SAND2015-3655C

An Adversary's View of Your Digital System

F. Mitch McCrory, Raymond C. Parks, Robert L. Hutchinson

Sandia National Laboratories[1], Albuquerque, New Mexico, United States

fmmccro@sandia.gov

Abstract. Guarding against cyber-attacks can be both challenging and expensive for any computer system defender. For nuclear power plants, the potential consequences of a successful cyber-attack make this task even more difficult and costly. To address these issues we advocate performing an adversary-based assessment of critical digital systems early in the design process to correct deficiencies and create a robust system. Through the lens of an adversary you could pinpoint how your critical digital systems could be attacked. This perspective is seldom exploited over the course of the design phase of a system, yet this is likely the most effective and least costly point at which to mitigate problems. During the design phase there is an opportunity to make changes to a design to improve its effectiveness and its ability to be secure throughout its lifetime. At this stage, there are unique opportunities to strengthen the architecture of the system for improved security, make security requirements overt, and possibly influence requirements of interfacing systems. This is also a key time to set the initial security benchmark for use with security metrics as they are developed and used over the lifecycle of the system. Creating adversary views of a system provides value over the full lifecycle of a system, but will pay the biggest dividend when done early in the development cycle.

Key Words: Nuclear Power Plants, Cyber, Adversary, Assessments

1. Introduction

This paper explores the merit of performing a security-focused, adversary-based assessment (ABA) early in the design lifecycle of a nuclear power plant (NPP). An ABA can be performed at any time in the lifecycle, but it provides the most effective opportunity for strengthening the security of the system, if performed early.

In the United States (US), 10CFR73.54 [1] requires all NPP operators to submit a cyber security plan that provides high assurance that, "digital computer and communication systems and networks are adequately protected against cyber attacks, up to and including the design basis threat." The US Nuclear Regulatory Commission (NRC) provides regulatory guide 5.71, *Cyber Security Programs for Nuclear Facilities* [2], that when followed, would be an NRC acceptable method for meeting the requirements of 10CFR73.54. The digital components and systems that fall under 10CFR73.54 include: safety-related and important-to-safety functions; security functions; emergency preparedness functions, including offsite communications; and support systems and equipment which, if compromised, would adversely impact safety, security, or emergency preparedness functions (for this paper, safety is used to infer safety, security, and/or emergency preparedness systems).

[1] Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND NO. 3026-XXXXP

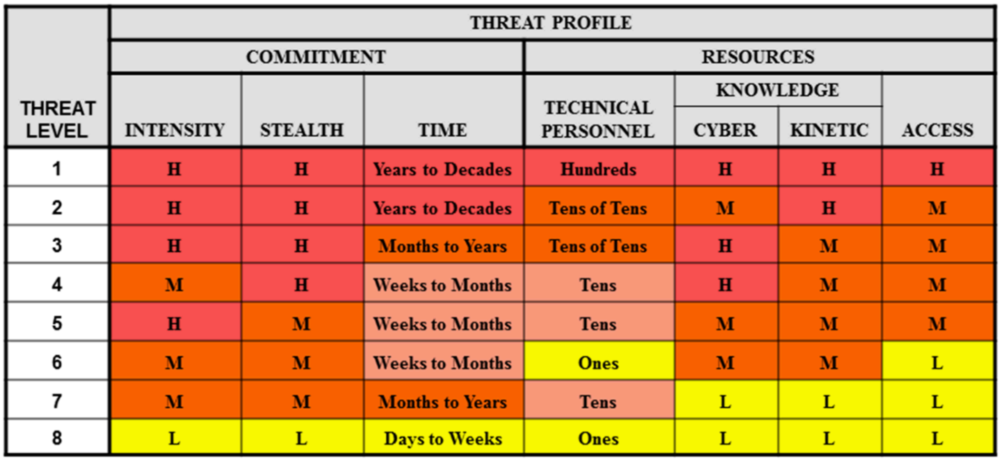
This paper discusses and explores what is meant by an ABA, why and when one should be performed, what are the important steps of an assessment, what is the output of the assessment, what the results tell you, and the limitations of the ABA.

