A Perspective on Risk Identification and Management for Complex Technical Projects

Scientific Computing Workshop
San Francisco, California
September 17-18, 2008

Dale Knutson, Division Director
Strategic Projects Division

PNNL-SA-62067

Pacific Northwest
NATIONAL LABORATORY

How Risk is Defined

- Classic definition of risk based on a theoretical model capturing the probabilistic evaluation of consequences and likelihood of occurrence
  - Risk = Probability x Consequence: Formally used to support
    - Safety Analysis
    - Engineered systems design evaluation
    - Formalized management tools supporting quantitative risk determination:
      - Results in “Tornado Diagrams,” “Confidence Curves,” and pictures that illustrate thresholds of “acceptable risk” for technical disciplines that lend themselves to such use

- What about projects that don’t lend themselves to the theoretical model?
  - Projects based on discovery
  - Projects based on assumptions of technical performance
  - Projects based on the knowledge of single individuals or very small teams?

Pacific Northwest
NATIONAL LABORATORY

How Risk is Defined

- How do we use practical methods to learn enough about project risk – and appropriate mitigation – when we don't have enough information to satisfy traditional quantitative methods?

  - Transform questions from **understanding consequences** into terms that challenge assumptions of performance:
    - *"If the assumed technical performance end state is not achieved (or achieved only partially), what does that do to the objectives of the project?"*
    - *"Is a transformational discovery close enough to reality that we should be taking it into account? If so, what consequences do we envision technically, financially, schedule?"*
    - *"Which element's of the project depends on a single individuals success? What if they leave?"*
    - *"Have we ever bought anything like this before? Are special approvals necessary to pull this off?"*

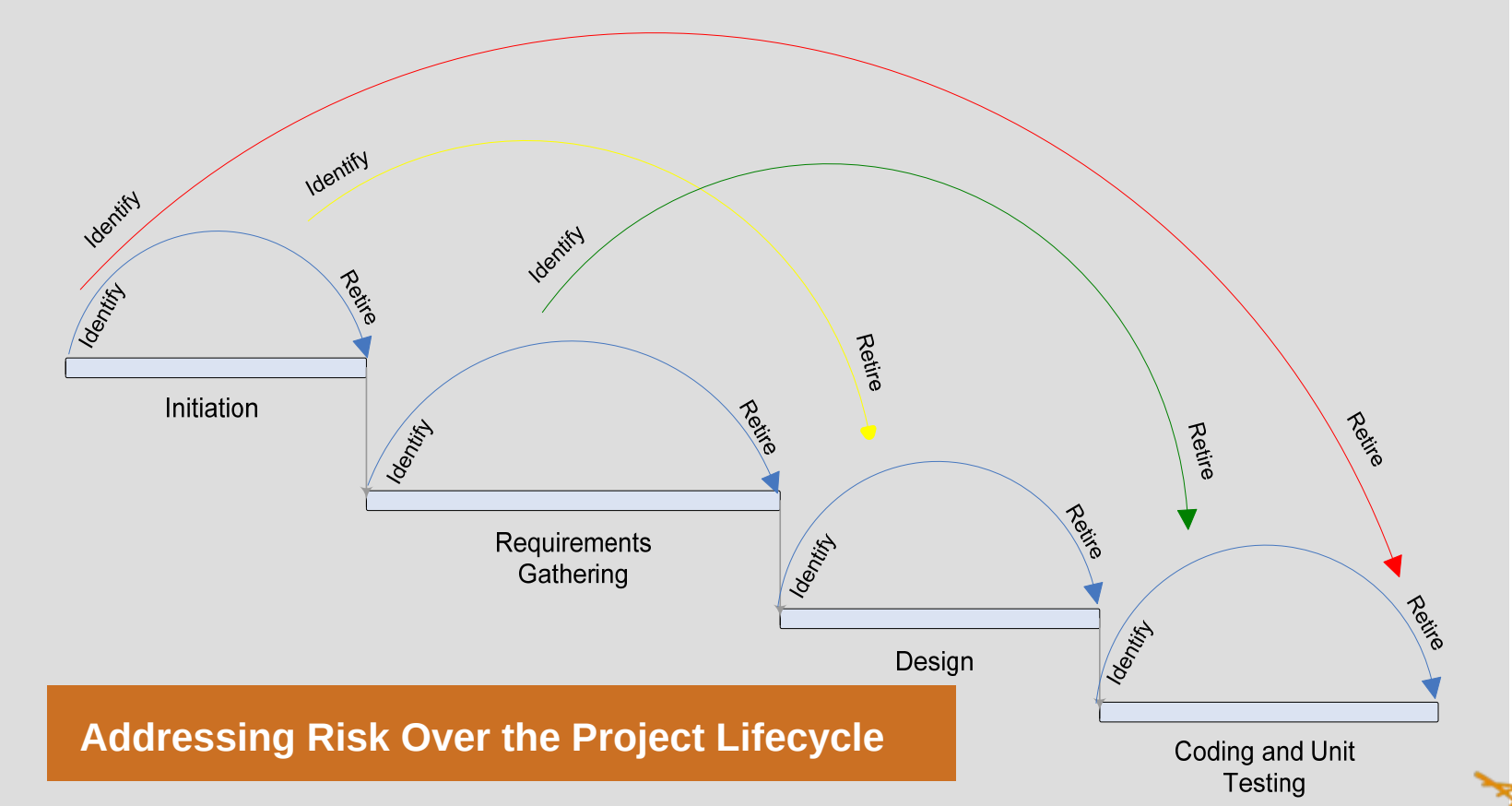
  - Transform questions from **understanding probability** into terms that question our sense of likelihood:
    - "*What is the likelihood that commercially available technology will get to the level we need?"*
    - *"What is the likelihood that investments in a development program will validate the technical performance assumptions?"*
    - *"What is the likelihood that a transformational discovery will occur to fundamentally change our technical approach?'*

  - Broaden question set to include the **integrated elements** of technical projects so we gain the perspective necessary to manage risk.

