Risk Management Techniques and Practice Workshop

WORKSHOP REPORT

December 2008

Compiled by Terri Quinn and Mary Zosel
Lawrence Livermore National Laboratory

LLNL-TR-409240

Workshop Steering Committee

Arthur Bland, ORNL; Vince Dattoria, SC/ASCR/DOE HQ; Bill Kramer, LBNL; Sander Lee, NNSA/ASC/DOE HQ; Terri Quinn, LLNL (workshop chair); Randal Rheinheimer, LANL; Mark Seager, LLNL; Yukiko Sekine, SC/ASCR/DOE HQ; Jeffery Sims, ANL; Jon Stearly, SNL; and Mary Zosel, LLNL (host organizer)

Workshop Group Chairs

Ann Baker, ORNL; Robert Ballance, SNL; Kathlyn Boudwin, ORNL; Susan Coghlan, ANL; James Craw, LBNL; Candace Culhane, DoD; Kimberly Cupps, LLNL; Brent Draney, LBNL; Ira Goldberg, ANL; Patricia Kovatch, UTenn/NICS; Robert Pennington, NCSA; Kevin Regimbal, PNNL; Randal Rheinheimer, LANL; Gary Skouson, PNNL; Jon Stearly, SNL; and Manuel Vigil, LANL

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Contents

I. Executive Summary

At the request of the Department of Energy (DOE) Office of Science (SC), Lawrence Livermore National Laboratory (LLNL) hosted a two-day Risk Management Techniques and Practice (RMTAP) workshop held September 18–19 at the Hotel Nikko in San Francisco.

The purpose of the workshop, which was sponsored by the SC/Advanced Scientific Computing Research (ASCR) program and the National Nuclear Security Administration (NNSA)/Advanced Simulation and Computing (ASC) program, was to assess current and emerging techniques, practices, and lessons learned for effectively identifying, understanding, managing, and mitigating the risks associated with acquiring leading-edge computing systems at high-performance computing centers (HPCCs).

Representatives from fifteen high-performance computing (HPC) organizations, four HPC vendor partners, and three government agencies attended the workshop.

The overall workshop findings were:

- Standard risk management techniques and tools are in the aggregate applicable to projects at HPCCs and are commonly employed by the HPC community
- HPC projects have characteristics that necessitate a tailoring of the standard risk management practices
- All HPCC acquisition projects can benefit by employing risk management, but the specific choice of risk management processes and tools is less important to the success of the project
- The special relationship between the HPCCs and HPC vendors must be reflected in the risk management strategy
- Best practices findings include developing a prioritized risk register with special attention to the top risks, establishing a practice of regular meetings and status updates with the platform partner, supporting regular and open reviews that engage the interests and expertise of a wide range of staff and stakeholders, and documenting and sharing the acquisition/build/deployment experience
- Top risk categories include system scaling issues, request for proposal/contract and acceptance testing, and vendor technical or business problems

II. Introduction

The RMTAP workshop, held Sept. 17–18, 2008, in San Francisco, CA, convened to assess current and emerging techniques, practices, and lessons learned for effectively identifying, understanding, managing, and mitigating risks associated with acquiring leading-edge computing systems at HPCCs. Sponsored by the DOE—jointly by the SC and the NNSA—and hosted by LLNL, the workshop was targeted at HPCC managers and key staff who are planning for leading-edge computational systems. A workshop steering committee from the major DOE computing centers and DOE headquarters (HQ) agreed on the abstract and specific goals for the workshop. Two major breakout discussion tracks were organized to address the specific questions associated with the theory and practice of risk management when applied to HPC. The workshop was attended by 64 HPCC representatives invited from the HPC community of DOE, the National Science Foundation (NSF), the Department of Defense (DoD), and major HPC platform vendors. (See Appendix D. Workshop Attendees.)

Workshop Abstract

HPC, by its very nature, is an exercise in multi-level risk management. Every aspect of stewarding HPCCs into the petascale era, from identification of the program drivers to the details of procurement actions and simulation environment component deployments, represents unprecedented challenges and requires effective risk management. The fundamental purpose of this workshop was to go beyond risk management processes as such and learn how to weave effective risk management practices, techniques, and methods into all aspects of migrating HPCCs into the next generation of leadership computing systems.

This workshop was a follow-on to the Petascale System Integration Workshop[1] hosted by Lawrence Berkeley National Laboratory (LBNL)/NERSC[2] last year. It was intended to leverage and extend the risk management experience of the participants by looking for common best practices and unique processes that have been especially successful. This workshop assessed the effectiveness of tools and techniques that are or could be helpful in HPCC risk management, with a special emphasis on how practice meets process. As the saying goes: "In theory, there is no difference between theory and practice. In practice there is."[3] Finally, the workshop brought together a network of experts who shared information as technology moves into the petascale era and beyond.

[1] Information about the Petascale System Integration Workshop can be found at http://www.nersc.gov/projects/HPC-Integration/.

[2] Information about NERSC can be found at http://www.nersc.gov/.

[3] Quote variously attributed to Yogi Berra, Jan L. A. van de Snepscheut, and even Albert Einstein.

Workshop Goals

The organizing committee agreed on the following goals for the workshop:

- Define and understand basic risk management terminology in the context of HPCCs
- Identify top risks and share existing practices and methodologies
- Understand how risk management practices and methodologies can be appropriately tailored to the level of risk associated with the activity
- Share lessons learned from past management of risk activities
- Discuss who should be responsible for various classes of risk and what are appropriate risk-sharing strategies
- Establish communication paths for managerial and technical staff at multiple sites to continue the discussion on effective risk management practices
- Present findings to DOE and other stakeholders to improve the process of high-performance computing system deployment

III. Workshop Format and Plenary Sessions

The RMTAP workshop agenda (see Appendix A. Workshop Agenda) was a combination of plenary sessions to provide a common perspective to all of the attendees and of breakout sessions for more detailed interactive discussion on different aspects of risk management theory and practice. The sections below describe the plenary sessions and the activities related to the breakout sessions.

Summary of Plenary Session 1

In the first plenary session, several government agencies (SC, NNSA, the NSF Computer & Information Science & Engineering (CISE), and the DoD) set the stage with a summary of their acquisition strategies and plans, and their expectations for the workshop. These agency presentations, as well as the plenary speaker presentations, are available on the workshop Web site (https://rmtap.llnl.gov).

Sander Lee described the system acquisition roadmap for the NNSA/ASC program, where three classes of systems were addressed:

- The advanced systems that push technology
- The capability systems where applications may use up to half of the systems
- The capacity systems

He said that each of these three classes of systems has different plans for risk management, and, in fact, advanced systems are a form of "risk mitigation." By completing these systems—pushing vendors to do things they would not usually do—the industry develops the more dependable systems needed for capability and capacity. The next NNSA system, Zia, will be sited at Los Alamos National Laboratory (LANL) in partnership with Sandia National Laboratory (SNL) as a New Mexico center. In addition, the selection is currently in progress for the Sequoia system at LLNL. Lee expected results from this workshop to feed into the NNSA planning process and documentation. Lee said that these are production systems. With such large acquisitions, it is important that acquisition executives have confidence that these systems will meet the needs of DOE customers. To the extent best practices can be captured, it helps stakeholders understand the differences between the HPC systems and traditional information technology. Risk management also impacts partnerships with industry to the extent a common language is used.

Dan Hitchcock, SC/ASCR/DOE HQ, reviewed the workshop addressing HPC facilities, which was held last year. Hitchcock noted that already some of the lessons from last year's workshop have been put into practice. He added that after collecting more risks, one must understand which risks need the most consideration. Like the ASC program, the SC has a spectrum of facilities, from advanced to production. There are always deployment issues, and reports from workshops like this are useful when HPCCs interact with stakeholders such as the Chief Information Officer.

Steve Scherr described investments for the DOD High-Performance Computing Modernization Program (HPCMO) in long-haul networking, software development, and

consolidated acquisitions. Last year, 125-T systems were installed, and they are growing toward petaFLOPS systems.

Steve Meacham described the two tracks of advanced systems being funded by the NSF CISE organization to provide an impact on the science possible on the NSF TeraGrid. NSF also manages the computing for the National Center for Atmospheric Research (NCAR) under a different funding portfolio.

The government agency presentations were followed by two invited presentations. David Featherman from Booz Allen Hamilton (BAH) gave a talk titled "The Risk Management Partnership—Achieving Project Success through Collaboration." He summarized how traditional approaches to risk management of large projects have changed from the idea that most risks are technical with cost and schedule handled by well-written contracts. Several complex, space-based research and development (R&D) projects illustrated the pitfalls of ignoring the interplay between technical, cost, and schedule factors. Featherman said, "Today, active project management of the risks associated with all of the factors can provide important information to critical decisions involving the project. The key is effective communication from project personnel up to the decision makers. The Rocky Flats clean-up project provides an example of how proactive project and risk management can be a big win. Initially, the project estimates were $37 billion over 70 years in an adversarial situation. This changed over time with "smart risk reduction" to a successful conclusion of the project after 14 years at the cost of $7 billion." The talk included discussion of risk terminology and identification processes separating known from unknown risk sources, as well as distinctions between internal and external risks.

The second invited presentation, titled, "A Perspective on Risk Identification and Management for Complex Technical Projects," was given by Dale Knutson from PNNL. Knutson's talk showed how the classic definition of risk as probability *x* consequence can be difficult to apply to complex technical projects. Knutson discussed how to transform questions of probability and consequence into terms that are more directly applicable. For example, rather than requiring numerics for probability, one can use more qualitative terms based on likelihood. Knutson said that instead of thinking of risk management as associated with scope, schedule, and budget, it is useful to think of it as the integration of these traditional aspects with communications; staff; environment, safety, and health; and procurement. DOE has project management orders that are basically "good practice." A complex project does not necessarily mean complex management tools. To conclude the talk, the application of risk management techniques was illustrated by applying it to some of the risks identified in last year's workshop.

Summary of Plenary Session 2

The second-day plenary session started off with a panel of platform vendors from Cray, HP, IBM, and SGI. Each panel member presented an overview on the complexities of risk management of the leading-edge computing systems from the supplier's perspective.

Kevin Stelljes, Cray, addressed the risks inherent to a small company in the HPC business. Stelljes said that HPC comes with a lot of business risks: small market, high development costs with front-loaded risk, and long sales/deployment cycles. For a small company, this poses many problems in cash flow, inventory control, and even staff management. The roles of the proposal and project managers in understanding the risks are critical, as are openness and partnership with the customer. Major deployment

challenges include dependency in third-party suppliers and facilities, and time to conduct full-scale testing.

Brad Blaine, HP services project management, addressed why projects fail. Blaine said that common causes are insufficient end-user involvement, unclear objectives and requirements, changing requirements, insufficient senior management support, and poor management of resources and expertise. All of these issues benefit from strong project management.

