A System-Theoretic Hazard Analysis Methodology for a Non-advocate Safety Assessment of the Ballistic Missile Defense System

Steven J. Pereira*
Missile Defense Agency, Washington, DC, 20301

Grady Lee† and Jeffrey Howard‡
Safeware Engineering Corporation, Seattle, WA, 98102

The Missile Defense Agency (MDA) is developing the Ballistic Missile Defense System (BMDS) as a layered defense to defeat all ranges of threats in all phases of flight (boost, mid-course, and terminal). The BMDS integrates into a single system a number of Elements that had been developed independently, such as SBIRS/DSP, Aegis BMD, and Ground-based Midcourse Defense (GMD). The Elements of the BMDS have active safety programs, but complexity, coupling, and safety risk are introduced by their integration into a single system. Assessing the safety of the integrated BMDS required analysts to come up to speed using existing Element project documentation, assess the safety risk of the system, and make recommendations regarding hazard mitigation and risk acceptance. This effort often required conducting hazard analyses to supplement existing Element analysis work; working with existing engineering artifacts; and making recommendations for hazard mitigations late in the system life cycle, when there is less flexibility for design changes. This paper presents a safety assessment methodology based on STPA (a systems-theoretic hazard analysis); the assessment methodology provides an organized, methodical, and effective means to assess safety risk and develop appropriate hazard mitigations regardless of where in the life cycle the assessment is started.

I. Introduction

Safety engineering efforts are most effective when begun early in the system life cycle. However, on some development programs, a significant portion of the safety engineering effort is deferred until late in the life cycle. There are a variety of possible reasons for this. Some development programs integrate already mature systems, such that safety for the new effort begins at a late life cycle stage for the pre-existing systems being integrated. Other programs spend much of their development as experimental or prototype systems, where safety is given a reduced role, until a decision to field the system drives an expanded system safety effort. There are some development programs conducted with inadequate initial safety effort that add emphasis at the end of the life cycle when the increased safety risk becomes apparent. Some programs have an independent assessment conducted to characterize the system's residual safety risk; these assessments face many of the same challenges as a safety effort deferred until late in the life cycle. On any such program, safety engineers must come up to speed using existing project documentation, assess the safety risk of the system, and make recommendations regarding hazard mitigation and risk acceptance. This effort often requires conducting hazard analyses where existing analysis work is missing or inadequate; working with existing engineering artifacts that may be erroneous, ambiguous, incomplete, or outdated; and making recommendations for hazard mitigations late in the system life cycle, when there is less flexibility for design changes. The methodology presented in this paper, based on STPA (a systems-theoretic hazard analysis), supports assessment of safety risk and development of appropriate hazard mitigations even when begun late in the system life cycle.

* Ballistic Missile Defense System Safety Officer, MDA/QS, 7100 Defense Pentagon, Washington, DC 20301.
† President, Safeware Engineering Corporation, 1500 Fairview Avenue East, Suite 205, Seattle, WA 98102.
‡ Safety Engineer, Safeware Engineering Corporation, 1500 Fairview Avenue East, Suite 205, Seattle, WA 98102.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
14 NOV 2006	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
A System-Theoretic Hazard Analysis Methodology for a Non-advocate Safety Assessment of the Ballistic Missile Defense System	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Missile Defense Agency, Washington, DC, 20301**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM202095, Proceedings of the 2006 AIAA Missile Sciences Conference Held in Monterey, California on 14-16 November 2006.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**UU**	**9**	

The Missile Defense Agency (MDA) employed this methodology to characterize the residual safety risk of the Ballistic Missile Defense System (BMDS). The Ballistic Missile Defense System is being developed as a layered defense to defeat all ranges of threats in all phases of flight (boost, mid-course, and terminal). The BMDS comprises a variety of components, including sea-based sensors on the Aegis platform, upgraded early warning radars (UEWR), the Cobra Dane Upgrade (CDU), Ground-based Midcourse Defense (GMD) Fire Control and Communications (GFC/C), a Command and Control Battle Management and Communications (C2BMC) Element, and Ground-based interceptors (GBI). Future block upgrades will introduce additional Elements into the BMDS, including Airborne Laser (ABL) and Terminal High Altitude Area Defense (THAAD). The Missile Defense Agency (MDA) has conducted safety assessments of the Limited Defensive Operations and Block 04 configurations of the BMDS and will continue to assess safety as the system evolves.

