Evolution of System Safety at NASA as related to Defense-in-Depth

Defense-in-Depth Inter-Agency Workshop

Rockville, MD
August 26, 2015

Homayoon Dezfuli, Ph.D.
NASA Technical Fellow for System Safety
Office of Safety and Mission Assurance
NASA Headquarters

Disclaimer and Acknowledgments

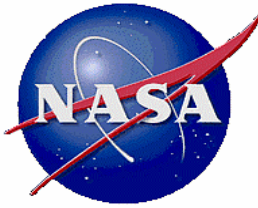
- Opinions expressed in this presentation are not necessarily those of NASA
- Most of the present discussion is based on work performed by the Office of Safety and Mission Assurance in conjunction with NASA System Safety Handbook, Volume 1 (NASA/SP-2010-580) and NASA System Safety Handbook, Volume 2 (NASA/SP-2014-612
- The presentation has benefited from valuable support from Mr. Chris Everett of Information Systems Laboratories

My Understanding of NRC's Defense-in-Depth Philosophy (Anchor Point)

- **Defense-in-depth is a basic element of the NRC's safety philosophy**
- **No single definition exists, but a review of the literature conveys a general consensus:***
  - **The defense-in-depth philosophy is a balance among accident prevention, accident mitigation, and the limitation of the consequences of an accident**
  - **Briefly stated, this philosophy**
    - requires high quality in the design, construction, and operation of nuclear plants to reduce the likelihood of malfunctions in the first instance
    - recognizes that equipment can fail and operators can make mistakes, thus requiring safety systems to reduce the chances that malfunctions will lead to accidents that release fission products from the fuel
    - recognizes that, in spite of these precautions, serious fuel damage accidents can happen, thus requiring containment structures and other safety features to prevent the release of fission products off site
  - ***In the defense-in-depth philosophy, the Commission recognizes that complete reliance for safety cannot be placed on any single element of the design, maintenance, or operation of a nuclear power plant.***

*SECY-13-0132, Enclosure 3: Defense-in-Depth Observations and Detailed History (ML13277A421)

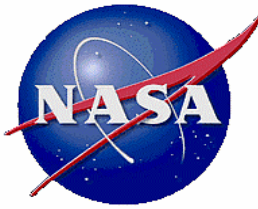
My Understanding of NRC's Motivations for Defense-in-Depth: Uncertainty



- **NRC statements make it clear that the driving motivation for defense-in-depth is to compensate for uncertainty:***
  - **Uncertainty due to lack of operational experience with new technologies and new design features, uncertainty in the type and magnitude of challenges to safety**
  - **Uncertainty and incompleteness in the knowledge of accident initiation and progression**
  - **Uncertainties in knowledge of plant behavior, component reliability, or operator performance that might compromise safety**
  - **Recognized lack of knowledge of nuclear reactor operations and the consequences of potential accidents**
  - **Unquantified and unquantifiable uncertainty in engineering analyses**
  - **Inadequacies, incompleteness, and omissions of risk analyses**

*SECY-13-0132, Enclosure 3: Defense-in-Depth Observations and Detailed History (ML13277A421)

An Agency-Level Perspective



- **"Defense-in-depth" is not in NASA's lexicon**
- **However, managing uncertainty *is* fundamental to NASA's mission**
- **At the Agency level, managing uncertainty is about more than the important task of developing, building, and operating safe systems**
- **It's also about:**
  - **Defining what constitutes adequate safety: *How safe is safe enough?***
    - Public (range, deorbit)
    - Crew (Probability of Loss of Crew (P(LOC)))
    - Environment (Earth environment, planetary protection)
    - Mission (Probability of Loss of Mission (P(LOM)))
  - **Establishing policies, requirements, standards, and guidance that result in adequate safety**
  - **Making informed risk acceptance decisions**
    - Understanding the risks
    - Understanding the uncertainties
    - Deciding whether or not the probability that the risk exceeds expectations/requirements is within the Agency's **risk tolerances**
      - Low risk tolerance for public safety
      - Necessarily higher risk tolerance for crew, asset safety

NASA's Unique Challenges

- **Spaceflight is an inherently *high-risk* endeavor**
  - Launch vehicle mission failure risk is currently in the ~10^{-2} rough-order-of-magnitude (ROM) range; in-space and reentry risks are also significant
  - Bottom line: Spaceflight requires risk takers to assume significant risk, which necessitates relatively high risk-tolerances in the pursuit of NASA's mission
- **Putting mass in orbit is expensive**
  - Safety systems and design conservatisms that involve significant mass increases can be prohibitively expensive
  - Spaceflight safety margins are necessarily thin
- **One-of-a-kind missions**
  - Actual risks are typically not accurately knowable
- **Increasingly *performance-based* acquisition models**
  - Relatively few system-level performance requirements are levied (e.g., P(LOC), P(LOM)), rather than relatively many deterministic (proxy) requirements (e.g., failure tolerance)
  - Safety is "non-observable," so system-level requirements must be V&V'd by other means in order to develop confidence that safety performance is (or will be) met
- **Commercial orbital transportation service acquisition**

How Safe Is Safe Enough?

