

SAND REPORT

SAND2002-1449
Unlimited Release
Printed May 2002

Level 1 Peer Review Process for the Sandia ASCI V&V Program: FY01 Final Report

Martin Pilch, Gary Froehlich, Ann Hodges, David Percy, Timothy Trucano, Jaime Moya

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation,
a Lockheed Martin Company, for the United States Department of
Energy under Contract DE-AC04-94AL85000.

Approved for public release; further dissemination unlimited.



Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

NOTICE: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof, or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865)576-8401
Facsimile: (865)576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: <http://www.doe.gov/bridge>

Available to the public from

U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800)553-6847
Facsimile: (703)605-6900
E-Mail: orders@ntis.fedworld.gov
Online order: <http://www.ntis.gov/ordering.htm>



SAND2002-1449
Unlimited Release
Printed May 2002

Level 1 Peer Review Process for the Sandia ASCI V&V Program: FY01 Final Report

Martin Pilch
Validation and Uncertainty Quantification Processes

Gary Froehlich and Ann Hodges
Software and Information Engineering

David Percy
Quality Engineering

Timothy Trucano
Optimization and Uncertainty Estimation

Jaime Moya
Validation and Qualification

Sandia National Laboratories
P. O. Box 5800
Albuquerque, New Mexico 87185-0638

Abstract

This report describes the results of the FY01 Level 1 Peer Reviews for the Verification and Validation (V&V) Program at Sandia National Laboratories. V&V peer review at Sandia is intended to assess the ASCI code team V&V planning process and execution. The Level 1 Peer Review process is conducted in accordance with the process defined in SAND2000-3099. V&V Plans are developed in accordance with the guidelines defined in SAND2000-3101. The peer review process and process for improving the Guidelines are necessarily synchronized and form parts of a larger quality improvement process supporting the ASCI V&V program at Sandia. During FY00 a prototype of the process was conducted for two code teams and their V&V Plans and the process and guidelines updated based on the prototype. In FY01, Level 1 Peer Reviews were conducted on an additional eleven code teams and their respective V&V Plans. This report summarizes the results from those peer reviews, including recommendations from the panels that conducted the reviews.

Acknowledgements

The authors thank Stephen E. Lott, Mary A. McWherter-Payne, Amalia R. Black, and Richard O. Griffith for formally reviewing this report. Their comments were appreciated and have improved the quality of this report.

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

Table of Contents

ACKNOWLEDGEMENTS	4
TABLE OF CONTENTS	5
LIST OF FIGURES	6
LIST OF TABLES	6
EXECUTIVE SUMMARY	7
1 INTRODUCTION	15
1.1 BACKGROUND, PURPOSE, AND SCOPE.....	15
1.2 INTENDED AUDIENCE.....	15
1.3 BENEFITS AND COST.....	15
1.4 OVERVIEW OF DOCUMENT.....	16
2 OVERVIEW OF THE LEVEL 1 PEER REVIEW PROCESS	17
2.1 LEVEL 1 PEER REVIEW PROCESS EVOLUTION.....	17
2.2 LEVEL 1 PEER REVIEW PROCESS.....	19
2.2.1 <i>Review Session</i>	21
2.2.2 <i>Review Panel Selection</i>	21
2.2.3 <i>Review Materials</i>	22
2.2.4 <i>Review Process</i>	22
2.2.5 <i>Review Deliverables</i>	24
3 DESCRIPTION OF FY01 LEVEL 1 PEER REVIEW SESSIONS	25
3.1 SESSION 1: MARCH 6-8, 2001.....	27
3.1.1 <i>Review Panel Selection</i>	27
3.1.2 <i>Review Materials</i>	27
3.1.3 <i>Review Deliverables</i>	27
3.2 SESSION 2: MARCH 28-29, 2001.....	28
3.2.1 <i>Review Panel Selection</i>	28
3.2.2 <i>Review Materials</i>	28
3.2.3 <i>Review Deliverables</i>	28
3.3 SESSION 3: MAY 15-17, 2001.....	29
3.3.1 <i>Review Panel Selection</i>	29
3.3.2 <i>Review Materials</i>	29
3.3.3 <i>Review Deliverables</i>	29
3.4 SESSION 4: SEPTEMBER 18-20, 2001.....	30
3.4.1 <i>Review Panel Selection</i>	30
3.4.2 <i>Review Materials</i>	30
3.4.3 <i>Review Deliverables</i>	30
4 RESULTS, ANALYSIS, AND RECOMMENDATIONS	31
4.1 V&V PLAN.....	31
4.1.1 <i>Conformance to V&V Guidelines</i>	31
4.1.2 <i>Strengths</i>	31
4.1.3 <i>Opportunities for Improvement</i>	32
4.2 V&V GUIDELINES.....	33
4.2.1 <i>Strengths</i>	33

4.2.2	<i>Opportunities for Improvement</i>	34
4.3	V&V LEVEL 1 PEER REVIEW PROCESS	37
4.3.1	<i>Strengths</i>	37
4.3.2	<i>Opportunities for Improvement</i>	37
4.4	GENERAL OBSERVATIONS	39
5	SUMMARY OF MAJOR RECOMMENDATIONS	41
5.1	CONTINUE USE OF V&V GUIDELINES AND PEER REVIEW PROCESS	41
5.2	FOCUS LEVEL 2 PEER REVIEW PROCESS ON INTEGRATED STS APPLICATIONS.....	41
5.3	INTEGRATE THE V&V PLAN PARTICIPANTS FOR THE MOST EFFECTIVE RESULTS.....	41
5.4	QUANTIFY METRICS AND ACCURACY REQUIREMENTS MORE RIGOROUSLY	42
5.5	DEVELOP ORGANIZATIONAL LEVEL GUIDANCE FOR SQE, VERIFICATION, VALIDATION, AND STOCKPILE COMPUTING.....	42
5.6	REVISE PEER REVIEWED V&V PLANS AND PUBLISH.....	42
5.7	REVISE V&V GUIDELINES 2.0 AND PEER REVIEW PROCESS 1.0 AND PUBLISH	42
	APPENDIX A. ACRONYMS AND REFERENCES	43
A.1.	ACRONYMS	43
A.2.	REFERENCES	44
	APPENDIX B. LEVEL 1 PEER REVIEW SCORING SUMMARIES	47
B.1.	SCORING SUMMARIES AND ANALYSES	47
B.2.	NOTEWORTHY PRACTICES.....	56

List of Figures

Figure 2-1.	Conceptual Evolution of the V&V Peer Review Process	18
Figure 2-2.	Sandia ASCI V&V Program Peer Review Levels.....	19
Figure B-1.	Overall Average of Scores by V&V Criterion.....	49
Figure B-2.	Overall Average of Scores by V&V Criterion (no Frameworks).....	51
Figure B-3.	Highest V&V Plan Average Score for each Criterion.....	52
Figure B-4.	Lowest V&V Plan Average Score for each Criterion.....	53
Figure B-5.	Example of Best V&V Plan Score vs. Overall Average Score	54
Figure B-6.	Comparison of Overall Average Scores for non-Integrated vs. Integrated Code Teams.....	55

List of Tables

Table 2-1.	Level 1 Peer Review Deliverables.....	24
Table 3-1.	Pre-Review Calendar: Session 1 Level 1 Peer Review Implementation Plan.....	25
Table 3-2.	Level 1 Peer Review Code Team Participants	26
Table B-1.	V&V Guidelines Criteria	47
Table B-2.	List of Noteworthy Practices.....	56

Executive Summary

This report describes the results of the Level 1 Peer Reviews of ASCI code team V&V plans during FY01. The results include programmatic review information related to the V&V plans and improvement suggestions related to the review process. This summary highlights the overall process, results from the reviews, and major recommendations derived from the review process.

PROCESS

The description of the Level 1 Verification and Validation (V&V) Plan Peer Review Process at Sandia National Laboratories is contained in Peer Review Process 1.0 [V&V-REVIEW-1_0]. The Level 1 V&V peer review at Sandia is intended to assess the ASCI code team V&V planning process and execution status. The peer review definition is designed to assess the V&V planning process in terms of the guidelines specified V&V Guidelines 2.0 [V&V-GUIDE-2_0].

The peer review process and process for improving the Guidelines are necessarily synchronized, and form parts of a larger quality improvement process supporting the ASCI V&V program at Sandia. The Level 1 Peer Review process is the first level of three peer review levels. Level 1 focuses on the programmatic aspects of creating an acceptable V&V Plan. Level 2 focuses on the technical results of executing the V&V Plan. Level 3 focuses on the use of the code as part of a stockpile certification calculation.

V&V plans are expected to demonstrate and document how the quality and fidelity of the code M&S (modeling and simulation) capability will satisfy specific stockpile programmatic requirements. Each code team, in partnership with their stockpile customers and designers, is expected to create, maintain, and execute these plans throughout the life cycle of the code.

The V&V Guidelines 2.0 specifies the general format and content of a V&V Plan. The V&V Plan defines the V&V information for modeling and simulation, in support of establishing the requirements, design, implementation, and/or qualification of a specific stockpile driver application. At the core of this information is a Phenomena Identification and Ranking Table (PIRT). The PIRT is a requirements management mechanism that links a stockpile application and its requirements to the phenomena required to be modeled/simulated in support of the application, and to specific verification tests and validation tests that provide assurance the implemented simulation model results adequately support the stockpile application.

The Peer Review Process 1.0 specifies the process for conducting V&V Peer Reviews. Review panel members were selected from customers (e.g., Defense Programs and Surety), designers and analysts familiar with the code, and the V&V Program office. Two or three code teams were combined into one review session. Core review panel participants reviewed all the plans and rotator participants reviewed only those specific to their areas of expertise. The four steps to the review process are: train the review panel; prepare for the review by reading and scoring the V&V Plan; conduct the review of a code team's V&V; and follow up with written reports to each code team, and a written report to the ASCI V&V Program Office summarizing the participating code team results and process improvement suggestions. Code teams were required to provide some action plan information to the ASCI V&V Program Office in response to the written report.

SESSIONS

An initial set of V&V Guidelines and a Level 1 Peer Review Process were developed in FY00 and used in a prototype Level I Peer Review involving two code teams and their V&V Plans. Results from that prototype were used to improve the guidelines and review process. The V&V Guidelines 2.0 and Peer Review Process 1.0 were published, and plans were made to conduct Level 1 Peer Reviews for eleven code team V&V Plans during FY01. Funding was provided through the ASCI V&V program for development of the V&V Plans and for the review of those plans. The following code teams participated in the FY01 reviews in four different sessions.

- Session 1 (March 6-8, 2001): SIERRA Framework, CALORE, FUEGO-SYRINX
- Session 2 (March 28-29, 2001): GOMA, HPEMS-XYCE
- Session 3 (May 15-17, 2001): ALEGRA Framework, EMPHASIS, CEPTRE-ITS
- Session 4 (September 18-20, 2001): NuGET, PEGASUS, ALEGRA-EMMA

RESULTS

V&V Plan Content:

Overall:

All Code Team V&V Plans followed the V&V Guidelines format. Exceptional V&V Plans were presented that satisfied nearly all of the programmatic requirements. More detailed specification of performance measures would improve most plans.

The Review Panels felt that the V&V Plans conformed to the format recommended by the Guidelines. The Code Teams made a valid attempt to identify their DP customer(s) in their Plans, although the Surety customers were not adequately identified when applicable. Many code teams inadequately addressed the Validation content. Performance measures for SQE, verification, and validation were inadequately addressed.

Strengths:

Through the PIRT mechanism, the V&V Plans have established a common terminology for all code teams, a common method for describing the identification and prioritization of customer requirements, and a capability to trace these requirements to implementation and testing. There were several strong V&V Plans and a large number of individual noteworthy approaches to describing the V&V planning information.

The Panels were impressed by the energy, effort, and professionalism that the Code Teams brought to the V&V planning and peer review process. Code Teams participated in the process in a professional and cooperative manner. All Code Teams conducted discussions of the feedback in an honest and accepting manner.

Most of the Plans were good. The Plan authors are aware of the basic requirements of V&V, are beginning to engage the DP and Surety customers, recognize minimal SQE requirements, and

demonstrate a basic understanding of computational modeling and V&V activities in a certification environment.

The Code Teams with prior experience in using their code for certification activities were much more aware of the issues, and addressed the stockpile computing issues more completely. An effort was made by most Plan authors to address software quality engineering (SQE) issues. Most of the major components of SQE were at least recognized and commented upon. The development of the PIRTs, VERTS, and VALTS were *very* effective for a couple of the V&V Plans. There were several strong plans and individual noteworthy practices, including:

- Team organization: integrated customer, analysts, experimentalists, code developers
- Identification of & justification for selection of various V&V Plan participating roles
- Document tree/reference use
- Stockpile Driver description and tie to M&S Requirements
- General PIRT Chart Representation and Scoring/Ranking system
- Good description of the prioritization methodology and its use
- Concept of extracting application-specific PIRTs from a more general Global PIRT
- SQE practices defined, being used, and found useful
- Question/answer table format for VERTS/VALTS representation/justification
- Concept of VERTS/VALTS "tiers" was expanded to cover levels of phenomena
- Mapping of V&V criteria to V&V Plan sections in which criteria discussed
- VERTS description and tie to PIRT
- VALTS description and tie to PIRT
- Metrics for assessing the success of VERTS/VALTS
- Methodology/application of PIRT to framework services - "services" IRT

Improvement:

The use of integrated teams with analysts and experimentalists in key author roles as well as code developers is suggested for all V&V Plans. The one common technical area that could use improvement is in the specification of V&V performance measures and success criteria.

Overall the Plans need to link more completely to customer schedules and milestones. Plans should include traceability between the PIRT and the code's capability releases. While the V&V Plans typically identify the experimenters, analysts, and various customers in a general sense, more specific links would be very helpful to readers of the Plans.

The current process appears to work much better with an analyst or phenomenologist as the lead author, than with a code developer alone in that role. Execution of different elements of the V&V Plans must belong to different people, hence the recommendation that *ownership* (as opposed to mere authorship) consist of code representatives, customer representatives, analysts, and experimentalists. The intent is that the various roles actually *participate* in V&V planning. It is essential that the owners of the V&V Plans have the authority and the resources to engage the necessary participants in the development of the V&V Plan.

Ties to needed experiments, and hence to required experimental accuracy, should be discussed more fully. This includes both experiments to provide input data (e.g., material properties), and validation experiments.

V&V Teams could compile fairly comprehensive lists of multiple stockpile drivers and related phenomena (as a few teams have already done). Then, each time a particular application is identified, a "slice" through such a tabulation would provide the input for a new prioritization activity, thus leading to a PIRT which forms the core of the new, application-specific V&V Plan.

The V&V Plans should strive to quantify the accuracy requirements for validation and verification in a more rigorous fashion, or reference where such quantification is to be described. There is inadequate treatment of success measures and acceptance criteria for verification and validation, as well as metrics to assess the effectiveness of SQE practices. Related to this is the specification of needed accuracy for experimental results. Although uncertainty quantification (UQ) is considered part of a standard ASCI product, in general very little, if any, V&V aspects of the UQ and its associated process have been discussed.

V&V Guidelines Version 2.0:

Strengths:

The common process and terminology with which V&V planning can be addressed was praised. The existing V&V Guidelines are effective, particularly the PIRT representation.

The existing guidelines are adequate and effective for M&S codes. Several of the V&V Plans were excellent examples of how these guidelines could and should be applied. Those teams where the members clearly share a common vision of what must be done and how to do it, were most successful in applying the guidelines. Such teams included code developers, analysts, experimentalists, and customers.

Code teams were compliant with the form and content suggested by the guidelines; with only a few exceptions, the guidelines were able to be read and interpreted by each of the code teams. Differences in following the guidelines tended to be related more to specific organizational structure and code team philosophies than a misinterpretation of the guidelines.

A common terminology for key aspects of V&V (e.g., Stockpile Driver, PIRT, VERTS, VALTS, SQE, Stockpile Computing) was established across a wide range of code team application domains and peer review participants; perhaps most important was the understanding provided to DP and Surety customers.

Improvement:

The existing V&V Guidelines need better descriptions of the GLOBAL PIRT concept, clarification of scoring and ranking concepts, and how risk issues can be integrated within the V&V Plans.

The Guidelines should encourage teams to reference generic (i.e., SNL ASCI Program level) guidance for Stockpile Computing, UQ, and SQE. Program-level guidance for Stockpile Computing needs to be developed. Teams should focus first and foremost on PIRT, VERTS, and VALTS.

The Guidelines should emphasize that the PIRT can be viewed as a tool for requirements development and representation, and hence is applicable even to frameworks and other areas such as visualization software. Frameworks have a “services” IRT (SIRT) instead of a PIRT; some explanation of how to apply the PIRT concept to frameworks and other model-based areas as an extension of requirements management would be useful.

Most panel members feel that the best approach to V&V planning is to have a documented information source with *all* the capabilities (and associated requirements) that a code must provide -- a Global PIRT. Then the V&V Plan for each specific, single stockpile driver would refer to this master document for ‘down-selection’ of the components required for that particular application.

The Guidelines discuss using a scoring system for the PIRT to help in performing the ranking or prioritization of phenomena. Many Panel members feel that adequacy scoring and importance ranking are different activities. A given phenomenon will be scored both for its importance and the adequacy of its current representation; these scores should then be used to rank the phenomena. Most teams only performed the scoring.

The Guidelines should ask for the V&V Plan to include risk issues. The Plans should identify risks and risk response strategies -- e.g., mismatch between stockpile-driver schedule and code availability for use. Development of a programmatic-level risk-management process that V&V Plans could reference (as with SQE, UQ, Stockpile Computing Guidelines) is suggested.