2. An Adversary's View

What is meant by an adversary's view of your system? When designing a system's security requirements, it is necessary to ask the question: "Secure from whom?" Without a good understanding of the commitment and resources of adversaries the system is designed to defend against, defenses can be inadequate or excessive – both leading to major costs and consequences. For NPPs, the level of adversary that needs to be protected against is a range from Recreational Hackers to Nation States [3]. The US nuclear power industry is required to design their physical protection and safeguards systems to prevent acts of radiological sabotage and to prevent the theft of special nuclear material [4]. The basis for the level of protection required for these systems is based on their design basis threat (DBT) [4], [5]. These requirements explicitly include protection against cyber-security related events.

Sandia National Laboratories' (SNL) Information Design Assurance Red Team (IDART^{TM}) [6] has developed and implemented a generic eight-level adversary matrix, Table I: Generic Threat Matrix (GTM), which provides various metrics for defining the capability of an adversary for use in an ABA. Methods such as this can be used to ensure that all stakeholders are communicating threat, vulnerability, and risk effectively. Because the cost (resources, manpower, etc.) of defending against a top-level adversary is expensive, it is important to understand early on in the design lifecycle what parts of the system are important to achieving an adversary's goal, as it is not reasonable (or possibly even necessary) to defend all parts of the system(s) at the same level. Adversary-based views of a system early in the design phase of the lifecycle can help ensure a system security architecture that participates in its defense and is relatively cost effective.

Table I - Generic Threat Matrix [7]



	THREAT PROFILE						
	COMMITMENT			RESOURCES			
					KNOWLEDGE		
THREAT LEVEL	INTENSITY	STEALTH	TIME	TECHNICAL PERSONNEL	CYBER	KINETIC	ACCESS
1	H	H	Years to Decades	Hundreds	H	H	H
2	H	H	Years to Decades	Tens of Tens	M	H	M
3	H	H	Months to Years	Tens of Tens	H	M	M
4	M	H	Weeks to Months	Tens	H	M	M
5	H	M	Weeks to Months	Tens	M	M	M
6	M	M	Weeks to Months	Ones	M	M	L
7	M	M	Months to Years	Tens	L	L	L
8	L	L	Days to Weeks	Ones	L	L	L

While IDART^{TM} uses the GTM as a basis for describing the adversary threat to a system under assessment, we rarely use it without customization and modification. Recently, we developed a potential representation of the Attacker Profiles defined in Nuclear Security Series (NSS) No. 17 [3]. Our customization includes the additional information in the GTM that is not present in the original document and tries to limit the uncertainty for some attacker profiles as shown in Table II - NSS Attacker Profiles Adversary Threat Matrix.

An adversary view of a system is intended to look at a system from the perspective of an adversary's point-of-view. One of the areas to try and avoid while designing security into a system is assuming that those whom might attack the system have similar sociological, economic, political, religious, and/or other values. This helps avoid mirroring, which in this application can be a subconscious or conscious judgement that an adversary might not take a course of action because the consequences are too grave to consider. Developing an ABA helps to account for the goals and objectives of a defined set of adversaries.

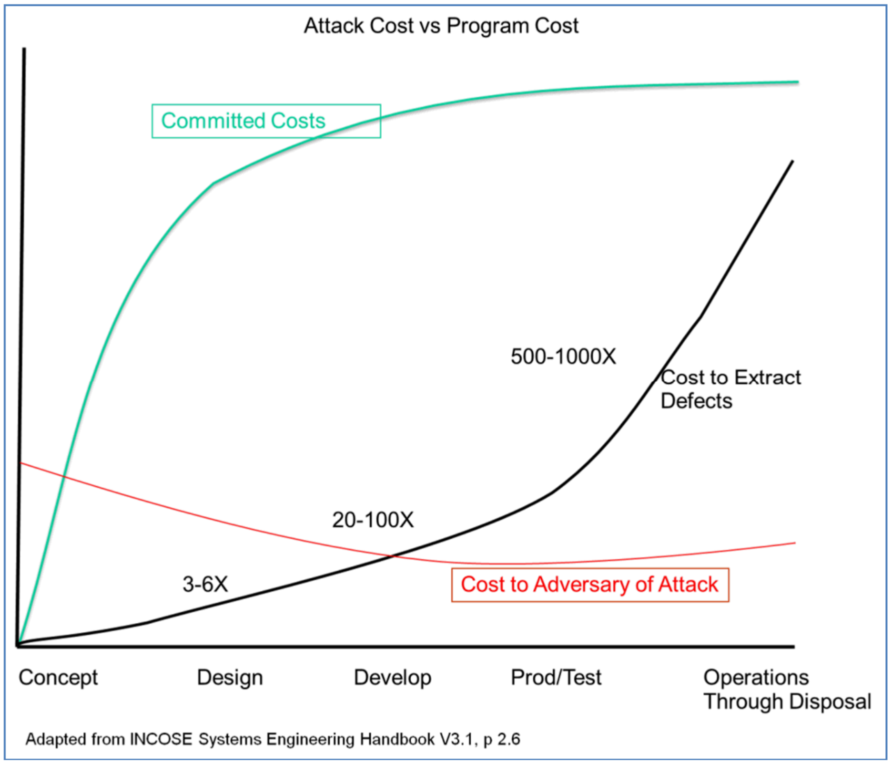
Table II – Potential NSS Attacker Profiles Adversary Threat Matrix