Conducting a safety assessment to characterize the residual risk of the integrated BMDS required the adaptation of an advanced hazard analysis methodology. Many Elements and components within Elements, such as the UEWR and CDU, are upgrades of fielded systems. The safety assessment must take into account legacy behavior, the safety impact of changes, and any safety risk created by reuse of the asset in a new environment. Moreover, the BMDS integrates into a single system a number of programs that had historically been developed independently, such as SBIRS/DSP, Aegis BMD, and Ground-based Midcourse Defense. The Elements of the BMDS have active safety programs, but considerable complexity, coupling, and safety risk is introduced by integrating the Elements into a single system. Successfully conducting a safety assessment required a hazard analysis methodology that 1) considers hazards and causes due to complex system interactions (more than just failure events); 2) provides guidance in conducting the analysis; 3) comprehensively addresses the whole of the system, including hardware, software, operators, procedure, maintenance, and continuing development activities; and 4) focuses resources on the areas of the system with the greatest impact on safety risk.

The MDA Non-advocate Safety Assessment of the BMDS used an assessment technique based on the Systems-Theoretic Hazard Analysis (STPA) technique developed by Professor Nancy Leveson of MIT (ref. 1). In this analysis, the system is viewed as a collection of interacting loops of control. For example, command authorities exert control over system operators by issuing guidance, providing training, and establishing tactics, techniques, and procedures (TTPs). Command authorities receive feedback in the form of reports and performance during training exercises. Operators exert control over software by inputting commands, and they receive feedback from displays and aural alerts. Software exerts control over other software and hardware by sending messages and commands and receives feedback in the form of measured values and status information. Safety is an emergent property of the system, arising from the interactions among software, hardware, and humans. Safety is maintained by placing constraints on the behavior of the system's components. Mishaps occur 1) when the constraints are insufficient to maintain safety and 2) when the controls present in the system are inadequate to enforce the safety constraints. The safety assessment methodology described herein provides information about where constraints may be insufficient to maintain safety, when controls may be inadequate to enforce safety constraints, and how mitigations can be developed to reduce safety risk with minimal impact on the system's design and operations.

The assessment begins with the hazards identified for the system. The analyst must ensure that appropriate top-level system safety constraints are in place to mitigate the hazards. Next, a control structure diagram is produced: the control depicts the components of the system and the paths of control and feedback. Using the control structure diagram as a guide for conducting the analysis, each control action is assessed for potential contribution to hazards. If the controls in place are inadequate, recommendations are developed for additional mitigations. If appropriate controls are in place, requirements, design, and verification documentation is examined to ensure that the system implements the control as intended. Each step of the assessment process is described below, using a Fictional Missile Interceptor System (FMIS) as an example – similar to programs within the BMDS, FMIS uses a hit-to-kill interceptor that destroys incoming ballistic missiles through force of impact. Although the example is not a real system within the BMDS, it is suitable for an example of how the safety assessment methodology is conducted and the results achieved at MDA.

II. Review System Hazards and System-level Safety Constraints

The first step of a system-theoretic safety assessment is to review the hazards identified for the system and ensure that appropriate system-level safety constraints are in place. By the time the system is being assessed for residual safety risk, the system hazard logs should already contain a list of the system's hazards. Although it is not expected that any new hazards will be discovered late in the system life cycle, the list should be reviewed for completeness and accuracy. If any new hazards are discovered, they are reported as part of the assessment's

findings and should be raised to management as quickly as possible. Additionally, safety analysts must verify that the hazard can be traced to one or more system-level safety constraints that mitigate the hazard.