Sohel Saiyed, program director for HPC architecture and delivery, IBM, described the (non-business) approach to managing risk associated with large system builds. Saiyed said that many formal risk management processes are used and there are many, many reviews during the process. Some of the key items, based on lessons learned, are that it is important to assign a strong project manager as soon as possible and to identify a high-level management sponsor. In the middle stages of the system development, it has been helpful to anchor all software development requirements to specific key contracts to help identify/avoid slippage. Another useful lesson in the later stages of the system build has been to assign tiger teams and have daily meetings of the technical team without layers of management present, so that issues can be discussed frankly. The right partnership attitude for a large contract is important; if IBM has signed up for many detailed requirements several years in the future, there needs to be partnership flexibility, such as recognition for overachieving as well as underachieving on these long-term requirements.

David Morton, SGI, addressed the issues related to financial risk associated with the large high-end systems. Morton said that these systems are inherently risky but can be profitable if proposed, planned, and executed well. From the supplier perspective, forecasting next-generation performance and extended schedules is difficult, bidding is expensive, and internal benchmarking and testing resources are also a challenge. Things that can help mitigate the financial risks are shorter contract times, limiting the number of options requested in request for proposals (RFPs), progress payment options, crisp acceptance criteria, and limited obligations beyond acceptance."

Workshop Breakout Tracks

The remainder of the workshop was organized around two tracks of breakout sessions. Each track was divided into four session topics, addressed by different breakout groups. Each topic was led by representatives from two institutions and facilitated by a note taker.

Track 1, Tailoring Risk Management to HPCCs, addressed the different aspect of formal risk management: risk ownership and analysis, risk identification and categories, risk management tools, and mitigation and contingency planning. These sessions addressed how well the formal methods associated with risk management fit HPC (the theory). Questions to address included:

- How well do the typical risks encountered in HPCCs fit the standard processes for (topic of session)?
- What are the key lessons learned?
- What are the best practice models for HPCC (in topic of session)?
- What are opportunities to share/improve practices and terminology?
- Any additional findings?

Track 2, Real Life Risk Experience, addressed the phases of acquisition and deployment of an HPC: from vision to contract, from contract award through the build, acceptance testing and integration, and management of the business risks. These sessions identified the top risks in different areas of HPC and discussed the application of risk management method (the practice). Questions to address included:

- What are the top risks (in area of session discussion)?
- For each risk, what are the successful (or not) mitigation strategies, what are the successful (or not) management techniques, what tailoring of techniques is helpful, and what gaps remain and how to cover?
- Any additional findings?

Following each track, each session reported their findings back to the full group of attendees. They also provided a written report of their discussions. These summary presentations for each breakout are included on the RMTAP Web site (https://rmtap.llnl.gov) and the detailed written reports are in Appendix B. Breakout Sessions and Reports.

Final Workshop Session

In the final workshop session, the issues and findings reported by the breakout sessions were listed and presented for a vote to identify the top findings in the opinion of the workshop attendees. The results of the vote are given in Appendix C. Analysis of Workshop Questionnaire. In the analysis of the issues and findings, it was found that they grouped into major themes. These themes, along with material from the breakout presentations and written reports, form the workshop findings in the next section of this document.

IV. Workshop Findings

This section presents and discusses the overall workshop findings originating from all sessions, plenary sessions, breakout groups, and the voting. Two votes were held for both the most important risks and the risk management best practices. More details on the findings can be found in the individual breakout group reports and the voting results. See Appendix C. Analysis of Workshop Questionnaires.

Finding #1: Standard Risk Management Techniques and Tools Are, in the Aggregate, Applicable to HPCC Projects and Are Commonly Employed by the HPCC Community

This finding was substantiated by essentially every breakout session. The tools and processes taught in project management classes and described in the DOE M 413.3-1[4] are indeed applicable to the HPC community, even though HPCC projects are complex. Today, HPCCs are employing these processes and tools with self-proclaimed success. Examples of tools in common use are the consequence-likelihood matrix used to analyze risks and the risk registers used to document and monitor risks. The well-documented risk management strategies of accept, avoid, mitigate, or transfer are relevant as well. Risks are routinely accepted by HPCCs when, for example, they deploy new releases of software, such as operating systems. Risks are often avoided by evaluating new technology before committing to using it. Warranties are a common method for HPCCs to transfer risk. Another widely used risk management technique in the HPCC community is continuous monitoring of risks to ensure the risks are being effectively addressed. More than one HPCC reported that risk management has become an integral part of their organization's project management process. Though much of the body of knowledge on risk management is being put into practice by HPCCs, there is room for the community to improve. One area for improvement is in risk identification. It was fairly common to hear risks being described by one word (for example, scalability or financial) instead of the more useful and formal way of describing risks as a cause and effect relationship using an if-then-resulting-in statement.

Finding #2: HPC Projects Have Characteristics that Necessitate a Tailoring of the Standard Risk Management Practices

The process of buying, integrating, and operating leading-edge computers is inherently a complex—and expensive—endeavor. The risk management approaches taken by HPCCs can differ in important ways. The government agencies that participated in this workshop are pursuing an aggressive path to petascale computing and beyond. The planned computers are, by necessity, either a one-of-a-kind computer or a first delivery of a new architecture, and they will be notable in their size—from tens of thousands to hundreds of thousand of processors—and their novel components. Therefore, both

[4] DOE M 413.3-1, *Program Management for the Acquisition of Capital Assets*, dated March 28, 2003.

vendors and HPC centers may have far from complete knowledge early in the process of the reliability and performance of the delivered computer. This lack of data limits the ability of HPCCs to make use of quantitative risk analysis techniques. As a result, HPCCs almost exclusively employ qualitative techniques. Other notable differences are the importance to the success of the project of delivering the computer on schedule and the difficulties associated with staying on schedule. The success of HPCCs depends upon their ability to offer their customers the use of a fully functioning, leading-edge computer in the customer-defined time frame. The vendors are also equally motivated to meet schedules and even to deliver early, if possible, since they have large assets tied up in the equipment, and their shareholders want them to book the revenues as soon as possible. However, the challenges are formidable in delivering these computers on time. The system complexity, the large numbers of components making up these systems, and the need for novel technologies challenge the vendor's ability to predict with precision the time needed to build the computer and have it pass the battery of acceptance tests. This puts stress on the partnership. For all these reasons, careful attention to risks associated with schedule is particularly important for HPCC projects.

Finding #3: All HPCC Acquisition Projects Can Benefit by Employing Risk Management but the Specific Choice of Risk Management Processes and Tools is Less Important to the Success of the Project

There was unanimous agreement by the attendees that risk management in some form contributed to the successful acquisition of leading-edge computers. Each organization linked their application of risk management to successes. Furthermore, the Track 1 breakout oral reports revealed many commonalities in how risk management is being practiced in the HPC community. For example, the discussions concerning risk management tools revealed a strong preference for choosing simple tools; complex tools are not necessary and can actually hinder progress. Complex tools are expensive, and tool gurus had to be hired to run them. Not only did this add direct costs to the project, but it added indirect costs, as the tool gurus need special assistance to overcome their lack of experience with HPCC projects. Conversely, there was little agreement on the best set of risk management tools for HPCCs. The breakout groups reported a variety of risk management tools in use at HPCCs, ranging from off-the-shelf tools, such as Pertmaster, to homegrown tools built on spreadsheets. Interestingly, no organization put forth their tool set as the best. Those organizations that used off-the-shelf tools were satisfied, as were those that used homegrown tools. In addition, there was no detectable direct relationship between the success of a HPCC project and the project's particular choice of tools. From the evidence given, one can conclude that the use of any particular tool is not of great consequence. Each HPCC should select the tools that will work the best for its center.

Finding #4: The Special Relationship between the HPCC and HPC Vendors Must Be Reflected in the Risk Management Strategy

Acquiring leading-edge computers is accomplished through a partnership between customers and vendors. This relationship is much more collaborative than the customer-vendor relationship when purchasing commodity computers. In the latter relationship,

the customer places an order and the vendor satisfies that order—a relatively low-risk transaction. Conversely, buying leading-edge computers is risky and at a level commensurate with the uncertainty inherent in building expensive, first-of-a-kind computers with costs on the order of $100M. Unless the acquisition budget is sufficient to pay the vendor to accept the risk, essentially funding a large contingency, the HPCC needs to share the risk with the vendor. This risk sharing is a fundamental component of the partnership. There are several successful strategies to deal with the shared-risk. These proven strategies are phasing the deliveries of hardware and software, holding frequent reviews, requiring a limited set of mandatory requirements in the RFP, and negotiating options, off-ramps, or go/no go decision points as contingencies.

Finding #5: Best Practices Findings (Based on Questionnaire Voting)

The workshop voting related to Track 1 Best Practices brought a clear theme: establish an effective risk-register procedure and then communicate, communicate, and communicate. A truism that was "discovered" in multiple Track 1 breakout groups and emphasized by the invited speakers to be of utmost importance is that HPCCs should embrace and apply risk management. Of much lesser importance are the specific tools and methods chosen.

Develop a Prioritized Risk Register with Special Attention to the Top Risks

An up-to-date risk register can facilitate the meetings and reviews, as well as serve as an effective aid to informed decision making at all levels of project management. Most sites use a hierarchical model in which the highest risks in each technical area are rolled up to a higher level and combined with organizational risks and risks affecting multiple technical areas, thus forming a high-level risk register. Top risks should be tracked throughout the project; risks are not static.

Establish a Practice of Regular Meetings and Status Updates with the Vendor Partner

The discussion of risk status and handling should be a part of weekly executive-level and technical project meetings. It is important to have the support of the entire management chain at both the HPCC organization and the vendor organization. Top-level management reviews of both successes and problem areas can strengthen the partnership. Problem escalation processes can help ensure the appropriate level of management attention. Regular discussions should not be limited to the direct vendor but should also include subcontractors.

Support Regular, Open Reviews that Engage the Interests and Expertise of a Wide Range of Staff and Stakeholders

To the extent possible, be open with acquisition process and system build reviews while recognizing that there are some parts of the process that require confidentiality. Engaging the knowledge of a wide range of stakeholders, users, and staff in

collaborative, non-hostile reviews can improve the opportunity to expose risks and identify appropriate mitigation strategies.

Document and Share the Acquisition/Build/Deployment Experience

The process of pushing the leading edge of computing systems does not stop with a single acquisition. Documentation of the processes and lessons learned helps establish best practices from which all centers can learn. There is no current formal mechanism for this sharing. This is a possible future HPC community action item.

Finding #6: Top Risk Categories (Based on Questionnaire Voting)

The workshop attendees easily reached consensus on naming the most important risk categories for HPCCs. These top categories are not linked to any particular vendor or HPCC site but were deemed to be generally true for all leading-edge computer acquisition efforts. As with the results from the best practices voting, like categories are grouped for presenting the results of the voting. More details are in Appendix C. Analysis of Workshop Questionnaires.

System Scaling Issues

The most common feature that characterizes a leading-edge system is integration of hardware, software, and applications that together are larger and faster than systems previously built. Experience shows that this is not just a case of wiring more or faster components together. Often, new approaches/algorithms are needed in many aspects of the system—sometimes in the hardware, such as system switches, and almost always in many aspects of the system software, where some serial process can kill performance. Even when the system delivered supplies the specified performance, there is still the challenge of adapting the primary applications of the key users to take advantage of the new system architecture to deliver the expected performance. A Track 1 item receiving a significant vote count highlights the problem: vendors often do not have the in-house resources for testing software at scale. This problem may be mitigated by offering test resources at an HPCC.