- **The trigger for dealing with the issue of "adequate safety" was the NASA Aerospace Safety Advisory Panel (ASAP) Recommendation 2009-01-02a:**
  - *"The ASAP recommends that NASA stipulate directly the acceptable risk levels—including confidence intervals for the various categories of activities (e.g., cargo flights, human flights)—to guide managers and engineers in evaluating "how safe is safe enough."*
- **NASA accepted the ASAP recommendation and committed to establishing safety thresholds and goals for human space flight**
  - **Safety threshold** expresses an initial minimum tolerable level of safety
  - **Safety goal** expresses expectations about the safety growth of the system in the long term
- **Additionally, because of spaceflight's high risk, NASA also recognized an ethical obligation to pursue safety improvements wherever practicable**
  - In other words, NASA systems should be **As Safe As Reasonably Practicable** (ASARP)
  - The ASARP principle applies regardless of meeting safety thresholds and goals
- **Threshold and goal values, as well as the level of ASARP application, are a function of *risk tolerances***

Adequate Safety

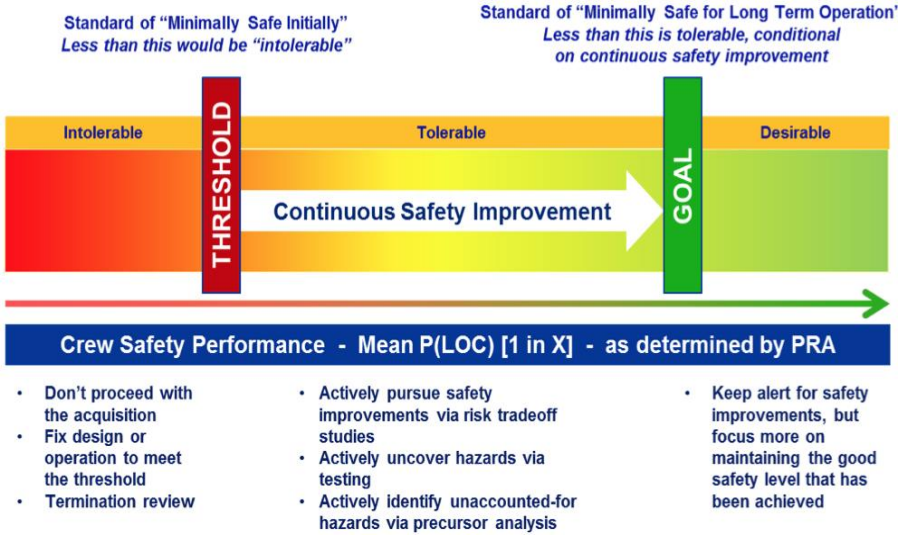
Adequate Safety

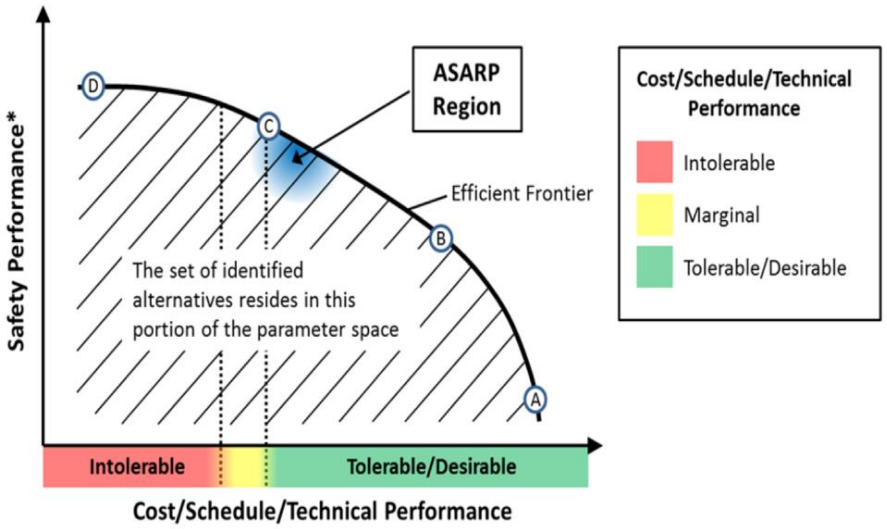
Meeting Minimum Levels of Safety

- Establish safety thresholds, safety goals, safety growth profiles
- Establish safety performance margins to account for UU risk
- Levy safety performance requirements and associated verification procedures (e.g., PRA, tests)
- Conduct verifications