		Commitment Category			Resources Category					
NSS Name	GTM #	Intensity	Stealth	Time	Technical Personnel	Computer Knowledge	Physical Security Knowledge	Nuclear Engineering Knowledge	Access	Motivation
Covert Agent	6	Medium	Medium	Weeks to Months	Ones	Medium	Medium	Medium	Medium	Theft of business information, technology secrets, personal information. Economic gain (information selling to competitors). Blackmail.
Disgruntled employee /user	6-	Low	Low	Weeks to Months	Ones	Medium	Medium	Medium	Medium	Revenge, havoc, chaos. Theft of business information. Embarrass employer/other employee. Degrade public image or confidence.
Recreational hacker	8	Low	Low	Days to Weeks	Ones	Low	Low	Low	Low	Fun, status. Target of opportunity. Exploitation of 'low hanging fruits'.
Militant opponent to nuclear power	3 to 7	Medium	Low	Months to Years	Tens	Medium	Medium	Medium	Low	Conviction of saving the world. Sway public opinion on specific issues. Impede business operations.
Disgruntled ex-employee /user	7+	Low	Low	Weeks to Months	Ones	Medium	Medium	Medium	Low	Revenge, havoc, chaos. Theft of business information. Embarrass employer/other employee. Degrade public image or confidence.
Organized Crime	4-5	Medium	Medium	Months to Years	Tens of Tens	Medium	Medium	Medium	Medium	Blackmail. Theft of nuclear material. Extortion (financial gain). Play upon financial and perception fears of business. Information for sale (technical, business or personal).
Nation State	1 or 2	High	High	Years to Decades	Hundreds	High	High	High	High	Intelligence collection. Building access points for later actions. Technology theft.
Terrorist	2 or 3	High	Medium	Months to Years	Tens of Tens	Medium	Medium	Medium	Medium	Intelligence collection. Building access points for later actions. Chaos. Revenge. Impact public opinion (fear).

Looking at the system from an adversary's point-of-view can generate questions that are not often taken into account, exposing potential attack paths that the system designers have overlooked. What is the motive of the adversary? Are there attractive attack scenarios that, if successful, would result in high consequence events? How easy or how hard would it be for an adversary to execute one or more attack scenarios and which, if any, mechanisms in the security architecture would impede or prevent the adversary from success?

3. Why and When Should an Adversary-Based Assessment be Performed

Experience has taught digital system security practitioners that guarding against cyber-attacks can be both challenging and expensive. For NPPs, the potential consequences of a successful cyber-attack are potentially catastrophic. To address these issues we advocate performing an ABA of critical digital systems early in the design process to correct deficiencies and create a robust system. Through the lens of an adversary you will pinpoint how your critical digital systems could be attacked. This perspective is seldom exploited over the course of the design phase of a system, yet this is likely the most critical and influential time over the lifecycle of the system to ask this question. During the design phase there is opportunity to make changes to a design to improve its effectiveness and its ability to be secure throughout its lifetime. At this stage, there are unique opportunities to strengthen the architecture of the system for improved security, make security requirements overt, and possibly influence requirements of interfacing systems. This is also a key time to set the initial security benchmark for use with security metrics as they are developed and used over the lifecycle of the system. The cost of developing an attack drops for the attacker during the design phase while committed costs rise rapidly as does the cost to extract defects and vulnerabilities for the defender as shown in Figure 1.