Request for Proposal/Contract and Acceptance Testing

The RFP may not accurately specify what the customer needs and/or can afford. This may take the form of overly aggressive specification, which leads to problems getting proposals or negotiating a contract for the specified dollars, or problems with late delivery. Alternatively, key requirements may be missing, such that the system contracted is missing features to meet user or facility requirements. There is also a collection of risks associated with the set of acceptance tests that are specified in the RFP.

Vendor Technical or Business Problems

When a contract is signed for a leading-edge system, by nature of the advanced technology required, it is often the case that few of the system components can be considered "off the shelf." Development of new technology takes time and resources that require a substantial business commitment by the platform partner. The unknowns of R&D results, as well as cost and schedule variability, all tax the ability of the vendor to deliver the system as contracted. The customer/vendor partnership must be structured in a way as to be open about such difficulties, and the partnership must be prepared to negotiate contractual compromises and to ensure ongoing commitment. One such compromise might be to accept alternative technologies to those originally bid if the technology is no longer viable.

Personnel Staffing and Interactions

An important element of project success is creating and maintaining teams at both the vendor customer sites that can handle the technical challenges, as well as the communication needed in a large, complex undertaking. There is competition for a limited number of well-qualified people, both at the leadership/management level and at the technical level. Competition comes both from other HPC sites and from the wider information technology marketplace.

Project Schedule

Procurement and deployment of a leading-edge HPC system differs somewhat from other high-technology efforts because the technology is changing so fast. The operational lifetime of a top system is often not much longer than five years; thus, deviations in the schedule to procure and deploy the system are visible. The risks may come from problems with the procurement process, from the rate of technology change, and from the project problems with the system build and deployment, including possible infrastructure and facility issues.

Sponsor Commitment

Most of the leading-edge systems are funded by the government with the requirement for multi-year funding. This requires continued support from the government agency and may even appear specifically in the text of legislation. To maintain the needed commitment for the project to full deployment, a significant number of different stakeholders must have sustaining support for the project duration.

Facilities and Operations

There were several items with lower overall votes that relate to general operational issues. Some of these related to issues outside of the control of the project manager, such as electrical rates, that still must be part of the overall project operational and financial planning. It is mandatory that the total cost of ownership be part of the project scope.

V. Conclusion

It is appropriate to evaluate the results of the RMTAP workshop against the original goals set for the meeting, as listed in Section II. The workshop successfully addressed most of the original workshop goals. The findings and breakout reports in this document give details related to the workshop goal to identify top risks and best practices. In addition, the undocumented dinner session on past experiences with risks contributed to sharing lessons learned, another goal. Organizations are now encouraged to be proactive about documenting these lessons for the benefit of future installations. This document, along with the workshop presentations posted on the RMTAP Web site, directly addresses the last goal—to present workshop findings to DOE and other stakeholders. The workshop did succeed in establishing communication paths among HPCC managers and staff. The management and technical staff who attended the workshop, representing the major HPC sites around the country, had not met before as a group. This resulted in many new introductions that will facilitate continued interactions.

The basic concepts and definitions of formal risk management were discussed as they apply to HPCCs; however, the group did not define any specific new terminology that might be applied to HPCCs or used in HPCC risk management planning documents. As a future activity, it may be useful to develop a set of common terms to encourage sharing best practices among HPCCs. The other goal not directly met was to identify ownership for various classes of risk, although some risk-sharing strategies were addressed.

As was the case with the previous workshop, the attendees expressed interest in continuing this series of workshops. Suggestions were solicited at the workshop for subjects for future meetings. Several excellent suggestions were received—a strong indication that DOE should host a third workshop.

This workshop documentation, along with received suggestions, has been forwarded to DOE for consideration when planning for next steps.

Appendix A. Workshop Agenda

Wednesday, Sept. 17

7:30–8:15	Breakfast and registration
8:15–9:45	Plenary opening session: Terri Quinn, Chair Welcome and introductions:, Dan Hitchcock, Sander Lee, Terri Quinn, Steve Meacham, Steve Sherr Invited speaker: David Featherman, BAH
9:45–10:00	Break
10:00–11:00	Plenary opening session (cont.) Invited speaker: Dale Knutson, PNNL Track 1 breakout charter: Randal Rheinheimer Separate to breakout rooms
11:00–12:00	Track 1 breakouts—Tailoring Risk Management to HPCCs
12:00–1:00	Lunch
1:00–2:45	Track 1 breakouts (cont.)
2:45–3:00	Break
3:00–5:15	Plenary session: Bill Kramer, Chair Summary of previous workshop Track 1 breakout reports and discussion
6:30	Working dinner
7:45	Dinner panel: Vince Dattoria, Chair Each organization addresses: What are some of the "unknown unknowns" that occurred during your recent installation/upgrade? What risks did you significantly over-plan or under-plan during your recent installation/upgrade?

Thursday, Sept. 18

7:30–8:00	Breakfast
8:00–9:15	Plenary panel—HPC Risks from Vendor Perspective: Mark Seager, Chair Kevin Stelljes, Cray; Brad Blaine, HP; Sohel Siayed, IBM; David Morton, SGI
9:15–9:30	Track 2 breakout charter: Jon Stearley Separate to rooms
9:30–12:30	Track 2 breakouts—Real Life Risk Experience
12:30–1:30	Lunch
1:30–3:30	Track 2 breakout reports: Jim Ang, Chair
3:30–3:45	Break
3:45–4:45	Plenary wrap-up session: Terri Quinn, Chair Workshop summary, report (discussion) Next steps: Dan Hitchcock

Appendix B. Breakout Sessions and Reports

Track 1: Tailoring Risk Management to HPCCs

Session 1: Risk Ownership and Analysis: How do you know which risks will "bite" you?

Session Leaders: Susan Coghlan, Kevin Regimbal

Participants: Susan Coughlan, ANL (lead); Kevin Regimbal, PNNL (lead); Brent Gorda, LLNL (note taker); Paul Cook, SGI; Bryan Embry, DOD; Jim Foster, TACC; Steve Meacham, NSF-HQ; Mike Showerman, NCSA; Jon Stearly, SNL; Bob Tomlinson, LANL; Francesca Verdier, LBNL; Manuel Vigil, LANL; Mary Zosel, LLNL.

Session Charter:

Session 1 will focus on risk ownership and analysis. In this session, the team will briefly discuss risk identification but will move quickly into an in-depth discussion of risk ownership and analysis. We will walk through the classic analysis process and methods, with a focus on where HPC is unique and what types of risk analysis tools and methods do and do not work for HPC projects. The discussion will cover methods for determining how to assess the risks once they are identified (including how to quantify, compare, and prioritize risks) and a discussion of how a project measures the likelihood or consequence to determine risk ranking. The discussion will also cover how analysis should take into consideration the various baselines (cost, schedule and technical).

Session Process and Discussion:

There were a few questions to help understand each other's backgrounds. About half the participants considered themselves primarily technical staff, half primarily project managers. Many of the participants have received at least some training specific to project management. This was followed by a discussion of the overall risk management process. The group proposed this set of activities as fairly typical of HPC risk management processes across the various agencies and centers:

- Risk planning: build risk management plan
- Identify risks: build risk inventory/register
- Analyze risks: qualitative and quantitative techniques
- Prioritize risks: typically falls out of analyze risks step
- Plan risk responses: plan risk mitigations
- Execute: project and pre-identified risk mitigations
- Evaluate risks: review risk inventory, implement mitigations as needed
- Document
- Present project and be reviewed
- Repeat

Breakout Question 1: How well do typical risks encountered in HPC projects fit the standard risk analysis approaches?

The standard risk analysis approaches are relevant and applicable to HPC projects. As with all project management tools, the project manager needs to select and apply the practices that make sense in the context of the specific project. Most HPC implementation projects involve integration of technology at a scale not previously implemented. Gathering statistically significant historical information to drive some quantitative analysis methods is exceedingly difficult. This leads to a heavy dependence on qualitative analysis approaches. The Program Evaluation and Review Technique (PERT) method of gathering the low, most likely, and high estimates for time and cost is one quantitative method routinely applied to HPC risks.

The collaborative nature of HPC projects involving participation between HPC centers and HPC vendors introduces a challenging set of risks to analyze. The Statement of Work (SOW) is the primary vehicle documenting the nature of the shared responsibilities. *Collecting and maintaining a set of SOW best practices could be helpful for project teams to understand the level of risk introduced by these types of collaborative efforts.*

Breakout Question 2: What are the key lessons learned?

The risk management process has significant benefits for the execution of HPC projects. The process of identifying, analyzing, and monitoring risks helps projects succeed even in cases where it is difficult to pinpoint specific examples in which a mitigation explicitly reduced the probability of a risk occurring.

Risk management tools vary from "meeting facilitation" approaches to Excel spreadsheets to full-blown software packages. *Teams should employ tools usable by the project team itself*. If the complexity of the tool requires a project team to 'outsource' use of the tool to an outside consultant, the effectiveness of the overall process may be diminished.

Access to historical data would greatly facilitate the quantitative analysis process. For example, it may be easy to predict six months in advance the probability that a processor will not be available on time. Having statistics relevant to HPC implementations that quantify the likelihood of delay six months in the future and the magnitude of the delay would be useful and provide solid basis for inserting schedule contingency into the project plan. Some of this data is available from industry research organizations such as Gartner and IDC (International Data Corporation) at a cost, but most HPC-relevant data would need to be compiled and maintained by the HPC organizations. There is question as to whether the benefit of having this information to support quantitative analysis would offset the cost to gather and maintain the data.

Risks can have positive or negative effects on the project. Positive risks (often called opportunities) deserve attention as part of the overall risk management process.

Breakout Question 3: What are the risk analysis best practices for HPC?

During the risk identification and risk analysis phases of the risk management process, an important question for the project team to consider is anything about the new system that is different from the current system. These differences probably represent areas of greatest uncertainty and deserve additional attention.

Perspectives outside of the project can be valuable in identifying gaps and providing opportunities for best practice exchange. *HPC centers should consider inviting outside participation/review at key points in the projects.*

Risks need to be examined regularly to maintain awareness and provide updates as the project progresses. *It is often particularly helpful to keep the top-10 risks close at hand for frequent visibility.*

Breakout Question 4: What are the opportunities for sharing or improving practices and terminology?

Collecting SOW best practices as discussed above can be useful both for understanding the risk and improving the contracts aspects of HPC project. This includes providing examples of how to describe and bound risks when collaborative efforts involve shared risk responsibilities.

Comparing and employing common elements in the likelihood and consequence matrix across HPC projects can help make the risk management process more concrete for HPC. This can be (and is today) done via personal networking and is one of the most critical and accurate forms of information available.

Session 2: Risk Identification and Analysis—the Classic Categories: Are we covering all the bases?