Being ASARP

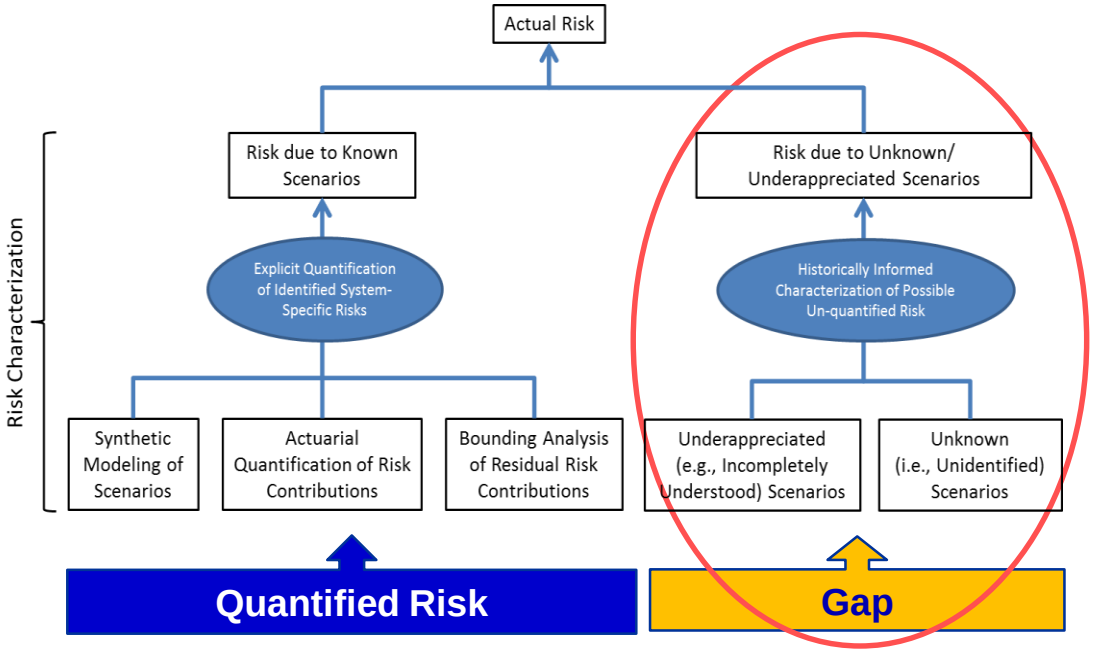
- Analyze a range of alternatives during major design, product realization, operations and sustainment decisions (i.e., risk-informed decision making (RIDM))
- Prioritize safety during decision making
- Implement design-for-safety strategies (e.g., hazard elimination, hazard control (e.g., Design for Minimum Risk (DFMR)), failure tolerance (e.g., redundancy/diversity), safing, emergency operations)
- Analyze and test (e.g., Hazard Analysis, Failure Modes & Effects Analysis and Critical Items List, PRA, qualification/acceptance testing)
- Monitor and respond to performance (e.g., precursor analysis, Problem Reporting and Corrective Action (PRACA), closed-loop risk management)
- Adhere to appropriate codes and standards
- Etc.