Examples of risks identified during the Level 1 Peer Review are:

1. Campaigns and DP milestones are not consistent.
2. Uncertainty quantification appears to be “falling through the cracks”.
3. Needed validation experiments may not be conducted in timeframe required to allow code use by DP.
4. In at least one case, it was critical for a code team to use Nuclear Weapon models from the two physics labs, and there was no apparent V&V evidence for these models.

Level I Peer Review Process Version 1.0:

Strengths:

The panels commented that without this year’s Level 1 peer review process, it is unlikely that the V&V Plans, and the code teams' understanding of V&V, would be at their current level of maturity. The involvement of DP and Surety customers in the review process was lauded and is critical to its success.

The peer review process forces teams to seriously think about what is required to perform adequate V&V. The current state of the Plans, the heightened awareness of the V&V issues related to coupled physics, and the justifiable concerns about roles and responsibilities, are all the

result of teams having confronted those issues for the purpose of developing their V&V Plans. Members on the various panels commented that without this year's Level 1 peer review process, it is unlikely that the V&V Plans, and the code teams' concomitant understanding of V&V, would be at their current level of maturity.

Selection of Panel members is considered to be very good, specifically with regard to the diverse representation of interests and expertise. It is especially commendable to include managers, customers, code subject-matter experts, and V&V/SQE expertise. The inclusion of DP/Surety customers on the review panels had an additional benefit; several of these panel members stated that they had a much better understanding of what ASCI was trying to achieve, an improved sense of confidence that we were "trying to do the right thing", and an appreciation for the difficulty of the problem.

The short turn-around time for deliverables (quick feedback, final report, etc.) is crucial to the Plan owners, and is also to be commended.

Improvement:

The peer review process should evolve to an application-centric approach: integrated teams developing integrated V&V Plans involving multiple phenomena and codes for STS environments.

The peer review process should make inclusion of a DP and/or Surety customer on the Panel a requirement rather than a suggestion. The value added by these panel members is exceptionally important. The Level 2 Peer Review process should be expanded to encompass the suggested application-centric approach to evaluating the technical content of the V&V Plan, and the resulting evidence. The Panels feel that Level 1 Reviews are still *essential* at the code-centric level, but that Level 2 and higher reviews are most relevant to application-centric V&V.

Although the peer review process cannot dictate specific Sandia organizational and process guidance, it should be emphasized that certain organizational structures and related verification and validation processes will improve the V&V Planning and execution. One key element would include having customers, experimentalists, analysts, code developers, and V&V specialists responsible for the V&V planning and execution within the suggested application-centric approach. Along with the definition of roles, the funding and management structure must permit code teams to acquire the resources necessary to implement the V&V requirements being imposed on them.

The panels were concerned that all the effort and results attained so far might be lost without appropriate sustainment support. This includes the publication of the existing V&V Plans, continued support for the development and evolution of new V&V Plans, particularly with an integrated application focus, and management commitment to continue with the Level 1, 2, and 3 peer reviews. The ASCI Program needs to provide adequate cross-cutting support for UQ efforts for all code teams, templates for many of the work products (e.g., design documents), resources for verification and validation experiments, and more personnel trained in SQE and V&V.

RECOMMENDATIONS

The following seven recommendations are from the accumulated panel reviews and represent the most important general suggestions from the reviews.

Recommendation 1. Continue Use of V&V Guidelines and Peer Review Process

The panels, in particular the customer representatives, felt that the customers should request continued use of V&V Plans and Peer Reviews as a basis for supporting a common approach for application-specific V&V activities. In addition, the customers should request the continued development of information as required by the V&V Guidelines 2.0 or its updated version.

Recommendation 2. Focus Level 2 Peer Review Process on Integrated STS Applications

The proposed Level 2 Peer Review, which will focus on technical content, should be applied primarily to integrated Stockpile-to-Target-Sequence (STS) applications. These reviews should focus on STS applications for the normal, abnormal, and hostile environments. In addition, the focus should be on coupling of multiple phenomena which would imply the integrated use of multiple codes. These integrated applications should develop V&V Plans following the V&V Guidelines 2.0 (as updated). The panel for such reviews should have strong customer participation.

Recommendation 3. Integrate the V&V Plan Participants for the Most Effective Results

The integration of code developers, analysts, experimentalists, and customers for the V&V Plan development and execution appears to provide the most effective results. The V&V Plans created by teams that had this recommended composition, and experience with stockpile certification, were the best with respect to programmatic organization and detailed content. In those cases where the center, department, and line organization infrastructure supported this integrated team concept, the review results were noticeably better.

It is essential that the owners of the V&V Plans have the authority and the resources to engage the necessary participants in the development of the V&V Plan.

Recommendation 4. Quantify V&V Success Criteria and Performance Measures

The V&V Plans should strive to quantify the accuracy requirements for verification and validation in a more rigorous fashion, or reference where such quantification is to be described. There is inadequate treatment of success criteria. This observation refers to acceptance criteria for verification and validation, as well as metrics to assess the effectiveness of SQE practices. Related to this is the specification of needed accuracy for experimental results. Although UQ is considered part of a standard ASCI product, in general very little, if any, V&V aspects of the UQ and its associated process were discussed.

Recommendation 5. Complete Organization-Level Guidance for SQE, Verification, Validation, and Stockpile Computing

ASCI program plans currently include activities for developing organization-level guidance for SQE, validation methods, and stockpile computing during FY02. It is recommended that this guidance be completed, and then referenced by the revised Peer Review Process. In addition, organization-level guidance should be provided for Verification.

Recommendation 6. Revise Peer Reviewed V&V Plans and Publish

It is recommended that each V&V Plan be revised as appropriate to incorporate suggestions from the code-specific peer review report and published as a SAND report. It is important to establish a baseline of information that will benefit future V&V Plans.

Recommendation 7. Revise V&V Guidelines 2.0 and Peer Review Process 1.0 and Publish

The V&V Guidelines 2.0 and Peer Review Process 1.0 should be updated to incorporate recommended improvements. Incorporate the suggested improvements and publish as updated versions.

1 Introduction

1.1 Background, Purpose, and Scope

In Fiscal Year 2000 (FY00), the Sandia ASCI V&V program (hereafter called the SNL V&V program) began to implement a formal three-level peer review process for the Accelerated Strategic Computing Initiative (ASCI) verification and validation (V&V) program at Sandia. Initial attention was devoted to V&V plans developed in FY00 for Sandia ASCI computational codes. Two documents, reference [V&V-GUIDE-1_0] and a draft of [V&V-REVIEW-1_0], were produced in FY00. A prototype Level 1 Peer Review was conducted in FY00 with two code teams. As a result of this prototype, the reference [V&V-GUIDE-1_0] was updated and published as reference [V&V-GUIDE-2_0], and the reference [V&V-REVIEW-1_0] was finalized. In addition, a formal lessons-learned memorandum provided official feedback to the V&V Program.

As a result of the Level 1 Peer Review prototype in FY00, plans were made to conduct a Level 1 Peer Review of V&V Plans produced by the remaining eleven code teams in FY01. The purpose of the present report is to summarize the results of the FY01 Level 1 Peer Review activities.

1.2 Intended Audience

This report is intended for use by the ASCI V&V Program as guidance for the future evolution of the V&V Peer Review Process within the Sandia Stockpile Stewardship Program. Defense Program (DP) and Surety System customers, as well as Modeling and Simulation (M&S) analysts, experimentalists, and code developers also have important roles to play in the V&V Program and future V&V Peer Review activities, and are expected to be users of this report.

1.3 Benefits and Cost

The results of the peer review process described in this report aim to contribute to the overall quality of the Sandia V&V program through the following results:

1. **Establish Status Information for the V&V Plans:** the peer review process provides metric information regarding the status of code team V&V activities to the Sandia V&V program, as well as other interested stakeholders; peer review results may also be applied to V&V program and individual code team process and product improvement.
2. **Help Direct V&V Planning:** the peer review process provides planning information for the Sandia V&V program; the process is coordinated with ASCI budget planning activities to aid in the development of program and code team implementation plans.
3. **Provide an Independent Assessment:** the peer review process is an independent assessment that can provide evidence to Defense Programs (DP) for the use of Sandia ASCI codes in weapon system qualification activities.
4. **Facilitate Future Coordination:** the peer review process contributes to coordination between the V&V program and experimental programs at Sandia, as well as DP.
5. **Improve The Review Process and V&V Guidelines:** the peer review process provides immediate feedback to individual code projects regarding their V&V plans and progress;

the process seeks to insure that scientifically credible V&V plans are produced, tasks are performed, and results are obtained for each ASCI code development project at Sandia.

The peer review process is coordinated with several specific elements of the ASCI V&V program at Sandia. These elements include:

1. Program guidelines for Sandia V&V plans in references [V&V-GUIDE-1_0] and [V&V-GUIDE-2_0];
2. V&V plans developed by individual ASCI code projects, e.g., reference [V&V-SACCARA];
3. Sandia V&V program V&V software quality engineering infrastructure and policy developed in FY00, see reference [DOE-ASCI-SQE], and FY01, see reference [SNL-ASCI-SQE]; and
4. V&V implementation planning that traditionally begins in April and is completed by the end of August of any given year.

The FY01 cost for development of the V&V Plans was approximately \$900K for eleven code teams. The FY01 cost for the publication update of the V&V Guidelines 2.0, Peer Review Process 1.0, and organization and conduct of the four peer review sessions plus the writing of this final report was approximately \$200K. Total cost to implement the peer review process in FY02 was approximately \$1.1M.

1.4 Overview of Document

Executive Summary provides a brief overview of the most significant aspects of this report, including seven general recommendations.

Section 1 provides an introduction to the scope and purpose of this document.

Section 2 summarizes the peer review process. Details can be found in references [V&V-REVIEW-1_0] and [V&V-GUIDE-2_0]

Section 3 provides a brief description of the four Level 1 Peer Review sessions conducted in FY01, including identification of the panel members, participating codes, and formal deliverables.

Section 4 summarizes Level 1 Peer Review results, analysis, and recommendations.

Section 5 provides a summary of the major recommendations, some of which are aggregates of multiple suggestions from several panels.

Appendix A includes a list of acronyms and important references.

Appendix B includes scoring summaries from the reviews and a list of noteworthy practices as identified by the various panels.

2 Overview of the Level 1 Peer Review Process

This section provides a brief overview of the Level 1 Peer Review process, its relationship to the Sandia Guidelines for V&V plans, and its relationship to the other two levels of the peer review process.

2.1 Level 1 Peer Review Process Evolution

It is helpful to distinguish the purpose of the ASCI applications program from that of the ASCI V&V program to best understand the goals in designing V&V guidelines and a peer review process.

ASCI Applications Program Purpose: To develop high performance computational tools and models to help manage the safety and reliability of the enduring nuclear stockpile.

ASCI V&V Program Purpose: To substantially increase the credible predictive content of high consequence modeling and simulation for science based stockpile stewardship.

The Sandia ASCI V&V Program mission is to establish high confidence in the use of our ASCI modeling and simulation tools by:

1. advocating the use of modern software engineering practices;
2. facilitating code verification; and
3. establishing a formal validation program.

The Sandia ASCI Program includes several code teams that develop a variety of computational codes to simulate physical phenomena and their interactions. Each code team may have one or more codes that they support. The approach is for each code to have a V&V Plan that describes the activities necessary to ensure the code can be used for its specific application. For the ASCI Program, the intended purpose is to support the Stockpile Stewardship Program for continued certification of existing weapons and planned upgrades. Certification of existing weapons and planned upgrades is dependent upon the general weapon and weapon-related processes as described in the Technical Business Practices, reference [TBP-SYS].

The Sandia V&V program has developed a reference [V&V-GUIDE-2_0] to provide overall guidance to code teams as to the expected content, activities, and evidence to be covered in their V&V plans. To assess progress on V&V plans, a three-level peer review process has been defined in reference [V&V-REVIEW-1_0]. A key goal of the peer review process is to assess conformance of V&V plans and technical results to the published Guideline requirements.

The first form of these guidelines, reference [V&V-GUIDE-1_0], was published as a Sandia report in December, 1999. The version of the peer review process discussed in this document is primarily based upon the content of these guidelines, the lessons learned from the FY00 V&V peer review prototype based on these guidelines, and the subsequent version of these guidelines, [V&V-GUIDE-2_0]. It is the intent that the peer review process description in this document will also evolve as lessons learned from prototype use.

A conceptual diagram of the content guidelines development and relationship to the three-level peer review, as shown in Figure 2-1, emphasizes some of the environmental factors that influence the implementation of the planning and peer review cycle:

1. Factors internal to the V&V program, such as changing requirements and a developing infrastructure for software quality engineering (SQE) and V&V.
2. The evolution of code specific V&V activities as work progresses over the duration of the program.
3. External program influences, most specifically the MAVEN experimental program and DP stockpile programs.

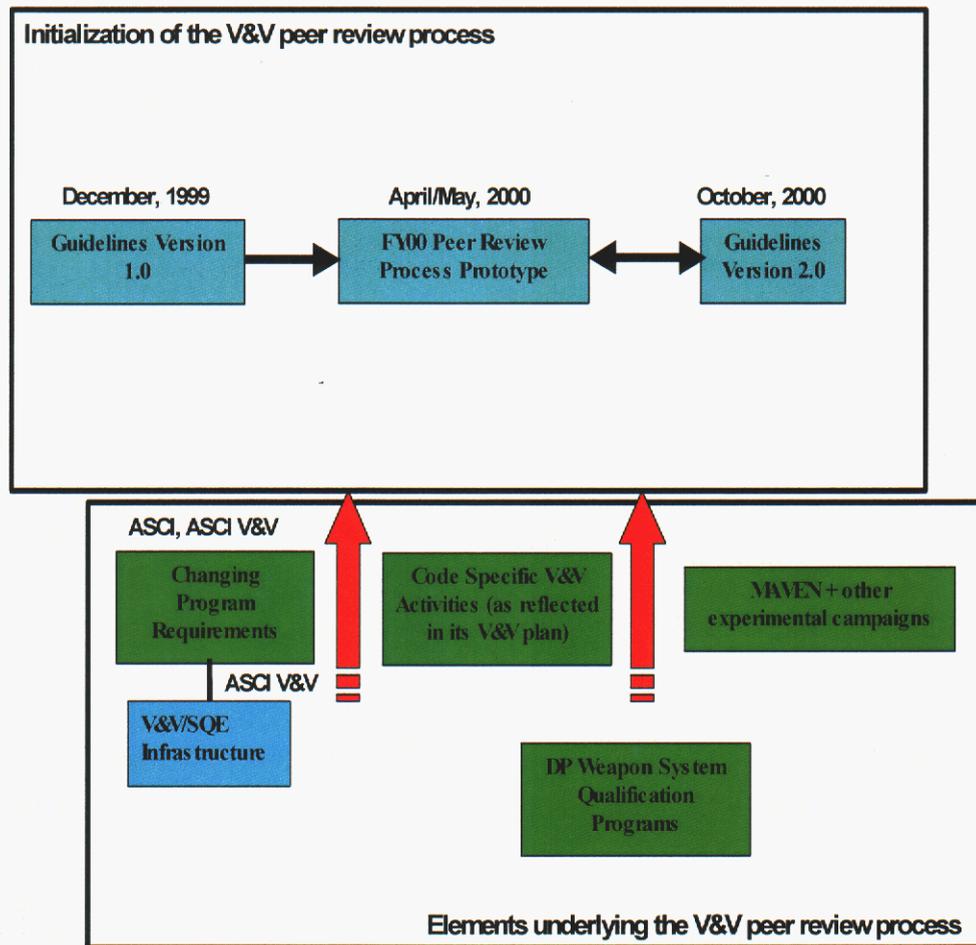


Figure 2-1. Conceptual Evolution of the V&V Peer Review Process

2.2 Level 1 Peer Review Process

This section summarizes the Level 1 Peer Review Process used during the FY01 V&V Plan assessments. The SNL ASCI V&V program three-level peer review process is illustrated in Figure 2-2. The levels are designed to reflect increasing depth of peer review for the V&V program at Sandia. All three of the levels are required to track the progress of the V&V program at Sandia. The levels are staged in time to reflect a promotional model of peer review. Codes (and integrated applications), represented by their V&V teams, are expected to undergo a Level 1 review before a Level 2 review, and one or more Level 2 reviews prior to a Level 3 review. Generally, the first level – **Level 1** - emphasizes a programmatic review. The second level – **Level 2** – emphasizes technical assessment of the content and implementation of the V&V program for selected ASCI codes, typically representing an integrated application. The third level – **Level 3** – emphasizes technical assessment of the performance of the V&V program for selected ASCI codes, typically in conjunction with a stockpile certification activity. This latter review level should be viewed as an essential component in a certification strategy for application of Sandia ASCI codes to stockpile problems. Therefore, DP and Surety at Sandia is a stakeholder and customer for this review. The peer review process seeks to address all of these areas – programmatic, technical and results - in terms of measurement of individual ASCI code team V&V efforts at Sandia.