Session Leaders: Kathlyn Boudwin, Patricia Kovatch

Participants: Kathlyn Boudwin, Oak Ridge National Laboratory LCF (lead); Patricia Kovatch, University of Tennessee NICS (lead); Tina Butler, LBNL (note taker); Thomas Bettge (NCAR), Vince Dattoria (SC/ASCR/DOE HQ), James Kasdorf (PSC), Thomas McKenna (PNNL), Tommy Minyard (TACC), David Morton (SGI), José Muñoz (NSF), Gary Skouson (PNNL), Brad Blaine (HP).

Session Charter:

Session 2 will focus on risk identification, with a discussion of risk ownership and risk analysis if time permits. In this session, the team will cover the standard categories, what areas fit the HPC and what ones do not, focusing on where the HPC community might be unique. We will move from the general categories to more focused and tailored categories with a goal of answering a number of questions, including: although the categories are broad, do they really cover HPC and what types of risks fall into each category? If the risk categorization topic is not sufficient to fill the time, the team will continue the discussion with risk ownership and risk analysis (how to measure and rank).

Process and Discussion:

With background information[5] as the initial basis of discussion, the group first discussed the various methods used to identify risks for past or on-going HPC projects. Generally, the team used the interviewing method as part of a brainstorming effort to identify risks. This method consists of "interviewing" experienced project participants, stakeholders, and subject matter experts. This interviewing technique might take place in a group setting or with one-on-one discussions with experts. The goal of the interviews is to identify specific project risks based on the experience and knowledge of those interviewed. The team also incorporated lessons learned from previous HPC installations with which they had been involved. Both of these brainstorming methods rely on the knowledge and experiences of each project team member and therefore can

[5] Background information was from *The Guide to the Project Management Body of Knowledge, Third Edition*.

fail when team members are inexperienced or the new project is different from previous ones. None of the more formal techniques (for example, Delphi technique) for risk identification had been used, but the general consensus was that as long as a wide net was cast drawing on the experiences of many people involved with HPC projects, the method was adequate for discovering possible risks. The group felt that there is a lot of HPC-specific knowledge in the community and that identifying risks and making this body of knowledge available to the community would be helpful.

This session had good representation from the HPC vendor community, and these team members were asked to talk about risk identification in industry. The industry participants reiterated the use of the brainstorming technique for identifying risks in their organizations. They used the SOW as the starting point in identifying their risks. The vendor group cautioned that transferring risk through the use of subcontracting mechanisms does not eliminate the risk. The transference of risk in this manner might even introduce new risks revolving around complexities in the relationship between the vendor and customer. A working partnership between vendor and customer seemed to foster overall risk reduction. A clear SOW and well-defined vendor deliverables were also suggested as a way to reduce the risk of contentious relations with vendors.

Breakout Question 1: "How well do the typical risks encounter in HPC fit the standard Risk Management models for risk identification?"

The categories in standard risk management models generally fit but are not specific enough in several key areas. These lists are most useful when used in conjunction with brainstorming exercises because they may trigger the identification of risks in areas not initially identified. They prompt the project team to think about the full spectrum of possible project-related risks. The suggested areas for more specificity were infrastructure, facilities, vendor viability, technology, scalability, customer readiness, system reliability measurements at scale, personnel (specific skills), data integrity, and documentation. Many of the HPC risks were subcategories of the technical risk category and further refinement of this broad category was suggested. Vendor viability also appeared to be a prominent and somewhat unique risk category to HPC.

Breakout Question 2: "What are the key lessons learned in risk identification?"

- Brainstorming is effective when used by experienced staff and incorporates a wide group of stakeholders and discipline experts.
- The process and outcomes should be documented. Historical records, including lessons learned, can be useful in identifying risks for the next HPC project.
- Risks should be review regularly by both customer and vendor. Do not think that the initial list of risks is the final list. Do not just go through the motions of risk management. The process can be useful as well as required.
- Not only risk identification and mitigation, but also a discussion of risk triggers, should be included in the risk process. How do you know when a risk occurs or when to initiate/or conclude a risk mitigation?
- Technical risks can often be mitigated by stepping away from the "bleeding edge" of technology. These risks should be evaluated in terms of the importance of the latest technology to the overall project goals.
- Something always breaks at scale. There will always be hardware problems. Inevitably something will not go according to plan.
- Remember that "risk filtering" occurs between the sub-contractor, contractor, and sponsor. An understanding of the tolerance for risk and risk reporting at each level is essential.

- HPC projects should leverage and cultivate open and productive vendor relationships.

Breakout Question 3: "What are the best practice risk models for HPC in the area of risk identification?

- Identify risks, mitigation strategies, and triggers, and then review them often.
- Incorporate risk into project discussions with all project participants (subcontractors, contractors, and sponsors). Make risk management part of the culture.
- Document the process, the risks, mitigation strategies, and the lessons learned.
- Foster open communication with all project stakeholders.

Breakout Question 4: "What opportunities exist to share or improve practices and terminology?"

The group identified several lessons learned from experiences achieved during the risk identification process. Brainstorming is an effective risk identification tools when it is used by experienced staff and incorporates a wide group of stakeholders and discipline experts. It is important to document the risk identification process, including the outcome of identified risks as well as risks which occurred but were not identified during the risk identification process. The risk register should be reviewed and updated regularly by project participants. The process of reviewing the risks listed in the risk register can be useful in the continuing risk identification and mitigation process.

The process of risk identification should also include the identification of risk triggers. The project should determine how they will know when a risk occurs or when to initiate or conclude a risk mitigation. Risk identification and mitigation should recognize that risks are inherent with the acquisition or development of "bleeding-edge" technology and strategies for risk mitigation should include the evaluation of the technology risk versus the benefit of being first. The group recognized that there are typically risks associated with "scaling," and that there will always be some hardware/software problems that should be considered when identifying project risks. All project participants should be aware that "risk filtering" occurs between the sub-contractor, contractor, and sponsor. An understanding of the tolerance for risk and risk reporting at each level is essential for successful risk management. HPC projects should leverage and cultivate open and productive vendor relationships to provide the best communication of risks during project planning and execution.

Several best practices for risk identification were also determined. The project participants should identify risks, mitigation strategies, and triggers, and then review them often. The discussion of project risks should be incorporated into the regular routine of the project participants (subcontractors, contractors, and sponsors). In this way, risk management will become part of the project culture. The project team should document the risk process, the risks, the mitigation strategies, and the lessons learned. Open communications with all project stakeholders should be fostered.

Opportunities exist in the HPC community to share and improve risk management practices and terminology. The HPC community could compile a list of generic risks and make this available to all HPC projects. A centralized Web site could be created to learn/share information about risks. This site could include project management risk terms for a better understanding of the formal risk management language. Additionally, inter-agency program managers should be encouraged to share information. The yearly SCXX conference could be used as a forum for discussing HPC risk management with a broad audience.

Summary:

The group was unanimous in its opinion that risk management has been useful in ensuring a successful HPC project. As the systems installed get larger and more complex, the discipline of managing risks will become more and more essential.

Session 3: Risk Management—Tools, Tips, and Tricks: "Please sign the register" or “There's nothing up my sleeve.”

Session Leaders: Jim Craw, Randal Rheinheimer

Session 3 Participants: Jim Craw, LBNL (lead); Randal Rheinheimer, LANL (lead); Pam Hamilton, LLNL (note taker); Candace Culhane, DoD; Ira Goldberg, ANL; Dale Knutson, PNNL; Charlie McMahon, LSU; Steve Louis, LLNL; Rob Pennington NCSA; and Yukiko Sekine SC/ASCR/DOE HQ.

Session Charter:

What risk management practices are required by DOE 413.3 and similar formal project management descriptions? Is there a common tool or set of tools currently in use by the participants? What additional tools are available and have been used in the HPC environment and found useful or not useful? What about practical techniques that are not associated with formal tools, such as reviews that escalate in frequency and visibility or requirement for interim deliverables? Are there any tools or techniques that are especially good for incorporating unknown risks?

Session 3 Process and Discussion:

Prior to the workshop, the following preparation questions were sent to contacts at all participating sites:

- What form of risk management documentation is done/required for your project/site (for DOE, this would be set out by DOE 413.3; you may have similar requirements for your organization), and does the answer change throughout the process?
- What formal tools (for example, Risk Radar) have you used or would you like to use to generate the risk management documents?
- Do you or your site have a best practice related to risk management tools that you are willing to share at the workshop?

At the breakout session, the general agenda was meant to be: 1) a review of the canonical areas of risk management, 2) a discussion of the relative relevance of formal risk management tools to each canonical area and how that relevance changes over time, and 3) a review of current risk management tools in use and a (pre-compiled) list of risk management software and processes, with a view to identifying gaps and best practices in applying the known tools and processes to the canonical areas of risk management.

Breakout Question 1: How well does the standard risk management model apply?

The breakout session immediately identified a huge variation in experience and application within the group, from risk management that is not really recognized as such, to formal application of custom risk management software within a highly formal project management infrastructure. Agreeing on a common risk management model against which to assess typical risks encountered in HPCCs, therefore, became impossible at the outset. However, the group agreed on the general principle that risk management tools are most useful, not for any of the particular canonical area of risk

management, but that increased formalism was most useful in managing the interdependencies of risks throughout a project.

Risk management models in use fell into two categories, a hierarchical model and a distributed model. Most sites use a hierarchical model in which the highest risks in each technical area are rolled up to a higher project management level and included with organizational risks and risks affecting more than a single technical area to form a high-level risk register. Typically, this is what gets reported to oversight bodies and is managed against a single SOW. The other model is distributed; each project component has its own SOW and SOW owner, and each manages its own risks within that context. Discussion of this model boiled down to appreciation of the project components having the focus and authority to effectively manage risk and concern that there is a large risk in missing risk interdependencies.

Within the hierarchical model, three levels of procedure emerged. The first has no formal recognition of risk management as a separate activity. The obvious criticism is that risks can be missed or fail to be managed if not called out explicitly. However, the site using this informal procedure has a history of successfully deploying large HPC clusters. The second procedural level was most common. This involves maintaining a hierarchical set of risk registers either through ad hoc tools such as Excel or via custom software, either RiskRadar or Pertmaster. The risk management process is not necessarily fully integrated into the overall project plan. Most of the sites operating at this level also used external reviews to validate either the risk management portion of projects or projects themselves. The highest procedural level has risk management fully integrated into the project process, but there still is no consensus on ad hoc versus custom software tools at this level of risk management formality.

The group consensus on how well typical HPCC risks are handled by existing risk management models consist of these two statements: *The processes currently being used are just adequate but have been sufficient to successfully field HPC systems in the past. Current processes do not lend themselves well to effectively reporting to the stakeholders.*

Breakout Question 2: What are the key lessons learned?

The key lessons learned in the discussion fell into two categories: what HPCCs can and would like to do and the barriers to doing so. The group agreed that HPCCs could learn from each other's experiences and methodologies, and that tools are needed to manage risks efficiently and report effectively to the stakeholders. The key lesson in this category is that there is no commonly agreed-upon process for risk management, so that **how** risk management is applied is much less important than **that** risk management is applied—thinking through risks at every level and phase of HPC project management is essential to success.