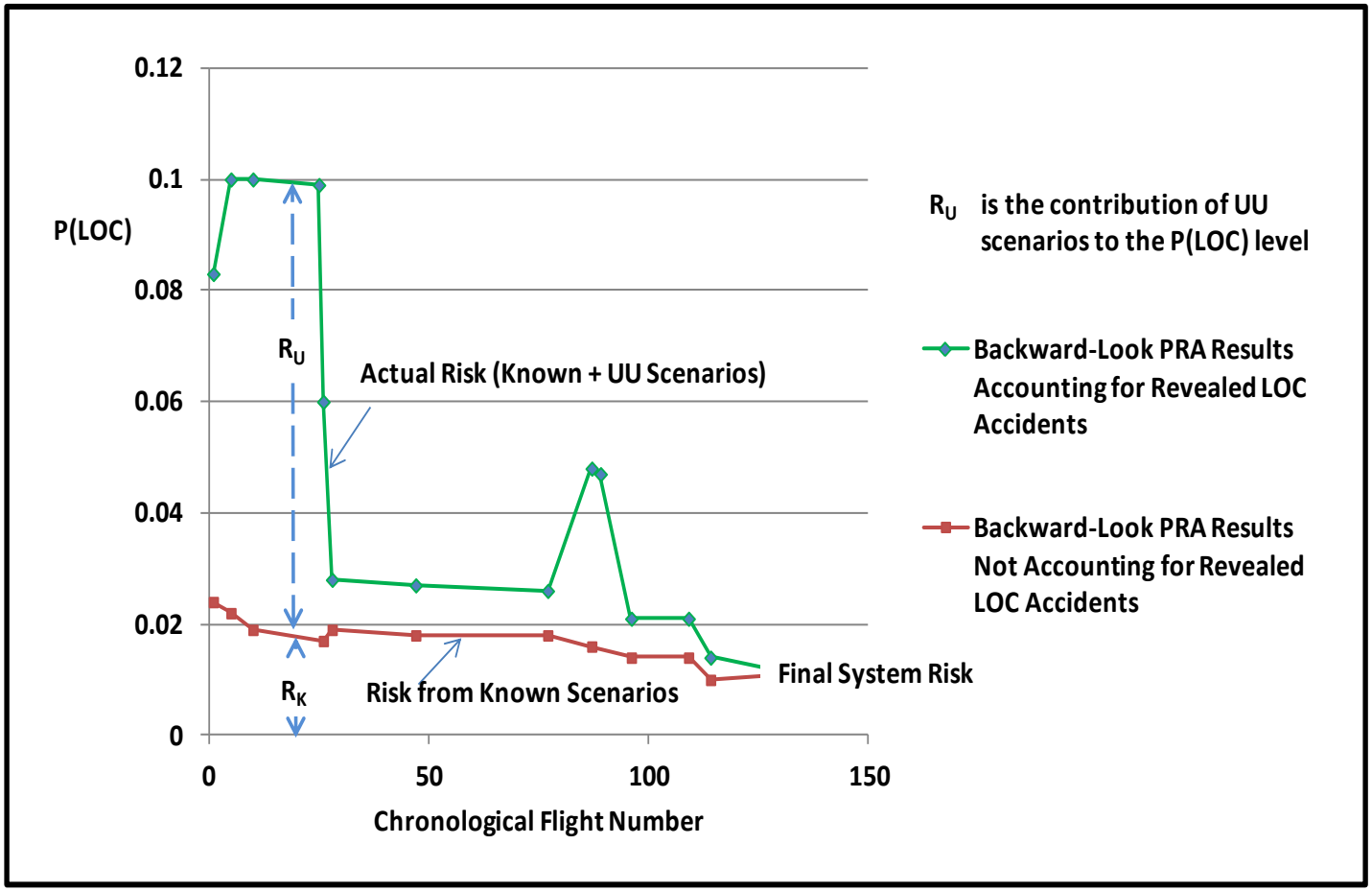


Risk Analysis Completeness

- **Safety goals and thresholds represent expectations about *actual risk*, from both known and unknown/underappreciated (UU) sources**
  - ***Known sources of risk*** are amenable to explicit quantification via synthetic, scenario-based methods of analysis (e.g., PRA), and actuarial methods (when sufficient data are available).
  - ***UU sources of risk*** are not amenable to synthetic analysis or direct actuarial characterization, yet are historically recognized as significant contributors to risk.
    - Tend to remain latent in the system until revealed by operational failures, precursor analysis, etc.
    - Tend to be most significant early in the system life cycle.
    - Disproportionally reflect **design flaws**, **organizational issues**, and **subsystem interactions**.



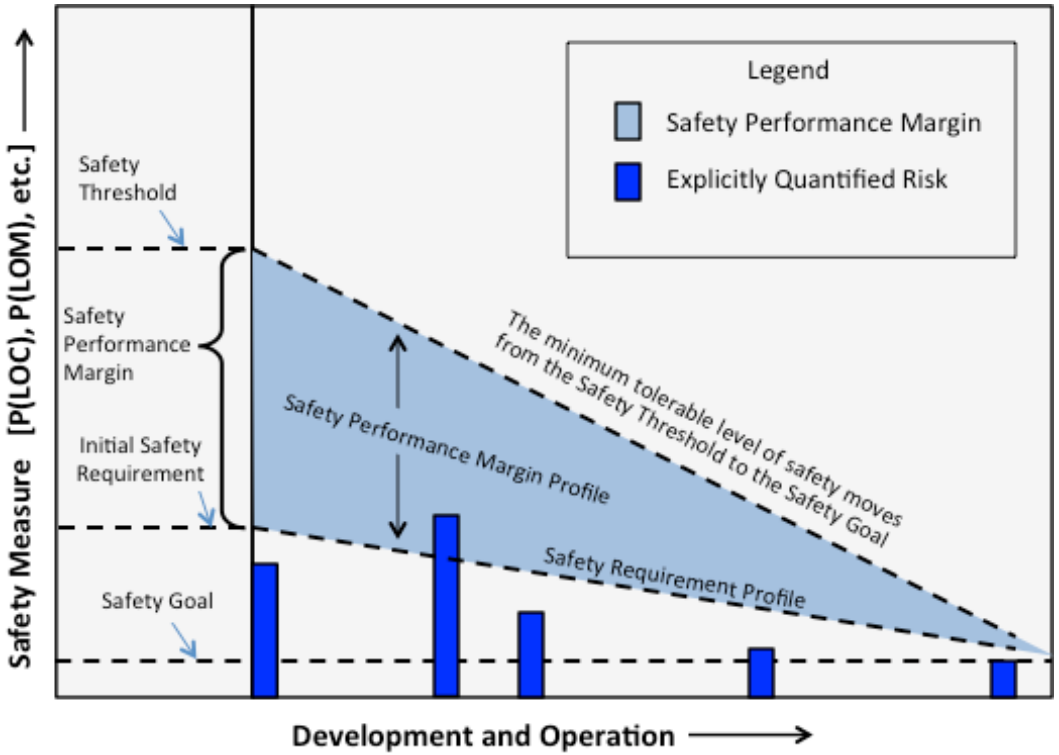
The Shuttle Risk Analysis Gap in Retrospect



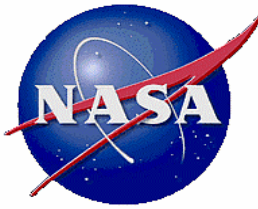
Source: *Shuttle Risk Progression: Use of the Shuttle Probabilistic Risk Assessment (PRA) to Show Reliability Growth, Teri L Hamlin et al. (AIAA, 2010) (downloadable from http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20110004917_2011004008.pdf)*

Safety Performance Margin

- One approach to accounting for the contribution of UU risk is to determine an appropriate *safety performance margin*, analogous to other types of margin (mass, cost, etc.), between the minimum tolerable levels of safety performance and the levied, verifiable (e.g., via PRA) safety performance requirements
- The safety requirements tighten over a defined timeframe in a manner consistent with operational learning and stakeholder expectations regarding the goal
- This provides a defined benchmark for scoping and assessing safety improvement efforts