Inadvertent launch is one of the hazards of the FMIS system:

The FMIS system inadvertently launches an interceptor missile.

This hazard traces to two requirements in the top level system specification:

3.7.2. The FMIS system shall make improbable the likelihood of occurrence of catastrophic hazards.

3.7.2.1 The FMIS system shall make improbable the likelihood of occurrence of inadvertent launch.

III. Define the Safety Control Structure

Once the hazards to be assessed have been reviewed, the analyst develops a diagram of the safety control structure of the system. The structure in figure 1, below, is generalized and does not represent any one particular system. However, it does show the information that a safety control structure diagram should capture. Each node in the graph is a human or machine component in a socio-technical system. Connecting lines show control actions used to enforce safety constraints on the system and feedback that provides information to the controlling entity. These lines are annotated with a description of the information reported or controls applied. Typically, there are two parallel management structures: one for oversight of system development and another for oversight of system operations.

Figure 2, below, shows a safety control structure diagram for the FMIS. A typical engagement would begin with warning from the Early Warning System. The fire control system would then task the radar to track the target. Once the system has adequate data to plan an intercept, the operators direct the system to engage its target. The fire control software develops interceptor tasking and sends it to the launch station. The launch station is responsible for the interface with the interceptor, transforming the interceptor tasking into a task load for the flight computer and controlling the launch sequence. The flight computer receives the task load, follows the launch sequence under the control of the launch station, and ignites the rocket motor.

The diagram shows the components of the system, the control actions each component exerts on the others and the feedback provided to controllers. For example, at the lowest level, the flight computer is responsible for arming and safing the interceptor hardware – these are control actions that influence the state of the hardware. These control actions must enforce the safety constraints defined to keep the system safe. Feedback to the flight computer is provided in the form of Built-in Test (BIT) results and status of whether the hardware is safe or armed. The view of the system as interacting control loops works as well for hardware, software, and human operators. The operators control the behavior of the fire control software by directing the software to change operating modes (between test, exercise, and live operations) and directing the system to engage targets.

When conducting an assessment, interface documentation is helpful in constructing the control structure diagram. Interface specifications and design documentation describe the messages exchanged between software components as well as the signals exchanged with hardware components. Typically, the embodiment of a control action implemented in software is some command message or signal sent to other hardware or software. Similarly, operator control actions can be identified by examining user manuals and user interface designs, focusing on what commands operators and give and the information imparted to operators to aid their decision-making. Note, however, that interface documentation is only a starting point. The control structure diagram is not a data flow diagram and cannot be constructed by copying message names from interface specs – the safety engineer constructing the diagram must omit irrelevant detail from messaging protocols and interface conventions to focus attention on the control and feedback.

IV. Identify Potentially Inadequate Control Actions

After the system control structure has been defined, the next step is to determine how the controlled system can get into a hazardous state. A hazardous state is a state that violates the safety constraints defined for the system. The assessment methodology views hazardous states as a result of ineffective control; therefore, the assessment proceeds by identifying potentially inadequate control actions. Inadequate controls fall into four general categories:

1. A required control action is not provided.
2. An incorrect or unsafe control action is provided.
3. A potentially correct or adequate control action is provided too early, too late, or out of sequence.
4. A correct control action is stopped too soon.