Three main barriers to formalizing risk management were identified, all of which fall under a general heading of insufficient expertise:

- Some organizations have not valued risk management formality highly enough to make it a priority to bring in expertise or to train staff.
- Risk- (and project-) management experts within the organizations do not, in general, have sufficient knowledge of HPC, making them of questionable utility.
- Custom risk management tools are related. Formal tools generally require HPC-external expertise to set up and maintain, moving the risk management process further from the subject matter experts who need to be intimately involved. Simpler tools can be more easily maintained within the HPC projects.

Breakout Question 3: What are the risk analysis best practices for HPC?

This question was energetically debated, given the range of practices. The group did not agree that there was enough information and experience to declare a best practice or set of practices, though the SC has recently implemented a standard project/risk management process that has the advantage of being standard.

Breakout Question 4: What are the opportunities for sharing or improving practices and terminology?

Continued sharing of experiences on the success of respective risk management processes was deemed to be the most useful extension of these discussions. The group also agreed it would be helpful to *form a generic HPC risk register as a template for new projects and a means of sharing risk categories that have been encountered in the community generally* but not necessarily at a given site.

Session 4: Risk Management—Mitigation and Contingency Planning: Know when to hold them and when to fold them.

Session Leaders: Bob Ballance, Kim Cupps

Participants: Robert Ballance (SNL, Lead), Kimberly Cupps (LLNL, Lead), Katie Antypas (LBNL, Note taker), Aaron Andersen (NCAR), James Ang (SNL), Ann Baker (ORNL), Christina Beldica (NCSA), Candy Culhane (DOD), James D'Aoust (SDSC), Brent Draney (LBNL), David Featherman (BAH), Ricky Kendall (ORNL), William Kramer (LBNL), Sander Lee (NNSA/ASC/DOE HQ), Matt Leininger (LLNL), Jonathan Monsein (DOD), Stephen Scherr (DOD), Sohel Saiyed (IBM), Kevin Stelljes (Cray), Douglas Tinnin (ANL), Francesca Verdier (LBNL), and Warren Yip (DOE Site Office).

Session Charter:

"The essence of risk management lies in maximizing the areas where we have some control over the outcome, while minimizing the areas where we have absolutely no control over the outcome and the linkage between cause and effect is hidden from us"—Peter Bernstein, *Against the Gods: The Remarkable Story of Risk*

This session goes to the heart of risk management in the context of HPC. How do traditional categories of risk mitigations apply specifically to HPC and the partnership nature of the platform acquisition? What's the difference between mitigations and contingency planning? Who owns the responsibilities for tracking, triggering mitigation actions, and deciding which strategies to use? How are unknowns, external risks and surprises factored in to the process? How does the ownership of activities change with the lifetime of the system?

Session Process and Discussion:

The first portion of the session was devoted to discussing how HPC risk management differs from other projects, and the terminology of risk management especially with respect to risk mitigation and contingency planning. Along with the standard Track 1 session questions, this group addressed three additional questions:

- What are the standard mitigation/contingency techniques for HPCCs?
- What triggers can be used for starting, boosting, or stopping a contingency or mitigation activity?
- What is the best way to gum up the risk management process?"

With regard to the third question, the group noted that there are two related behaviors that can compromise the risk management process: lower-level managers do not communicate up and down the tree, and upper-level managers ignore or accept risk inappropriately.

Breakout Question 1: "How well do the typical risks encounter in HPC fit the standard Risk Management models for risk identification?"

The group noted several differences in risk planning for HPC operations as contrasted to industry norms: facilities tend to be one-of-a-kind and unique; funding and staffing for projects tend to be incremental and unstable (as contrasted to fully funded projects in other areas); many stakeholders are able to influence design, methods, complexity; schedule risk in HPC is high because lifetime of a system is short; and HPC projects have a high visibility both inside and outside their home institutions.

Regarding terminology, the group held a spirited discussion around the differences between mitigations, contingency planning, and workarounds, and concluded that the term "contingency" was most often used to denote a financial category and should not be used loosely as a synonym for a workaround or an alternative course of action that would be used to avoid a risk.

The group also reviewed the work of Kaplan and Garrick[6], in which the naïve characterization of "Risk" was defined by the equation:

$$Risk = Hazard/Safeguard$$

This approach provided two different avenues into the discussion of HPCC risks. The group noted that this way of thinking about risk brings in the mitigation strategy, along with residual risk.

Finally, before moving into directed activities, the group was presented with several other thought-provoking questions:

- What triggers can be used for starting, boosting, or stopping a contingency or mitigation activity?
- What is the best way to gum up the risk management process?
- What fraction of the acquisition budget should be allocated to risk mitigation and contingency activities?
- What happens when risk mitigation activities become, themselves, risks?
- To what extent are risks and control techniques independent?

Breakout Question 2: "What are the key lessons learned in risk identification?"

The core of the session revolved around two rounds of intensive brainstorming and discussion. The first round centered on the question of "What risk categories, other than those provided by the conference organizers, apply to HPC systems?" While this exercise was a prelude to the focus on mitigation best practices, a number of other risks to HPCC systems were identified and are briefly summarized below.

- Parts/components/subcontractors: parts not compatible, supplier changes roadmap, and older components no longer available/vendor out of business/product changes
- General/long term: key personnel/experts leaves, and vendor goes out of business changes

[6] Kaplan and Garrick, "On the Quantitative Definition of Risk", 1981

- Contracts: disputes between contractor and subcontractor delay
- System integration: hardware, center facility not ready, power/cooling/other facility changes, and shake-out of system takes too long, becoming schedule risk
- Performance: backups take too long; system passes acceptance tests, but performance degrades over time; system lacks consistency; performance is poor, but one cannot assign performance problem to any system component (this is related to the complexity of the system)
- Usage and Applications: systems become unstable when all users/apps on system, system upgrades cause performance decrease, users do not adapt to new system/higher concurrencies, system workload changes/unanticipated applications, user demands (software or raw hours) exceed capabilities, too much/too little demand (system over or under utilized), and insufficient resources to help users port/improve applications

Breakout Question 3: "What are the best practice risk models for HPC in the area of risk identification?

The second round of brainstorming centered on the question of "What are possible best practices for risk mitigation and management?" Overall, the group of 17 participants suggested over 70 potential best practices. The group then spent the remaining time discussing, grouping, and categorizing the practices. Necessarily, the groupings are not unambiguous, so for instance, discussions about vendor relations may appear in more than one place.

Within the domain of project management, the group developed the following potential best practices: define authority to make decisions; employ a wide range of staff/users/expertise; involve a diverse group of people from RFP through acceptance and beyond; support frequent open collaborative/non-hostile review; document roles and authority in project plan; require risk handling status as part of the weekly executive-level project management status meetings; maintain a defendable budget; develop a risk management working group to share risks/opportunities/lessons learned across projects in the same portfolio/enterprise; obtain and maintain solid management support; and define triggers to initiate/boost/stop mitigation efforts.

With respect to risk management in general, the working group suggested the following as best practices: pull team of technical experts together to craft safeguard strategy; remember that not all risks are negative; transition build-time tests to operations test suite; use acceptance tests and friendly users to assess progress; keep some functions in house, such as facility mechanics/electricians, to maintain flexibility (it is risky to outsource everything); use an "observational approach" as opposed to complete site characterization; and integrate computer security concepts and personnel early in acquisition process.

Vendor relationships provided a rich source for suggestions: offer resources to vendor partners for system update testing at scale—vendors do not have resources for full-scale testing in-house; identify weak points in vendor-supplied software ahead of time and fill those gaps with in-house or third-party participants; bring in partners with unique expertise; and make vendors responsible for providing three years of spare parts independent of failure rates. Other ideas relating to vendors are discussed in subsequent paragraphs.

Contracts, procurement, and acceptance also provided many opportunities for discussion. The group suggested a number of possible best practices: proactively engage vendors; establish contractual performance requirements; allow vendor flexibility to exchange performance for schedule; consider advanced and milestone payments; have,

know, and understand plan B; use real applications for acceptance testing—Linpack is not useful! Other suggestions included: integrate performance measures over time; make performance on equal playing field with usability, effectiveness, consistency, and reliability; transfer/share risk to vendor; share development and share source code; locate vendor risk that the customer can easily mitigate; keep communication lines open with vendor/explore all options; understand vendor's risk and track in risk register; understand vendor's cost as best you can; ensure multiple vendors can participate in procurements and bids; define benchmarks appropriate to test system/factory test; purchase a maintenance plan from vendor; try to include several funding options specified in contract to mitigate funding changes; establish off-ramps in contract to share risk—either party can get out a certain points—no harm no foul; and mandate few mandatory requirements in RFP to reduce possibility of no successful bids.

Communications and sharing are essential to reducing risk. Suggestions included: establish frequent status updates/meeting with vendor; hold regular progress meetings with vendor site with both management and technical teams; talk not only with direct vendors but also subcontractors and be sure to explore all options; communicate with other centers; ideally, make problem reports with vendors public; exploit user group communication (although a proposal to host a Birds of a Feather session at SCXX was rejected); share and document lessons learned; and share software, including diagnostics.

Facilities and operations provided several ideas for long-term risk reduction: negotiate future operating costs if possible; consider hiring a commissioning agent as a third party; and use outside engineers, perhaps not available in house—they can serve as an independent party but might be cost-prohibitive.

User involvement is critical to ensuring a viable procurement. Suggestions in this area included: form teams of interdisciplinary scientists to guide the deployment; build science-driven systems (as opposed to architecture-driven systems); improve communication between users and centers; communicate with future users early and often; set expectations for areas such as multi-core, lower memory nodes, and system performance; encourage users to scale codes/try new strategies; map user applications to appropriate resources; and match users to platforms. These suggestions deal with both managing user expectations as well as gathering guidance from users to shape the platform.

There were four suggestions that did not easily fall into any of the above categories: strive for a balanced system; add more resources; cut back on scope; and track technology by using other agency/program resources to help. Take advantage of what other agencies are doing and the services that they provide.

Finally, a review of the identified risks and the proposed best practices uncovered several risks that were not directly addressed by any single proposed best practice and would warrant further discussion: applications may not be ready when the platform comes online, and users will not move to new systems; and key people depart and go to other centers.

Several general strategies became apparent by the conclusion of the session: maintain a diversity of systems; use parallel development to reduce risk; use outside experts; and work closely with end-users and even more closely with vendors.

Breakout Question 4: "What opportunities exist to share or improve practices and terminology?"

As a whole, the group offered the ideas developed in response to Question 3 as content for sharing to a wider community.

Track 2: Real Life Risk Experience

Session 1: From Vision to Contract: "So you want to buy a mega-WHAT?!"

Session Leaders: Candace Culhane, Jon Stearley

Participants: Candace Culhane, DOD (lead); Jon Stearley, SNL (lead); Steve Louis, LLNL (note taker); Aaron Andersen, NCAR; James Ang, SNL; Katerina Antypas, LBNL; Paul Cook, SGI; Pam Hamilton, LLNL; Dale Knutson, PNNL; David Morton, SGI; José Muñoz, NSF; Kevin Regimbal, PNNL

Session Charter:

Complying with all the rules, regulations, and paperwork associated with big-ticket acquisition is not for the fainthearted. All acquisitions start with the requirement phase, and successful ones lead to contract signature. This session will cover the many steps in between, including budget approval, technology survey, market survey, competitive procurement, request for proposal, proposal evaluation, source selection, and the legal review gauntlet. It is vitally important for risks (both technical and business related) to be identified and analyzed as early in the project as possible to gain approval and to build in contingencies. A failure point for large projects is when risks are not considered up front.