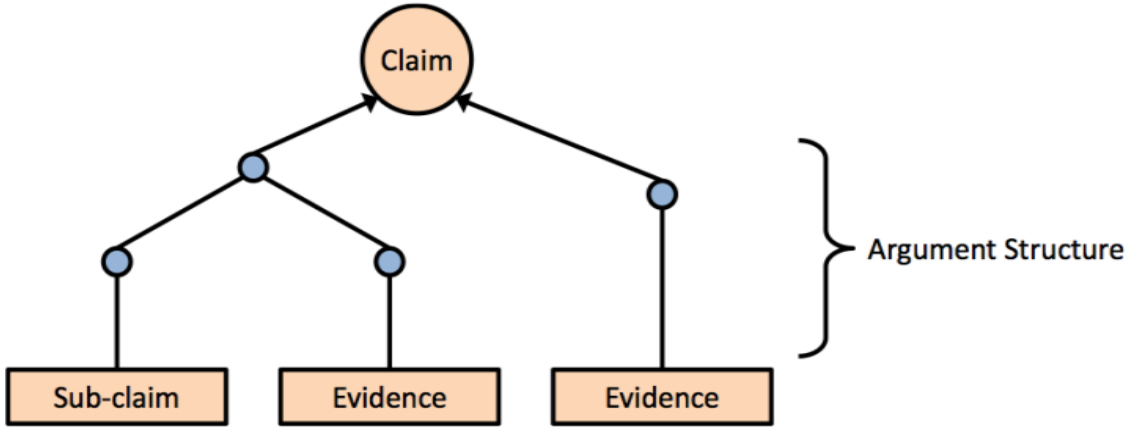
Informed Risk Acceptance Decision-Making



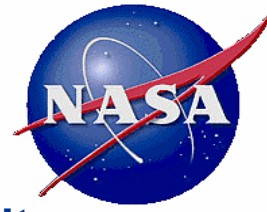
- **In order to be adequately informed, risk acceptance decision-making must go beyond the risk analysis**
- **A holistic "case" must be made that the system is adequately safe**
  - **Substantiation that UU risks are adequately managed via application of the ASARP principle:**
    - Minimize the presence of UU scenarios (e.g., via margin, programmatic commitments)
    - Maximize discovery of UU hazards (e.g., via testing, liberal instrumentation, monitoring, and trending, anomaly investigation, Precursor Analysis, use of best safety analysis techniques)
    - Provide broad-coverage safety features (e.g., abort capability, safe haven, rescue)
  - **Substantiation that the known risk (calculated by PRA) is within the specified safety performance requirement**
    - Known risks are managed by applying controls that are designed to mitigate identified accident scenarios

The Risk-Informed Safety Case (RISC)

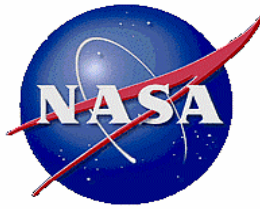
- **The Risk-Informed Safety Case (RISC) is a coherent and evidentiary statement of how safe we are (or will be) at a given stage of the life cycle**
- **RISC is a specialization of the "safety case" construct. The term "risk-informed" is used to emphasize that adequate safety is the result of a deliberative decision making process that involves an *assessment of risks*, and strives for a proper balance between safety and performance in other mission execution domains.**
- **The RISC is the totality of the "uncertainty story" about the actual safety performance of the system**
  - Presented and defended by the provider at key decision points
  - Involves serious consideration of things that live outside traditional risk models (e.g., organizational and management factors)