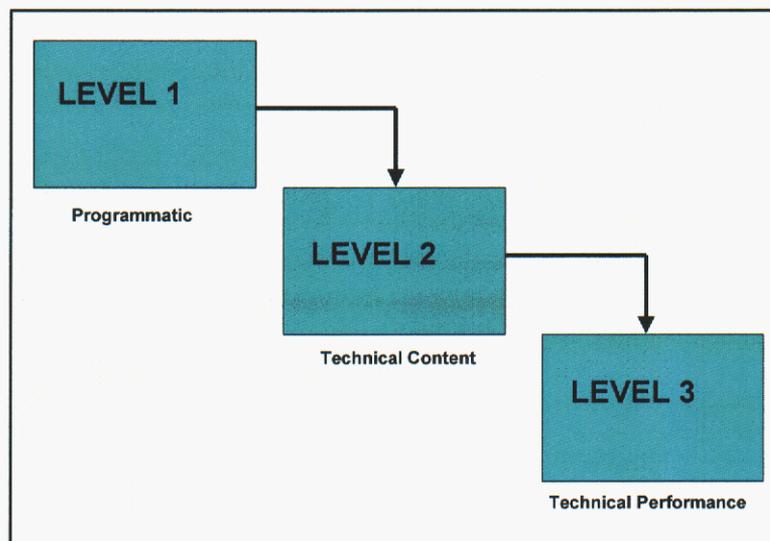


Figure 2-2. Sandia ASCI V&V Program Peer Review Levels

The Level 1 peer review process is a programmatic review of the development, maintenance, and execution of a Sandia application/code team V&V plan. The goal of this review is acceptance of an individual application's/code team's V&V plan by the ASCI V&V program at Sandia. An independent internal panel performs the assessment. The primary materials required for performing a Level 1 assessment are the V&V plans for the codes being assessed that have been

written to conform to the most current version of the reference [V&V-GUIDE-2_0]. For FY01, the primary metric used to assess the individual code plans is conformance to the reference [V&V-GUIDE-2_0]. The goal of the Sandia V&V program was that all ASCI-funded code projects at Sandia undertake Level 1 reviews by the end of FY01.

The key objectives of the Level 1 review are to assess:

1. status of the code-specific V&V activities with respect to development and maintenance of a documented V&V plan conforming to the most current published guidelines;
2. implementation status of that plan; and
3. programmatic performance or progress on that implementation.

The main areas of specific V&V plans targeted for assessment of conformance are:

1. understanding of Sandia DP stockpile requirements relevant to critical code applications and their use as drivers for the code specific V&V program;
2. development of a Phenomenology Identification and Ranking Table (PIRT);
3. implementation of a Software Quality Engineering (SQE) verification approach, and conformance of that approach to relevant programmatic SQE requirements associated with the ASCI program, references [DOE-ASCI-SQE] and [SNL-ASCI-SQE];
4. existence of a Verification Test Plan;
5. existence of a Validation Test Plan; and
6. development of guidelines for stockpile computing applications of the code.

The precise format and content of conformance criteria has evolved from reference [V&V-GUIDE-1_0] to reference [V&V-GUIDE-2_0]. These specific criteria are extracted from reference [V&V-GUIDE-2_0] and listed in Appendix B. These criteria detail V&V plan content requirements in each of the six specific areas listed above. Assessing the conformance of V&V plans to these criteria is the key element in the Level 1 peer review. Content in these conformance criteria can be measured through either quantitative or qualitative means. The particular approach performed for the FY01 assessments was:

1. Qualitative and general feedback on the content of the code team's V&V plan in each of the six general content categories above.
2. Scoring of the individual criteria on a three point scale:
 - 0** - no content present in the current plan;
 - 1** - content is present but inadequate for complete assessment;
 - 2** - substantial content is present.

A mix of qualitative and quantitative assessment of conformance criteria is important.

The Level 1 review also provides a top-level measurement of the implementation status of the code team's V&V plan by reviewing the performance self-assessment with code team personnel in terms of the six general category areas.

It is also important to develop some information on the code team's V&V plan performance. The Level 1 review uses the performance self-assessment by the code team. By briefly examining performance, the Level 1 review achieves a sanity check of the required alignment of

the code team's V&V plan with the evolving Sandia V&V program, as well as the overall DOE ASCI V&V program.

Implementation and performance results based on the code team's V&V plan is not the major goal of a Level 1 assessment. Level 2 and Level 3 peer reviews will concentrate more fully on these areas. For a Level 1 review, the goal is to determine progress achieved to date, at the time of the review, toward accomplishing the objectives of the code team's V&V plan.

2.2.1 Review Session

Each code team prepared a specific V&V Plan for review. In order to make the reviews more efficient, up to three code teams were grouped into a review session. The Review Panels for the code teams consisted of some participants who were common across all the code teams involved in a session and some participants who participated only in a specific code team V&V Plan review. In addition, to add some continuity to the sessions, some panel members (primarily the process observer and core member facilitator/SQE expert) were common to all sessions. During FY01, four review sessions were conducted. These specific review sessions are described in Section 3.

2.2.2 Review Panel Selection

Members of a peer review panel are selected by the Sandia V&V program to perform the formal Level 1 assessment of a given code team's V&V plan. Suggestions from the code teams as to their direct DP/Surety customers and internal experts were solicited and used. The membership of the panels, as somewhat evolved from the general guidance in reference [V&V-PR-Guide-1_0], included:

1. **Core Members:** A minimum of three panel members (chair, facilitator/SQE, and code team customer) were chosen as *core* panel members. These members had the responsibility for reviewing all (typically two or three) codes that were included in a Level 1 review session. These three people were selected from the Sandia V&V program, supporting infrastructure support personnel, the ASCI Applications Program Office, and possibly from DP/Surety. One member from this group served as the chairperson of the panel. A common "facilitator" that had extensive experience with M&S, V&V, and SQE was used on all panels. It is desirable that one of the core panel members have software development or software engineering experience.
2. **Rotator Members:** Up to two members of the panel were selected by the specific code project under review. Their duties extend only to that specific project. These participants are called *rotators* for the core panel. If possible, two functions were targeted for the rotators. One function of the rotators was to serve as a subject matter specialist in the technical focus area of the code under review. This subject matter expert must be independent of the specific code development project. The second function of the rotators was to represent the customer community for the code. In particular, this representation could be as an analyst who is independent of the code project but likely to be a user of the code. Or, this representation could be as a DP/Surety customer for the modeling capability of the code. Ideally, the rotator who represents a customer presence for the review panel represents the main stockpile application community that serves as the focus

of the code team's V&V plan under review. Sometimes this function was provided by a core team member.

3. **Observer Members:** The V&V program also placed one or more observers on the panel who provided programmatic, process support, and facilitator functions for panel deliberations. For the FY01 panels, the V&V Program Manager served as a general programmatic observer, the V&V SQE team lead (or designee) served as the process observer, and the scribe/facilitator function originally assigned to the observer member was assumed by a common core team member facilitator for all panels.

The chairperson of the review panel was responsible for organizing and managing the delivery of the review outcomes for that particular panel. The chairperson was also responsible for facilitating panel deliberations. For the FY01 reviews, a common facilitator assisted the panel chair in these activities. In addition, this facilitator was assigned the responsibility of developing first drafts for written work products of the panel.

The Sandia V&V program formally tasked the Level 1 review teams in a letter or memorandum from the V&V Program manager. All Level 1 review personnel were trained in the methods and expected outcomes of the peer review process.

2.2.3 Review Materials

For the Level 1 review process, the following material was provided by each code project at the beginning (prior to the panel training session) of the review process:

1. **V&V Plan:** One electronic copy of the existing current V&V plan for the specific code under review. For FY01, the panels assessed this document for compliance with reference [V&V-GUIDE-2_0].
2. **Performance Self-Assessment:** One electronic copy of a written performance statement which concentrates on the topics outlined in Appendix A of the reference [V&V-REVIEW-1_0].

The electronic input from a code team was delivered before beginning training of the applicable review panel in most instances. Occasionally this delivery occurred shortly after the training, which seemed to cause very little problem. Typically the training was held no later than one-two weeks before the scheduled date of the Level 1 review of the code.

2.2.4 Review Process

Training (1/2 day):

Once selected, personnel representing the Sandia V&V program trained the panel. Training lasted approximately one-half day for each panel, although the training time tended to decrease from one-half day to about two hours by the end of the fourth session.

Preparation (2 days per code):

At the time of training (or shortly thereafter) the panel members received written copies of the needed review materials from the participating code teams. Rotators received material only relevant to their specific code. The core panel members received material for all of the codes they were responsible for reviewing. Prior to the day of the panel review for a code the panel members

performed individual assessments of this material, including scoring of the V&V criteria. Based on experience with the FY01 reviews, a level of effort of approximately one to two days by each reviewer was required to assess the material for one code.

Review Day (1 day per code):

The actual review day consisted of a full working day per code project. The overall schedule for a review day was as follows.

Two hours:	Panel initial deliberations
Two hours:	Code presentation and panel Q&A
Three hours:	Panel deliberations and preparation of same-day outbrief
One hour:	Panel outbriefs code team

The initial two hours of the review was devoted to transforming individual assessments and understanding into a common panel assessment and understanding. In addition, panel member input critical for developing the panel work products for the review was developed.

The individual assessments and initial panel deliberation lead to further questions and issues. These questions and issues were discussed during the code presentation phase in which the code team presented information about the review materials. The total two hours allotted for this discovery phase was roughly one hour for the code team and one hour for the review panel questions and issues. The oral presentation was also an opportunity to provide further guidance on the structure of the code team’s written V&V plan. The panel did not present assessment conclusions (individual or group) during the oral presentation and Q&A time.

Following the oral presentation and Q&A session, the panel had a working lunch and prepared general strength and improvement opportunity themes. Supporting evidence was derived from the individual V&V conformance criteria scores, oral discovery information, and general knowledge of the panel members. The themes and supporting evidence was drafted into a vugraph outbrief presentation. As time permitted plans for future follow-up meetings, written report development, and responsibilities were also discussed.

The final activity of the review day was the outbrief presentation to the code team participants. This one hour period was led by the chair although the facilitator member was given the responsibility for presenting the outbrief. Other panel members participated in discussions during the presentation. The objective of this presentation was to provide a quick look feedback to the code team that highlighted the panel assessment of the V&V Plan conformance to reference [V&V-GUIDE-2_0]. Some clarification and interaction was encouraged, but detailed discussion was reserved for the follow-up written report to the code team.

Follow-Up (2 weeks):

Each panel was responsible for delivering several work products for each code. A summary discussion of these work products (Section 2.2.5) is given below. Some time was allocated for discussion and development of initial drafts of these work products by the review panel during previous process steps. One of the work products was the same day feedback briefing to the code team to give them overall panel findings. The same day feedback briefing was the primary input for the more detailed written work products. A written report was delivered to the code team, typically within approximately one-two weeks of the review day. A written report was also delivered to the V&V program office within approximately two weeks of the review day that

summarized general conclusions across all the code teams that participated in that panel’s peer review. In addition, the panel for each code team was available for a follow-up meeting to discuss any clarification of the written report information and possible insight into the development of code team action plans in response to the review. No code team requested a follow-up meeting. Each code team was then responsible for providing to the V&V program office an action plan on their response to the assessment report and a feedback report concerning strengths or areas of improvement for the peer review process. Feedback was done verbally and captured in writing as part of the V&V Program written report.

2.2.5 Review Deliverables

A high priority was placed on achieving timely relevant outcomes from this peer review process. The outcomes, or work products, of the Level 1 review consist of the items shown in Table 2-1. The panel review team was responsible for the first three products, both the code team and the panel for the fourth product, and the code team for the fifth and sixth products.

Table 2-1. Level 1 Peer Review Deliverables

Product	Prepared By	Deliver To	Description	Template
#1 - Quick Look Feedback	Review Panel Team	Code Team	Panel provides outbrief presentation to the code team at the end of the review day	See Fig 3-3, of reference [V&V-REVIEW-1_0]
#2 - Code Team Written Assessment Report	Review Panel Team	Code Team	Panel provides a written feedback report to each code team typically one week after the date of the formal review. This information will be shared with the Sandia V&V program office.	See Fig 3-4, of reference [V&V-REVIEW-1_0]
#3 - V&V Program Written Assessment Report	Review Panel Team (primarily core members)	SNL V&V Program Office	Panel provides a written feedback report on important findings of the review panel to the Sandia V&V program office on a fast time scale (typically two weeks after the panel’s last code review). The chairmen of the review panel is responsible for delivery.	See Fig 3-5, of reference [V&V-REVIEW-1_0]
#4 - Follow-Up Meeting With Code Team	Review Panel Team	Code Team	An oral follow-up meeting with the code team will be held by the panel to discuss the contents of the written feedback report.	Not requested by FY01 code teams.
#5 - Code Team Action Plan	Code Team	SNL V&V Program Office & Code Review Panel	The code team will develop a formal action plan (memo or more extensive written response) to the findings of the panel. This work product is delivered to the program office and members of the review panel for that code.	Verbal communication and implemented through V&V and Code Team IPs
#6 - Code Team Feedback Report	Code Team	SNL V&V Program Office	The code team provides feedback on the peer review process to the V&V program office.	See Fig 3-6, of reference [V&V-REVIEW-1_0]

3 Description of FY01 Level 1 Peer Review Sessions

Prior to each review session, the V&V Program provided extensive coordination with each of the selected code teams and the potential members of the review panel. An implementation plan for each review session was constructed with a day-by-day list of activities that covered approximately one month in time prior to the actual review day. As an example, the calendar of activities contained in the Session 1 peer review implementation plan is presented in Table 3-1.

Table 3-1. Pre-Review Calendar: Session 1 Level 1 Peer Review Implementation Plan

Plan Date	Actual Date	Activities	Responsible Personnel
01/29/01	01/29/01	• Establish Potential Review Panel Team	Dave Peercy
		– Contact potential panel members and send out e-mail request for panel member participation	Marty Pilch
01/29/01 - 01/31/01	01/29/01 - 02/05/01	– Contact Potential Review Panel Team Discuss review process and schedule Obtain commitment or other potential panel members Determine training date and best three days of 5-9March	Dave Peercy/Marty Pilch
02/02/01	02/07/01	– Finalize Review Panel Members, Training Date, Review Dates Send e-mail to all panel members w/schedule Send e-mail to all code team leads/V&V Plan owners w/schedule • Distribute references [V&V-GUIDE-2_0] and [V&V Peer Review Process V1.0] to each panel member for early prep.	Dave Peercy/Marty Pilch Dave Peercy
02/02/01	01/29/01	– Reserve conference room needed for review week. 880/C38F: March 5 (>10am); March 6,7,8,9 (all day) • Reserve conference room needed for training 892/297: February 21 (8:30-12pm)	Dave Peercy
02/02/01	02/07/01	– Establish lunch & refreshments for panel members Morning: coffee, bagels, juice Lunch: sandwiches & sodas Afternoon: sodas for panel and code team	Dave Peercy
02/16/01		– Complete Training Notebook for Panel Members V&V Plan & Performance Self-Assessment Training materials, forms, templates • Send reminder of training session to panel members	Gary Froehlich/Dave Peercy V&V Plan Owners Gary Froehlich/Dave Peercy Gary Froehlich/Dave Peercy
02/21/01		– Complete Half-day Panel Review Team Training (8:30-12:00 in 892/297) – Review training materials – Review code team inputs – Review schedule and deliverables (discuss potential dates for written report follow-up effort)	Gary Froehlich/Dave Peercy and All review panel members
02/21/01 - 03/05/01		• Review panel preparation Review V&V plans vs Reference [V&V-GUIDE-2_0] guidelines Review Performance Self-Assessment Perform informal scoring of V&V Plans & complete comments • Send electronic copy of the scoring summaries to Panel Facilitator by COB 03/05/01	All review panel members All review panel members, optional for observers, members score only those codes they will be reviewing.
03/05/01		• Compile electronic scoring for analysis use by the panel members during review week	Gary Froehlich/Dave Peercy

The codes/code teams for each review session are summarized in Table 3-2 with a short description of the code's application domain. Each session is briefly described in the subsections of this section.

Table 3-2. Level 1 Peer Review Code Team Participants

Session	Code Name	Application Domain/Description
Prototype	SACCARA	Sandia Advanced Code for Compressible Aerothermodynamics Research and Analysis. SACCARA provides high-fidelity, 3D compressible fluid mechanics/aerothermodynamics modeling, simulation, and analysis capability for three categories of systems: gravity bombs, reentry systems, and glide bombs/cruise missiles.
Prototype	PRONTO/JAS PRESTO/ADAGIO	Solid mechanics codes. PRONTO/JAS legacy and PRESTO/ADAGIO are respective evolution within the SIERRA Framework.
1	SIERRA Framework	Physics Framework: Sandia Integrated Environment for Research and Robust Analysis Services Framework that provides various common services and interfaces to libraries for physics solid mechanics applications modeled by various code teams.
1	CALORE	Thermal Simulation Code. Thermal transport and thermally induced phenomena are of significant importance in all phases of DP weapon production and deployment, including design, fabrication, testing and certification, manufacturing and assembly, transportation, storage, and deployment.
1	FUEGO-SYRINX	Fire Simulation Code: FUEGO means fire in Spanish. SYRINX is a water nymph in Greek mythology. FUEGO is the flow/combustion mechanics module and SYRINX is the radiation transport module that is embedded in the SIERRA software architecture. Codes used to address abnormal thermal environments for nuclear weapons safety considerations.
2	HPEMS/XYCE	High Performance Electrical Modeling and Simulation/ChileSpice with XYCE as newest model version. Characterization of nuclear weapon electrical systems from first principles in all environments over a 50-year lifetime.
2	GOMA	Spanish for rubber, gum, or elastic. Strongly coupled multi-physics transport simulation for manufacturing, material aging, and electrochemical performance applications.
3	ALEGRA Framework	ALEGRA - Arbitrary Lagrangian Eulerian General Research Applications Services Framework that provides various common services and interfaces to libraries for physics application classes modeled by various code teams.
3	EMPHASIS	Electromagnetic Physics Analysis System. EMPHASIS is suite of electromagnetics codes, the V&V Plan was for Eiger, the frequency-domain portion of EMPHASIS. EMPHASIS has a milestone to perform frequency-domain simulations of the W76-1 system response to a variety of STS normal EM environments.
3	CEPTRE-ITS	Coupled Electron Photon Transport for Radiation Effects – Integrated Tiger Series (SNL radiation transport package) V&V Plan has multiple aspects for CEPTRE code and ITS code. Used in hostile radiation environments. ITS is used to determine energy deposition, which can drive both mechanical response and transient electrical response. For electrical components, CEPTRE, is use to predict charge deposition which drives electromagnetic pulse effects.
4	ALEGRA EMMA	ALEGRA - Arbitrary Lagrangian Eulerian General Research Applications – Electro-Mechanical Modeling in Alegra: shock physics application code.
4	NUGET	Neutron Generator simulation code. The NuGET code is being developed to provide a high fidelity code for characterizing the coherent nuclear environments and neutron damage response from hostile nuclear events and from fratricide encounters.
4	PEGASUS	Neutron tube particle based simulations. Simulation of chemically reacting plasma/neutral systems in both 2D and 3D geometries from the collisional limit to collisionless systems.