Session Process and Discussion:

Four top risk areas were identified:

- Problematic RFP documents
- Sluggish or overlong procurement processes
- Overestimated level of stakeholder "buy-in"
- Unplanned technology change between bid and award

Each of these risks is described in more detail below. Additional risk areas were also identified during the Session 1 breakout discussion and are listed at the end of this section.

Request for Proposal Risks

Inaccuracies, omissions, and over-aggressiveness in the RFP represent major risk. For example, forgetting to include key user requirements or necessary facility requirements can be devastating to the vendor response and evaluation process. RFP benchmarks not accurately representing the real user workload are another risk example. An RFP can be too aggressive in describing the desired schedule, technology, customization, or cost. The misalignment of requested technology versus available funding is also a key risk. The bottom line is that the acquired system might end up being built incorrectly and, therefore, not fulfill its intended purpose.

Mitigation strategies: Successful mitigation strategies to address the above risks include adequate consultation with vendors, users, and computer center staff to understand vendor technology roadmaps, specific programmatic requirements, and site facility requirements. Distributing one or more draft RFPs for comment has proven useful in mitigating RFP risk. Another practice is to set risk levels and associated strategies to be commensurate with the class of computer system in question (such as, advanced architecture, capability platform, or capacity platform). The ability to fund technology development through additional vehicles not tied directly to the platform build contract

can also help as a mitigation strategy. Examples are the DoD High Productivity Computing Systems Program (HPCS) and NNSA PathForward programs. The identification and development of accurate benchmarking kernels and system tests is also an important mitigation strategy.

Management strategies: Successful management techniques to address RFP risks varied according to federal agency. Some agencies recommended pairing experienced staff with less experienced staff for mentoring purposes. Revising and iterating on multiple RFP drafts based on internal, external, and peer review comments was suggested, where allowed. There was concern that an overly heavyweight earned value management system (EVMS) project management process could easily derail the best of intentions. More lightweight and more appropriately tailored methods for project oversight at the right level of rigor should be given adequate thought.

Gaps: The primary "gap" identified for RFP risk was the inability to define or observe any kind of practical "fast-track" mechanism for easier incremental procurement of HPC systems (referred to in this session colloquially as *supercomputing purchases "by the yard"*).

Procurement Schedule Risks:

The acquisition process may be become unnecessarily extended or delayed much longer than anticipated for many reasons, including: the involvement of less inexperienced staff in the procurement; too many programmatic, project, or legal reviews; and potentially lengthy vendor protests. Schedule delays present an unwanted hazard for cost increases not desirable for either the buyer or the bidder.

Mitigation strategies: Successful mitigation strategies to address schedule risk can include well-designed technology surveys or requests for information (RFIs). The ability to create draft RFPs, as described above, can also help mitigate risk. A common lament from the agencies present in this session was the seemingly inevitable practice of scheduling many critical procurement and contract decisions near the major end-of-calendar-year holidays.

Management strategies: Successful management techniques include following an "adapt and reuse" philosophy wherever possible (this can be applied to RFPs, benchmarks, review teams, and even lawyers). Some agencies have created dedicated HPC procurement officers/staff and acquisition processes to streamline and accelerate schedules. Of course, this usually comes with extra overhead costs. Pre-briefing the key decision makers can also help eliminate or reduce schedule risk. Again, the use of adequately tailored project management oversight methods can help (for example, appropriately light-weight EVMS).

Gaps: A key gap associated with this risk is an inability to affect legal negotiations between government and vendors when they are harder and more complex than they should be. Some suggestions for closing other significant gaps are more effective escalation processes for key contract approvals; an easy-to-use, fast-track, sole source authority, where appropriate; and the inclusion of influential HPC advocates throughout the procurement decision chain.

Overestimating Stakeholder Buy-In:

This risk can take the form of overestimated end-user buy-in, overestimated laboratory and HQ management commitment levels, as well as uncontrollable congressional funding fluctuations that are outside the control of the programmatic management structure. These risks have consequence for both the buyer and the bidder.

Mitigation strategies: Successful mitigation strategies are numerous and will vary according to the stakeholder. For users, it will be helpful to understand the diversity of

user needs and priorities, establish more accurate user expectations when their needs and priorities do not necessarily match reality, and to provide incentives for those users to gear up for new technologies and capabilities, especially when the provided capabilities do not perfectly match the expectations. For laboratory and HQ management, it will be helpful to hold more pre-briefs for those decision makers to help them understand the wider impact of the acquisition and procurement process. Mitigation strategies for congressional funding fluctuations are much harder to come by.

Management strategies: Successful management techniques include the practice of developing and presenting compelling arguments that the solutions proposed are indeed meeting the identified needs. The ability to effectively communicate strategic plans to stakeholders is also a useful technique. Tailoring the strategy and arguments to effectively communicate the distinguishing value to stakeholders is critical.

Gaps: Gaps identified were the inability to control congressional appropriations language, and the realization that a broad external community consensus (for example, for the presumed technical approach) does not always guarantee sufficient buy-in from stakeholders.

Risk of Technology Changes:

This is defined as the risk of significant, and perhaps unplanned, technological change between the time of the bid and the time of the final contract signing. This risk can manifest itself in decisions to change the underlying technology away from what was originally bid to what can actually be delivered (and either forced upon vendors by elimination of subcontracted component availability or chosen by the vendor after recognizing that the bid solution is likely to be inadequate). This risk affects both the buyer and the bidder.

Mitigation strategies: Successful mitigation strategies are for both the buyer and bidder to stay abreast of supplier technology roadmaps, along with flexible, less constraining RFP language to allow for technology refreshes (or substitutions). Strong and clear acceptance criteria in the contractual language can also help mitigate the occurrence or impact of this risk.

Management strategies: The best management technique from session participant perspectives was to maintain good ongoing buyer/bidder relationships, including close communication and interaction between the time of the bid and the contract signing.

Gaps: An identified gap related to this risk is a decreasing ability for the HPC community to influence technology trends due to the ever-increasing size of commodity marketplace. Continuing to work with HPC vendors to stress the importance of the highest end systems may help but cannot guarantee success.

Other Risks:

Several other risks were discussed. Since guidelines to the session chairs stressed the identification of the top three to five risks, the risks below, while still important, were not analyzed in as much detail as the identified top four risks previously discussed.

- Overestimation of internal computing center or external vendor capabilities
- Protests: losing bidder protests, resulting in long (and painful) re-evaluation
- Lack of qualified reviewers and users due to time and non-disclosure agreement requirements
- Inexperience or ignorance of full acquisition process, including approval process (experienced people mentor inexperienced people on purchases)

- Inappropriate sharing of information; this is low probability but high impact: possible protest and re-bid process may be the result (use strict non-disclosure agreements)
- Vendor mistakes on export control issues (non-compliance on requirements)
- Market survey questions may prolong the pre-contract phase and RFP release
- Acquisition staff protection from unreasonable bidding period extensions (last-minute and/or unreasonable vendor requests for bidding extension)

Other comments/findings:

Session attendees thought it was important to document past performance of bidders for situations where the capabilities of a vendor to successfully execute might be open to serious question (past performance could outweigh bid price). A related concern was the need for early, efficient mechanisms to screen out unqualified bidders while still satisfying fairness requirements. An unqualified bidder was deemed to be a vendor that could not possibly develop, build, or deliver the system as described in the RFP.

Session 2: Management of System R&D from Contract Award through the Build: "Moore's Law meets Murphy's Law"

Session Leaders: Ann Baker, Manuel Vigil

Participants:

Manuel Vigil, LANL (lead); Ann Baker, ORNL (lead); Matt Leininger, LLNL (note taker); Cristina Beldica, NCSA; James D'Aoust, SDSC; Vince Dattoria, SC/ASCR/DOE HQ; Jim Foster, TACC; Mark Seager, LLNL; Francesca Verdier, LBNL; and Mary Zosel, LLNL.

Session Charter:

Leading-edge systems often have a long lead-time and require a significant amount of new technology to be developed and deployed. This session will address how risks are identified, tracked, and managed during this phase when the system is still in the vendor's hands. What are the major risks in this phase of the acquisition and who owns them? This session will cover real-life experiences and examples regarding risks, including tips and suggestions on how to prepare for situations that will occur, usually at inopportune times. Lessons learned may help with managing risks. (Yes, there will be surprises.)

Session Process and Discussion:

First, the group identified and defined the top four risks, in the following order:

- Encountering technical and/or business problems by the vendor. These problems could be experienced in a variety of areas, such as hardware/software development, unknown technology, changing roadmap, performance not meeting requirements, component pricing variability, parts availability and schedule issues, vendor strength/breadth/depth, and integration issues.
- Schedule deviation caused by general slippages, something better coming along, and facility availability issues.

- The system encountering scaling issues. Examples include inability for early scaling testing, software does not scale as expected, usability at scale, original equipment manufacturer (OEM) components that do not function at scale, and issues with applications scaling.
- Funding turbulence.

Risk mitigation and management techniques: Once the risks were defined, the next task was to think about the most effective risk management/mitigation techniques to address these risks. One technique is to trust and verify, checking with the vendors to confirm the progress, status, and all pertinent information. Another important technique is that of formal risk monitoring, including ranking of risks by likelihood and impact and defining mitigations for those risks with higher consequence and probability, actively tracking the top 10–20 risks. This also includes receiving, approving, and monitoring the vendor's risk plan. Frequent vendor communications and monitoring is critical to project success, as is regular communication between customer and vendor executives and high-level executive buy-in from both sides. Early delivery of a test system to perform internal testing is key, but building and testing at the factory at scale is an even better solution. Pre-delivery testing to include applications testing serves to mitigate risk. Interim option execution and interim go/no-go decision points, both in the RFP and in the SOW, were suggested as risk techniques that work well. Sharing of information among HPC sites, allowing visibility of progress/failure, and sharing of solutions/problems among the sites, are also good risk management and mitigation techniques.

Standard risk management techniques: The next question is "For the top risks, how well does risk management practice fit?" Session participants believed that not all funding profiles allow for funding contingency, but that most other parts of the standard process work as tools for mitigation. Schedule or scope contingency should be used where applicable. The DOE HQ management-reporting model hinders programmatic flexibility and that increases risk, for example, get-well plans that must be done for early success. The government bureaucracy can insert time delay into processes that must occur quickly.

Risk management key lessons: There were several key lessons to be shared from this. HPCCs must be willing to pull the trigger of their risk-mitigation strategy on the date as planned. They must establish the vendor partnership early and explicitly share the risk appropriately. The build phase will be easier if technology risk is reduced by taking advantage of things such as incremental deliveries, open source, and commodity parts. Software stability and scaling testing always get squeezed, so HPCCs must guard against this vigorously. In addition, communicate effectively and often, and prepare applications and end users for change.