Pacific Northwest
NATIONAL LABORATORY

How Risk is Managed

- Risk management is not static and occurs through all stages of the project's life cycle. New risks materialize as a project enters each phase of work and old non relevant risks are retired



Pacific Northwest NATIONAL LABORATORY

Managing Risks (continued)

- "Mitigation" means implementing a chosen method to address the risk. Method categories include:
  - **Acceptance** – If you understand the risk and have sufficient information to address the potential impacts, the project can chose to accept the risk by:
    - Increasing the amount of contingency allocated to a budget
    - Increasing the amount of time allocated to complete a task
  - **Avoidance** – Take an action that prevents the risk from being realized
    - Technology development plans or prototyping
    - Insure against the risk
  - **Transfer** – Move the risk burden to another entity
    - Use a fixed price acquisition strategy
    - Use programmatic R&D focused on de-scoped elements of the project to clarify the technical outcome in a timely manner and reduce the risk sufficiently to accept the scope back into the project

Managing Risks (continued)

- "Monitoring" includes regularly assessing (e.g. monthly) the status of identified risks to ensure that they are being addressed and to evaluate the effectiveness of the method chosen to address them.
  - Use a risk registry method
  - Establish retirement dates
  - Maintain project priorities based on the assessment of risk
    - What gets done first is always a function of the risk elements that need to be mitigated earliest.

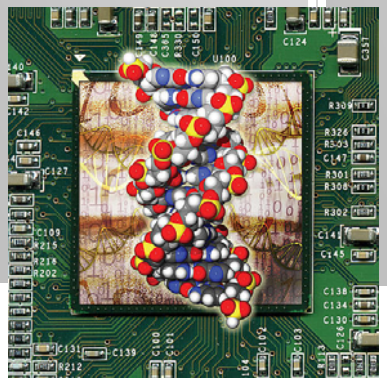
WBS Number:	Date:
1. Project Element, Step, Or Activity:	
2. Risk Type: **Cost_____Schedule___Technical__** Describe Potential Problem **(Impact Event)** :	
3. Likelihood Of Occurrence **(unlikely, likely, very likely) :**	
4. Unmitigated Consequence:	
5. Baseline Impact / Consequence Level **(Low, Medium, High) :**	
6. Risk Categorization **(High Risk, Medium Risk, Low Risk) :**	
7. Likely Causes:	
8. Descriptive Mitigation Actions **(Accept, Transfer, Avoid)**	
9. Responsible Individuals :	
10. Expected Closeout Date: **(assume mitigation actions are achieved as requested)**	

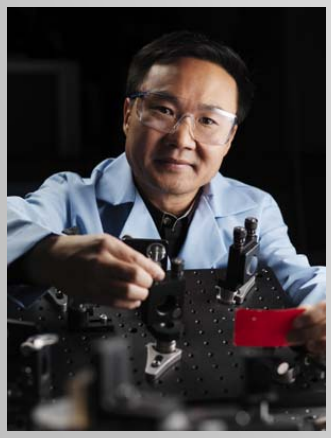
Pacific Northwest
NATIONAL LABORATORY

Managing Risks (continued)

- Common project risk categories – use a project centric model:

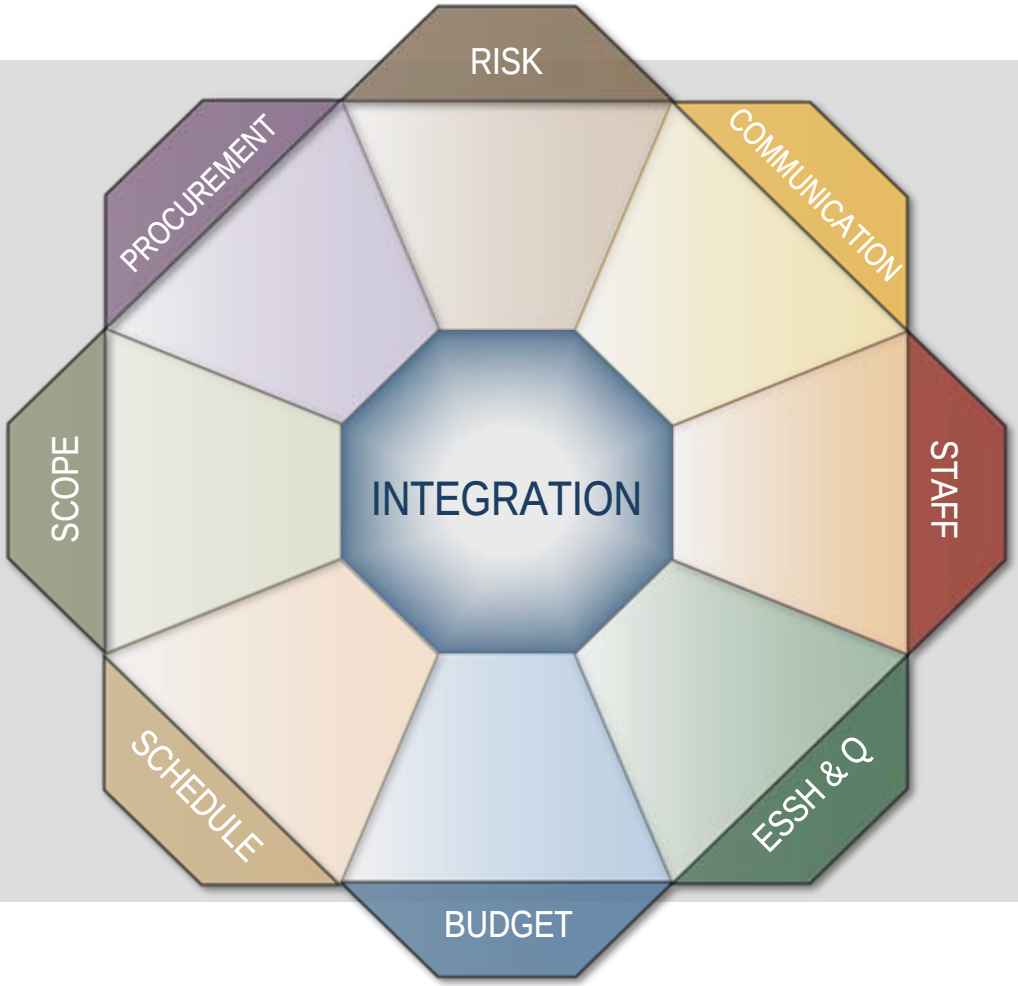
Scope Risks	Schedule Risks	Management Risks	Quality Risks	Cost Risks
▶ Unclear scope requirements ▶ Overly optimistic technical approach ▶ Scope creep ▶ Not using a charge control process	▶ Overoptimistic schedule estimates ▶ Resources not available ▶ Wrong resources assigned ▶ High personnel turnover ▶ Dependencies not fully identified	▶ Micro-management ▶ Organizational instability ▶ Unresponsive management	▶ Missing, unclear, changing specifications ▶ Lack of customer involvement ▶ Use of un-proven technology ▶ Lack of reviews	▶ Price escalations ▶ Overoptimistic cost estimates ▶ No prior cost data available for cutting-edge work