3.1 Session 1: March 6-8, 2001

The code teams that participated in Session 1 included: SIERRA Framework, CALORE, and FUEGO-SYRINX.

3.1.1 Review Panel Selection

The peer review panel for this review is in the following table. Information current as of review.

Function (codes to review)	Name	Org	Phone	E-mail	Location bldg/room
Chair (all)	Richard Griffith	09117	505-844-8232	rogrijf@sandia.gov	880/B34G
Facilitator/SQE (all)	Gary Froehlich	06535	505-284-3930	gkfroeh@sandia.gov	6585/1801
Core (all)	Bill Moffatt	02114	505-844-7062	wcmoffa@sandia.gov	836/1236
Rotator (SIERRA & FUEGO-SYRINX)	Mike Bohn	12333	505-844-4598	mpbohn@sandia.gov	836/1061
Rotator (CALORE)	Brad Altman	12333	505-844-5784	bsaltma@sandia.gov	836/1139
Rotator (SIERRA)	Kevin Copps	09143	505-844-4521	kdcopps@sandia.gov	880/Y10K
Rotator (FUEGO-SYRINX)	Paul Desjardin	09116	505-844-7740	pedesja@sandia.gov	880/B30C
Rotator (CALORE)	Shawn Burns	09111	505-844-6200	spburns@sandia.gov	880/Y23C
Programmatic Observer (all)	Mike McGlaun	09140	505-845-8236	jmmcgla@sandia.gov	880/A-35B
Programmatic Observer (all)	Marty Pilch	09133	505-845-3047	mpilch@sandia.gov	880/C38E

3.1.2 Review Materials

For the Level 1 review process, the following material was provided by each code project at the beginning (prior to the panel training session) of the review process:

1. V&V Plans:

SIERRA Framework: *"Sierra Framework Verification and Validation Plan"*

CALORE: *"Draft CALORE Verification and Validation Plan for Weapon-in-a-Fire Simulations, Version 1.1"*

FUEGO-SYRINX: *"SIERRA/FUEGO and SIERRA/SYRINX Verification and Validation Plan"*

2. Performance Self-Assessment:

Each plan had an associated performance self-assessment.

3.1.3 Review Deliverables

- SIERRA_quicklook.ppt and SIERRA_Code-SpecificSummaryReport.doc
- CALORE_quicklook.ppt and CALORE_Code-SpecificSummaryReport.doc
- FUEGO-SYRINX_quicklook.ppt and FUEGO-SYRINX_Code-SpecificSummaryReport.doc
- FY01_Session1_ProgramOfficeFinalReport.doc

3.2 Session 2: March 28-29, 2001

The code teams that participated in Session 2 included: GOMA and HPEMS/XYCE

3.2.1 Review Panel Selection

The peer review panel for this review is in the following table. Information current as of review.

Function (codes to review)	Name	Org	Phone	E-mail	Location bldg/room
Chair (all)	Amalia Lopez	09133	505-844-8057	arlopez@sandia.gov	880/X23D
Facilitator/SQE (all)	Gary Froehlich	06535	505-284-3930	gkfroeh@sandia.gov	6585/1801
Core (all)	Chuck Harjes	01639	505-845-7106	hcharje@sandia.gov	962/1481
Rotator (HPEMS/XYCE)	Doug Weiss	02333	505-845-9134	drweiss@sandia.gov	891/4501
Rotator (GOMA)	Tom Baer	09114	505-845-8912	tabaer@sandia.gov	880/Y-101
Programmatic Observer (all)	Marty Pilch	09133	505-845-3047	mpilch@sandia.gov	880/C38E
Process Observer (GOMA)	Dave Peercy	12326	505-844-7965	depeerc@sandia.gov	892/296D
Process Observer (HPEMS)	Ann Hodges	06534	505-844-6284	alhodge@sandia.gov	6585/1101

3.2.2 Review Materials

For the Level 1 review process, the following material was provided by each code project at the beginning (prior to the panel training session) of the review process:

1. V&V Plans:

GOMA: *“Verification and Validation of Encapsulation Flow Models in GOMA”*

HPEMS/XYCE: *“High Performance Electrical Modeling and Simulation Software Normal Environment Verification and Validation Plan”*

2. Performance Self-Assessment:

Each plan had an associated performance self-assessment.

3.2.3 Review Deliverables

- GOMA_quicklook.ppt and GOMA_Code-SpecificSummaryReport.doc
- HPEMS-XYCE_quicklook.ppt and HPEMS-XYCE_Code-SpecificSummaryReport.doc
- FY01_Session2_ProgramOfficeFinalReport.doc

3.3 Session 3: May 15-17, 2001

The code teams that participated in Session 3 included: ALEGRA Framework, EMPHASIS, and CEPTRE-ITS.

3.3.1 Review Panel Selection

The peer review panel for this review is in the following table. Information current as of review.

Function (codes to review)	Name	Org	Phone	E-mail	Location bldg/room
Chair (all)	Mary McWherter-Payne	09115	505-844-8500	mapayne@sandia.gov	880/X23N
Facilitator/SQE (all)	Gary Froehlich	06536	505-284-3930	gkfroeh@sandia.gov	6585/1801
Core (EMPHASIS, CEPTRE-ITS)	Kaz Oishi	02113	505-844-0159	koishi@sandia.gov	836/1197
Rotator (CEPTRE-ITS)	Bill Barrett	15344	505-845-7596	whbarre@sandia.gov	962/4467
Rotator (EMPHASIS)	Joe Kotulski	01642	505-845-7955	jdkotul@sandia.gov	962/1100
Rotator (CEPTRE-ITS)	Clif Drumm	15345	505-845-7953	crdrumm@sandia.gov	962/3050
Rotator (ALEGRA Framework)	Kevin Brown	09231	505-844-0475	khbrown@sandia.gov	880/Y32E
Programmatic Observer (all)	Marty Pilch	09133	505-845-3047	mpilch@sandia.gov	880/C38E
Process Observer (all)	Dave Peercy	12326	505-844-7965	depeerc@sandia.gov	892/296D

3.3.2 Review Materials

For the Level 1 review process, the following material was provided by each code project at the beginning (prior to the panel training session) of the review process:

1. V&V Plans:

ALEGRA Framework: *"ALEGRA FRAMEWORK V&V Plan for the EMPHASIS code team modeling and simulation scenarios for the W76-1 Life Extension Program STS Electromagnetic Radiation (EMR) environment"*

EMPHASIS: *"EMPHASIS: ASCI V&V Plan for Electromagnetic Radiation Environment"*

CEPTRE-ITS: *"CEPTRE and ITS Verification and Validation Plan"*

2. Performance Self-Assessment:

Each plan had an associated performance self-assessment.

3.3.3 Review Deliverables

- ALEGRA_quicklook.ppt and ALEGRA_Code-SpecificSummaryReport.doc
- EMPHASIS_quicklook.ppt and EMPHASIS_Code-SpecificSummaryReport.doc
- CEPTRE-ITS_quicklook.ppt and CEPTRE-ITS_Code-SpecificSummaryReport.doc
- FY01_Session1_ProgramOfficeFinalReport1.doc

3.4 Session 4: September 18-20, 2001

The code teams that participated in Session 4 included: NuGET, PEGASUS, and ALEGRA-EMMA.

3.4.1 Review Panel Selection

The peer review panel for this review is in the following table. Information current as of review.

Function (codes to review)	Name	Org	Phone	E-mail	Location bldg/room
Chair (all)	Mary McWherter-Payne	09115	505-844-8500	mapayne@sandia.gov	880/X23N
Facilitator/SQE (all)	Gary Froehlich	06536	505-284-3930	gkfroeh@sandia.gov	6585/1801
Core (all)	Bob Paulsen	02109	505-844-0132	rapauls@sandia.gov	836/1217
Rotator (NUGET)	Gary Harms	06422	505-845-3244	gaharms@sandia.gov	6585/2306
Rotator (NUGET)	Bill Barrett	15344	505-845-7596	whbarre@sandia.gov	962/4467
Rotator (PEGASUS)	Robert Koss	02564	505-284-2455	rjkoss@sandia.gov	891/2056
Rotator (ALEGRA/EMMA)	John Ludwigsen	12333	505-844-8581	jsludwi@sandia.gov	836/1141
Rotator (ALEGRA/EMMA)	Steve Montgomery	02561	505-844-9062	stmontg@sandia.gov	891/1028
Rotator (ALEGRA/EMMA)	Allen Robinson	09231	505-844-6614	acrobin@sandia.gov	880/D34H
Observer (ALEGRA/EMMA)	Aamer Haque*	DOE/ NNSA/ DP-15	202-586-7289	Aamer.Haque@nnsa.doe.gov	DOE*
Programmatic Observer (all)	Marty Pilch	09133	505-845-3047	mpilch@sandia.gov	880/C38E
Process Observer (all)	Ann Hodges	06546	505-844-6284	alhodge@sandia.gov	6585/1101

* - unable to attend due to travel restrictions

3.4.2 Review Materials

For the Level 1 review process, the following material was provided by each code project at the beginning (prior to the panel training session) of the review process:

1. V&V Plans:

NuGET: *"NuGET Verification and Validation Plan: Version 0.99"*

PEGASUS: *"Verification and Validation Plan for the Codes LSP and ICARUS (PEGASUS)"*

ALEGRA-EMMA: *"ALEGRA-EMMA Verification and Validation Plan for Modeling and Simulation Scenarios involving the FerroElectric Neutron Generator Power Supply"*

2. Performance Self-Assessment:

Each plan had an associated performance self-assessment.

3.4.3 Review Deliverables

- NuGET_quicklook.ppt and NuGET_Code-SpecificSummaryReport.doc
- PEGASUS_quicklook.ppt and PEGASUS_Code-SpecificSummaryReport.doc
- ALEGRA-EMMA_quicklook.ppt and ALEGRA-EMMA_Code-SpecificSummaryReport.doc
- FY01_Session4_ProgramOfficeFinalReport.doc

4 Results, Analysis, and Recommendations

Results and analyses are presented in the following three areas:

- V&V Plan;
- V&V Guidelines; and
- V&V Level 1 Peer Review Process

Throughout these areas there are many observations and suggestions/recommendations. References are made, when appropriate, to analysis information in Appendix B that support the various suggestions and recommendations.

4.1 V&V Plan

The following observations and suggestions provide a general set of results that apply appropriately to all, most, many, or some of the V&V Plans. There are usually some exceptions to nearly any statement, but the sense of major issues is presented.

4.1.1 Conformance to V&V Guidelines

- Overall, the Review Panels felt that the V&V Plans conformed to the format recommended by the Guidelines.
- The concept of a “documentation tree”, referencing information that is or will be in other documents, was used by most Plans. This was recommended by the Guidelines.
- The Code Teams made a valid attempt to identify their DP customer(s) in their Plans. The Code Teams neglected to identify Surety customers, or whether that was applicable. The framework code teams (SIERRA and ALEGRA) appropriately identified code teams as their customers, with the associated customer base.
- Many code teams inadequately addressed the Validation content (see Appendix B.1), and some gave only cursory acknowledgement to SQE in a general sense, although following the guidelines for the most part. Performance measures for SQE, verification, and validation were inadequately addressed (see Appendix B.1).

4.1.2 Strengths

- The Code Teams brought an energy, effort and professionalism to the V&V planning and peer review process. The Panels were impressed by the effort that the Code Teams gave to the V&V planning and peer review process. Code Teams participated in the process in a professional and cooperative manner, although there were some honest disagreements voiced about the process and the guidelines. Discussions of the feedback were conducted in an honest and accepting manner by all Code Teams.
- Most of the Plans were good. The Plan authors are clearly aware of the basic requirements of V&V; they have or are beginning to engage the DP and Surety customers; they recognize minimal SQE requirements; and they demonstrate a basic understanding of computational modeling and necessary verification activities that are needed in a certification environment.

- Most of the Code Teams acknowledged the demands of computational modeling in a certification environment, evidenced by the attention that was paid to issues of reproducibility, traceability, and configuration management.
- The Code Teams with prior experience in using their code for certification activities were much more aware of the issues and addressed the stockpile computing issues more completely (see Appendix B.1).
- An effort was made by most Plan authors to address software quality engineering (SQE) issues. Most of the major components of SQE were at least recognized and commented upon.
- The development of the PIRTs, VERTS, and VALTS were *very* effective for the NuGET Plan. The NuGET PIRT and the process leading to its development could serve as a model for other physics codes. There were also several other strong plans with various practices of noteworthy mention (See Appendix B.2).

4.1.3 Opportunities for Improvement

- Overall the Plans need to tie more completely to customer schedules and milestones. Then, working backward from customer milestones, credible V&V milestones could and should be developed, and included in the prioritization scheme of the PIRTs. Plans should include traceability between the PIRT and the code's capability releases.
- While the V&V Plans typically identify the experimenters, analysts, and various customers in a general sense, more specific links would be very helpful to readers of the Plans. The Plans should provide specific identification of the Plan owner, the *process* owner, their customers (DP, and Surety if applicable), and experimentalists. These links should be to specific organizations, and to the specific individuals who are currently responsible. The volatility of organization numbers is understood, but more readers are familiar with those (and their histories) than with organization names. The same is largely true of individuals. The NuGET Plan provides an excellent example that other Teams should be encouraged to emulate.
- The current process appears to work much better with an analyst or phenomenologist as the lead author, than with a code developer alone in that role. The reason is that this analyst is more qualified than the code developer to identify and prioritize the phenomena that need to be addressed.
- It is very clear that execution of different elements of the V&V Plans must belong to different people, hence the recommendation that *ownership* (as opposed to mere authorship) consist of code representatives, customer representatives, analysts, and experimentalists. The current requirement (PR1) has in many cases become a 'box-checking' exercise, in which authors' names are included to meet the requirement. The intent is that the various roles actually *participate* in V&V planning.
- Ties to needed experiments, and hence to required experimental accuracy, should be discussed more fully. This includes both experiments to provide input data (e.g., material properties), and validation experiments.
- If codes are intended to interface at some point, and someone will eventually need to be evaluating the V&V of this coupling, would it make sense for someone other than the code teams to choose an appropriate stockpile driver for all parties? Or, should code teams choose a driver that delivers maximum M&S value for the effort invested, regardless of coupling

issues? **A suggestion:** V&V Teams could compile fairly comprehensive lists of multiple stockpile drivers and related phenomena (as a few teams have already done). Then, each time a particular application is identified, a "slice" through such a tabulation would provide the input for a new prioritization activity, thus leading to a PIRT which forms the core of the new, application-specific V&V Plan.

- It is essential that the owners of the V&V Plans must have the authority and the resources to engage the necessary participants in the development of the V&V Plan.
- The V&V Plans should strive to quantify the accuracy requirements for validation and verification in a more rigorous fashion, or reference where such quantification is to be described. As a whole, there was a notable lack of depth and detail in the discussion of these metrics in the V&V Plans. Regarding acceptance criteria, quantitative comparisons are preferable, where possible, to visual comparison of plots. It is understood that if no metrics are given, the Panels will ask for them; if you do give metrics, the Panels will ask for justification. The *real* questions are: How do you know when you're 'done'? How good is good enough?
- There is inadequate treatment of success measures. This observation refers to acceptance criteria for verification and validation, *as well as metrics to assess the effectiveness of SQE practices*. Related to this is the specification of needed accuracy for experimental results.
- Although uncertainty quantification (UQ) is considered part of a standard ASCI product, in general very little if any V&V aspects of the UQ and its associated process have been discussed.
- In general, Plans have had little or no discussion of software metrics. Metrics for SQE are intended to provide *quantifiable* evidence that software development processes are (or are not) adding value, so that developers need not rely solely on their intuition regarding those processes.
- In general, acceptance criteria for third-party software are not discussed in the Plans. This is an important concern, and needs more attention.

4.2 V&V Guidelines

The following conclusions provide a general set of observations of strengths and recommendations for improvement that apply to the existing version of the V&V Guidelines, reference [V&V-GUIDE-2_0].

4.2.1 Strengths

- The existing guidelines are adequate and effective for M&S codes. They are not "broken". What may be "broken" instead of the guidelines is an understanding of roles and responsibilities across multiple participants and departments. Several of the V&V Plans were excellent examples of how these guidelines could and should be applied. Those teams where the members clearly share a common vision of what must be done, and how to do it were most successful in applying the guidelines. Such teams included code developers, analysts, experimentalists, and customers.
- Code teams were compliant with the form and content suggested by the guidelines; with only a few exceptions, the guidelines were able to be read and interpreted by each of the code teams. Differences in following the guidelines tended to be related more to specific

organizational structure and code team philosophies than a misinterpretation of the guidelines.

- A common terminology for key aspects of V&V (e.g., Stockpile Driver, PIRT, VERTS, VALTS, SQE, Stockpile Computing) was established across a wide range of code team application domains and peer review participants; perhaps most important was the understanding provided to DP and Surety customers.