Adapted from INCOSE Systems Engineering Handbook V3.1, p 2.6

Figure 1, Attack Cost vs. Program Cost

4. Important Steps of an Adversary-Base Assessment

An ABA can be done in many different ways. One method that Sandia uses in the performing an ABA is based on Sandia's IDART^{TM} methodology and includes planning, data collection, characterization, analysis, report generation, and engagement. Figure 2 shows the methodology the Sandia IDART^{TM} uses when performing an ABA. The inputs, constraints, process, and output of each stage are defined briefly in the discussion below. While these bullets are not detailed or all-inclusive of everything performed in each process step, they provide the reader a basic understanding of the information needed for each step, basically what the process entails, and the output for each stage.

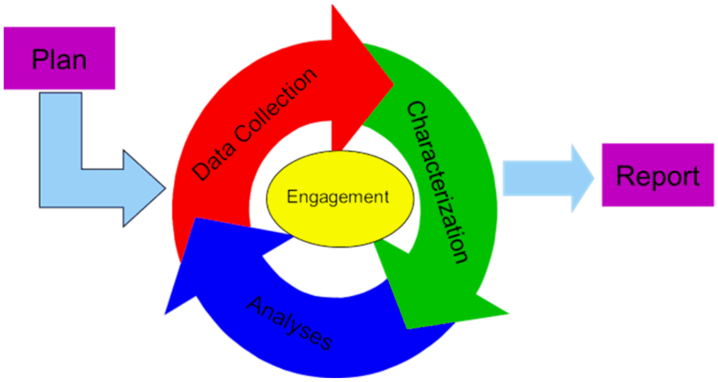


Figure 2, IDART™ Methodology

4.1 Planning

The planning phase of an ABA can contribute the most to the success or failure of the entire effort. The objective is to confirm the customer's problem and define the focus, scope, and scale of the red teaming effort.

The inputs to this phase are:

- The customer's perceived problem
- Analysis of how ABA would best work to satisfy that problem using the Red Teaming for Program Managers (RT4PM) [8] process, if done in advance
- Any red team metrics considerations identified in advance
- Initial funding profile
- Customer requirements independent of the perceived problem
- Customer negotiations
- Non-disclosure agreement
- Statement of work
- Boundaries of the assessment

The planning constraints include:

- The customer's budget and timeline
- Customer's enterprise mission, culture, and policies
- The threat(s)/adversary description(s) the red team will model
- The awareness of the threat

The actual planning process takes the inputs and:

- Jointly identifies the customer's concerns, aims, and expectations
- Determines the type(s) of red teaming to be performed (if not already done with RT4PM)
- Determines the responsibilities of the red team and the logistics of the effort
- Develops a project plan

The outputs of the planning phase can include:

- A detailed plan for the effort, to include agreements, resource requirements, and capture plans

- A concise, balanced problem statement(s)
- A concise statement of the customer's nightmare consequences
- The rules of engagement (ROE)
- Specific threat model, derived from generic threat model
- Identification of primary team and subject-matter experts (SMEs)

4.2 Data Collection

The objective of data collection is to collect the data required to characterize the customer's mission and its associated systems.

The inputs of data collection include:

- The red team plan including the scope and scale of the effort
- The nightmare consequences
- Target system context(s)
- The interfaces between the target system(s) and external systems
- System documentation
- Open-source information
- Data from site visits

The constraints on data collection include:

- The budget and timeline
- The relative availability of the information
- Agreements between the customer and the red team
- The level of cooperation between the red team and defenders/operators/blue team as dictated by the objectives

The data collection process may involve:

- Identifying the likely data sources (virtual, paper, human)
- Eliciting the data from these sources
- Validating the data – as-built versus as-designed
- Reviewing system documentation and asking for more as needed
- Conducting anonymous open source intelligence searches
- Undertaking external and internal engagements to collect data
- Conducting personnel interviews

The outputs of the data collection process are:

- System description(s) and mission(s)
- Concept of operations
- Additional nightmare consequences
- Raw data to be collated, analysed, and categorized into views

4.3 Characterization

The objective of characterization is to assemble the collected data into views, which the red team can use to understand and analyse the system. The team uses the views to identify potential vulnerabilities in the system.