Pacific Northwest
NATIONAL LABORATORY

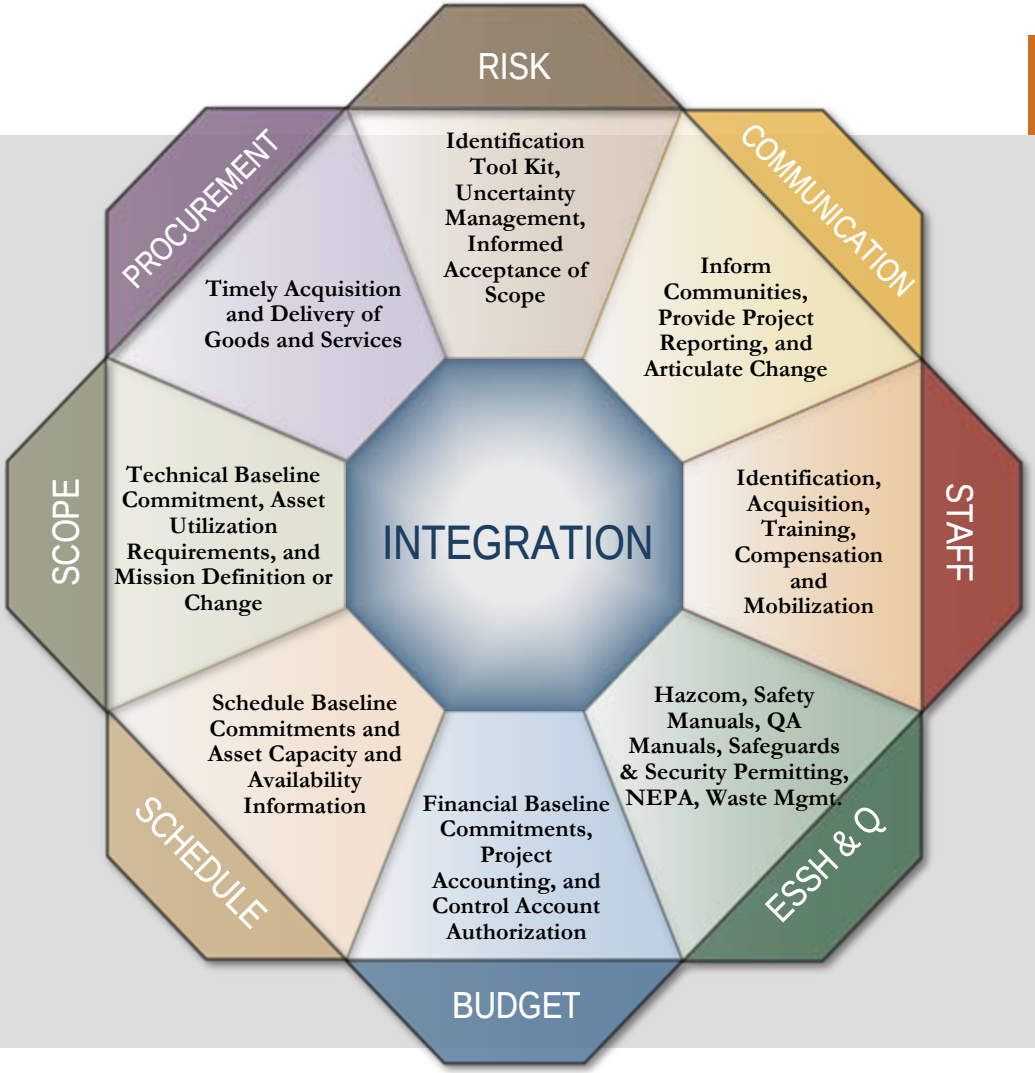
Project Centric Models Looks Like:



- The Project Manager's medallion is a project centric model that has been proven adaptable across multiple environments:
  - Project Delivery
  - Project Training
  - Enterprise Management
- Each segment of the medallion can contribute to risk identification and management

Pacific Northwest NATIONAL LABORATORY

How to Adapt the Model to Our Needs



Put the model together:

- Broaden questions to understand consequences and likelihood of failure for our performance assumptions in each segment - provide a complete risk picture.
- Integrate management actions to mitigate or transfer risk consequences and minimize likelihood of failure
- Risk management tools can help us – or complicate our lives if we apply them inappropriately
- Remember: simple projects – simple tools, complex projects – more sophisticated tools, but keep it useful and relevant

Pacific Northwest NATIONAL LABORATORY

What does a simple risk management tool look like for a complex project?

- Risk likelihood categories:
  - **Very Likely (V) —** Consequence is likely to occur with a high probability, i.e. greater than or equal to 90% based on best technical judgment.
  - **Likely (L) —** Consequence is likely to occur with a reasonable probability, i.e. greater than or equal to 50% and less than 90% based on best technical judgment.
  - **Unlikely (U) —** There is less than 50% chance that this consequence will occur based on best technical judgment.

What does a simple risk management tool look like for a complex project?

Consequence Category	Definition		
	Cost: Impact on project contingency	**Schedule**: Impact on project schedule	**Technical**: Impact on performance
Marginal (M)	≤ $500k	None	Minor degradation, Performance falls below upper end of goal; end objectives can still be met
Significant (S)	> $500k, but ≤ $1M	Impacts significant lower level performance commitment milestones defined by project leadership	Moderate performance shortfall, but workarounds available; Performance might impact end objectives if not mitigated.
Critical (C)	> $1M	Impacts primary contractual or Client commitment milestones defined in formal project documentation	End objectives will not be met (essential performance is not met); if alternatives exist the project has difficulty accepting the alternative, or no alternatives exist

Thresholds Are Negotiated With Your Client

Pacific Northwest NATIONAL LABORATORY

What does a simple risk management tool look like for a complex project?