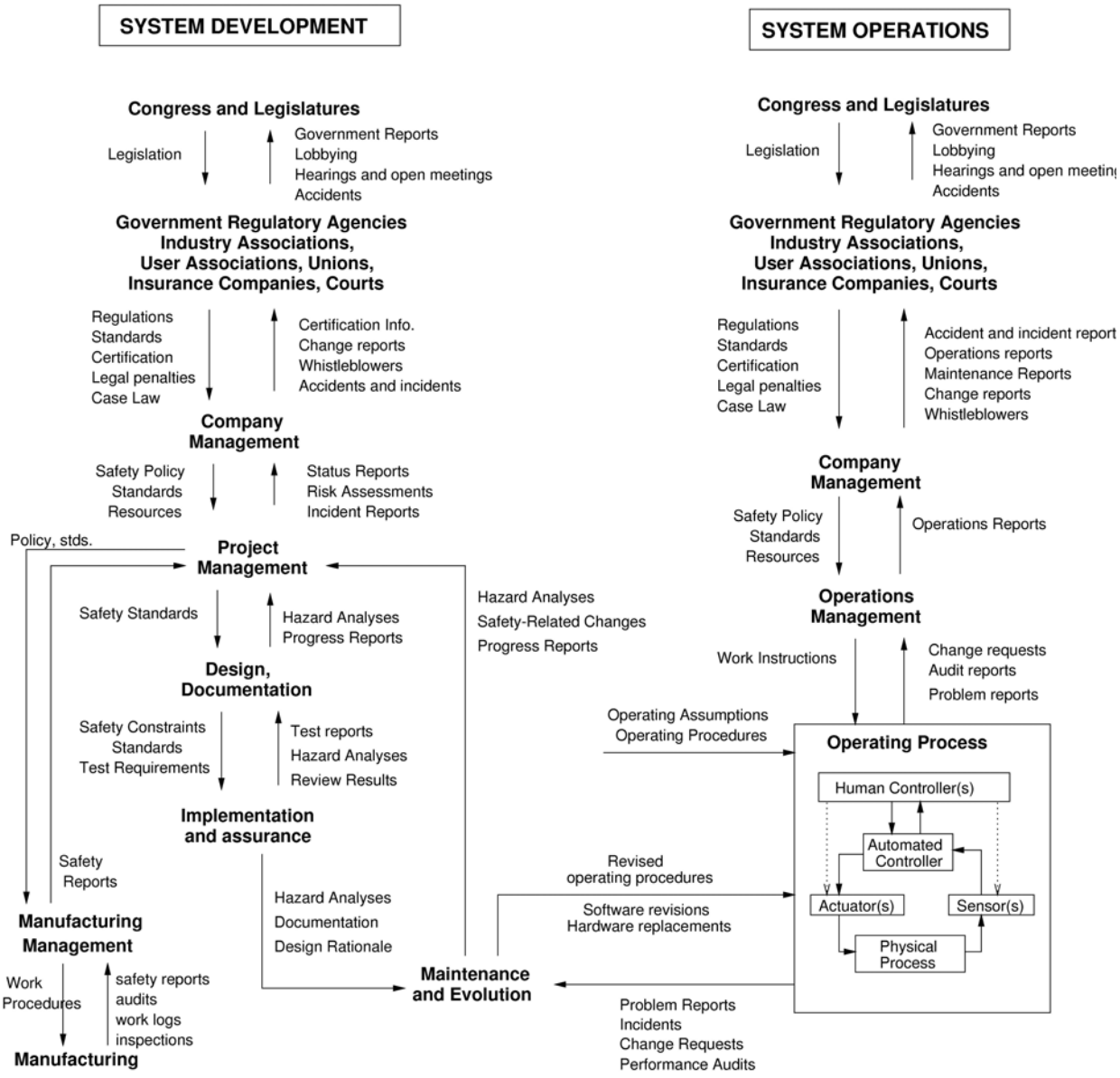

Figure 1. Generalized Safety Control Structure Diagram (Ref. 2)

Control actions may be required to handle component failures, environmental disturbances, or dysfunctional interactions among the components. Incorrect or unsafe control actions may also cause dysfunctional behavior or interactions among components. This view of control flaws applies equally well to automated system behavior, operational procedure, and management organizations.