Mapping into Defense-in-Depth

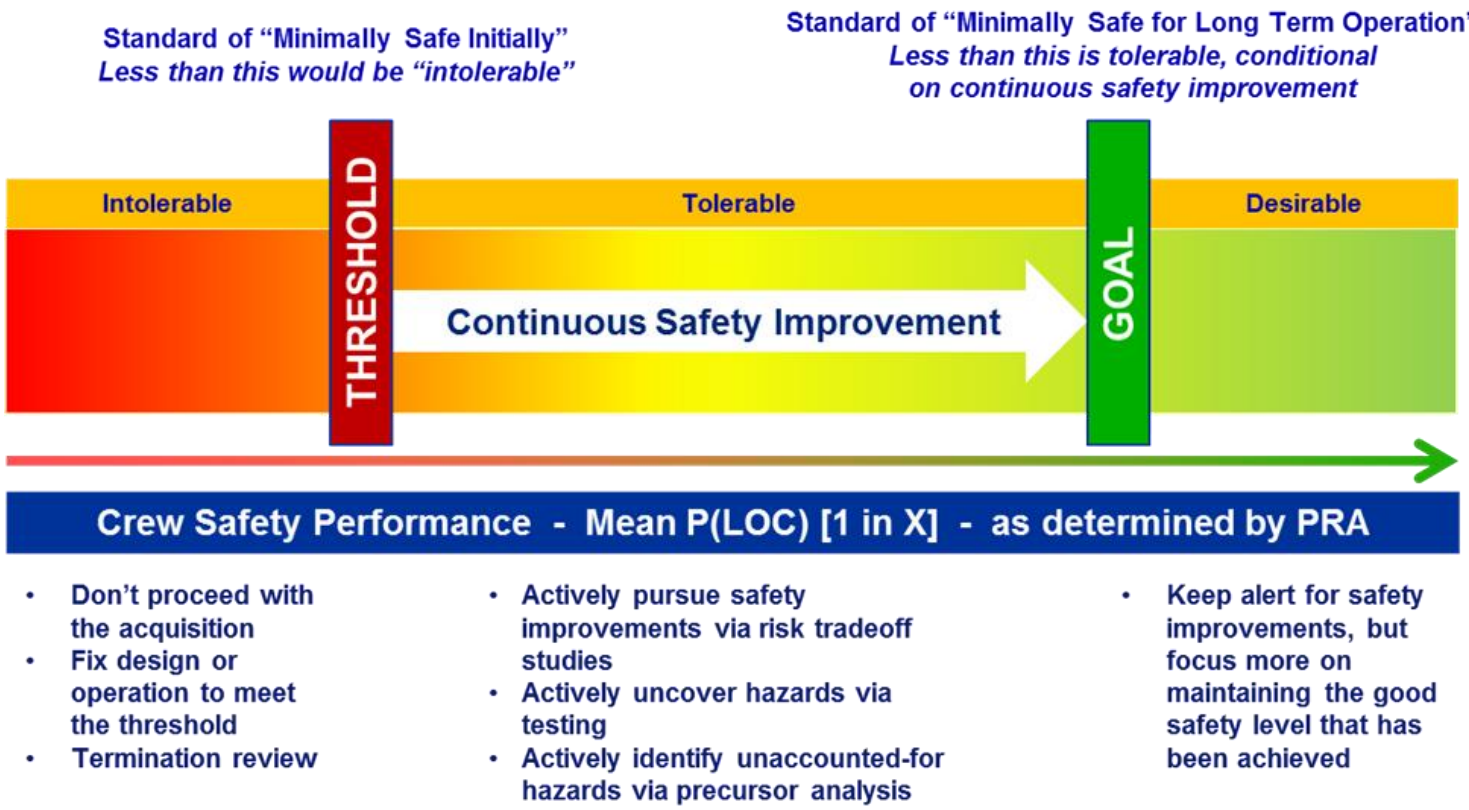


- **Defense-in-depth is not a part of NASA's lexicon, nor is it an explicit element of the NASA system safety framework**
- **Operationally, NASA employs a range of safety strategies that maps into defense-in-depth**
  - ***Prevention***: QA, testing, training, certification, lifecycle reviews, anomaly resolution, margins (where practicable), V&V, hazard/risk analysis, etc.
  - ***Fault management***: Redundancy/diversity, Integrated Vehicle Health Management (IVHM), safe modes, etc.
  - ***Accident mitigation***: Abort systems, flight termination systems, redundancy/diversity (campaign/program-level, e.g., International Space Station (ISS) resupply)
- **The ASARP principle:**
  - Is an ethical response to the high-risk nature of spaceflight, rather than a principle of distributed reliance
  - However, ASARP does *imply* defense-in-depth (i.e., prevention, fault management, accident mitigation) operationally
- **The RISC goes beyond traditional system-centric risk analysis to address the totality of the "uncertainty story" about the actual safety performance of the system**
  - Quality of models, qualifications/experience of people, management and organizational factors, etc.