There were also a few risk management success stories shared around the table. The Roadrunner technical assessment positive review was a result of careful project planning and risk management. ASCI Blue Pacific came in three months early and achieved 20 percent over the performance requirement as a result of careful risk management. One site stated that they did not execute an option when they realized the risk was too high. A financial risk was managed by changing node architectures for ASC Purple, delivering on-time and meeting programmatic requirements for significantly less money. LBNL/NERSC used risk management processes to implement Compute Node Linux nine months earlier than scheduled.

Gaps: Finally, session participants discussed the areas of HPCC risk management techniques with the most opportunity for improvement. The overall acquisition and deployment process for large HPC systems is a high-risk area and needs to be managed

accordingly. However, many of the processes imposed do not allow agility. EVMS and other processes are considered a hindrance and increase the risk of failure. They can insert delays at critical decision points that need to be made quickly.

Session 3: Acceptance Testing and Integration: "How to get the system installed and stay out of jail."

Session Leaders: Brent Draney, Gary Skouson

Participants:

Brent Draney, LBNL (lead); Gary Skouson PNNL (lead); Douglas Tinnin, ANL (note taker); Thomas Bettge, NCAR; Brad Blaine, HP; Tina Butler, LBNL; Susan Coghlan, ANL; James Craw, LBNL; Kimberly, Cupps, LLNL; Bryan Embry, DOD; David Featherman, BAH; James Kasdorf, PSC; Ricky Kendall, ORNL; William Kramer, LBNL; Gary Mack, LBNL; Charlie McMahon, LSU; Randal Rheinheimer, LANL; Stephen Scherr, DOD, Mike Showerman, NCSA; Kevin Stelljes, Cray.

Session Charter:

If you are installing a leading-edge system, by definition it is unlikely that you have had any chance to "road test" the final product before this point, and there are bound to be engineering issues remaining from the design and development phase. This session will address how to make sure that you get something you can accept while keeping in mind the requirements of the contract and the SOW. How can you effectively test a system that is being built and the vendor is asking you to pay an invoice? What tests/benchmarks can you effectively run at the factory before delivery? What tests/benchmarks can be run on a small phase 1 system? What benchmarks have to wait for the fully integrated system at your site? How do you develop benchmarks that are meaningful for your expected workload? What happens when proposed system software or hardware does not materialize or does not meet the requirements in the SOW? We will discuss real-life experiences with what to do with conflicts between reality and the contract requirements as we try to get a system installed and running.

Session Process and Discussion:

There are many risks associated with fielding HPC systems that are, by their very nature, extremely complex. These computational systems are the antithesis to the Keep It Simple Sweetie (KISS) model of reducing risk. As a consequence, these systems will often fail in complex and unexpected ways, but the broad categories of these failures are well understood. Most of these risks are identifiable in the vision-to-contract and contract-to-build phases of procurements and are significant enough to require mitigation. The primary purpose of acceptance testing is to determine what risks identified in previous phases have actually occurred so that the mitigations can be implemented. Conversely, if a risk can be accepted, then there is little benefit to having it clarified with an acceptance test. The chances of other risks may be greatly lowered if functionality can be clearly demonstrated.

There are, however, a small number of risks specific to the window-of-acceptance testing and integration. The session participants discussed both the set of risks identified with acceptance testing and their mitigations, as well as risks specific to acceptance testing and integration.

Identified Top Risks:

The following are the top risk categories that may occur *prior to the time of acceptance testing*. Design of the acceptance testing can be a key mitigation strategy or a way to identify if a problem has actually occurred and if an appropriate action may be needed. A good rule of thumb is that almost every risk that is germane to HPC can be followed with "at scale":

- Scalability. The number-one risk identified in the session and backed up through later straw polls is scalability. Scalability risk takes many forms and include hardware, operating system, application, and reliability. The importance of this risk warrants a further breakdown:
    - Hardware scalability problems frequently seem to be associated with the communication interconnect but may also be caused by other in-node components such as memory and input/output bandwidth.
    - Operating system scalability most often shows up in communication (message passing interface) buffer management, associated libraries, and compilers ability to generate efficient programs from portable code.
    - Application scalability is most often determined by the order of algorithms and communication patterns used in the science domains. Often, the next bottleneck in application scalability on a given system is not known until previous bottlenecks are removed.
    - Reliability scalability is a major issue due to the highly parallel nature of the applications run on HPC systems. What would normally be a highly redundant system will have a single application fail if any of thousands of individual nodes fail. This greatly increases the risks related to component failure and can significantly reduce the reliability of the system. A large HPC system can be the most effective parts reliability tester and the aggregate component test time may significantly exceed what a manufacturer has performed.
- Component risks. The group identified that critical components supplied by subcontractors have both schedule and technical risks that may get overlooked. Central processing unit delivery schedule and power supply holdup time are two examples where components have adversely impacted projects. These risks cannot be mitigated by acceptance tests. These types of risks are independent of the quantity of the components used in the HPC systems; hardware scaling reliability is discussed in a later section.
- Integration risks. HPC systems are rarely (if ever) made completely from components manufactured by one company, and any company large enough to do so will have similar integration issues itself from the diverse divisions contributing to the system. The primary vendor often takes on the role of an integrator and the risks associated with the role. In many cases, the full-system integration is performed at the customer site.

Acceptance Testing and Integration Risks

These risks are specific to acceptance testing and integration. They are either only relevant to this time period or associated with integrating the HPC system and, therefore, when identified, must be resolved before proceeding. The following are the top three risks identified:

- Customer-related risks. Customer issues and delays, such as site infrastructure (seismic, power, and cooling), interfacing to existing networks, and effectively

mounting external file systems, are common risks that need to be addressed. Requiring HPC systems to operate at the extreme ends of the ASHRAE standards may also pose significant risks to reliability. Finally, the customer may not be ready to receive the HPC system when the vendor is ready to deliver it, and this most often happens over a holiday shutdown.

- Vendor-related risks. Vendor issues and delays can include shortages of critical parts to both build the system and to cover infant mortality. Vendors may not have adequate support staff available during integration and availability testing, or support staff availability may be highly variable.
- Ending acceptance. Overly aggressive acceptance schedules may not allow sufficient time to resolve all issues identified with acceptance testing, but the acceptance time period cannot be indefinite either. Acceptance tests may not clearly cover the risks they are designed to identify and can either lead to ambiguity or a misclassification of a risk's probability.
- Hidden problems. Additionally, small problems associated with risks that are low or easily mitigated may hide large problems that are of significant consequence.

Risk Mitigation and Management Strategies: The most effective mitigation strategy for the risks discussed above is an early and thorough acceptance test of the HPC system. Some key strategies are to procure an early test system, perform factory tests on as large a system as possible, use a phased plan for installation, and to focus on the largest scale possible with a wide variety of testing software that closely models the expected workload.

- **Early Test System:** A test system has been critical to many sites so that on-site staff can work on an effective integration plan and resolve local compatibility, configuration, file system, and environmental issues before delivery. After delivery, a test system may be used to replicate problems in a controlled environment and test upgrades and new features before moving them to production HPC systems.
- **Factory Tests:** Vendors normally perform factory tests on small systems, but it is crucial to include large-scale (as large as possible) factory tests with benchmarks that mirror the expected workload. Most of the critical problems appear at scale with a real workload and not synthetic tests (Linpack).
- **Phased Installation:** Using a phased installation plan allows the system to be scaled in a controlled manner closer to the factory testing footprint. In the ideal case, the largest factory tested blocks are retested at the customer site and then integrated after verification. This reduces the chances of confounding problems and focuses troubleshooting around integration and scaling. After each phase is integrated, focus is on running benchmarks at the largest scale possible to detect problems not previously seen. Using as wide a set of benchmarks as necessary to cover the expected workload is also necessary.

Talented Staff: Hiring and cultivating talented staff is critical to a successful HPC system procurement. The problems most vexing are the ones no one has experienced before. There are likely to be many issues that arise, and it is critical to divide scaling and other problems among many people for better focus on individual issues.

Session 4: Managing HPC Business Risks: "Herding cats and dollars" or "Where DOES the buck stop?"

Session Leaders: Ira Goldberg, Rob Pennington

Participants:

Ira Goldberg, ANL (lead); Rob Pennington, NCSA (lead); Bob Tomlinson, LANL (note taker;) Robert Ballance, SNL; Kathlyn Boudwin, ORNL; Patricia Kovatch, UTenn/NICS; Tommy Minyard, TACC; Jonathan Monsein, DOD; Terri Quinn, LLNL; Sohel Saiyed, IBM; Yukiko Sekine, SC/ASCR/DOE HQ; Francesca Verdier, LBNL; Warren Yip, DOE Site Office

Session Charter:

Business risks are numerous, critical, and challenging. How are vendor interdependencies best dealt with (including equipment, software, schedules, and failure rates)? How to ensure that you can actually house the systems, meet security plans, and afford to operate the systems? How to estimate and plan for lifetime issues like maintenance, staffing, cooling, and electrical costs? How are other parts of the infrastructure impacted, such as data archives and networks? We will discuss how to think about, capture, and measure major business risks... so there is a next time.

Session Process and Discussion:

The nature of risk and how it is framed are essential to the discussion of specifics. How does one decide which risks to consider and which not? How does one prioritize? People who study human behavior have learned and continue to learn about how humans make decisions and inject biases, and under what circumstances people are more prone to make errors.

The types of business risks considered were natural environment, business environment, human processes, products, financial markets, and operations. A business risk was considered to be anything that could interrupt the flow of the project. These risks can be internal to the project, external to the project but within the organization, or external to the organization. They may or may not involve interdependencies.

Environmental

Environmental risks arise from both natural and man-made causes. Examples of natural hazards are earthquakes, tornados, and lightning strikes. Man-made hazards considered were fires, terrorist attacks, and cyber attacks. Participants generally agreed that, at some level, environmental risks are adequately addressed with existing plans. Environmental risks are generally considered at the local (site) level. Continuity of the scientific programs is not considered unless it is mission critical.

Financial

The group identified two sources of financial risks. First, the project may end up with a funding profile that does not match the project's spending plan—a cash flow problem. Second, the vendor may find itself in financial trouble, thus threatening its ability to meet contractual obligations. Uneven cash flow situations can be dealt with non-standard leases ("lumpy lease payments"). In these leases, the payments are not regular. Another method of dealing with uneven cash flows is with cash reserves (if any exist) and spending adjustments. Another strategy would be to renegotiate the contract to align the payments with the funding profile.

In the HPC community, vendor viability has been a concern as HPCCs have experienced the downfall of several HPC vendors over the years. A central reason these vendors fail

is they have problems managing cash flow. When the vendor does not have much cash reserve, their financial position can be improved with the use of partial payments instead of one final payment at computer acceptance. This can be accomplished by negotiating the contract milestone payments. Another possible mitigation strategy is for the community to work together to assure that the few HPCC vendors in business remain healthy business enterprises.

Human processes

Human process risks involve management support (primarily upper management), maintaining an effective project team, and good communications and expectations. Human processes involve a number of issues such as complexity, judgment, relationships, bias, and uncertainty.