Likelihood of occurrence	Consequence		
	Marginal	Significant	Critical
Very likely	Medium	High	High
Likely	Low	Medium	High
Unlikely	Low	Low	Medium

- The combination of likelihood and consequences in the matrix satisfies the classical definition of risk – using qualitative methods.
- The red / yellow / green establishes thresholds and prioritization of management attention.
- For those areas that require management attention, the project can do one of three things to "mitigate risk:"
  - Accept the risk
  - Avoid the risk
  - Transfer the risk

Pacific Northwest
NATIONAL LABORATORY

Three ways to Mitigate Risk — Acceptance

- If the conditions surrounding the risk are sufficiently well understood, the project can accept the risk:
  - For most technical elements an evaluation of contingency percentage (typically increase) is appropriate based on the nature of the risk conditions.
    - Examples include:
      - Schedule commitment to allow a known commercial technology to emerge, i.e. new chip release or OS release,
      - Budget and schedule commitment associated with the cost of rework, i.e. university-led development, or best efforts relationships with other national laboratories
  - If a risk acceptance strategy is used:
    - Increase the contingency value (cost and or schedule) associated with the risk element.
    - Remove the element from the risk registry and track it as part of routine project performance.

Three ways to Mitigate Risk – Avoid

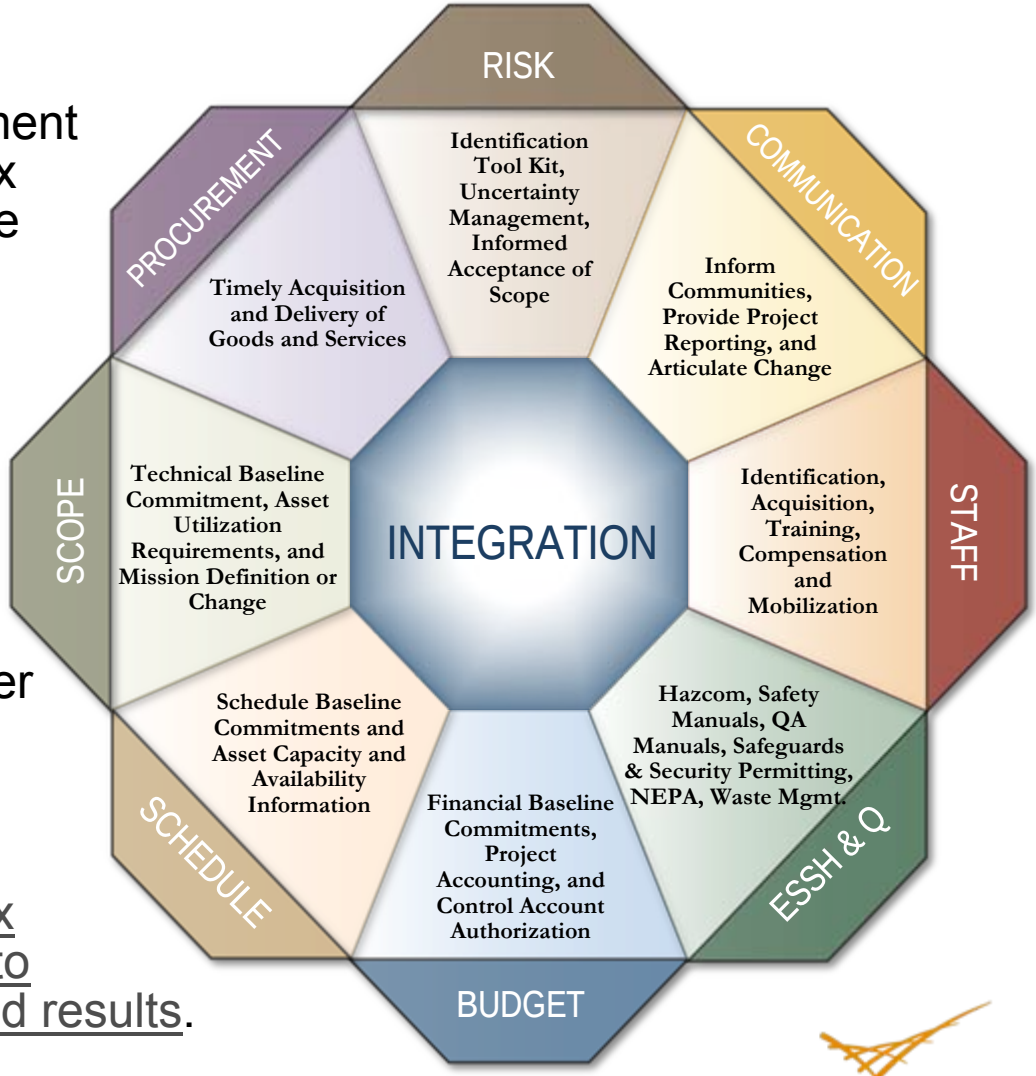
- If conditions surrounding the risk element are not clearly understood, but lend themselves to resolution, the project can attempt to avoid / mitigate the risk.
  - For most project elements, an investment in R&D focused on reducing potential consequences or in reducing likelihood is appropriate.
  - Risk elements become part of a technology development plan and would be "accepted" into the project as the risk is ultimately reduced or the baseline becomes robust enough to "accept" the risk.
    - Examples include:
      - Prototype development to avoid production risks,
      - Reducing novel technology ideas to more common practice,
      - Automation of manual tasks, etc.
  - Avoidance strategy requires program integration actions:
    - Tightly link the programmatic R&D strategy to the project delivery schedule, i.e. parallel software development in time to support commissioning tests.
    - Drive program priorities based on project needs, include risk in project's risk registry, but don't include efforts in the baseline until after results can be "accepted."

Three ways to Mitigate Risk – Transfer

- If conditions surrounding risk element are too random or occur in a manner better suited to specialized commercial practices, the project can attempt to transfer the risk.
  - Risk elements become part of a contracting and procurement strategy, make versus buy evaluation's), or novel contracting strategies and insurance options.
    - Examples include:
      - Long-term computer lease actions with known upgrade requirements
      - Third-party acquisition of floor space and maintenance
      - Insurance programs for safety risks, etc.
  - Transfer strategy requires additional planning effort:
    - Plan for additional time and money required to address approval cycles and incremental costs for specialized services that help implement the strategy
    - Reduce project contingency associated with the risk element | once it has been effectively transferred
    - Maintain risk element in risk registry until such time as it has been “retired” via transfer.