BACKUP

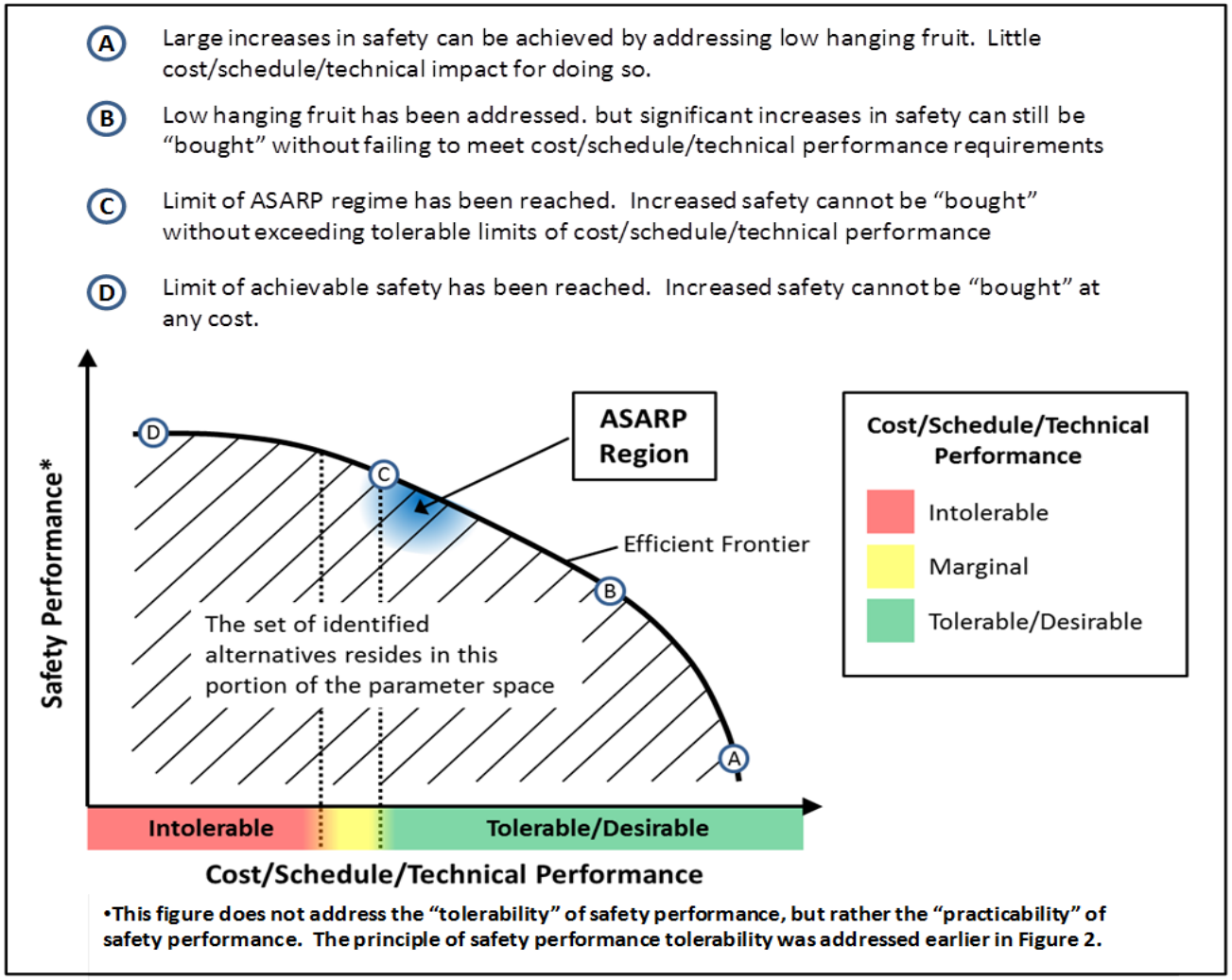
Meeting or Exceeding a Minimum Tolerable Level of Safety



- Minimums may be applied to any safety performance measure, e.g., Probability of Loss of Crew (P(LOC)), Probability of Loss of Mission (P(LOM)), Probability of Loss of Vehicle (P(LOV)), Expected Casualty (E_c).

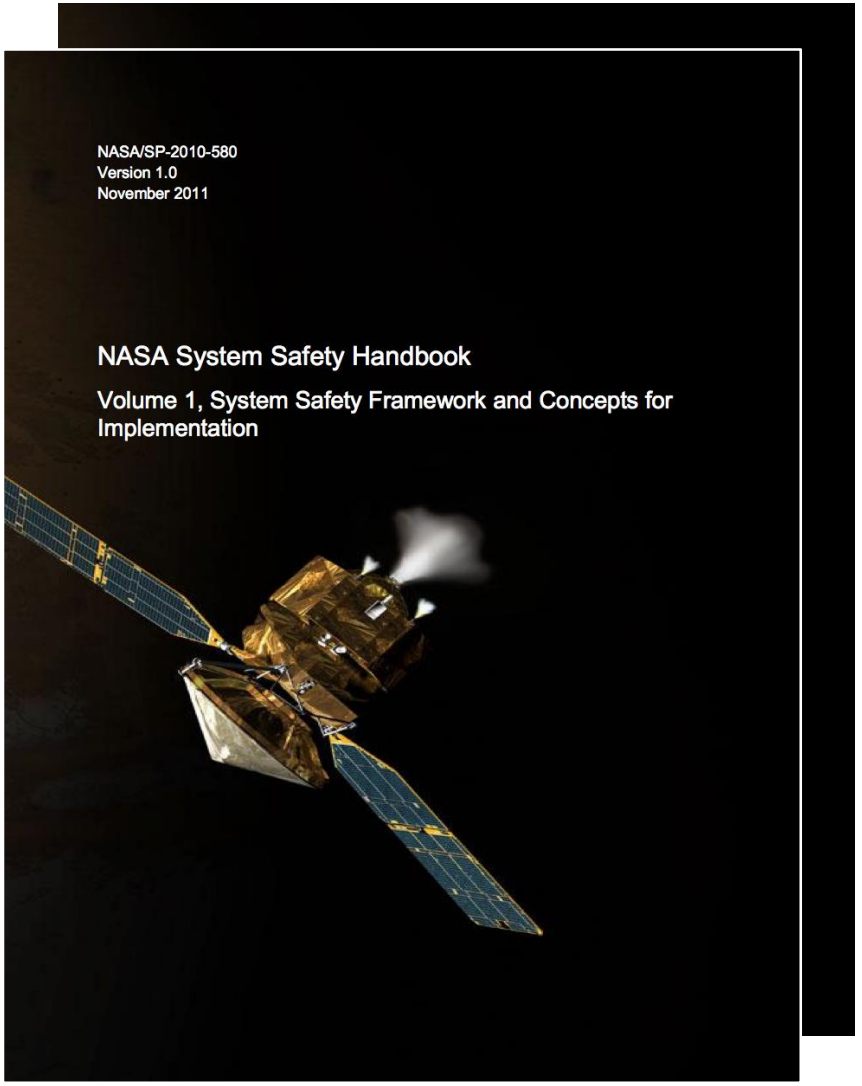
As Safe As Reasonably Practicable (ASARP)