The importance of these risks was illustrated during the incident at Three Mile Island. The report[7] concluded that, "The subsequent investigations and lawsuits disclosed a seemingly endless story of incompetence, dishonesty, and cover-ups before, during, and after the event." When the performance is examined in light of other accidents, "The performance of all concerned was about average." Complexity increases potential for error—"*any* part of a system might interact with other parts in unanticipated ways" (probabilities need not be independent).[8]

Human process risks include the embedded issue of uncertainty. In the face of uncertainty, people still make judgments (and create a mental model). There is no guarantee that the model, composed in an environment of uncertainty, will actually be congruent with the real world.

It is possible to lessen these risks by maintaining an effective project team, improving communication, and setting and communicating clear expectations. Management support can be improved with good communications and making sure the project is visible to the right decision makers. Maintaining an effective project team entails attention to professional development and executing a good succession plan.

Products

Leading-edge systems make use of the newest technology. It is often true that the products that were developed using this technology are often first deployed in the HPC systems. Many factors can influence risks associated with these products. Examples include materials, technology, design, production, tolerances, and maintenance. Each can involve tradeoffs with the others. Possible risks associated with the use of the latest products are problems with quality control; competition for the product, which may be in high demand; and designing and deploying infrastructure for the leading-edge systems that use these untried products.

Managing quality control in "bleeding-edge" projects can be improved by using a step-wise process. It is important to get as much information as early as possible, for example, early system access and testing plans. It is also useful to obtain agreements to run on other sites' systems and become more involved with the vendor development process.

Competition for scarce "bleeding-edge" products can impact projects, but the group did not have a good solution.

---

[7] Normal Accidents: Living with High-Risk Technologies, Charles Perrow, Princeton University Press, 1999, p. 16.

[8] Ibid., p. 21.

Creating infrastructure for the next "bleeding-edge" system can best be accomplished with improved planning, including capacity planning.

There was general concern regarding reliability of various vendors, including components and vendors inability to test at scale. Vendors could also move to other countries, be acquired, or go out of business. There are additional vendor risks associated with contracts including types of contracts, sources, and change management.

Financial Markets

Financial markets have become a larger source of risk than in the past due to the size (expense) of the machines and fluctuations in interest and electricity rates. The interest rate markets are influenced by default risk and market risk. Within the market, the supply and demand for loanable funds determine the real interest rate. Demand and supply of funds are influenced by many factors, including economic strength, inflation, and future expectations. Current experience shows that credit markets can melt and impact borrowing availability and rates. Electricity rates can also vary significantly and change power costs dramatically for HPCs. Some electric rates have varied up to 60 percent over the last year.

The electric rate risk is ongoing, so it is useful to model and check electricity rate variation regularly. Reserves can be used to mitigate. Fluctuations in utility rates can be addressed through long-term contracts, financial market contracts, and improved development/implementation of greener machine rooms.

Interest rates can vary significantly and impact financing costs. Average interest rates have varied by more than a factor of two for the five-year Treasury bonds over the last year. Average commercial paper rates have varied over three-hundred basis points (three percent) during the last year, and spreads (the difference between Treasury rates and market rates paid) have also varied dramatically. The significance of the potential changes means it is useful to model and check interest rate variation regularly prior to locking rates involving large lease principals. Financial market contracts and locking rates early will reduce interest rate risk. The risk is retired when rates are locked.

Competition on interest rates can be increased and finance costs potentially lessened by separating the purchasing and financing decisions. Money is a commodity and competition is good, allowing for savings from shorter lease terms and lower interest rates. LBNL/NERSC and Argonne National Laboratory (ANL) split the financing decision from the machine purchase and shopped for financing separately through third-party leasers. For ANL, competition and declining rates ended up saving approximately nine-million dollars in interest over the terms of the first generation Argonne Leadership Computing Facility, or ALCF-1, leases.

Operational

Operations involve risks in planning for and meeting operational expectations of the delivered system. Operational risks can include a number of issues such as the transition process, operational planning, long-term funding, customer support, and appropriate staffing. There are also risks associated with facility upgrades and deliveries aligning (across multiple programs/agencies) with the end of the government fiscal year.

The critical decision process, all four steps, as a formal contract initiation/phasing mechanism, can be used to reduce funding risk.

Operating risks can be reduced by improved planning for and meeting of operational expectations of the delivered system.

Appendix C. Analysis of Workshop Questionnaires

In the workshop questionnaire, each attendee had the option of casting eight votes from a list of breakout findings. The analysis of these votes grouped into the themes presented below for Track 1 and Track 2.

Track 1 Best Practices Finding (Based on Voting)

Develop a prioritized risk register with special attention to the top risks.

21	*Develop a Risk Register*
21	*Watch list – top 10 risks*
2	*Track risks continuously; risks are not static*
15	*Define triggers to initiate/boost/stop mitigation efforts*
11	*Use milestones to define trigger/decision points to start/stop risk mitigation*
14	*Require risk-handling status as part of the weekly executive level PM status meetings*

Establish a practice of regular meetings and status updates with the platform partner.

35	*Status updates / meeting with vendor*
14	*Require risk-handling status as part of the weekly executive level PM status meetings*
10	*Talk not only with direct vendors, but subcontractors and explore all options*
18	*Obtain and maintain solid management support*

Support regular, open reviews that engage the interests and expertise of a wide range of staff and stakeholders.

24	*Support frequent open collaborative/ non-hostile reviews*
23	*Employ a wide range of staff/Stakeholders/users/expertise*
18	*Obtain and maintain solid management support*

Document and share the acquisition/build/deployment experience.

20	*Communicate with other Centers*
19	*Share and document lessons learned (e.g., Knowledgebase, post-mortems)*
17	*Lessons learned/Postmortems are helpful*
6	*Share software, including diagnostics*

Track 2 – Top Risk Categories (Derived from Voting)

System Scaling Issues

- 25 *SYSTEM ENCOUNTERS SCALING ISSUES [Inability for early scaling testing, Software does not scale as expected, Usability at scale, OEM provisions don't function at scale, Applications scaling]*
- 22 *SCALABILITY – OS, Hardware, Applications, etc.*
- 8 *Integration of multiply vendor pieces at scale*
- 13 *Track 1 Best Practice mitigation vote that fits with this risk category: Offer resources to vendor partners for system update testing at scale; vendors do not have resources for full scale testing in house.*

RFP/Contract and Acceptance Testing

- 22 *RFP RISKS: Inaccuracies, omissions, and over-aggressiveness in the RFP (missing key user or facility requirements). Benchmarks do not accurately represent workload. [buyer] RFP too aggressive in schedule, technology, customization, or cost. [buyer] / Misalignment of technology vs. the dollars*
- 7 *Overly aggressive acceptance schedule*
- 7 *Small problems may hide large problems*
- 2 *Track 1 Best Practice mitigation vote that fits with this risk category: Offer resources to vendor partners for system update testing at scale; vendors do not have resources for full scale testing in house.*
- 3 *Vendor responsibilities of non-vendor supplied pieces during acceptance*

The Vendor Encounters Technical or Business Problems

- 26 *VENDOR ENCOUNTERS TECHNICAL/BUSINESS PROBLEMS [HW/SW development, Unknown technology, Changing roadmap, Performance does not meet requirements, Component pricing variability, Parts availability and schedule issues, Vendor strength, breadth, and depth, Integration issues]*

Personnel Staffing and Interactions

- 21 *HUMAN [Management support, Maintaining effective project team, Communications and expectations]*

Project Schedule

13 *If there is a SCHEDULE DEVIATION [Slippages, Something better comes along, Facility availability]*

11 *SLUGGISH PROCUREMENT PROCESS: If the acquisition process goes too long (inexperienced people, program, project, and legal reviews, protests…), opportunity cost increases [buyer, bidder]. Tech survey, RFI, Draft, …*

5 *TECHNOLOGY CHANGES BETWEEN BID AND CONTRACT SIGN [buyer, bidder]*

1 *Schedule delay by customer of system delivery*

Sponsor Commitment

11 *If there is FUNDING TURBULENCE*

7 *OVERESTIMATE STAKEHOLDER BUY-IN: management commitment levels. [buyer, bidder] e.g., stakeholders: HQ, Lab, Vendor, Users, Congress / Funding fluctuations outside the control of the program manager*

Facilities and Operations

6 *OPERATIONS [Planning for and meeting operational expectations of the delivered system]*

4 *Site infrastructure and interfacing of the systems*

6 *PRODUCTS AND SERVICES [Managing quality control in bleeding edge projects Competition for bleeding edge components. Creating infrastructure for the next bleeding edge system.]*

1 *Vendor support during integration at scale is highly variable*

2 *Operating at ASHRAE (air conditioning) standards*

1 *FINANCIAL. [Cash flow for the project. Vendor viability]*

2 *FINANCIAL MARKETS [Fluctuations in utility rates, Fluctuations in financial markets]*

Appendix D. Workshop Attendees

Aaron Andersen, NCAR
James Ang, SNL
Katerina Antypas, LBNL
Ann Baker, ORNL
Robert Ballance, SNL
Cristin, Beldica, NCSA
Thomas Bettge, NCAR
Brad Blaine, HP
Kathlyn Boudwin, ORNL
Tina Butler, LBNL
Susan Coghlan, ANL
Paul Cook, SGI
James Craw, LBNL
Marty Crawley, HP
Candace Culhane, DoD
Kimberly Cupps, LLNL
James D'Aoust, SDSC
Vince Dattoria, SC/ASCR/DOE HQ
Brent Draney, LBNL
Bryan Embry, DOD
David Featherman, BAH
Jim Foster, TACC
Ira Goldberg, ANL
Brent Gorda, LLNL
Pam Hamilton, LLNL
Daniel Hitchcock, SC/ASCR/DOE HQ
James Kasdorf, PSC
Ricky Kendall, ORNL
Dale Knutson, PNNL
William Kramer, LBNL
Patricia Kovatch, UTenn/NICS
Sander Lee, NNSA/ASC/DOE HQ
Matt Leininger, LLNL
Steve Louis, LLNL
Gary Mack, LBNL
Michel McCoy, LLNL
Thomas McKenna, PNNL
Charlie McMahon, LSU
Steve Meacham, NSF, HQ
Tommy Minyard, TACC
Jonathan Monsein, DoD
José Muñoz, NSF, HQ
David Morton, SGI
Robert Pennington, NCSA
Terri, Quinn, LLNL
Kevin, Regimbal, PNNL
Randal, Rheinheimer, LANL
Sohel, Saiyed, IBM
Stephen, Scherr, DoD
Barry Schneider, NSF, HQ
Mark Seager, LLNL
Yukiko Sekine, SC/ASCR/DOE HQ
Mike Showerman, NCSA
Gary Skouson, PNNL
Jon Stearley, SNL
Kevin Stelljes, Cray
Douglas Tinnin, ANL
Robert Tomlinson, LANL
Francesca Verdier, LBNL
Manuel Vigil, LANL
Warren Yip, DOE-Berkeley Site Ofc
Mary Zosel, LLNL

Administrative Support

Tracy Berkich, LLNL

Laura Farro, LLNL

Lori McDowell, LLNL