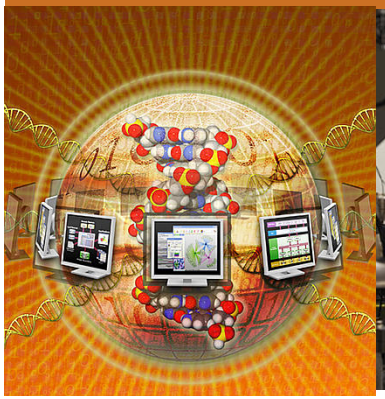
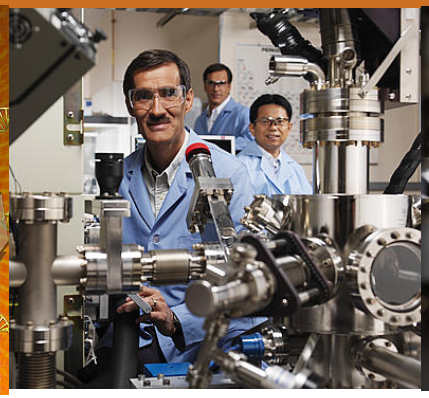
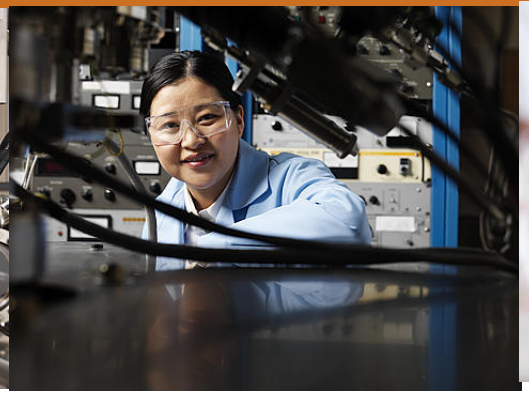
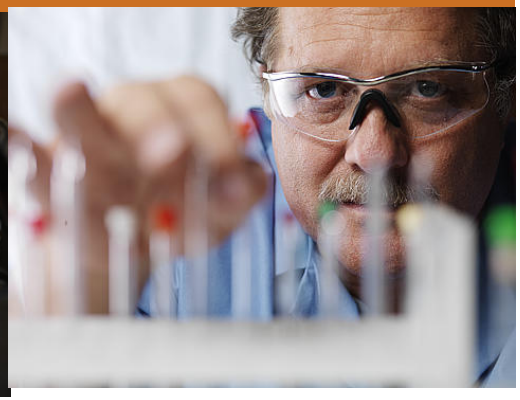
Risk Identification and Management – Conclusion

- Traditional risk and risk management definitions can be met on complex technical projects using qualitative means.
- Risk management is an appropriate tool for use in delivery of complex technical projects.
- Accept, avoid / mitigate, or transfer strategies apply equally well to complex technical projects.
- Just because you have a complex project does not mean you need to use complex tools to achieve good results.



Pacific Northwest NATIONAL LABORATORY

Questions?

Pacific Northwest
NATIONAL LABORATORY

Petascale Infrastructure Systems Integration Project (PISIP)

Scope Summary

There is a great deal of enthusiasm for making large scale system integration a full-fledged partner with other major thrusts supported by the agencies in the definition, design, and use of a petascale systems. Integration technology and the issues affecting optimization of that technology should have a full "seat at the table" as petascale and exascale initiatives and programs are planned. This project is developed to improve the state of the art for large scale systems integration by increasing the scientific productivity of these systems.

Current computer operations sites have a great deal of expertise in overall systems integration; but, the ability to leverage the expertise between sites is inadequate. Many issues get in the way of sharing information, including:

- available time and effort,
- sharing proprietary information
- technology gaps and inconsistency in standard use
- geographic dispersal
- conflict and priorities resolution between commercial, academic, and programmatic benefits
- benefit recognition when establishing integrated infrastructure technology, programmatic research, user services, and hardware procurement priorities.

This project is established to resolve and minimize the impact of these issues as petascale and exascale computing initiatives emerge.

Introduction

As high-performance computing vendors and supercomputing centers move toward petascale computing, every phase of these systems, from the first design to the final use, presents unprecedented challenges. Activity is underway for requirements definition, hardware and software design, programming models modifications, methods and tools innovation and acquisitions. After systems are designed and purchased, and before they can be used at petascale for their intended purposes, they must be installed in a facility, integrated with the existing infrastructure and environment, tested and then deployed for use. **Unless system testing and integration is done effectively, there are risks that large scale systems will never reach their full potential.**

Specifically, the PISIP will:

- Identify challenges and issues involved in the installation and deployment of large scale HPC systems
- Provide the infrastructure to deploy best practices for installing large-scale HPC systems into scientific petascale facilities
- Deploy methods to assure system performance and function continue after initial testing and deployment
- Resolve systematic issues and research issues for vendors, sites and facilities that would improve the speed and quality of deployment
- Provide a platform to share tools and methods that are helpful in expediting the installation, testing and configuration of HPC systems
- Establish and increase the communication paths for technical staff at multiple sites that make HPC installations more effective
- Make recommendations to DOE and other stakeholders to improve the process of HPC system deployment

PISIP Work Breakdown Structure

WBS 1.0 Improve the ability to record and process log data.

WBS 1.1 Improve handling and use of log data describing system use and status.

1.1.1 Develop common log data models and data formats that enable integrated analysis tools and easier sharing

1.1.2 Develop consolidation of events tool

1.1.3 Increase the amount of "Intelligent" logging – ala the AMR tool – that does broad analysis with automated detailed focus as needed.

1.1.4 Increase the site's and vendors' ability to do forensic analysis

WBS 1.2 Share tools, system integration techniques, and operational data across sites. Enable the ability to look for larger trends across six existing sites who have agreed to share current data. Ensure the end product enables understanding different systems and code behavior.

1.2.1 Define and deliver a common/standard format for tools and techniques.

1.2.2 Establish a conversion protocol convert proprietary and/or current formats to the standard format.

1.2.3 Capture and deliver a data model that addresses the entire log file analysis in the standard format and ensure data management, curation, and data protection are addressed.

WBS 1.3 Build a repository of log files across sites as a prerequisite to building data management tools. Enable eight sites to share system log data with appropriate controls, hosted at an academic institution, with communication protocols established to support seamless participation.

1.3.1 Define and establish the format for data models

1.3.2 Establish the level and fidelity of information to be captured and translate that level into performance criteria for bandwidth and storage

1.3.3 Incorporate extensible data model criteria to address unanticipated / new data availability, i.e. include" sanitize logs" research in order to release them to the general academic community for further contributions.