1. The flight computer could omit the required control action of safing the interceptor hardware if a design error in the interface allowed the launch station to remove power from the flight computer before the abort sequence completes.
2. The launch station could issue an unsafe control action if it commands a real interceptor to arm when the expectation is that it is communicating with an interceptor simulator.
3. The operators might provide a correct control action too late if they transition to weapons hold too late, after an interceptor has been inadvertently launched.

4. The command authority might stop a correct control action too soon if a reduction in training to maintain operator proficiency increases the risk that the operators cannot act to prevent an inadvertent launch.

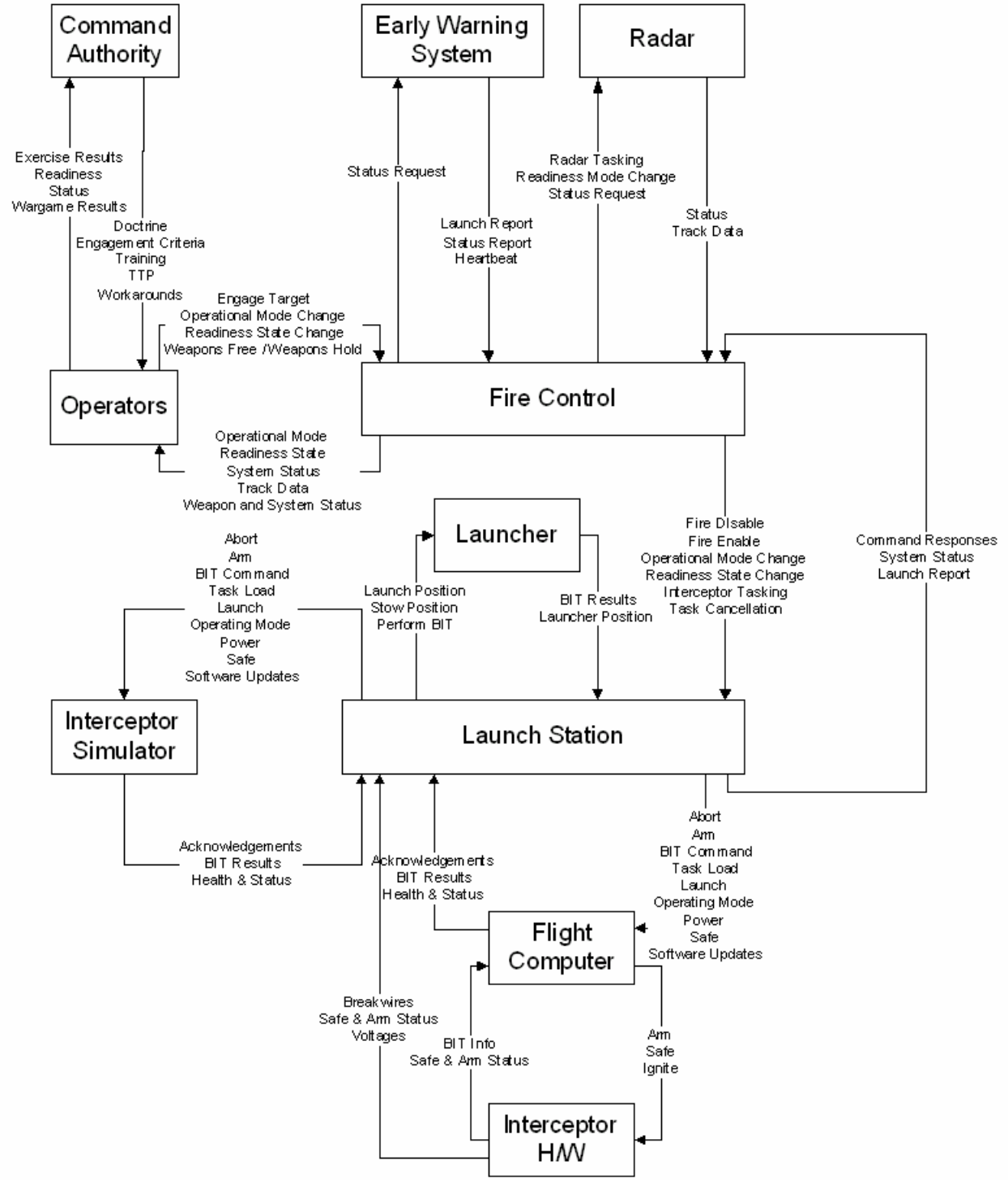

Figure 2. Safety Control Structure Diagram for FMIS

Note that these inadequate control actions may or may not be present in the actual system. These are hypotheses that must be confirmed or rejected based on investigation into the behavior of the system as it has been designed and built. To ensure a complete assessment, each control action must be investigated in turn. When conducting the Non-advocate Safety Assessment of the BMDS, the project team created worksheets to structure the investigation. Entry blanks prompted analysts to consider required control actions that were missing between each two components in the system. For each control action present in the system, entry blanks on the worksheet prompted consideration of how that control action might be omitted, unsafe, mistimed, or stopped to soon. A worksheet fragment for the *Interceptor Tasking* control action from the fire control software to the launch station would look like:

Fire Enable Missing

The fire enable control action directs the launch station to enable the live fire of interceptors. Prior to this control action, the launch station will return an error when sent interceptor tasking and discard the task messages. If this control is omitted, a launch will not take place. Although potentially a mission assurance concern, this is would not be a potentially inadequate control contributing to the hazard of inadvertent launch.

Fire Enable Provided Incorrectly

If the fire enable command is provided to a launch station incorrectly, the launch station will transition to a state where it accepts interceptor tasking and can progress through a launch sequence. In combination with other incorrect or mistimed control actions (see interceptor tasking), this could contribute to an inadvertent launch.

Fire Enable Too Early, Too Late, or Out of Sequence

A late fire enable command will only delay the launch station's ability to process a launch sequence, which will not contribute to an inadvertent launch.

