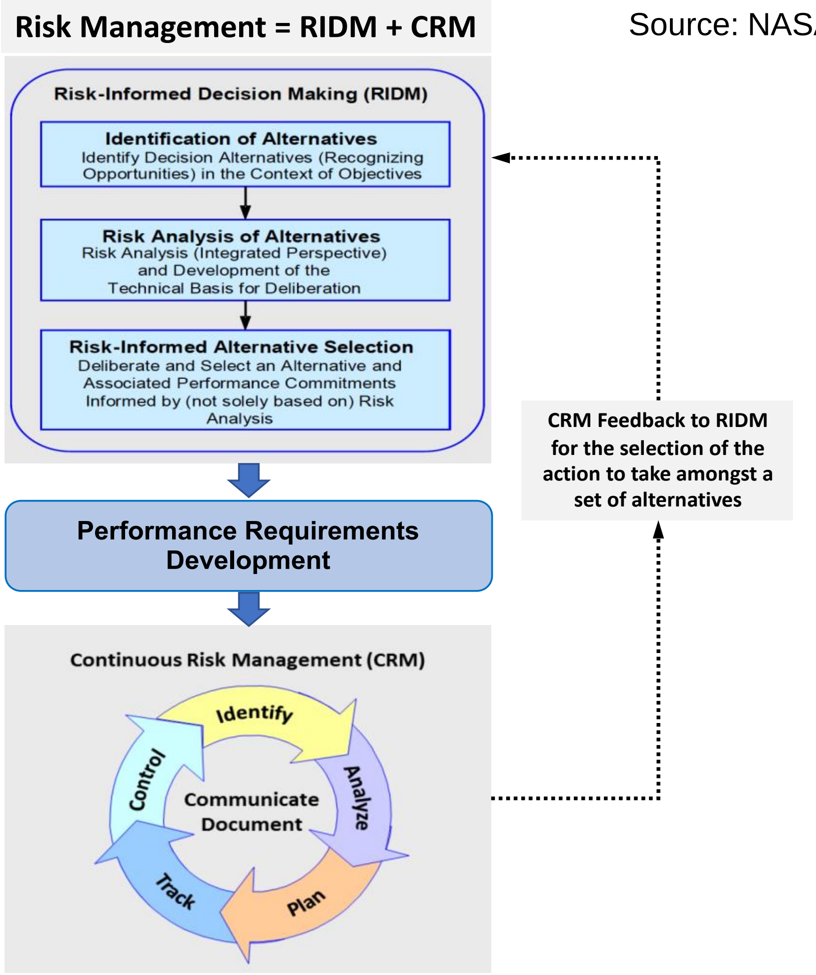
An assurance deficit will have an associated risk that is identified, assessed, and managed if not mitigated by other means

Source: NASA

Risk-Informed Decision Making Continuous Risk Management

- Risk-Informed Decision Making (RIDM)
  - Uses a diverse set of performance measures and other considerations to inform decision making
  - Acknowledges the role that human judgment plays in decisions, and that technical information cannot be the sole basis for decision making
  - Cumulative judgment provided by experienced personnel is an essential element for effectively integrating technical and nontechnical factors to produce sound decisions when faced with multiple competing objectives
- Continuous Risk Management (CRM)
  - Management of risks associated with implementation of designs, plans, and processes
  - CRM provides a disciplined environment for continuously assessing
    - What could go wrong
    - Determining which issues are important to deal with
    - Implementing strategies for dealing with them.

Source: NASA



Effective risk and issue management starts with requirements, as requirements determine almost everything about the risks that need to be managed

Resources

- Defense Science Board (DSB) Task Force, “Report on Cyber Supply Chain,” March 2017
- Department of Defense (DoD) Instruction 5000.02, “Operation of the Defense Acquisition System,” January 7, 2015, Incorporating Change 2, Effective February 2, 2017
- Department of Defense (DoD), “Risk, Issue, and Opportunity Management Guide for Defense Acquisition Programs,” Office of the Deputy Assistant Secretary of Defense for Systems Engineering, January 2017
- Department of Defense (DoD), “Standard Practice System Safety,” MIL-STD-882E, May 11, 2012
- Department of Defense (DoD), “Risk, Issue, and Opportunity Management Guide for Defense Acquisition Programs,” Office of the Deputy Assistant Secretary of Defense for Systems Engineering, January 2017
- International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) 15026-1:2013, “Systems and software engineering -- Systems and software assurance - Part 1: Concepts and Vocabulary,” November 2013
- International Organization for Standardization/International Electrotechnical Commission/Institute of Electrical and Electronics Engineers (ISO/IEC/IEEE) 15288:2015, “Systems and software engineering — Systems life cycle processes,” May 2015
- National Aeronautics and Space Administration (NASA), “System Safety Handbook Volume 1, System Safety Framework and Concepts for Implementation,” NASA/SP-2010-580, Version 1.0, November 2011
- National Aeronautics and Space Administration (NASA), “System Safety Handbook Volume 2: System Safety Concepts, Guidelines, and Implementation Examples,” NASA/SP-2014-612, Version 1.0, November 2014
- National Aeronautics and Space Administration (NASA), “Risk-Informed Decision Making Handbook,” NASA/SP-2010-576, Version 1.0, April 2010
- National Aeronautics and Space Administration (NASA), “Risk Management Handbook,” NASA/SP-2011-3422, Version 1.0, November 2011
- Homayoon Dezfuli, et al., “The Role of NASA Safety Thresholds and Goals in Achieving Adequate Safety,” Probabilistic Safety Assessment and Management (PSAM 12), June 2014
- Richard Hawkins, et al., “A New Approach to Creating Clear Safety Arguments,” University of York and University of Virginia, 2010
- M. McEvilley, G. Vecellio, “Strategic Vision for Safety and Security in Weapon Systems Engineering,” MITRE Technical Report/MTR 180261, July 2018

DoD Research and Engineering Enterprise
Solving Problems Today – Designing Solutions for Tomorrow










DoD Research and Engineering Enterprise
https://www.acq.osd.mil/chieftechnologist/

Defense Innovation Marketplace
https://defenseinnovationmarketplace.dtic.mil

Twitter
@DoDInnovation

For Additional Information

Melinda Reed

ODASD, Systems Engineering

571.372.6562 | melinda.k.reed4.civ@mail.mil

Michael McEvilley

The MITRE Corporation

703.472.5409 | mcevilley@mitre.org