1.3.4 Deliver analysis tools supporting statistical analysis with first deliverable being a Simple Event Correlator – an open source package – to analyze events.

1.3.5 Establish and deliver a suite of improved methods to visualize the system state.

WBS 2.0 Enable multi-version system software on petascale systems.

WBS 2.1 Test system software at scale and be able to quickly go back and forth between versions.

2.1.1 Establish the necessary system software transition requirements to ensure effective switching between software levels.

2.1.1.1 Ensure commercial practice delivers a mechanism to isolate system level changes to support limited test environments

2.1.1.2 Ensure that the shared file system can be partitioned without replicating hardware costs.

2.1.2 Demonstrate the integrated solution between two sites (up to 6 as funding and interest dictate)

WBS 2.2 Deliver more robust hardware and software diagnostics

2.2.1 Deliver diagnostics with root cause predictive capability without establishing dedicated mode operations.

2.2.2 Deliver field replaceable action identification based on predictive diagnostics for both software and hardware problem calls

2.2.3 Create the communication linkage between commercial vendor, system administration, and hardware operations to effectively learn these diagnostics and optimize their use

2.2.4 Define and deliver language for commercial system performance RFP's that establishes the "standard" of performance in MTTR.

2.2.5 Create compatible on-line diagnostics that are as comprehensive as off-line diagnostics.

2.2.5.1 Provide the ability to analyze soft errors in a predictive mode rather than response mode.

2.2.5.2 Establish test diagnostics as part of an acceptance test, reflecting realistic test mode parameters with "test failures" to challenge diagnostic performance.

WBS 2.3 Improve I/O Activity and Analysis Tools

2.3.1 Deliver "standardized" storage system tools that are equivalent to HPM/PAPI that monitor CPU performance data.

2.3.1.1 Diagnose code performance for application I/O improvement

2.3.1.2 Tune / Balance overall system performance and workload for data systems designer use

2.3.2 Deliver multi-layer visualization to correlate I/O system data. Monitor and assess overall system behavior, true aggregate performance and conflicting use and demands using tools such as IOR and AMR I/O benchmarks.

WBS 3.0 Peer Interactions and Best Practices

WBS 3.1 Establish a community of practice user group for facility managers to brainstorm integrated solutions prior to design. Look at *Instructions to America* that creates communities of practice on many topics.

WBS 3.2 Create a trusted environment for the sharing of "problem area concerns", i.e. reliability, software and hardware problems, security issues and user issues.

WBS 3.3 Create a "trusted" routine format and expectation for summary of acceptance test results and lessons learned experience

WBS 3.4 Expand facility specific conference opportunities and lessons learned related forums for integrated large system performance.

3.4.1 Create a journal edition devoted to Large Scale System Integration.

3.4.2 Create an on-line community, euphemistically labeled "HPC_space.com" utilizing mechanisms such as twikis or blogs. Ensure the "rules of engagement" are carefully considered to address the needs of vendors, systems, areas of interest and sites.

WBS 4.0 Tools for Understanding Performance and Function for storage (disk) management and reliability, and other Key components.

WBS 4.1 Deliver tools to better move, manage, and distill information from data.

4.1.1 Deliver data insight tools that operate in real time with application improvements.

4.1.2 Integrate the performance of small data writes and new ways to handle very large numbers of files. Optimize advancements made by vendors and others who are dealing with Information Lifecycle Management, including hierarchies of data movement.

4.1.3 Provide the ability to monitor large numbers of storage and fabric devices as well as tools for disk management.

4.1.4 Create tools for understanding interconnects such as PAPI/IPM and performance profiling tools for other programming models such as PGAS languages.

WBS 4.2 Establish and institutionalize use of Parallel Debugger Tools

Parallel debuggers can be a significant help with integration issues as well as help users be more productive. It is not clear that sites understand how users debug their large codes. Although this topic was partly covered at a DOE-sponsored workshop on petascale tools in August 2007, it is not clear there is a connection with the Integration Community

List of priorities from attendees based on a ranking algorithm.

Description	Summary ranking
Provide methods for regular interaction between peer centers to avoid potential pitfalls and identify best practices.	187
Consolidated event and log management with event analysis and correlation. Must be an extensible, open framework. Common format to express system performance (at a low level), health, status, and resource utilization (e.g <machine name="franklin">; <cab	142
Tools for ongoing "Intelligent" syslog / data reduction and analysis.	123
Develop methods to combine machine and environment monitoring.	113
Ability to have multiple versions of the OS that can be easily booted for testing and development. Ability to do rolling upgrades of the system and tools. Can you return to a previous state of the system? Very broad impacts across kernel, firmware, etc.	113
Develop standard formats and ability to share information about machines and facilities (Wiki?). Community monitoring API (SIM?)	112
Develop better hardware and software diagnostics	108
Develop tools to track and monitor message performance (e.g. HPM/PAPI for the interconnect and I/O paths, hybrid programming models)	105
Create better parallel I/O performance tests and/or a Parallel I/O test suite	101
Funded Best Practices Sharing (Chronological list of questions and elements of project plan, Top 10 Risk Lists, Commodity equipment acquisition / performance)	99
Better ways to visualize what the system is doing with remote display to a system administrator for more holistic system monitoring.	96
Visualization and Analytics tools for log and performance data. Tool that can analyze the logs, perhaps drawing a parallel to Bro, which can analyze network traffic for anomalies that indicate attacks, etc	90
Invite peer center personnel to review building design plans to achieve as much input as possible and so reviewers can benefit as well.	86
Tools for storage (Disk) Systems management and reliability	85
Develop/identify computer facility modeling and analysis software (Air flow, cooling system, etc. – e.g Tileflow)	81
Develop automated procedures for performance testing	81
Share problem reports with all sites – a vendor issue rather than a site issue.	80
Improved parallel debugger for large scale systems including dump analysis	77
Develop tools to verify the integrity of the system files and to do consistency checking among all the pieces of the system.	67
Develop accurate methods for memory usage monitor/OS intrusion	65
Tools to monitor performance relative to energy power draw	63
Failure data fusion and statistical analysis	60
Develop realistic interconnect tests	59
Scalable configuration management	58
Job failure due to system failure should be calculable. In house tools are being used in some cases to try and correlate batch system events with system events.	57
Share (Sanitized) Project Plan experience among sites as part of project closeout – RM activities required	56
Implement facility sensor networks in computer rooms including analysis and data fusion capabilities	55

