Jet Propulsion Laboratory
California Institute of Technology

4800 Oak Grove Drive
Pasadena, CA 91109-8099

28 May, 2010
Refer to: 930 TDL:csh #8

Mr. William Knopf
Program Executive
Solar System Division
National Aeronautics and Space Administration
Washington, DC 20546

Bill:

Attached in "Word" format is the System Review Board Report from the Planetary Data System (PDS) System Review held March 22 – 24 at the Business Offices in Washington DC.

Please let me know if you have any questions or comments.

Terry D. Linick
Multimission Ground Systems and Services Program Office
Jet Propulsion Laboratory

TD:csh
Attachtment

Cc: David Heather, ESA, ESTEC
Dave Linick, JPL, Board Chair
Andrew Schain, NASA HQ
David Korsemeyer, ARC
Jan Merka, GSFC
Peter Shames, JPL

PDS 2010 System Review Board Report

The Planetary Data System (PDS) System Review was held on March 22nd, 23rd, and 24th, 2010 at the JPL Offices in Washington, D.C. The Board membership was:

David Heather, ESA, ESTEC
David Korsemeyer, ARC
Dave Linick, JPL, Board Chair
Jan Merka, GSFC
Andrew Schain, NASA HQ
Peter Shames, JPL

The intent of the System Review was to assess the architecture and design of the next generation of the PDS (PDS-4) and to assure that the implementation plans, and the deployment and transition planning is adequate and appropriate. This is reflected in the stated review objectives:

1. Assess technical architecture and ensure that it is responsive to the needs, drivers and requirements for the PDS over the next decade
2. Ensure PDS has a design that is responsive to the architecture and PDS drivers
3. Assess the implementation plan (schedule, resources, phasing)
4. Ensure that PDS has a deployment and infusion plan for PDS-4 that includes PDS nodes and missions
5. Assess the transition plan from PDS3 to PDS4 operations
6. Provide overall technical and project management recommendations

REVIEW BOARD FINDINGS

The findings of the Review Board are presented in the following sections of this report:

- Executive Summary
- Assessment Against Review Objectives
- Individual Board Member Comments (attachment 1)
- Requests for Action (attachment 2)
- Summary Debriefing to the PDS management Council (attachment 3)

Executive Summary:

The consensus of the Board was that the defined objectives of the Review were achieved with the following exceptions:

- The level of detail specified in the requirements that were presented was insufficient for the Board to assess the adequacy of the resources (budget and schedule)
- The transition strategy did not clearly define when new missions and data providers will be required to use the PDS-4 format

In the view of the Board, the proposed architecture for PDS-4 appears to be very appropriate to meeting the existing requirements and is exceptionally robust with respect to being evolvable as future requirements are identified and as relevant technologies progress. Therefore, this architecture should serve the PDS well for the foreseeable future.

In addition, the Board noted that the technical expertise applied to the task appears to be excellent. The presenters were knowledgeable of both the PDS domain and the applicable information system technologies. The Board further noted that the teaming among the nodes, and particularly between the system node and the discipline nodes, was apparent and represents a significant improvement. Overall, the Board was very complimentary of the quality and competence of the effort that has been applied to date.

The principal issues and concerns were:

1. **The engineering rigor applied to PDS 2010 should be better specified and appropriate to the task.** This System Review served a useful and necessary purpose in exposing the proposed PDS architecture to critical review. However, since it was not formulated as a review defined by NPR 7120.8 and was missing some key elements, it was difficult to assess progress-to-date against a defined yardstick. This results in some level of uncertainty about the rigor that will be applied to the review of future progress.

   To address this, the Board recommends that an appropriately tailored NPR-7120.8 be applied to PDS-4 and to subsequent reviews to assure the proper level of rigor. This standard should be applied to each of the six defined implementation projects that comprise the overall task.

   Particular emphasis should be given to the specification of requirements and their expression in architecture and design. These need to be documented at a sufficient level of detail to assure that all of the stakeholders (users, nodes, and the NASA sponsor) are aligned, and that the end product is clearly defined so that progress can be crisply measured. In addition, the implementation should include the appropriate suite of standard reviews, including requirements reviews, PDRs, CDRs, etc.