A fire enable command sent too early could open a window of opportunity for inadvertently progressing toward an inadvertent launch, similar to an incorrect fire enable. The degree of risk this contributes depends both on the likelihood of the inadequate control and how early the control action is carried out.
In the worst case, a fire enable command might be out of sequence with the fire disable command. If possible in the system as designed and built, the system could be left capable of processing interceptor tasking and launching when not intended.

Fire Enable Stopped Too Soon

The fire enable command is a single command sent to the launch station to signal that it should allow processing of interceptor tasking. It is not a continuous control like steering a rudder. Therefore, it does not make sense to talk about fire enable in terms of stopping too soon.

There are several advantages provided by this method of assessment. The template above does not contain any information about how potentially inadequate control actions might be present in the design or construction of the system. That portion of the analysis is conducted next (see below). For the moment, the effort is focused entirely on the possible impact of an inadequate control: how could it potentially contribute to a hazard? This divide and conquer strategy eases the analytical burden on the safety engineer conducting the assessment.

Secondly, the assessment provides guidance in how to proceed. One drawback of many hazard analyses is that there is little or no guidance on what to consider during the analysis. For example, fault tree analysis provides no assistance in determining what contributing events should be included in the tree. The completeness and relevance of the fault tree boxes is dependent entirely on the experience and insight safety engineer conducting the analysis. This assessment methodology gives analysts guidance on what factors to consider.

Lastly, there is a finite scope to the analysis. One of the questions often asked by engineers using other analysis methods is, "When is the analysis complete?" The analyses themselves do not provide any guidance for determining when the effort is complete. Typically, the analyst must use personal experience alone to determine

when the hazard has been broken down into sufficiently fine-grained causal events. With this systems-theoretic safety assessment methodology, the assessment is complete when all of the control actions have been analyzed.

V. Determine How Potentially Inadequate Control Actions Could Manifest in the System and Develop Mitigations

The previous step of the assessment will yield a set of potentially inadequate control actions. If present system, these inadequate controls will provide a means for the system to enter a hazardous state. In this step of the assessment, the analyst determines whether and how the potentially hazardous control actions can occur.