4.2.2 Opportunities for Improvement

- Guidelines should encourage teams to reference generic (i.e., SNL ASCI Program level) guidance for Stockpile Computing, UQ, and SQE. Someone must be tasked to develop this guidance (SQE is done; UQ and Stockpile Guidance are being planned). Some aspects of each of these areas will likely remain code specific. Teams should focus first and foremost on PIRT, VERTS, and VALTS.
- Program-level guidance for Stockpile Computing should be developed. DP & 12300 want confidence in the stockpile-computing process. This includes confidence in the codes themselves, but also confidence that the codes were properly applied to the problems of interest. Some specifics, such as required training for use of a given code, will necessarily remain the responsibility of the individual teams, but much of the guidance can be ‘factored out’ and applied to all teams. This guidance really refers to issues of traceability and reproducibility, as they apply to "official" calculations. How do you prove to your customer, or an auditor, that you did in fact use the executable you claim to have used? Or that said executable was compiled from the claimed source version? What are the technical training requirements for users of the code, or of the code results? What are *all* the components of a product *set*? What are the environmental requirements (e.g., operating system, platform, stability)? How should one handle "restarts" during an official certification calculation? How do you demonstrate that your inputs were properly verified and validated prior to use? How is all this information captured and archived? Etc. Virtually all Plans have interpreted this section to apply *solely* to training users of the code. Stockpile Computing Guidance should also address the following: What artifacts are put into CM? What becomes part of the body of evidence?
- The Guidelines are very specific to M&S codes, at least with respect to some of the terminology. This is especially true of the PIRT concept. However, for M&S codes, the phenomena are also a format used to develop requirements. In fact, the PIRT is a specific representation of the general quality approach for requirements representation called Quality Function Deployment (QFD). If the PIRT is viewed as a tool for requirements development and representation, then the guidelines remain applicable even to frameworks, visualization software, etc. Frameworks have a “services” IRT (SIRT) instead of a PIRT. Some explanation of how to apply the PIRT concept to frameworks and other model-based areas as an extension of requirements management would be useful. SIERRA provided a good meta model for how to interpret the guidelines for frameworks and ALEGRA provided a superb example of applying the SIERRA model. This interpretation should be considered for inclusion in any revision to the guidelines. Perhaps the guidelines should provide a definition or description of what characteristics a “requirement” has, i.e., non-ambiguous, traceable, testable, and so forth. The guidelines might also offer examples of requirements *sources*,

such as STS, SLEP, Military Characteristic specifications, Vulnerability Committee, Nuclear Safety studies, customers, weapon designers, and so forth.

- DP customers felt a V&V Plan's focus should be on a single problem of interest to them. Others such as analysts felt that if synergy between multiple drivers exists, then teams should take advantage of it. Most Panel members (all 4 sessions) feel that the best approach to V&V planning is to have a documented information source with *all* the capabilities (and associated requirements) that a code must provide -- a Global PIRT. Then the V&V Plan for each specific, single stockpile driver would refer to this master document for 'down-selection' of the components required for that particular application. To some extent, some code teams in each session did this. This approach permits rapid construction of individual V&V Plans, as well as prioritization *across* applications. As an example of this reusability – a PIRT could be “tweaked” to change it from a W78-component PIRT to a W76-component PIRT.
- The Guidelines discuss an adequacy scoring system for the PIRT, yet give examples of importance ranking (low, medium, high); many Panel members feel that importance ranking and adequacy scoring are different activities; if the guidelines suggest that *ranking* is adequate, then the requirement to evaluate the *scoring* system (PIRT3) should perhaps be reworded; the Guidelines should differentiate between how to score, and how to perform ranking – teams are stopping after scoring; add schedule and difficulty to the ranking criteria - most Teams only included importance (need) and capability.
- The Guidelines should ask Teams to raise risk issues in the V&V Plans. The Plans should identify risks and risk response strategies; e.g., mismatch between stockpile-driver schedule and code availability for use. There needs to be a programmatic-level risk-management plan that Plans could refer to (as with SQE, UQ, SCG, etc.). Examples of risks identified during the Level-1 Peer Review are:
 1. Campaigns and DP milestones are not consistent. For example, C7, W78, and RES do not share consistent milestones.
 2. Uncertainty quantification appears to be “falling thru the cracks”.
 3. Validation money is functionally allocated, not code specific.
 4. Needed validation experiments may not be conducted in timeframe required to allow code use by DP.
 5. There is currently no V&V of NW models from the 2 physics labs -- in particular, EDM72, NWM80, NWM96
- Clarify the meaning and intent of the VAL3 and VAL5 evaluation areas; both the code teams and the review panels felt that these are very confusing as currently stated. Some Panel members feel they say the same thing; others feel they mean to address the distinction between two general types of data-quality requirements. One is determined by the accuracy needed for input to the codes (e.g., material properties) and the resulting impact any uncertainties will have on code output. The second is related to the impact of measurement uncertainties on the ability to make meaningful comparisons between observed and predicted behavior. It is worthwhile to distinguish between these, but VAL3 and VAL5 are currently unclear. In fact, content criteria VAL1 through VAL7 all come from a single bulleted statement each, in the guidelines document. If the code developers and the plan reviewers are to understand them, perhaps a lot more explanation of what is intended/expected would be good. (See also SQE, SQE1, PR1, PR2, and PIRT2).

- By applying the Guidelines, there is a clearer understanding of the strengths and weaknesses. Some specific opportunities for improvement in the guidelines include:
 1. Provide a clearer identification of roles and responsibilities across multiple participants and departments.
 2. Explicitly define what a stockpile driver is; e.g., is it a specific intended use or application of the code, to support stockpile computing.
 3. Identify the purpose of the V&V Plan; e.g., it is a contract between the code team and their customer, as well as the intended audiences (e.g., code-team members, external reviewers).
 4. Consider adding Surety as a customer (not solely DP); the requirement to list DP customers is too restrictive; equally important might be Surety customers, component groups, etc; the guidelines should be revised to include these other possibilities.
 5. In general, acceptance criteria for third-party software are not discussed in the Plans. This is an important concern, and needs more attention; guidelines for verifying “third-party” models, such as ProE models, should be peer reviewed as well.
 6. Suggest specifying in the guidelines to include the stockpile driver in the title of the V&V Plans for application specific V&V Plans.
 7. The Guidelines ask that customers and constraints be identified (DP1); however, the Guidelines (and the Panels) are quite unclear on what ‘constraints’ refers to, exactly; revise the guidelines to clarify this requirement.
 8. The *process* for planning V&V activities should be described within the Plans; in other words, the Plans should describe who met, how often, whether customers had input, etc.; perhaps the guidelines should be more explicit on this point; the Guidelines should make it clear that, ideally, verification happens *before* validation.
 9. Plans should show the high-level “concept of operations” for a code, or group of codes, and their application to a stockpile problem; what is the “workflow” - what is done first, and then what follows, and so forth.
 10. The guidelines discuss an adequacy scoring system for the PIRT, yet give examples of importance ranking (low, medium, high). Many Panel members feel that adequacy ranking and importance scoring are different activities; if the guidelines suggest that *ranking* is adequate, then the requirement to evaluate the *scoring* system (PIRT3) should perhaps be reworded.
 11. Plans should indicate what kinds of numerical models are involved (e.g., finite-element, finite volume, gridless, etc.) to permit assessment of the *adequacy* of the VERTS.
 12. Every V&V Plan should include a mapping matrix from the V&V criteria to sections within the plan that cover the criteria; it would be helpful if the V&V plan authors were asked to identify where in the plan each criterion was addressed (section number, page number, or something); perhaps the “Summary Scoring Matrix” could be modified with the second column used to give the location where the criterion is addressed (instead of a score); the Guidelines should indicate that V&V Plans are required to provide this mapping to the criterion requirements (PR1, VAL2, etc.).
 13. Every Plan should provide a Glossary, and a list of acronyms, initializations, and abbreviations; this should become a guidelines requirement.

14. Clarify the meaning and intent of the VAL3 and VAL5 evaluation areas; both the code teams and the review panels felt that these looked like the same thing.
15. There is a word missing from DP2 in the Scoring Sheet Template.

4.3 V&V Level 1 Peer Review Process

The V&V Level 1 Peer Review Process was also evaluated by each Review Panel. In addition, the Code Teams were provided the opportunity to comment on this process. By extension, recommendations for the next level (Level 2) of the peer review process structure and overall SNL ASCI program organization and process improvements were provided. This section identifies key aspects that were reported during the various peer review sessions.

4.3.1 Strengths

- The peer review process forces teams to seriously think about what is required to perform adequate V&V. The current state of the Plans, the heightened awareness of the V&V issues related to coupled physics, and the justifiable concerns about roles and responsibilities, are all the result of teams having confronted those issues for the purpose of developing their V&V Plans. Members on the various panels commented that without this year's Level 1 peer review process it is unlikely that the V&V Plans and the code teams' concomitant understanding of V&V, would be at their current level of maturity. Plan authors also have stated this repeatedly.
- Selection of Panel members is considered to be very good, specifically with regard to the diverse representation of interests and expertise. It is especially commendable to include managers, customers, code subject-matter experts, and V&V/SQE expertise. The inclusion of DP/Surety customers on the review panels had an additional benefit: several of the panel members stated that they had a much better understanding of what ASCI was trying to achieve, an improved sense of confidence that we were "trying to do the right thing", and an appreciation for the difficulty of the problem.
- Virtually every Panel member has observed that having an experienced member, *i.e.*, one who has been on many Panels, is invaluable. There was some consensus that having *more than one* full-time, core member (e.g., facilitator, chair, and perhaps even a 'permanent' pool of customer representatives) would reduce training requirements and flatten the learning curve that each Panel has faced. However, such participation in the Peer Review process would have to be viewed as valuable by the respective members' managers and by the ASCI Program.
- The short turn-around time for deliverables (quick feedback, final report, etc.) is crucial to the Plan owners, and is also to be commended.

4.3.2 Opportunities for Improvement

- Specific opportunities for improvement in the peer review process include:
 1. Inclusion of a DP and/or Surety customer on the Panel should be a requirement rather than a suggestion for the Review Panel members.
 2. Make the Peer Review templates more useable by pre-formatting with spaces and bullets.

3. Provide more explicit guidance to the code teams as to their oral presentation format; e.g., follow the V&V Plan outline, and include requested information from the Performance Self-Appraisal; the Panel Q&A follows more readily with that format.
 4. Expand the Level 2 Peer Review process to encompass the suggested application-centric approach.
- Specific ASCI program guidance for improving organizational structure and related processes that would facilitate an improved peer review process include:
 1. Management within the ASCI program must develop and communicate a clear and consistent vision of the evolution of V&V activities, and of the roles and responsibilities that permit those activities to be performed. The “code team” does not consist only of the people writing code; it is made up of the analysts who will use the tool, the experimentalists that will provide validation data, the code developers, and SQA representatives. The entire “code team” is responsible for the V&V plan and activities, not just the line manager of one element of the code team. All four of these elements must be clearly represented on a “V&V team” for it to successfully meet the V&V guidelines, and the leadership of the team must recognize this need. Along with the definition of roles, the funding and management structure must permit code teams to acquire the resources necessary to implement the V&V requirements being imposed on them.
 2. What is needed at the next level (application-centric focus) is a clear project organization for an application (e.g., a normal/abnormal/hostile milestone, weapon in a crash/fire, etc.) that pulls together elements from across many departments and disciplines (e.g., analyst, experimentalist, codes, V&V). From this project, the development of a V&V Plan could be accomplished, subject to the Peer Review process just like code-centric V&V Plans. V&V Plans produced via this mechanism (which for a given stockpile application will include several codes and the associated PIRTs, models, VERTS, and VALTS) should still be subject to Level 1 Peer Review, just like code-centric V&V Plans. In addition, these application-centric V&V Plans would be subject to Level 2 and Level 3 Peer Reviews. The Panels feel that Level 1 Reviews are still *essential* at the code-centric level, but that Level 2 and higher reviews are most applicable to application-centric V&V.
 3. Along with the definition of roles, the funding and management structure must permit code teams to acquire the resources necessary to implement the V&V requirements being imposed on them. In some cases the current management structure creates barriers to a “code team” because the staff representing the four elements listed above reside in multiple departments. Our current structure can also make unclear which line manager is responsible for the “code team”.
 4. Elements of the V&V process can provide products that are extremely valuable to the Center in coordinating work and resources. For example, the identification of stockpile drivers and the phenomena identification and ranking tables yield specific information on code development needs and experimental validation needs that can be used to prioritize investments under ASCI Apps, Experimental Programs, and Laboratory Research and Development Programs. A process that combines the results of the V&V reports for the different code projects and provides them to Center management to guide programmatic

decisions might be valuable. Otherwise this information may sit in isolated V&V reports and may not be used beyond the confines of the individual code projects.

5. Some thought should be given to capturing tangible benefits now, to hedge against losing all this work (i.e., V&V Plans, PIRTS, understanding of needed V&V activities, experimental requirements, etc.) if future funding goes away.
6. Code specific feedback should become input for the next phase of V&V Plan development and review. The Panels and Code Teams would like to see a task relating to closure/response on V&V Program feedback to the ASCI V&V Program Office.
7. The Program needs to provide adequate cross-cutting support for:
 - o UQ efforts for all code teams.
 - o templates for many of the work products (e.g., design documents).
 - o resources for verification and validation experiments.
 - o more personnel trained in SQE and V&V.

4.4 General Observations

It should be made absolutely clear that Code teams are doing V&V for their customers, not for the ASCI V&V Program. It is essential that the owners of the V&V Plans have the authority and the resources to engage the necessary participants in the development of the V&V Plan. Along with the definition of roles, the funding and management structure must permit code teams to acquire the resources necessary to implement the V&V requirements being imposed on them.

The panels feel that what is needed at the next level (application-centric focus) is a clear mechanism for pulling together elements from across many departments and disciplines (e.g., analysts, experimentalists, code developers, SQE/V&V experts). This mechanism must address funding and authorization issues. V&V Plans produced via this mechanism (which for a given stockpile application will include several codes and the associated PIRTS, models, VERTS, and VALTS) should still be subject to Level 1 Peer Review, just like code-centric V&V Plans. In addition, these application-centric V&V Plans would be subject to Level 2 and Level 3 Peer Reviews. The Panels feel that Level 1 Reviews are still *essential* at the code-centric level, but that Level 2 and higher reviews are most applicable to application-centric V&V.

Code teams have a perception that needless, burdensome SQE practices are going to be imposed upon them. Good practices, *i.e.* those that demonstrably add value, are strongly encouraged, and will help ASCI applications achieve success. However, teams must be made to understand that SQE practices should be evaluated and adopted *commensurate with the size (and operational style) of the code team*. Program-level guidance *and support* should also be available to code teams for SQE-related activities. As these become available, teams should take advantage of them, and should cite in their V&V Plans the practices they use.

Configuration Management appears to be piecemeal. A programmatic or corporate solution, one that *meets the needs of both V&V and code teams*, should be provided. CM needs to include the concept of movement from Unclassified to Classified. CM and other practices should be commensurate with scope and rigor applicable to teams. The issue of backups (CM of tools, including the CM tool) is apparently nowhere addressed.

Organization-level guidance on Verification should be provided (e.g., grid-convergence analysis, Method of Manufactured Solutions).

The Panels felt that the definition of Validation is too narrow, in that it does not allow for validation in the absence of relevant data. The actual issue is probably not necessarily to extend the definition of Validation. Perhaps there should be an "accreditation/qualification" acceptance step included in the Stockpile Computing process where expert judgment is applied to existing evidence (e.g., verification, validation, certification, qualification, or the lack thereof) to determine whether existing M&S evidence is sufficient. Accreditation must be applied when resources aren't available for needed experiments. Who owns these "accreditation" activities is a subject for Stockpile Computing guidance.

There may be intellectual-property issues with respect to a Commercial Off-The-Shelf and Government-Off-The-Shelf code that is integrated into another code or a framework. This has implications for CM, security models, and procurement.

5 Summary of Major Recommendations

The following recommendations are from the accumulated panel reviews and represent the most important panel suggestions from the reviews.

5.1 Continue Use of V&V Guidelines and Peer Review Process

The panels, in particular the customer representatives, felt that the customers should request continued use of V&V Plans and Peer Reviews as a basis for supporting a common approach for application-specific V&V activities. It was recommended that DP and Surety customers of M&S support for the stockpile applications should request the continued development of information as required by the V&V Guidelines 2.0 or its updated version. The code teams and integrated STS applications would be more responsive to continuing the development of V&V Plans with such a request. The panels' consensus was that the approach to developing the information required in the V&V Plans was excellent and should be continued. The panels' consensus was that the Level 1 Peer Reviews, as conducted, were valuable and well organized. The success of continued reviews is dependent upon a well-defined Level 2 Peer Review process that can focus on an integrated approach and technical content.

The existing V&V Plans and results are adequate to do significant FY02 V&V planning and execution. All panels (and most of the code teams) have voiced feedback that the process was beneficial in establishing improved communication among the various stakeholders, defining better technical content, and developing a common method/terminology for future use.

5.2 Focus Level 2 Peer Review Process on Integrated STS Applications

The proposed Level 2 Peer Review, which will focus on technical content, should be applied primarily to integrated Stockpile-to-Target-Sequence (STS) applications. These reviews should focus on STS applications for the normal, abnormal, and hostile environments. In addition, the focus should be on coupling of multiple phenomena which would imply the integrated use of multiple codes. These integrated applications should develop V&V Plans following the V&V Guidelines 2.0 (as updated). The panel for such reviews should have strong customer participation. The development team would include all the necessary participants from the customer, designer, analyst, experimentalist, code team, and V&V program. The effort should be run as a project with schedule, effort, and performance defined by the project.