- **ASARP reflects a mindset of continuous safety improvement regardless of the current level of safety.**

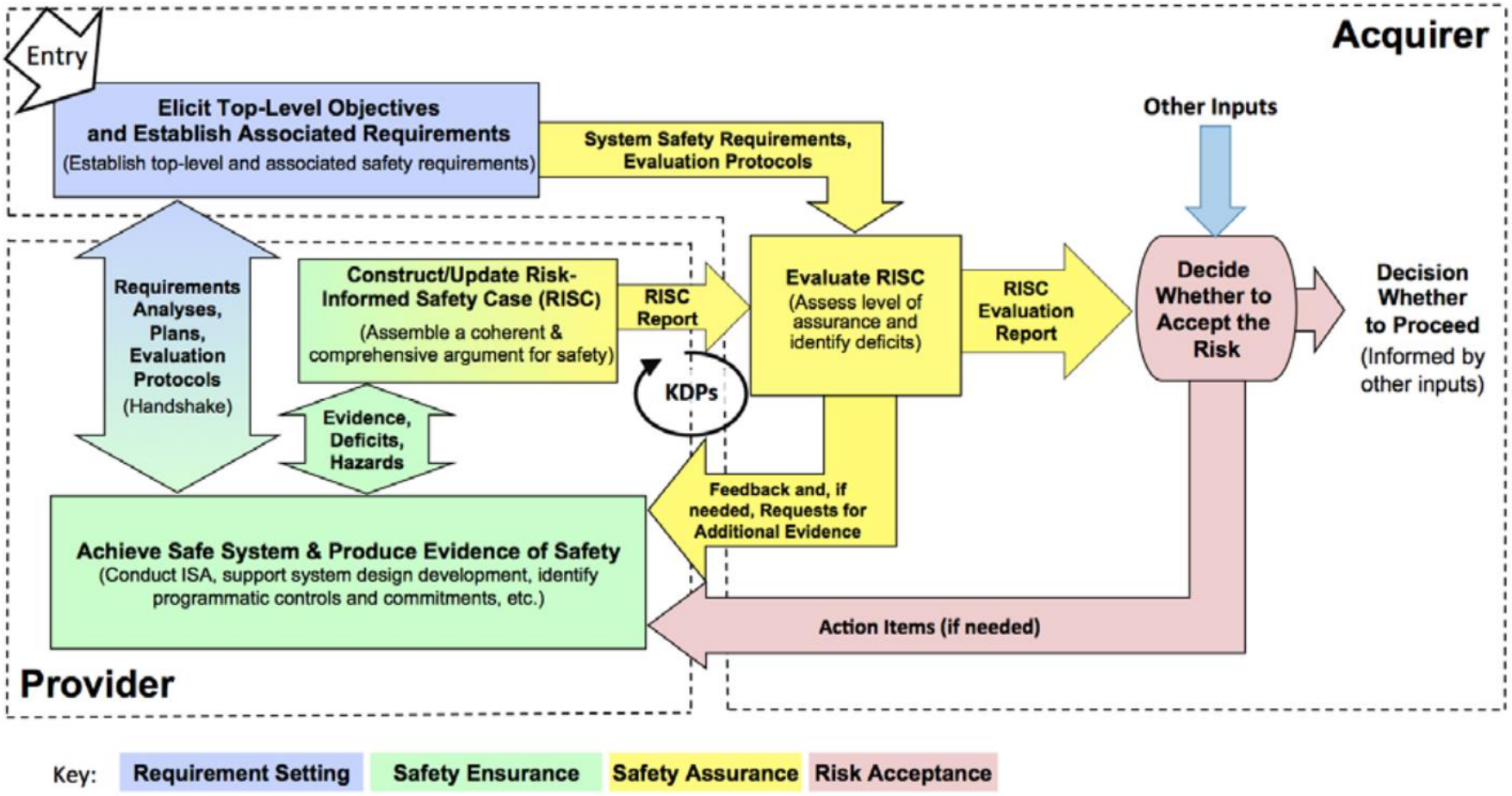


NASA System Safety Handbook Vols. 1 & 2

- **Presents an objectives-driven system safety framework:**
  - **Safety requirements setting**
  - **Safety *ensurance***
    - Responsibility of the organization *providing* the system/service
    - Active participation in designing for safety and in the reduction or elimination of risks
    - Produces the RISC
  - **Safety *assurance***
    - Responsibility of the organization *acquiring* the system/service
    - Evaluates the RISC to support risk acceptance decisions
  - **Risk acceptance**
    - Decision whether to accept the risk
    - Responsibility of the organization *acquiring* the system/service



NASA System Safety Framework



From NASA/SP-2014-612, NASA System Safety Handbook Volume 2: System Safety Concepts, Guidelines, and Implementation Examples