Develop accurate performance models for non-existent systems including full system including I/O, interconnects, software	55
Develop tools to measure real Bit Error Rate (BER) for interconnects and I/O	52
Developing improved methods of backing up a systems of this size	52
Develop *matrix* to correlate the four major facility categories (Space, Power, Cooling, Networking) with each of these phases of a facility for Greenfield (new), Retrofit, Relocate, Upgrade for planning purposes.	48
Create the ability to fully simulate a full system at scale without having to build it (e.g UCB RAMP)	46
Develop ways to receive real data for power consumption from vendors	42
Hard partitioning of systems so one partition cannot impact another. Virtualized the I/O partition so that it can be attached to multiple compute partitions.	42
Have external participation in proposal and plan reviews	42
Fund studies about systemic facility design (What voltage? AC or DC? CRAC Units, Etc.) including the level of detail needed for monitoring?	41
Statistical analysis of logs can detect pending failures. Is deployed by Google, Yahoo to address reliable and adaptive systems requirements. Currently based on http:, but may be extensible to this situation.	40
Coordinate scheduling of resources in addition to CPUs	39
Share WBS, RM, Communications Plans, etc. among sites	37
Improve Project Management expertise in organizations	36
Create a framework with general acceptance by most of the community consisting of a set of tests that provide decision points on the flow chart for debugging	17
Benchmarks in new programming models (UPC, CAF, …)	17
Reporting that shows just the important (different) areas in the machine	15
Develop the methods and roles for statistician with PSI projects	6
Configuration management tools can manage this component-level knowledge base, but will be executed on a site-by-site, case-by-case basis.	2

Petascale Infrastructure System Integration Project (PISIP) Risk Registry

WBS Number:	Date:
Project Element, Step, Or Activity:	
2. Risk Type: Cost Schedule Technical Describe Potential Problem (Impact Event) :	
3. Likelihood Of Occurrence (unlikely, likely, very likely) :	
4. Unmitigated Consequence:	
5. Baseline Impact / Consequence Level (Low, Medium, High) :	
6. Risk Categorization (High Risk, Medium Risk, Low Risk) :	
7. Likely Causes:	
8. Descriptive Mitigation Actions (Accept, Transfer, Avoid)	
9. Responsible Individuals :	
10. Expected Closeout Date: (assume mitigation actions are achieved as requested)	

Petascale Infrastructure System Integration Project (PISIP) Risk Registry

Petascale Infrastructure System Integration Project (PISIP) Risk Registry

WBS 1.2 **Share tools, system integration techniques, and operational data across sites.** Enable the ability to look for larger trends across six existing sites who have agreed to share current data. Ensure the end product enables understanding different systems and code behavior.

WBS Number: *1.2.3*	Date: September 17, 2008

Project Element, Step, Or Activity:

Define and deliver a data model that addresses the entire log file analysis in the standard format and ensure data management, curation, and data protection are addressed.

2. Risk Type: Cost Schedule Technical

Describe Potential Problem (Impact Event) :

Data models at infrastructure sites do not lend themselves to routine formats and / or logs do not contain all elements of format data requiring significant rework or new structures.

3. Likelihood Of Occurrence (**Unlikely**, **Likely**, **Very Likely**) :

Very Likely

4. Unmitigated Consequence:

Rework of common log data input requires significant schedule extension and possible reconstruction of data files due to incompatible or corrupted data use requiring manual intervention and input correction.

5. Baseline Impact / Consequence Level (**Marginal**, **Significant**, **Critical**) :

Significant

6. Risk Categorization (**High Risk,** **Medium Risk**, **Low Risk**) :

High Risk

7. Likely Causes:

Proprietary commercial practice or claims of parochial needs prevent seamless integration of common log formats and tools

8. Descriptive Mitigation Actions (Accept, **Transfer**, Avoid)

Good potential to transfer to a subcontracted action once a common format is established

9. Responsible Individuals : *TBD*
10. Expected Closeout Date: (assume mitigation actions are achieved as requested) *TBD*

Petascale Infrastructure System Integration Project (PISIP) Risk Registry

WBS 1.2 **Share tools, system integration techniques, and operational data across sites.** Enable the ability to look for larger trends across six existing sites who have agreed to share current data. Ensure the end product enables understanding different systems and code behavior.

WBS Number: *1.2.3*	Date: September 17, 2008

Project Element, Step, Or Activity:

Define and deliver a data model that addresses the entire log file analysis in the standard format and ensure data management, curation, and data protection are addressed.

2. Risk Type: Cost Schedule Technical

Describe Potential Problem (Impact Event) :

Data models at infrastructure sites do not lend themselves to routine formats and / or logs do not contain all elements of format data requiring significant rework or new structures.

3. Likelihood Of Occurrence (**Unlikely**, **Likely**, **Very Likely**) :

Very Likely

4. Unmitigated Consequence:

Rework of common log data input requires significant budget increase to address cost of labor in support of data file reconstruction due to incompatible or corrupted data use requiring manual intervention and input correction.

5. Baseline Impact / Consequence Level (**Marginal**, **Significant**, **Critical**) :

Critical

6. Risk Categorization (**High Risk,** **Medium Risk**, **Low Risk**) :

High Risk

7. Likely Causes:

Proprietary commercial practice or claims of parochial needs prevent seamless integration of common log formats and tools

8. Descriptive Mitigation Actions (Accept, **Transfer**, Avoid)

Good potential to transfer to a subcontracted action once a common format is established

9. Responsible Individuals : *TBD*
10. Expected Closeout Date: (assume mitigation actions are achieved as requested) *TBD*

Petascale Infrastructure System Integration Project (PISIP) Risk Registry

WBS 1.1 Improve handling and use of log data describing system use and status.

WBS Number: *1.1.1*	Date: September 17, 2008

Project Element, Step, Or Activity:

Develop common log data models and data formats that enable integrated analysis tools and easier sharing

2. Risk Type: Cost Schedule Technical

Describe Potential Problem (Impact Event) :

Data models at infrastructure sites do not lend themselves to routine formats and / or logs do not contain all elements of format data requiring significant rework or new structures.

3. Likelihood Of Occurrence (**Unlikely**, **Likely**, **Very Likely**) :

Very Likely

4. Unmitigated Consequence:

Rework of common log data input requires significant schedule extension and possible reconstruction of data files due to incompatible or corrupted data use requiring manual intervention and input correction.