2. **Better define the balance between the centralized and decentralized approach.** The Board recognizes that the PDS-4 architecture needs to strike a balance between the benefits of having autonomous discipline nodes (ability to apply discipline-specific expertise, development of node-optimized processes, minimal bureaucracy and overhead, etc.) and the benefits of a centralized approach (consistency across the system, reduced duplication and ambiguity, support to cross-discipline queries and global searches, etc.). The general architecture that was

presented supports a broad range of the spectrum of possibilities. The Board feels that the system design should be more specific in defining where PDS-4 will fall along that spectrum and further recommends that some migration toward a stronger system-level (centralized) approach is warranted. This could be made specific through the definition of governance processes and the specification of system-level requirements. In particular, requirements that define system-wide standards and common tools, including those related to global data searches, common data definitions, etc., may provide significant benefits without compromising the benefits of the discipline-centric view. In addition, provisioning for semantic rather than syntactic heterogeneity would also enhance search and query services with more robust/complex indexes.

3. **The design and management of the information models and data dictionaries needs to be better defined.** These are key components of PDS-4 and it is important that the specification for these models be clear and unambiguous, otherwise the nodes are each likely to interpret them differently. In addition, the management and validation of changes to these items is a critical component of the operations of the PDS and needs to be carefully thought through and specified in the governance model, including how to resolve semantic conflicts while allowing rapid updates.

4. **Clarify the priority of and resources allocated to the international component.** The Board recognizes the desirability of being able to accommodate the extension of PDS to address international collaboration, but this is not a specified requirement. The degree to which this is factored into the design and the amount of resources allocated to addressing this collaboration should be worked with the sponsor and specifically defined. Given the current priorities, a reasonable approach would be to assure that nothing in the design precludes future collaboration/integration, but that no significant resources be allocated toward enabling it.

5. **Clearly specify when, and under what conditions, missions and new data providers must be compliant with the new PDS-4 format.** The Board recommends development of clear guidelines for new missions that specifies when a new mission must be compliant with the new specifications and what is required of them technically as far as data formats and/or integration of a new node with the system.

Additional details regarding these issues and concerns are contained in the individual Board member comments (attachment 1) and the “Requests for Action” (attachment 2).

Assessment Against Review Objectives:

A general summary of the Board’s assessment of the System Review against the review objectives is shown below. Although there were two general liens identified, the Board agrees that no follow-up review is necessary and the Team should proceed toward implementation of PDS-4. Since the Board was not able to assess the resource adequacy (objective 3), the Program Executive may wish to engage

a mechanism to validate that the budget is commensurate with the requirements. This might best be assessed after the level 2 and 3 requirements have been defined.

1. Assess technical architecture and ensure that it is responsive to the needs, drivers and requirements for the PDS over the next decade

   Achieved.

2. Ensure PDS has a design that is responsive to the architecture and PDS drivers

   Achieved.

3. Assess the implementation plan (schedule, resources, phasing)

   Partly achieved. The requirements detail presented was not sufficient for the Board to evaluate the resource adequacy.

4. Ensure that PDS has a deployment and infusion plan for PDS-4 that includes PDS nodes and missions.

   Mostly achieved. The definition of when new missions and data providers are required to use the PDS-4 format was not specified.

5. Assess the transition plan from PDS3 to PDS4 operations

   Achieved.

6. Provide overall technical and project management recommendations.

   See specific Board member comments and RFAs.

Attachment 1

Board Member Comments

Dave Korsmeyer

Here are my overview comments from the review. I have nested them within the overall review objectives as presented.

In general this was pretty good. PDS4 seems like a solid response to address the presented requirements, but I had some critiques/comments/questions that were not answered yet.

PDS Review Objectives (and sub comments per objective)

1) Assess technical architecture and ensure that it is responsive to the needs, drivers and requirements for the PDS over the next decade

- Only saw the PDS level 1 requirements (would have liked to see L2/3 discussed some)

- L1 is being met. Overall Needs and basic drivers being met.