5.3 Integrate the V&V Plan Participants for the Most Effective Results

The integration of code developers, analysts, experimentalists, and customers for the V&V Plan development and execution appears to provide the most effective results. The current process appears to work much better with an analyst or phenomenologist as the lead author, than with a code developer alone in that role. The V&V Plans created by teams that had the recommended integration and experience with stockpile certification were the best for programmatic organization and detailed content. In those cases where the center, department, and line organizations infrastructure supported this integrated team concept, the review results were noticeably better. Code Teams that had active participants in all the indicated roles provided more detailed customer requirements, a more thorough ranking and prioritization method for the

resulting phenomena requirements, and created more complete VERTS and VALTS with detailed traces to the recommended testing tiers and the phenomena requirements.

It is very clear that execution of different elements of the V&V Plans must belong to different people, hence the recommendation that *ownership* (as opposed to mere authorship) consist of code representatives, customer representatives, analysts, and experimentalists. The current requirement (PR1) has become a ‘box-checking’ exercise, in which authors’ names are included to meet the requirement. The intent is that the various roles actually *participate* in V&V planning. It is essential that the owners of the V&V Plans must have the authority and the resources to engage the necessary participants in the development of the V&V Plan.

5.4 Quantify Metrics and Accuracy Requirements more Rigorously

There is inadequate treatment of success measures. This observation refers to acceptance criteria for verification and validation, as well as metrics to assess the effectiveness of SQE practices. Related to this is the specification of needed accuracy for experimental results. Although uncertainty quantification (UQ) is considered part of a standard ASCI product, in general very little if any V&V aspects of the UQ and its associated process have been discussed.

The V&V Plans should strive to quantify the accuracy requirements for verification and validation in a more rigorous fashion, or reference where such quantification is to be described. In general, Plans have had little or no discussion of software metrics. As a whole, there was a notable lack of depth and detail in the discussion of metrics in the V&V Plans. Regarding acceptance criteria, quantitative comparisons are preferable, where possible, to visual comparison of plots. Answers to the following questions should be more apparent: How do you know when you’re ‘done’? How good is good enough?

5.5 Develop Organizational Level Guidance for SQE, Verification, Validation, and Stockpile Computing

ASCI program plans currently include activities for completing organizational level guidance for SQE, validation methods, and stockpile computing during FY02. It is recommended that this guidance be provided and referenced by the peer review process. In addition, organization-level guidance should be provided for Verification (e.g., grid-convergence analysis, Method of Manufactured Solutions).

5.6 Revise Peer Reviewed V&V Plans and Publish

It is recommended that each V&V Plan should be revised as appropriate to incorporate suggestions from the code-specific peer review report and published as a SAND report. It is important to establish a baseline of information that will benefit future V&V Plans (at the code or application level).

5.7 Revise V&V Guidelines 2.0 and Peer Review Process 1.0 and Publish

The V&V Guidelines 2.0 in reference [V&V-GUIDE-2_0] and Peer Review Process 1.0 in reference [V&V REVIEW-1_0] should be updated to incorporate recommended improvements. Incorporate the suggested improvements and publish as updated versions. The Sections 4.2 and 4.3 provide details of the suggested changes.

Appendix A. Acronyms and References

A.1. Acronyms

ASCI	Accelerated Strategic Computing Initiative
CM	Configuration Management
DOE	Department of Energy
DP	Defense Programs
DSW	Directed Stockpile Work
FY	Fiscal Year (October 1 – September 30)
IP	Implementation Plan
M&S	Modeling and Simulation
MC	Military Characteristics
MMS	Method of Manufactured Solutions
PIRT	Phenomena Identification Ranking Table
QFD	Quality Function Deployment
SCG	Stockpile Computing Guidelines
SIRT	Services Identification Ranking Table
SLEP	Stockpile Life Extension Program
SNL	Sandia National Laboratories
SQE	Software Quality Engineering
SSP	Stockpile Stewardship Program
STS	Stockpile to Target Sequence
TBP	Technical Business Practice
UQ	Uncertainty Quantification
V&V	Verification and Validation
VALTS	Validation Test Suite
VERTS	Verification Test Suite

A.2. References

- [TBP-SYS] TBP-SYS, Issue G, "*Technical Business Practices System*," NWC Technical Business Practice, February 15, 2001.
- [V&V-GUIDE-1_0] T. G. Trucano and J. L. Moya, "*Sandia ASCI V&V Plan Guidelines – Version 1.0*," Sandia National Laboratories Report, SAND99-3098, December 1999.
- [V&V-GUIDE-2_0] M. L. Pilch, T. G. Trucano, J. L. Moya, G. Froehlich, A.L.Hodges, and D. E. Peercy, "*Sandia ASCI V&V Plan Guidelines – Version 2.0*," Sandia National Laboratories Report, SAND2000-3101, January 2001.
- [V&V-REVIEW-1_0] M. L. Pilch, T. G. Trucano, D. E. Peercy, A. L. Hodges, E. R. Young, and J. L. Moya, "*Peer Review Process for the Sandia ASCI V&V Program: Version 1.0*," Sandia National Laboratories Report, SAND2000-3099, Albuquerque, NM January 2001.
- [DOE-ASCI-SQE] Department of Energy, "*ASCI Program Software Quality Engineering: Goals, Principles, and Guidelines*," DOE/DP/ASC-SQE-2000-FDRFT-VERS2, Albuquerque, NM February 2001.
- [SNL-ASCI-SQE] J. D. Zepper, K. M. Aragon, M. A. Ellis, K. A. Byle, D. S. Eaton, "*Sandia National Laboratories ASCI Applications Software Quality Engineering Practices*," Version 1.0 November 15, 2001.
- [V&V-ALEGRA_FRAMEWORK] D. E. Carroll, M. F. Pasik, C. D. Seidel, C. D. Turner, A. L. Hodges, G. K. Froehlich, D. E. Peercy, "*ALEGRA Framework Verification and Validation Plan for the EMPHASIS Code Team Modeling and Simulation Scenarios for the W76-1 Life Extension Program STS Electromagnetic Radiation (EMR) Environment*," Draft Sandia National Laboratories Report, May 1, 2001.
- [V&V-ALEGRA_EMMA] C. J. Garasi, D. E. Carroll, R. J. Lawrence, S. T. Montgomery, A. L. Hodges, G. K. Froehlich, and D. E. Peercy, "*ALEGRA-EMMA Verification And Validation Plan for Modeling and Simulation Scenarios Involving the Ferroelectric Neutron Generator Power Supply, Version 01.05*," Draft Sandia National Laboratories Report, September 5, 2001.
- [V&V-CALORE] V. J. Romero, S. R. Tieszen, R. R. Lober, R. J. Cochran, D. D. Dobranich, and T. Y. Chu, "*CALORE Verification and Validation Plan for Weapon-in-a-Fire Simulations Version 1.1*," Draft Sandia National Laboratories Report, February 20, 2001.
- [V&V-CEPTRE ITS] L. Lorence, D. E. Beutler, W. H. Barrett, C. R. Drumm, W. C. Fan, B. C. Franke, R. P. Kensek, T. W. Laub, and J. L. Powell, "*CEPTRE and ITS Verification and Validation Plan, Version 1.0*," Draft Sandia National Laboratories Report, May 1, 2001.

- [V&V-EMPHASIS] R. E. Jorgenson, W. A. Johnson, L. K. Warne, H. G. Hudson, and K. Oishi, "EMPHASIS: ASCI Verification and Validation Plan for Electromagnetic Radiation Environment," Draft Sandia National Laboratories Report, April 30, 2001.
- [V&V-FUEGO_SYRINX] S. R. Tieszen, A. R. Lopez, C. D. Moen, T. Y. Chu, V. F. Nicolette, W. Gill, S. P. Burns, and W. C. Moffatt, "*SIERRA/FUEGO and SIERRA/SYRINX Verification and Validation Plan Version 2.0*," Draft Sandia National Laboratories Report, February 20, 2001.
- [V&V-GOMA] L. A. Mondy, R. R. Rao, P. R. Schunk, P. A. Sackinger, D. A. La-breche, and D. B. Adolf, "*Verification and Validation of Encapsulation Flow Models in GOMA, Version 1.1*," Draft Sandia National Laboratories Report, March 14, 2001.
- [V&V-HPEMS] S. D. Wix, C. W. Bogdan, J. P. Marchiondo, A. V. Nunez, and D. R. Weiss, "*High Performance Electrical Modeling and Simulation Software Normal Environment Verification and Validation Plan, Version 1.0*," Draft Sandia National Laboratories Report, March 1, 2001.
- [V&V-NUGET] P. J. Griffin, P. J. Cooper, R. S. Pacheco, and W. H. Barrett "NuGET Verification and Validation Plan: Version 0.99," Draft Sandia National Laboratories Report, September 4, 2001.
- [V&V-PEGASUS] M. E. Riley, R. J. Buss, R. B. Campbell, M. M. Hopkins, P. A. Miller, W. R. Wampler, A. R. Moats, and M. A. Gallis "Verification and Validation Plan for the Codes LSP and ICARUS (PEGASUS)," Draft Sandia National Laboratories Report, September 4, 2001.
- [V&V-PRONTO_JAS] T. D. Hinnerichs, "*Validation Plan for the B61-3/4/7/10 Radar Nose Assembly Crush Model*", Draft Sandia National Laboratories Report, April 4, 2000.
- D. K. Parrish and M. L. Blanford, "*Solid Mechanics Software Verification Plan*," Draft Sandia National Laboratories Report, April 5, 2000.
- [V&V-SACCARA] M. L. McWherter-Payne and K. Salari, "*SACCARA Verification and Validation Plan, Version 1.0*," Sandia National Laboratories Report, SAND2000-1346, June 2000.
- [V&V-SIERRA_FRAMEWORK] D. L. Knirk, K. M. Aragon, and H. C. Edwards, "*SIERRA Framework Verification And Validation Plan, Version 01.21*," Draft Sandia National Laboratories Report, February 21, 2000.

(This page is intentionally blank except for this statement)

Appendix B. Level 1 Peer Review Scoring Summaries

This Appendix provides scoring information to indicate basic conformance across code teams (without reference to which code teams or which session is involved) to the V&V Guideline criteria, as well as identification of noteworthy practices recognized by the Panels during the reviews. Noteworthy practices do not include the acceptable practices that most V&V Plans reflected, so there should be no reflection on Code Teams that are not listed. In addition, it is noted that there are probably other noteworthy practices that were not recognized due to the specific focus of the reviews. Information in this appendix is intended to provide some indication of areas where overall strengths and areas of improvement might be targeted relative to the specific criteria in the V&V Guidelines, and practices that may be useful to use at an organizational level for future applications. The information and charts in this section are presented in the spirit of identifying general areas of concern and particularly good practices (recognized by code team).

B.1. Scoring Summaries and Analyses

Each panel member contributed scores of 0,1,2 for each of the V&V Plan criterion listed in Table B-1 for each of the V&V Plans for which the panel member was a reviewer. Although it is not appropriate to disclose individual V&V Plan scores, it is informative to look for trends or indications in the scores as to criteria that scored particularly well or poorly. It is also appropriate to stratify the scoring to determine if there are general code team organization methods or best practices that might have contributed to particularly high or low scores.

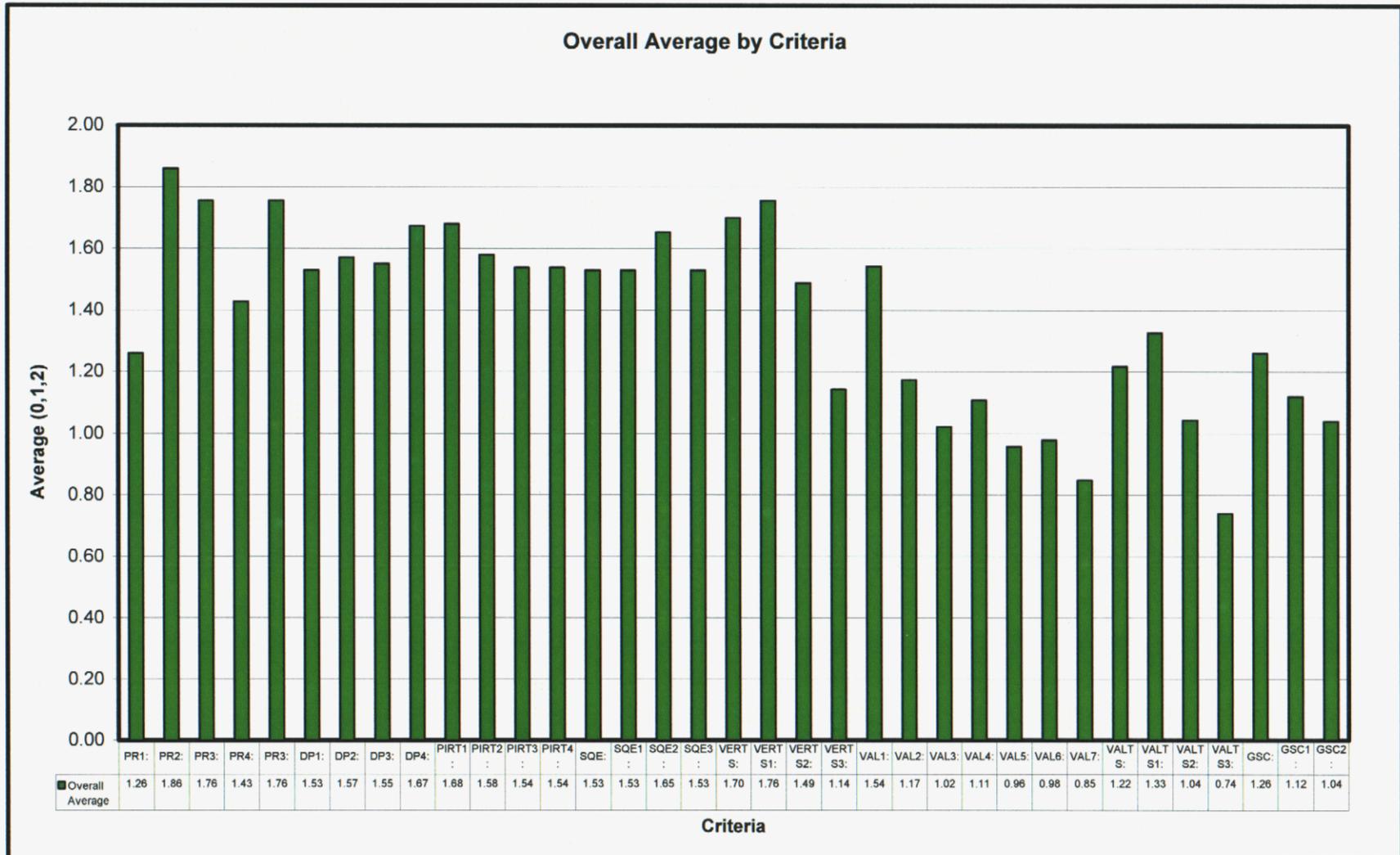
Table B-1. V&V Guidelines Criteria

1.	General Guidelines
PR1	The V&V plan authorship includes the V&V process owner, and experimenter, and a DP customer representative.
PR2	The plan is compatible with the format specified in the Version 2.0 guidelines.
PR3	A single stockpile driver for the V&V plan is identified.
PR4	The V&V planning process is described.
2.	Stockpile Drivers and DP Customer Requirements
PR3	A single stockpile driver for the V&V plan is identified.
DP1	The appropriate customers and constraints associated with the stockpile driver have been identified.
DP2	The detailed stockpile requirements been extracted from the stockpile driver.
DP3	The stockpile requirements have been mapped to M&S needs and requirements.
DP4	There are sufficient requirements to allow the development of a useful PIRT.
3.	Phenomena Identification And Ranking Table (PIRT)
PIRT1	The PIRT is present in the V&V plan

PIRT2	The PIRT process methodology is described.
PIRT3	Phenomena in the PIRT are ranked by a rational scoring system.
PIRT4	Current M&S status (capability) for each of the phenomena in the PIRT is presented.
4.	Software Quality Engineering (SQE)
SQE	Content required by DOE and Sandia SQE policy documents is present.
SQE1	Software V&V
SQE2	Software Engineering
SQE3	Project Management
5.	Verification Testing (VERTS)
VERTS	A Verification Test Suite has been constructed and documented.
VERTS1	The structure and logic of the VERTS is addressed.
VERTS2	The construction of the VERTS is addressed.
VERTS3	The acceptance metrics for the VERTS are addressed.
6.	Validation Plan (VALTS)
VAL1	The relation between validation activities and the PIRT is defined.
VAL2	The data requirements associated with the validation activity are described.
VAL3	The known dependence of assessment of the fidelity of the code physical models upon experimental data is described.
VAL4	The opportunities and obligation for experimental investigation of the assumptions underlying the implemented physical models are discussed.
VAL5	The dependence of the code models upon experimentally measured quantities is discussed.
VAL6	The prioritized experimental needs are described.
VAL7	The anticipated use of technologies like uncertainty quantification in the validation activities is described.
VALTS	A Validation Test Suite is constructed and documented.
VALTS1	The structure of the VALTS is described. The structure is compatible with the PIRT and the four-tiered approach described in these guidelines.
VALTS2	The construction of the VALTS is addressed.
VALTS3	The acceptance metrics for the VALTS are addressed.
7.	Stockpile Computing Guidance (STG)
GSC	Guidance for stockpile computing using the code is provided.
GSC1	Technical guidance for code application is provided.
GSC2	Process guidance for code application to the associated stockpile driver is provided.

The following Figure B-1 is the overall average of the scores within each of the V&V Criterion areas. Scores above 1.5 are considered to be good. Scores below 1.0 are considered to be areas for general improvement.

Figure B-1. Overall Average of Scores by V&V Criterion



The following subjective bounds provide a measure of how well the overall set of V&V Plans evaluated satisfied the V&V Guidelines criteria.