STPA – the hazard analysis on which the assessment method is based – makes use of process models of component behavior for each of the components in the control structure. The use of formal models allows mitigations features and the impact of various types of faults in other components of the control structure to be evaluated during the hazard analysis process. The use of formal models in the hazard analysis process is guided by a set of generic control loop flaws. Because accidents result from inadequate control and enforcement of safety constraints, the process that leads to accidents can be understood in terms of flaws in the system development and system operations control structures in place during design, implementation, manufacturing, and operation. These flaws can be classified and used during accident analysis to assist in identifying all the factors involved in the accident or during hazard analysis and other accident prevention activities to identify required constraints. Figure 3, below, shows a general classification.

1. Inadequate Enforcement of Constraints (Control Actions)
- 1.1 Unidentified hazards
- 1.2 Inappropriate, ineffective, or missing control actions for identified hazards
  - 1.2.1 Design of control algorithm (process) does not enforce constraints
    - Flaw(s) in creation process
    - Process changes without appropriate change in control algorithm (asynchronous evolution)
    - Incorrect modification or adaptation
  - 1.2.2 Process models inconsistent, incomplete, or incorrect (lack of linkup)
    - Flaw(s) in creation process
    - Flaws(s) in updating process (asynchronous evolution)
    - Time lags and measurement inaccuracies not accounted for
  - 1.2.3 Inadequate coordination among controllers and decision makers (boundary and overlap areas)

2. Inadequate Execution of Control Action
- 2.1 Communication flaw
- 2.2 Inadequate actuator operation
- 2.3 Time lag

3. Inadequate or missing feedback
- 3.1 Not provided in system design
- 3.2 Communication flaw
- 3.3 Time lag
- 3.4 Inadequate sensor operation (incorrect or no information provided)

Figure 3. A Classification of Control Flaws Leading to Hazards (Ref. 2)

In each control loop at each level of the socio-technical control structure, unsafe behavior results from either a missing or an inadequate constraint on the process at the lower level or inadequate enforcement of the constraint leading to its violation. Because each component of the control loop may contribute to inadequate control, classification starts by examining each of the general control loop components and evaluating their potential contribution: (1) the controller may issue inadequate or inappropriate control actions, including inadequate handling of failures or disturbances in the physical process; (2) control actions may be inadequately executed, or (3) there may be missing or inadequate feedback. These same general factors apply at each level of the socio-technical safety control structure, but the interpretations of the factor at each level may differ. Where a human or organization is involved, it is necessary to evaluate the context in which decisions are made in order to understand the types and reasons for potentially unsafe decisions to be made and to design controls or mitigation measures for them.

The formal models of component behavior constructed for an STPA analysis and the generic list of control loop flaws are also of significant benefit in conducting an assessment late in the system life cycle. However, some assessment efforts do not construct component models. In these cases, the assessment continues by applying the

generic control loop flaws in a review of existing project documentation. The analyst looks for documentation that the potentially inadequate control action has been designed out of the system, or if present, is adequately mitigated. The assessment must consider system requirements, design, and verification to ensure that appropriate mitigations were required for the system, built into the system, and verified to function correctly. During the assessment of the BMDS, this information was summarized on analysis worksheets, adding to the information compiled when identifying potentially inadequate controls. Continuing the example for the FMIS system from above, the template would look like:

Fire Enable Provided Incorrectly

If the fire enable command is provided to a launch station incorrectly, the launch station will transition to a state where it accepts interceptor tasking and can progress through a launch sequence. In combination with other incorrect or mistimed control actions (see interceptor tasking), this could contribute to an inadvertent launch.