- Would have liked a presentation on the mapping of the PDS 4 design and functions against the L1/2/3 requirements and assessment of improvements over PDS3. (I saw the provided reference document)

- Q: Is the prime goal long term storage or access to data? Seems fuzzy in some of the discussions.

2) Ensure PDS has a design that is responsive to the architecture and PDS drivers

- Architecture and design were presented as "in sync" and virtually synonymous. I don't know which one drove the other. This was not clear.

- How does the PDS4 design compare and contrast with the LMMP work happening in ESMD?

- We did not see alternative designs discussed or what downselect process occurred. That would have been enlightening to assess whether the design was as robust as possible.

- {I read the provided reference papers, but would have liked to hear discussion what where the key decision points and results or the architectural down select}

3) Assess the implementation plan (schedule, resources, phasing)

- Little justification on the schedule (which seems reasonable)

- Little justification on resources. Was the work scaled to fit the budget or the budget scaled to fit the work? This needs to be answered.

4) Ensure that PDS has a deployment and infusion plan for PDS 2010 that includes PDS nodes and missions

- PDS2010 includes the nodes. Who were the 3rd party reviewers of this plan other than this board?

5) Assess the transition plan from PDS3 to PDS4 operations

- Reasonable plan for the transition from PDS 3 to PDS 4. Who were the 3rd party reviewers of this plan other than this board?

6) Provide overall technical and project management recommendations

- Are each of the subprojects going to have Preliminary Design Reviews, CDRs, and then delivery reviews?

- Would like to see some more rigor in the PDS Project management with respect to the technical review process.

- Dave

Peter Shames

To: PDS Review Board and Team:

Before addressing the issues identified during this review I want to take the time to acknowledge the excellent work that this whole team has done, on a rather limited budget, in developing what I believe is a superb architectural framework. One of the axioms in system design is that the only thing that is certain is change, whether it is accommodating new missions, adding new nodes, evolving with the technology underpinnings, or dealing with shifting winds of funding and programmatics. To deal with this fact of life this architecture is making use of the best current practices in the field and is also building in the necessary layering, isolation and integration mechanisms. It is also "future proofing" through clever use of information models, interoperable protocols, isolation from underlying technologies and portability across multiple platforms. There is little here to find flaw with and much to praise.

It must also be acknowledged that the PDS 2010 project is being developed within a very challenging organizational environment, where the system must be designed and operated within a federated organizational and funding structure. This was not apparent to me earlier, but it appears that control over the complete PDS development, deployment, and operation does not rest solely within the PDS project, rather it is a federation of a central node and a set of rather autonomous, independently funded and operated, discipline nodes. In some ways the stated goals for the system are in direct tension with the way that the project is organized:

- On-line services allowing users to access and transform data quickly from anywhere in the system
- A highly reliable, scalable computing infrastructure that protects the integrity of data, links the nodes into an integrated data system, and provides the best service to both data providers and users

One path to success, and needed in any case, is for all the stakeholders to work together for the common good. From the materials presented and from the interactions among the partners in this venture it is apparent that the members of the organization are pulling together to get the job done, kudos to the whole team. And special kudos to Dan Crichton for leading this effort so ably. I will note, however, that given the very loose nature of the organizational structure there may need to be some further clear guidance from the program office if these goals are really to be achieved. Some potential stressors were visible even within this short set of interactions.

One of the other aspects of this worthy of note is that there is quite clear evidence of forethought, consideration of the future growth and requirements on the system, and it is evident that the design is responsive not just to known requirements but also to principles of flexibility and adaptability to deal with the future unknown deltas that will occur in any such long lived system. This was evident in the design materials themselves and in many of the answers provided in response to review board questions. While the team may not yet have all of the answers, it is quite clear that they have thought about most of the questions and at least have some well formed opinions about where the best answers might lie. This all augurs well for success of this project and I look forward to seeing it evolve.