Average Score ≥ 1.5	Criterion well- satisfied across the V&V Plans
Average Score ≥ 1.0 & < 1.5	Criterion partially satisfied across the V&V Plans
Average Score < 1.0	Criterion not satisfied across the V&V Plans

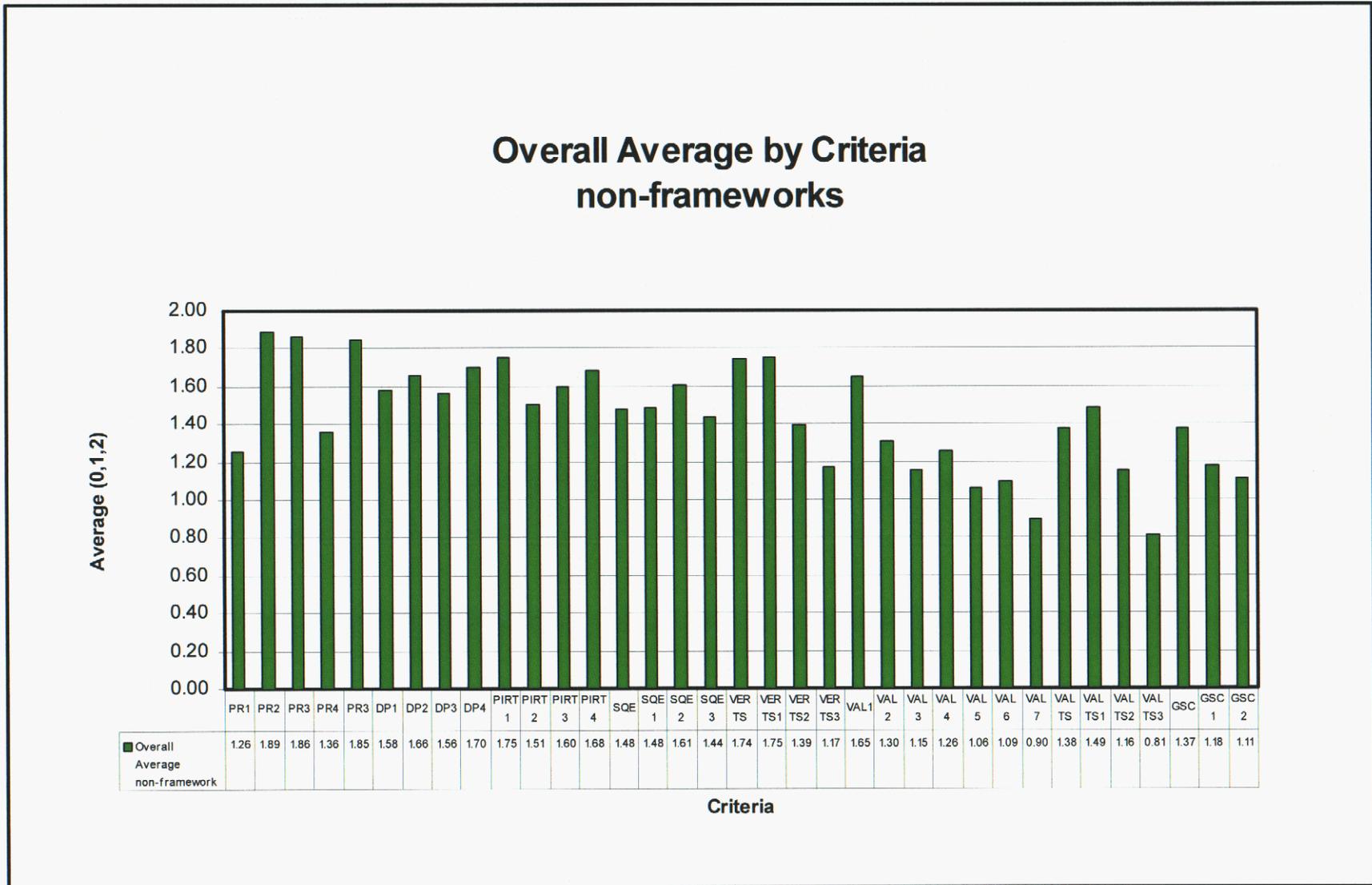
Using these general bounds and the overall averages as indicated in Figure B-1, the General Guidelines (except for PR1), Stockpile Drivers, Phenomena Identification Ranking Table criteria, and Software Quality Engineering criteria all scored in the "well-satisfied" category. The Verification Plan/Test Suite category was split between well-satisfied and partially satisfied. The Validation Plan/Test Suite category was essentially "not satisfied". The Stockpile Computing Guidance category was on the low side of partially satisfied. Hence, the general areas of programmatic concern are the VALTS and the Stockpile Computing Guidance. Along with the relatively lower score for PR1 (integrated team of authors for the V&V Plan), these scores reflect the general lack of an organization structure within many of the teams constructing the V&V Plans that includes customers, analysts, and experimentalists as an integral part of the team. It also reflects a lack of understanding about how the M&S codes are going to be used as part of the stockpile computing program. However, there were a few code teams who did have this organizational structure and who did have experience with using their codes for stockpile certification activities. These code teams were knowledgeable about the validation and stockpile computing issues and scored well in these areas.

Analyzing the specific criteria, the lowest criteria were VAL3, VAL5, VAL6, VAL7, and VALTS3. VAL3 and VAL5 were acknowledged as being confusing, but the consensus of the panels was that experimentally measured data, dependencies, and priorities (VAL6) were not adequately described. Likewise, for VAL7, the use of Uncertainty Quantification (or in general error analysis of the results) was inadequately described. VALTS3 was the lowest and indicated a general lack of adequate acceptance/success metrics for validation. One can also observe that VERTS3 was also fairly low for the acceptance/success metrics for verification. All together, the low scores for these criteria indicate that there is inadequate understanding across the V&V Plans of adequate measures of "goodness" for the verification and validation activities.

It is important to remember that these scores indicate programmatic information that may or may not be directly linked to how well the actual verification and validation activities of a code team are being accomplished. However, the panels indicated strong support for the construction of such V&V Plans for integrated applications as an indication to the customer that their stockpile computing needs are being adequately addressed. Analysis and review of the actual V&V activities through methods such as the Level 2 and Level 3 Peer Reviews was considered essential to provide convincing evidence that the integrated application does satisfy their stockpile computing needs.

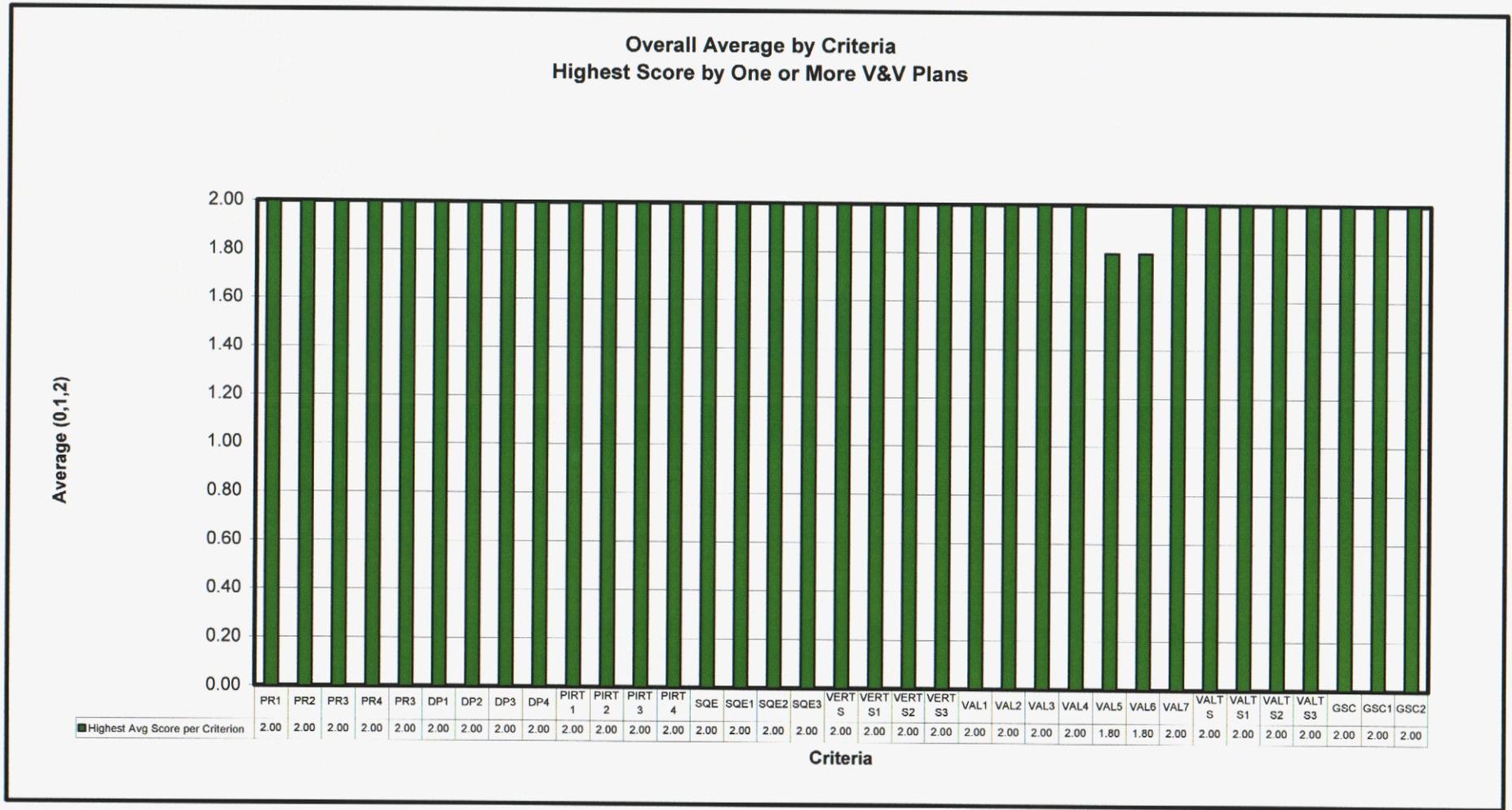
The following Figure B-2 is the overall average scores excluding the Frameworks (ALEGRA, SIERRA). This figure illustrates a slightly better score for the Validation part since the Frameworks were not targeted for Validation Plan and Stockpile Computing Guidance criteria except through the services provided to the code teams that use the Frameworks.

Figure B-2. Overall Average of Scores by V&V Criterion (no Frameworks)



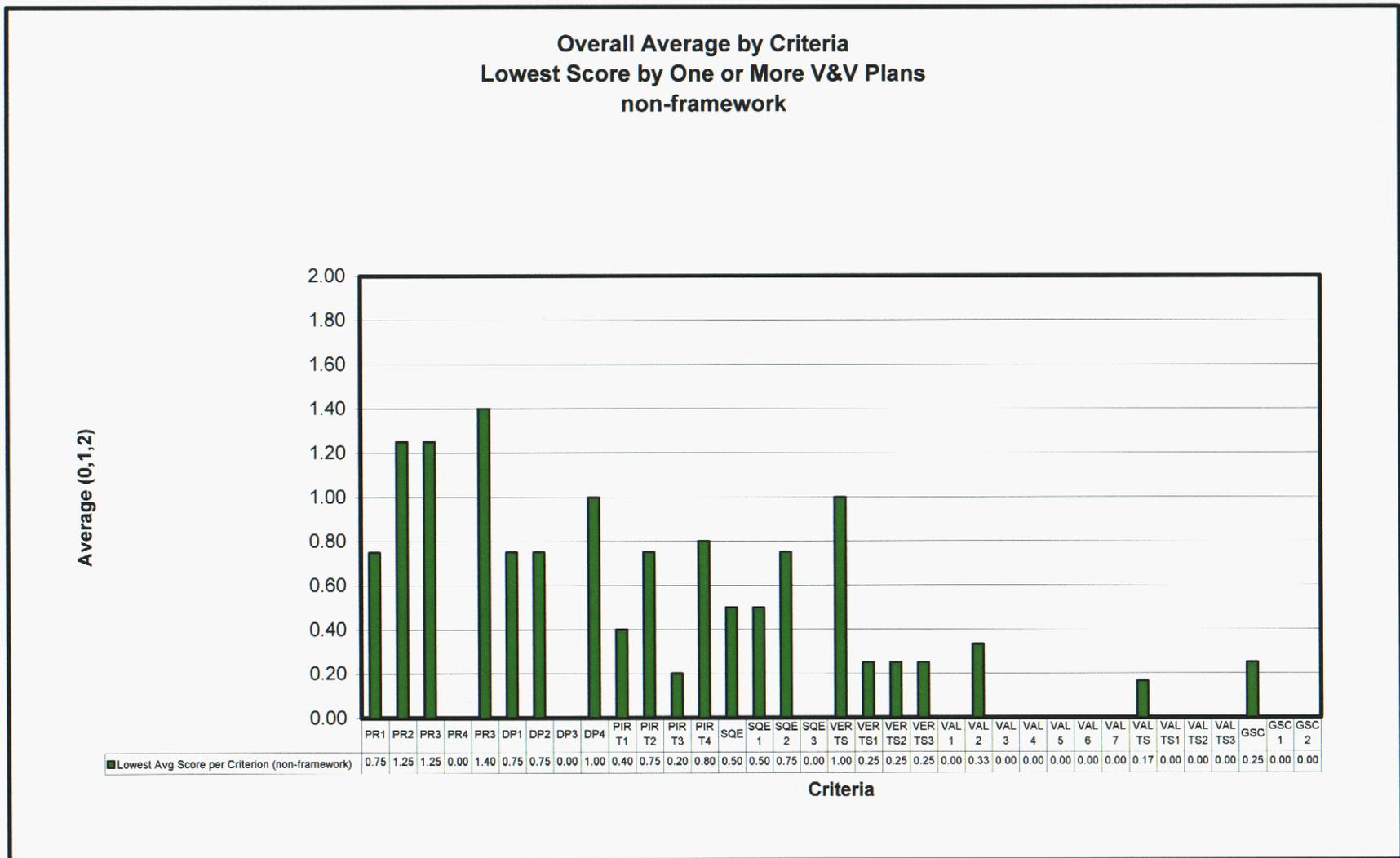
The following Figure B-3 shows the highest average score given for each criterion for one or more code teams. The scores indicate that for every criterion, at least one code team's V&V Plan scored very well.

Figure B-3. Highest V&V Plan Average Score for each Criterion



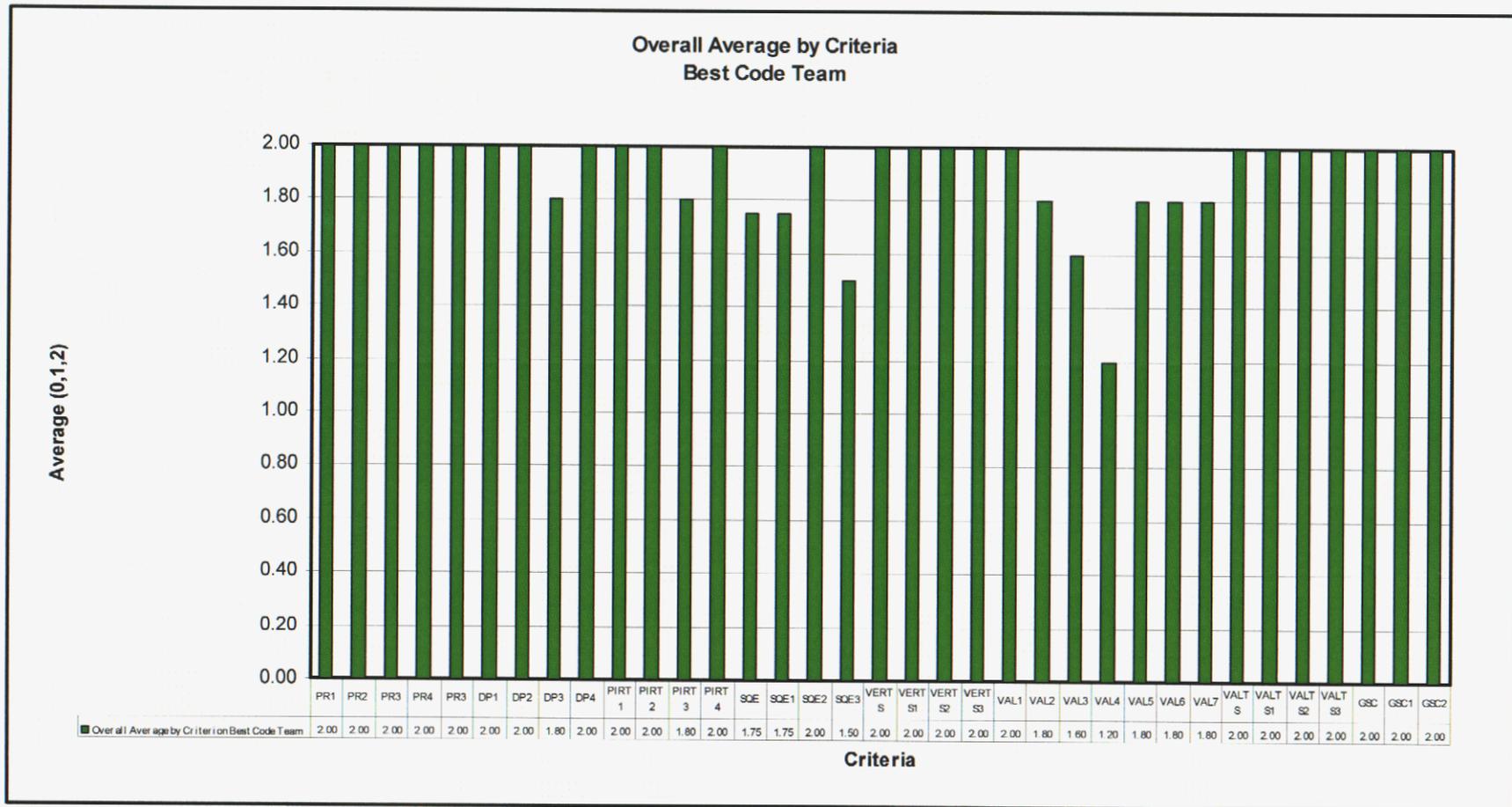
The following Figure B-4 shows the lowest average score given for each criterion for one or more code teams. The scores indicate that for every criterion, at least one code team's V&V Plan did not score well. "Blank" column is actually a zero average score. Non-framework codes were used so as to not bias the validation and stockpile computing scores.