The inputs of characterization include:

- Collected data
- Existing system diagrams
- Consequences

- Target system dependencies and interdependencies
- Target system description
- Target system mission
- Critical success factors

The constraints of characterization include:

- The budget and timeline
- The relative availability of the information
- Agreements between the customer and the red team
- The level of cooperation between the red team and defenders/operators/blue team as dictated by the objectives

The process of characterization has the following steps:

- Determine the necessary views based on the nature of the target system, the red team's requirements, and the available data
- Distribute view creation tasks to team
- Choose existing view types or develop new types as needed.
- Choose an appropriate communications medium
- Characterize target system dependencies
- Identify the target system's critical success factors, or the actions, factors, and assumptions required for the system to fulfill its mission
- Validate views with customer

The outputs of the characterization phase are:

- One or more of system, physical/spatial, functional/logical, temporal, lifecycle, and consequence views
- Other views as indicated by target subject
- Single points of failure
- High-value nodes (from the attacker's perspective)
- Assumptions and questions related to the views constructed by the red team

4.4 Analysis

This is the meat of the process but it requires all the previous steps for success. The objective of this phase is to analyze the system using the views to identify, explore, and prioritize possible attacks.

The inputs of analysis are:

- The system views
- SME inputs

The constraints on analysis include:

- The available time and budget
- The adversary model and the associated capabilities and characteristics
- The assessment ROE, if applicable (for attack selection)

The process of analysis involves these steps:

- Perform consequence analysis using adversary goals for guidance
- Assess target system's performance in face of attack
- Analyze for vulnerabilities (and for strengths)
- Brainstorm attacks and develop attack metrics
- Apply necessary attack filters
- Identify enabling attack resources to support engagements, if applicable

- Validate attack steps for viability against system design
- Review adversary metrics for reality in light of system design
- Save complete attack graph
- Apply adversary capability screen(s) to produce new attack graph(s)
- Apply ROE filter to attack graph – identify in-bounds and out-of-bounds attacks
- Identify and document critical attack paths
- Present the identified set of attacks to the customer
- Assess the attacks for mitigation options
- Develop attacks for engagement (optional)

The output of the analysis phase includes:

- A fully characterized and screened set of attacks that meet the attacker's goals
- An attack graph (preferred) or tree, textual descriptions, and/or attack flow charts
- System strengths, weaknesses, and mitigation strategies for identified attacks of concern

When an ABA is performed early in the design phase as advocated for in this paper, the analysis part of the IDART^{TM} process is instrumental in informing the design process on how the proposed architecture might be attacked or otherwise manipulated by an adversary. An example of how part of this phase can be used to inform the designer is shown the generic attack graph shown in Figure 3. On the right side of the graph are listed the potential adversary-based goals. On the left side of the graph are the potential starting access points of the adversary based on the understanding of the DBT and the adversary capabilities. The middle part of the attack graph shows the steps required by the adversary to move from access to goal. Analysis of these steps often identifies architectural weaknesses that need to be addressed, the best location to place a security sensor, where additional security access requirements need to be added, and other security design considerations that need to be considered.

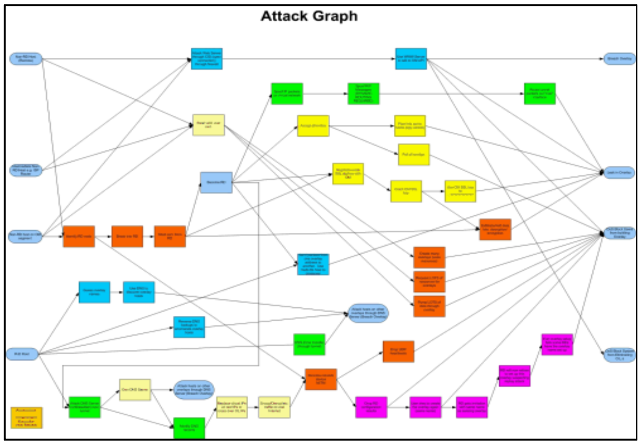


Figure 3, Hypothesized Attack Graph

Difficulty metrics, based on the adversary model used, can be applied to the various steps of the attack graph to help generate the relative risk of each attack path through the graph. These difficulty metrics can be manipulated to ask, "What if?" questions or to do sensitivity analysis of the assumptions used to generate the metrics.