5. Baseline Impact / Consequence Level (**Marginal**, **Significant**, **Critical**) :

Significant

6. Risk Categorization (**High Risk,** **Medium Risk**, **Low Risk**) :

High Risk

7. Likely Causes:

Proprietary commercial practice or claims of parochial needs prevent seamless integration of common log formats and tools

8. Descriptive Mitigation Actions (Accept, **Transfer**, Avoid)

Good potential to transfer to a subcontracted action once a common format is established

9. Responsible Individuals : *TBD*
10. Expected Closeout Date: (assume mitigation actions are achieved as requested) *TBD*

Petascale Infrastructure System Integration Project (PISIP) Risk Registry

WBS 1.1 Improve handling and use of log data describing system use and status.

WBS Number: *1.1.1*	Date: September 17, 2008

Project Element, Step, Or Activity:

Develop common log data models and data formats that enable integrated analysis tools and easier sharing

2. Risk Type: Cost Schedule Technical

Describe Potential Problem (Impact Event) :

Data models at infrastructure sites do not lend themselves to routine formats and / or logs do not contain all elements of format data requiring significant rework or new structures.

3. Likelihood Of Occurrence (**Unlikely**, **Likely**, **Very Likely**) :

Very Likely

4. Unmitigated Consequence:

Rework of common log data input requires significant budget increase to address the cost of labor caused by reconstruction of data files due to incompatible or corrupted data requiring manual intervention and input correction.

5. Baseline Impact / Consequence Level (**Marginal**, **Significant**, **Critical**) :

Critical

6. Risk Categorization (**High Risk,** **Medium Risk**, **Low Risk**) :

High Risk

7. Likely Causes:

Proprietary commercial practice or claims of parochial needs prevent seamless integration of common log formats and tools

8. Descriptive Mitigation Actions (Accept, **Transfer,** Avoid)

Good potential to transfer to a subcontracted action once a common format is established.

9. Responsible Individuals : *TBD*
10. Expected Closeout Date: (assume mitigation actions are achieved as requested) *TBD*

Petascale Infrastructure System Integration Project (PISIP) Risk Registry

WBS 1.1 Improve handling and use of log data describing system use and status.

WBS Number: *1.1.1*	Date: September 17, 2008

Project Element, Step, Or Activity:

Develop common log data models and data formats that enable integrated analysis tools and easier sharing

2. Risk Type: Cost Schedule Technical

Describe Potential Problem (Impact Event) :

Data models at infrastructure sites do not lend themselves to routine formats and / or logs do not contain all elements of format data requiring significant rework or new structures.

3. Likelihood Of Occurrence (**Unlikely**, **Likely**, **Very Likely**) :

Very Likely

4. Unmitigated Consequence:

Data logs don't talk to each other as well as they could (or should) requiring manual intervention and input correction

5. Baseline Impact / Consequence Level (**Marginal**, **Significant**, **Critical**) :

Marginal

6. Risk Categorization (**High Risk,** **Medium Risk**, **Low Risk**) :

Medium Risk

7. Likely Causes:

Proprietary commercial practice or claims of parochial needs prevent seamless integration of common log formats and tools

8. Descriptive Mitigation Actions (Accept, **Transfer**, Avoid)

Good potential to transfer to a subcontracted action once a common format is established

9. Responsible Individuals :

TBD

10. Expected Closeout Date: (assume mitigation actions are achieved as requested) *TBD*

Petascale Infrastructure System Integration Project (PISIP) Risk Registry

WBS 1.2 **Share tools, system integration techniques, and operational data across sites.** Enable the ability to look for larger trends across six existing sites who have agreed to share current data. Ensure the end product enables understanding different systems and code behavior.

WBS Number: *1.2.3*	Date: September 17, 2008

Project Element, Step, Or Activity:

Define and deliver a data model that addresses the entire log file analysis in the standard format and ensure data management, curation, and data protection are addressed.

2. Risk Type: Cost Schedule Technical

Describe Potential Problem (Impact Event) :

Data models at infrastructure sites do not lend themselves to routine formats and / or logs do not contain all elements of format data requiring significant rework or new structures.

3. Likelihood Of Occurrence (**Unlikely**, **Likely**, **Very Likely**) :

Unlikely

4. Unmitigated Consequence:

Rework of common log data input is not technically challenging and does not introduce new risk elements. It is not likely to affect final performance characteristics.

5. Baseline Impact / Consequence Level (**Marginal**, **Significant**, **Critical**) :

Significant

6. Risk Categorization (**High Risk,** **Medium Risk**, **Low Risk**) :

Medium Risk

7. Likely Causes:

Proprietary commercial practice or claims of parochial needs prevent seamless integration of common log formats and tools

8. Descriptive Mitigation Actions (Accept, **Transfer**, Avoid)

Good potential to transfer to a subcontracted action once a common format is established

9. Responsible Individuals : *TBD*
10. Expected Closeout Date: (assume mitigation actions are achieved as requested) *TBD*

Petascale Infrastructure System Integration Project (PISIP) Risk Registry

WBS 2.1 Test system software at scale and be able to quickly go back and forth between versions.

WBS Number: 2.*1.1.1*	Date: September 17, 2008

Project Element, Step, Or Activity:

Ensure commercial practice delivers a mechanism to isolate system level changes to support limited test environments

2. Risk Type: Cost Schedule Technical

Describe Potential Problem (Impact Event) :

Petascale software tools are fundamentally too large to support simultaneous version utilization and there is no technical solution to support real time switching at scale. The problem is so intractable that industry will not accept this as commercial practice.

3. Likelihood Of Occurrence (**Unlikely**, **Likely**, **Very Likely**) :

Unlikely

4. Unmitigated Consequence:

Commercial practice does not deliver system changes necessary to isolate system level changes and test at scale between versions. The result is extensive individual version test and correlation between successive test runs(??)

5. Baseline Impact / Consequence Level (**Marginal**, **Significant**, **Critical**) :

Significant

6. Risk Categorization (**High Risk,** **Medium Risk**, **Low Risk**) :

Low Risk

7. Likely Causes:

The physical size of petascale version data set switching represents a technical limitation that can not be addressed without transformational discovery in data management and tools (???)

8. Descriptive Mitigation Actions (Accept, Transfer, **Avoid**)

Potential to avoid as a linked outcome to a programmatic initiative focused on transforming the means and tools of petascale version management and switching.

9. Responsible Individuals : *TBD*
10. Expected Closeout Date: (assume mitigation actions are achieved as requested) *TBD*