The fire control computer is intended to send the fire enable command to the launch station upon receiving a weapons free command from an FMIS operator and while the fire control system has at least one active track. According to the requirements and design specifications, the handling of the weapons free command is straightforward. Although the specification requires an "active" track, it is more difficult to determine what makes a track active. Interviews with the development staff clarified that activity criteria are specified by the FMIS operators according to their operational procedures. The software supports declaring tracks inactive after a certain period with no radar input, after the total predicted impact time for the track, and/or after a confirmed intercept. It appears one case was not well considered: if an operator deselects all of these options, no tracks will be marked as inactive. Under these conditions, the inadvertent entry of a weapons free command would send the fire enable command to the launch station immediately, even if there were no threats to engage currently tracked by the system.

The FMIS system undergoes periodic system operability testing using an interceptor simulator that mimics the interceptor flight computer. Hazard analysis of the system identified the possibility that commands intended for test activities could be sent to the operational system. As a result, the system status information provided by the launch station includes whether the launch station is connected only to missile simulators or to any live interceptors. If the fire control computer detects a change in this state, it will warn the operator and offer to reset into a matching state. However, there is a small window of time before the launch station notifies the fire control component of the change during which the fire control software might send a fire enable command intended for test to the live launch station.

Note that in the example above, neither of the causal factors identified by the assessment involved component failures. In both cases, all the components involved were operating exactly as intended; however, the complexity of their interactions led to unanticipated system behavior. Component failure can be a cause of inadequate control over a system's behavior, and this assessment methodology does include those possibilities. However, many analysis techniques, particularly those based on reliability failure, consider only failure events, not the effects of complex system interactions.

The assessment worksheets include enough information for someone familiar with the system being assessed to understand the results. Facts supporting the analysts' conclusions are backed up with references to relevant system documentation. Once inadequate control actions are identified within the system, the analyst develops a recommendation for mitigating the safety risk. Recommendations could include hardware or software design changes to improve safety, changes to maintenance or test procedures, or workarounds for the system's operators. Ideally, mitigation suggestions are developed in conjunction with the design engineers responsible for the system.

VI. Summary

Successfully conducting a safety assessment of the Ballistic Missile Defense System required a hazard analysis methodology that 1) considers hazards and causes due to complex system interactions (more than just failure events); 2) provides guidance in conducting the analysis; 3) comprehensively addresses the whole of the system, including hardware, software, operators, procedure, maintenance, and continuing development activities; and 4) focuses resources on the areas of the system with the greatest impact on safety risk. The safety methodology described above views safety as maintained by adequate control over a system's behavior. Component failure

events are only one cause of inadequate controls identified by the analysis. Throughout the assessment effort, the methodology provides guidance to analysts; safety engineers are not required to fill in a blank page using personnel experience alone. Similarly, the methodology assists engineers in scoping their effort. The entire system, including software, hardware, and operators are included in the analysis, but the effort is bounded – once all the control actions have been examined, the assessment is complete. Lastly, as the control structure is developed and the potentially inadequate control actions identified, it is possible to prioritize according to which control actions have the greatest role in keeping the system from transitioning to a hazardous state.

The Missile Defense Agency conducted a Non-Advocate Safety Assessment of the Ballistic Missile Defense System using the safety assessment methodology described above. The BMDS is composed of a number of independently developed Elements, many of which had a lengthy history of development and operations before being integrated. The STPA safety assessment methodology based on systems theory provided an orderly, organized fashion in which to conduct the analysis. The effort successfully assessed safety risks arising from the integration of the Elements. The assessment provided the information necessary to characterize the residual safety risk of hazards associated with the system. The analysis and supporting data provided management a sound basis on which to make risk acceptance decisions. Lastly, the assessment results were also used to plan mitigations for open safety risks. As changes are made to the system, the differences are assessed by updating the control structure diagrams and assessment analysis templates.

References

[1]Leveson, Nancy. "A New Approach to Hazard Analysis for Complex Systems." International Conference of the System Safety Society, Ottawa, August 2003.

[2]Leveson, Nancy. "A New Accident Model for Engineering Safer Systems." Safety Science, Vol. 42, No. 4, April 2004.