4.5 Engagement

Engagements can happen at any point after planning and before reporting. They can support any or all of the other activities, such as during data collection or analysis. The objective of engagements is to perform system-testing activities that supply needed data, support or refute a hypothesis, demonstrate the feasibility or consequences of an attack, verify one or more vulnerabilities, or test one or more mitigations. When performing the ABA during the design phase of a system, the engagement process is often limited since physical systems are not likely available to physically test adequacy of the design.

The inputs to any engagement include:

- Engagement objectives: what the customer wants to learn/achieve
- Engagement objectives: to collect data, characterize the system, or to analyze system
- Engagement resources, constraints, and participants

The constraints on engagement include:

- Time and budget
- Location (travel)
- The adversary model
- Specific authorizations
- Ensuring operational security
- Risks associated with engaging an operational system

The process of engagement encompasses the following steps:

- Clearly define the purpose of the engagement
- Plan, develop, and, as necessary, test the engagement activities
- Perform the engagement, collect the resulting data, and report results to the customer
- A comprehensive operational plan that includes the rules of engagement, attack descriptions, targets, schedule, points of contact, recovery instructions/guidance, operational security, and authorizations. This is the initial output of the process before the actual activity

The outputs of the engagement process include:

- Customer inbrief/outbrief as appropriate, engagement metrics – results, lessons-learned
- New information regarding the nature of a vulnerability or the behavior of the target system
- Potential mitigations

4.6 Report Generation

The assessment is useless unless the red team communicates their findings to the customer in a way that is actionable. The objective of reporting is to report the red team's findings to the customer clearly and intuitively.

The inputs to reporting include all materials generated during the IDART process, including chronologies and data logs.

The constraints on reporting include:

- The available time and budget
- The customer's background and preferences

The process of reporting includes these steps:

- Determine the type of information the customer requires

- Collect the materials generated during the process
- Write a report that answers the customer's security questions

The output of reporting is a report tailored to the customer's needs that adequately documents the red team's findings.

5. Result Interpretation, Limitations, and Conclusion

The results of the ABA should be used to identify where potential design modifications or additional security requirements need to be made to meet the security requirements of the NPP as dictated by the DBT. The results of the assessment can often be used to identify the need for additional resources or to prioritize the limited resources available. It is important to note that ABAs are snapshots in time and that if the threat changes, portions of the assessment maybe invalidated and need to be redone. If the system design is modified based on the ABA, it can be useful to rework part of the ABA against the affected design. While we advocate for performing an ABA as early in the design lifecycle as possible, this process should be done periodically to account for changes in the system as it is used and operated, to account for changes in threat and technology, and whenever it is needed to validate the current security posture of the system.

NPP safety systems are subject to a broad range of evolving threats and, given the consequences, must be able to complete their safety functions through failure and attack. ABA may be the only NPP security approach that can: (1) address evolving threats, and (2) do so efficiently and effectively. We advocate for the routine and continuous use of ABA as a basic security tool for threat-informed risk management. ABA must be applied throughout the entire lifecycle; only then can we effectively anticipate evolving threats and decide how to best address those threats.

REFERENCES

[1] US Code of Federal Regulations, Title 10, Part 73.54, *Protection of digital computer and communication systems and networks,* Mar. 27, 2009.

[2] Regulatory Guide 5.71, "Cyber Security Programs for Nuclear Facilities", U.S. Nuclear Regulatory Commission, January 2010.

[3] International Atomic Energy Agency, Computer Security at Nuclear Facilities IAEA Nuclear Security Series No. 17, IAEA, Vienna (2011)

[4] US Code of Federal Regulations, Title 10, Part 73.1, *General Provisions*, Nov. 28, 1979.

[5] NRC DBT definition: http://www.nrc.gov/reading-rm/basic-ref/glossary/design-basis-threat-dbt.html

[6] The Information Design Assurance Red Team (IDART^{TM}): http://www.idart.sandia.gov/

[7] SAND 2007-5791: http://energy.gov/sites/prod/files/oeprod/DocumentsandMedia/14-Categorizing Threat.pdf

[8] Red Teaming for Program Managers: http://www.idart.sandia.gov/methodology/RT4PM.html