Figure B-4. Lowest V&V Plan Average Score for each Criterion



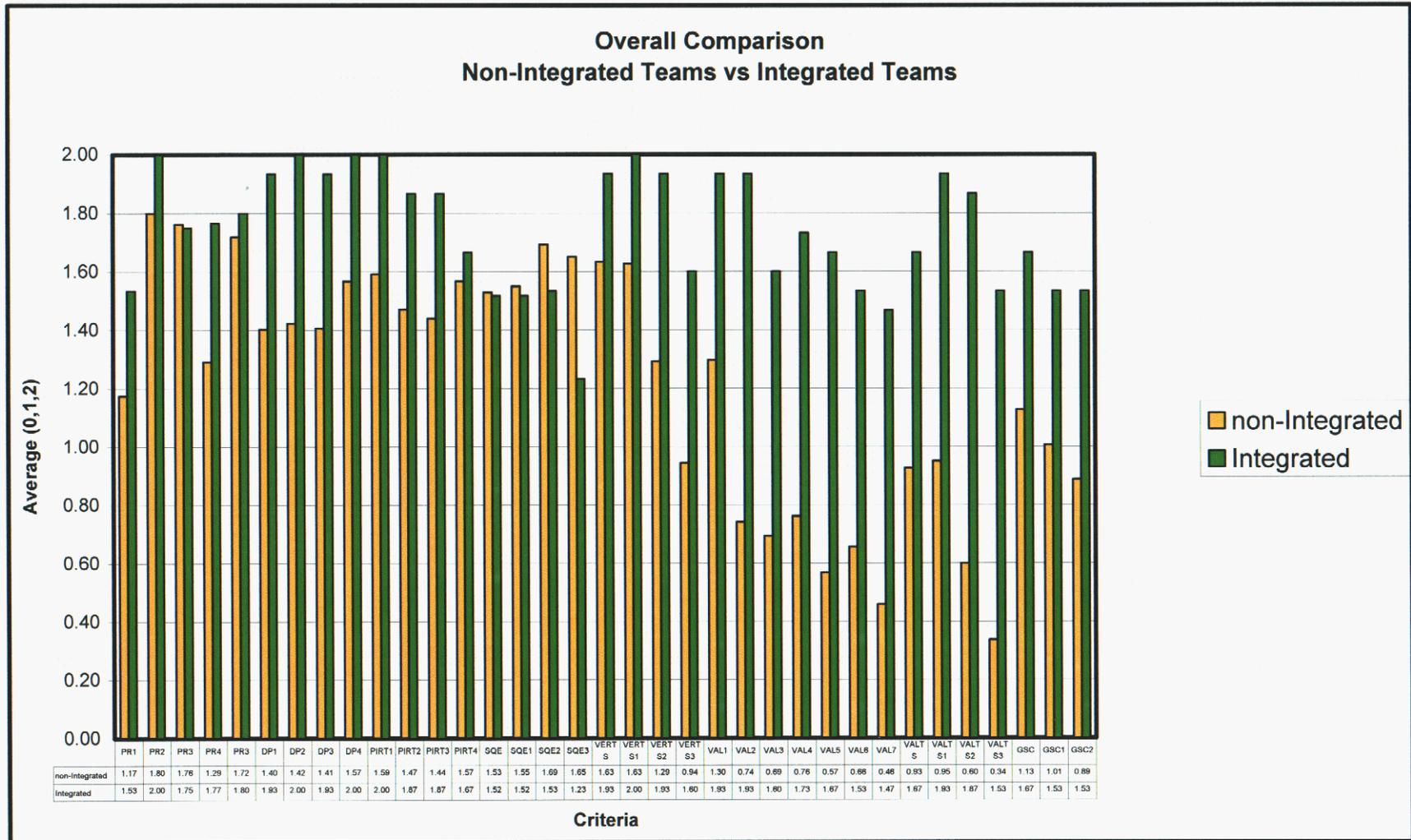
The following Figure B-5 provides an example of the best code team V&V Plan average scores. This provides some evidence that the V&V Guidelines as defined in reference [V&V-GUIDE-2_0] can be followed exceptionally well. Furthermore, this particular team had a lot of Level 2 evidence of having actually conducted verification and validation activities, so the maturity of the code/application is significant in being able to adequately satisfy V&V requirements. This also provides some evidence that the other codes, as used within specific stockpile computing activities, can adequately satisfy V&V requirements as these codes become more mature.

Figure B-5. Example of Best V&V Plan Score vs. Overall Average Score



The following Figure B-6 provides a comparison of the overall average of scores for those code teams that did not have an integrated (customer, experimentalist, analyst, code developer, V&V) organization versus those code teams that had more of an integrated organization. With only a few exceptions (notably in SQE), the integrated organization scored much better, particularly in the Validation Plan and Stockpile Computing Guidance criteria.

Figure B-6. Comparison of Overall Average Scores for non-Integrated vs. Integrated Code Teams



B.2. Noteworthy Practices

There were several noteworthy practices as recognized by the responsible panel and itemized in Table B-2. These practices may assist code teams and applications in improving the overall approach to development of V&V Plans and conducting V&V activities in the future. Such practices typically represent a rather unique way to implement a part of the V&V Plan. Normal good practices are not identified in this table. It is recognized that there are probably other noteworthy practices that have not been captured in this table.

Table B-2. List of Noteworthy Practices

Area	Noteworthy Practice Description	Code Team(s)
Stockpile Driver/Customer	<ul style="list-style-type: none"> • Identification of customers • Identification of and justification for selection of various V&V Plan participating roles • Stockpile Driver description and tie to M&S Requirements 	<ul style="list-style-type: none"> • CEPTRE-ITS (Tables 1&2) • NuGET • NuGET
PIRT	<ul style="list-style-type: none"> • General PIRT Chart Representation and Scoring/ranking system • Methodology/application of PIRT to framework services • Concept of extracting application-specific PIRTs from a larger & more general Global PIRT 	<ul style="list-style-type: none"> • CEPTRE-ITS, NuGET • SIERRA & ALEGRA • HPEMS, EMPHASIS
SQE	<ul style="list-style-type: none"> • Discussion of framework services (SIERRA) used by code team's V&V Plan. • List of software development recommendations • SQE practices defined, being used, convinced of their usefulness • Promotion concept (development, design analysis, qualification, release); software as product 	<ul style="list-style-type: none"> • FUEGO-SYRINX • GOMA • HPEMS, NuGET • CEPTRE-ITS, NuGET
VERTS	<ul style="list-style-type: none"> • Questions and tables format for presenting verification and validation tests • Metrics for assessing the success of VERTS • Good description of the prioritization methodology and its use • Concept of "tier" is expanded to cover levels of phenomena • VERTS description and tie to PIRT 	<ul style="list-style-type: none"> • EMPHASIS • FUEGO-SYRINX • GOMA • CEPTRE-ITS • NuGET
VALTS	<ul style="list-style-type: none"> • Metrics for assessing the success of VALTS and validation experiments • VALTS description and tie to PIRT 	<ul style="list-style-type: none"> • FUEGO-SYRINX • NuGET
Stockpile Computing Guidance	<ul style="list-style-type: none"> • General description of issues, concerns, approach for use of code for stockpile certification. 	<ul style="list-style-type: none"> • NUGET
Performance Self-Appraisal	<ul style="list-style-type: none"> • Evaluation criteria matrix mapping criteria (e.g., PIRT1) to the relevant section of the Plan. 	<ul style="list-style-type: none"> • CEPTRE-ITS (performance self-appraisal part)
Other	<ul style="list-style-type: none"> • Team organization: integrated customer, analysts, experimentalists, code developers. • Document tree/reference use • Plans in place for the sustainment of the code and the V&V Plan • Color-coded cross-reference among the PIRT, VERTS, VALTS and qualification status 	<ul style="list-style-type: none"> • NUGET, CEPTRE-ITS, FUEGO-SYRINX • HPEMS • GOMA • NuGET (Table 12)

Distribution

EXTERNAL DISTRIBUTION

Department of Energy, NNSA (4)
1000 Independence Ave., SW
Washington, DC 20585

Attn: D. E. Bird, NA-115
Attn: K. D. Greenaugh, NA-115.3
Attn: T. Y. Chu, NA-115.3
Attn: William Reed, NA-114

Los Alamos National Laboratory (4)
Mail Station 5000
P.O. Box 1663
Los Alamos, NM 87545

Attn: Joseph Kindel, MS B259
Attn: Ken Koch, MS F652
Attn: Douglas Post, MS F661
Attn: David Tubbs, MS B220

University of California (4)
Lawrence Livermore National Laboratory
7000 East Ave.
P.O. Box 808
Livermore, CA 94550

Attn: R. Christensen, MS L-160
Attn: Richard Klein, MS L-023
Attn: Roger Logan, MS L-125
Attn: Cynthia Nitta, MS L-096

Argonne National Laboratory (1)
MCS Division
Bldg. 221, Rm. C-236
9700 S. Cass Ave.
Argonne, IL 60439
Attn: Paul Hovland

Texas Institute of Comp. Mechanics (1)
University of Texas at Austin
Austin, TX 78712
Attn: Tinsley Oden

SANDIA INTERNAL

#	MailStop	Org	Name
1	MS 1056	1111	W. R. Wampler
1	MS 1118	1423	M. E. Riley
1	MS 1152	1642	W. A. Johnson
1	MS 1152	1642	R. E. Jorgenson
1	MS 1152	1642	M. L. Kiefer
1	MS 1152	1642	J. D. Kotulski
1	MS 1152	1642	M. F. Pasik
1	MS 1152	1642	D. B. Seidel
1	MS 1152	1642	C. D. Turner
1	MS 1152	1642	L. K. Warne

1	MS 1152	1643	H. G. Hudson
1	MS 1194	1644	H. C. Harjes
1	MS 1186	1674	R. B. Campbell
1	MS 1186	1674	R. J. Lawrence
1	MS 0525	1734	C. W. Bogdan
1	MS 0525	1734	R. B. Heath
1	MS 0525	1734	J. P. Marchiondo
1	MS 0525	1734	A. V. Nunez
1	MS 0525	1734	P. V. Plunkett
1	MS 0525	1734	S. D. Wix
1	MS 0885	1802	G. S. Heffelfinger
1	MS 0888	1811	D. B. Adolf
1	MS 0429	2100	J. S. Rottler
1	MS 0427	2104	P. A. Sena
1	MS 0482	2109	R. A. Paulsen
1	MS 0483	2112	R. J. Koss
1	MS 0483	2112	R. S. Pacheco
1	MS 0479	2113	B. S. Altman
1	MS 0479	2113	K. Oishi
1	MS 0479	2113	W. J. Tedeschi
1	MS 0481	2114	W. C. Moffatt
1	MS 0481	2114	M. A. Rosenthal
1	MS 0509	2300	M. W. Callahan
1	MS 0537	2333	D. R. Weiss
1	MS 0521	2561	S. T. Montgomery
1	MS 0769	5800	D. S. Miyoshi
1	MS 1146	6423	P. J. Cooper
1	MS 1146	6423	P. J. Griffin
1	MS 1146	6423	G. A. Harms
1	MS 1146	6423	K. O. Reil
1	MS 0451	6500	S. G. Varnado
1	MS 1138	6532	E. R. Young
1	MS 1137	6536	S. M. Eaton
10	MS 1137	6536	G. K. Froehlich
5	MS 1137	6536	A. L. Hodges
1	MS 1137	6536	M. A. Tebo
1	MS 1137	6536	C. M. Williamson
1	MS 9042	8728	C. D. Moen
1	MS 9003	8900	K. E. Washington
1	MS 9217	8915	M. L. Koszykowski
1	MS 0836	9100	M. R. Baer
1	MS 0841	9100	T. C. Bickel
1	MS 1393	9100	T. Y. Chu
1	MS 0824	9110	R. A. Ratzel
1	MS 0826	9113	W. L. Hermina
1	MS 0827	9113	M. A. Gallis
1	MS 0834	9114	M. M. Hopkins
1	MS 0834	9114	J. E. Johannes
1	MS 0834	9114	D. A. La-breche
1	MS 0834	9114	L. A. Mondy
1	MS 0834	9114	R. R. Rao
1	MS 0834	9114	P. R. Schunk
1	MS 0825	9115	M. A. McWherter-Payne
1	MS 0825	9115	W. H. Rutledge

1	MS 0825	9115	K. Salari	1	MS 0847	9226	P. Knupp
1	MS 0836	9116	D. D. Dobranich	1	MS 0318	9227	P. D. Heermann
1	MS 0836	9116	E. S. Hertel	1	MS 0318	9230	P. Yarrington
1	MS 0836	9117	R. J. Buss	1	MS 0819	9231	E. A. Boucheron
1	MS 0836	9117	R. O. Griffith	1	MS 0819	9231	K H. Brown
1	MS 0836	9117	A. R. Moats	1	MS 0819	9231	K. G. Budge
1	MS 0847	9117	D. K. Parrish	1	MS 0819	9231	S. P. Burns
1	MS 0847	9120	H. S. Morgan	1	MS 0819	9231	D. E. Carroll
1	MS 0555	9122	M. S. Garrett	1	MS 0819	9231	C. J. Garasi
1	MS 0893	9123	R. M. Brannon	1	MS 0819	9231	A. C. Robinson
1	MS 0557	9125	T. J. Baca	1	MS 0819	9231	M. K. Wong
1	MS 0847	9126	T. D. Hinnerichs	1	MS 0820	9232	P. F. Chavez
1	MS 0847	9126	R. A. May	1	MS 0820	9232	R. M. Summers
10	MS 0824	9130	J. L. Moya	1	MS 0316	9233	S. S. Dosanjh
1	MS 0836	9132	P. E. Desjardin	1	MS 0801	9300	A. L. Hale
1	MS 0836	9132	W. Gill	1	MS 0660	9519	M. A. Ellis
1	MS 0828	9132	L. A. Gritz	1	MS 0136	9713	P. E. D'Antonio
1	MS 0836	9132	V. F. Nicolette	1	MS 0139	9900	M. O. Vahle
1	MS 0828	9132	S. R. Teiszen	1	MS 0139	9902	P. J. Wilson
1	MS 0828	9133	A. R. Black	1	MS 0139	9904	R. K. Thomas
1	MS 0828	9133	K. J. Dowding	5	MS 0139	9905	S. E. Lott
1	MS 0828	9133	W. L. Oberkampf	1	MS 0428	12300	D. D. Carlson
1	MS 0557	9133	T. L. Paez	1	MS 0428	12301	V. J. Johnson
10	MS 0828	9133	M. Pilch	1	MS 0829	12323	J. M. Sjulín
1	MS 0828	9133	V. J. Romero	1	MS 0638	12326	M. A. Blackledge
1	MS 0828	9133	M. P. Sherman	1	MS 0638	12326	D. L. Knirk
1	MS 0557	9133	A. Urbina	10	MS 0638	12326	D. E. Percy
1	MS 0847	9133	W. R. Witkowski	1	MS 0492	12332	D. R. Olson
1	MS 1135	9134	S. R. Heffelfinger	1	MS 0405	12333	M. P. Bohn
1	MS 0835	9140	J. M. McGlaun	1	MS 0405	12333	T. R. Jones
1	MS 0835	9141	R. J. Cochran	1	MS 0405	12333	J. S. Ludwigsen
1	MS 0835	9141	S. N. Kempka	1	MS 0830	12335	K. V. Diegert
1	MS 0835	9141	R. R. Lober	1	MS 1170	15310	R. D. Skocypec
1	MS 0835	9141	P. A. Sackinger	1	MS 1179	15340	J. R. Lee
1	MS 0847	9142	S. W. Attaway	1	MS 1179	15341	D. E. Beutler
1	MS 0847	9142	M. L. Blanford	1	MS 1179	15341	B. C. Franke
1	MS 0847	9142	M. W. Heinstein	1	MS 1179	15341	R. P. Kensek
1	MS 0847	9142	S. W. Key	1	MS 1179	15341	T. W. Laub
1	MS 0847	9142	J. S. Peery	1	MS 1179	15341	L. J. Lorence
1	MS 0847	9142	G. M. Reese	1	MS 1179	15341	J. L. Powell
1	MS 0827	9143	K. M. Aragon	1	MS 1159	15344	W. H. Barrett
1	MS 0826	9143	K. D. Copps	1	MS 1166	15345	C. R. Drumm
1	MS 0827	9143	H. C. Edwards	1	MS 1166	15345	W. C. Fan
1	MS 0847	9143	G. D. Sjaardema				
1	MS 0826	9143	J. D. Zepper	1	MS 9018	8945-1	Central Technical
1	MS 0321	9200	W. J. Camp				Files
10	MS 0819	9211	T. G. Trucano	2	MS 0899	9616	Technical Library
1	MS 1109	9212	R. J. Pryor	1	MS 0612	9612	R&A Desk
1	MS 1110	9213	L. G. Benavides				for DOE/OSTI
1	MS 1110	9214	R. B. Lehoucq				
1	MS 1110	9214	D. E. Womble				
1	MS 0310	9220	J. A. Ang				
1	MS 0847	9220	R. W. Leland				
1	MS 1110	9223	N. D. Pundit				
1	MS 1110	9224	D. W. Doerfler				