Best regards, Peter Shames

Andrew Schain

Information Architecture:

Data Dictionary – Is there a formal specification for the data dictionary and a process for keeping it updated? It appears that stewardship/curation of the data dictionary is being planned as a human intensive/human reliant activity. In other words, humans, not machines would be responsible for resolving semantic conflict or changes in meaning to nodal data dictionaries or to the PDS-wide global dictionary. If true, is the premise that there will be (a) few updates, (b) few conflicts, or (c) sufficient lead-time and execution time? Will this approach be viable if plans for increased and more frequent missions and with greater involvement of non-US Government participants are realized? Otherwise, are more automated mechanisms (to at least the first order of abstraction) being pursued to identify semantic conflict? Will terms from PDS 3 be part of the dictionary? Recommendation: Consider having the data dictionary contain terms from previous versions of PDS. Consider a mechanism to computationally resolve semantic conflicts that incrementally becomes a richer service.

Downstream dictionary customers - Will the dictionary output be available for other systems (e.g. LMMP, Moon in Google Earth) to use as part of a query or search service? Recommendation: If not already being planned, consider a WSDL service for downstream customers.

Naming resolution for XML schemas - It isn't clear if the data dictionary will be updated as a result of taking class and attribute names from XML schemas and ingesting/resolving them to the data dictionary. If this is true, will this be an automated or manual process? A similar condition applies to node customers who may want to look at the data dictionary and make a change or an additional meaning. How will this take place? If a node customer looks in the PDS 4 data dictionary and sees a term that is already used, but not in a way that is appropriate for the node, can they augment it, is it resolved in another way – and how long will it take? Recommendation: Target specific delivery and response goals in assisting those who must generate XML schemas as part of their PDS submission. Consider mechanisms to identify differences in logic or attributes of classes in XML schemas and process to resolve them.

Comment: A mechanism to fully or partially resolve semantic conflicts computationally is appropriate. Provisioning for semantic rather than syntactic heterogeneity would also enhance search and query services with more robust/complex indexes.

Information Models: As the corpus of information grows what is the planned approach for how PDS evolves information models? Recommendation: Plan this activity in your project plan.

RDF & OWL - There is a strategy for expressivity and RDF seems to be the choice for the information model and the data dictionary, yet the documentation seems to be fairly shy about that. Recommendation: Be specific.

Change Requests – is there a target metric for (1) responding or (2) resolving? Are metadata standards are restricted to project or mission or discipline? Metadata labels – goes thru a preliminary stage. Recommendation: Develop a concept of operations inclusive of certain types of change requests (e.g. conflict resolution) within your project plan.

Provenance standards & attributes, - is needed but not currently in place. What is the approach for obtaining and attributing a base set of provenance information and when will this be needed? Recommendation: Consider adopting as a starting point the provenance attributes identified in the NASA Information Architecture Policy Guide.

The NASA Information Architecture Policy Guide – Are the plans for PDS v4 compliant with the guide and or are there any recommend changes to the NASA IA policy guide? Recommendation: Any additions or holes in the NIA should be identified and sent to the author prior to the end of FY 2010.

Comment - Chart 30 "rdf is a standard model for data interchange" – you might want to say: is a powerful mechanism for doing data interchange or is a targeted standard (there simply may not be any cases of current use).

Query Service Strategy and Scope: - is underspecified, will there be a common query engine or a common query service that nodes can use to build additional query browsers or interfaces? Recommendation: Specify the architecture and make it part of a focused review.

Comment: Change all your references of "system of registries" as it might be mistaken for the Ames project "System of Registries" that professes to use similar technologies.

Storage Architecture:

Developing a Storage Strategy – For the next decade, it is likely that the storage demands will increase not only because of the number of new missions carrying instruments, but because of increases in data density and richness. Currently the strategy appears to be keeping the storage at the nodes. This may make it difficult to adopt lower cost solutions that are both environmentally and computationally advantageous. During the review it was indicated that LRO data would increase the overall storage baseline of PDS by an order of magnitude. Has analysis been performed (or is it planned to be performed) to assure that the storage requirements are met for the next three years and has it been appropriately budgeted for the out years? Increased storage needs for PDS will have a direct impact on the requirements for archival / preservation storage. Is a common approach being considered? Recommendation: Perform the analysis to determine advantages, disadvantages, trades for a uniform storage solution that provisions analytic cluster/cloud services (e.g. Hadoop), and provisions for archival needs.

Comment - A communication Strategy to PDS customers that are either instrument team's planners or mission planners/simulators (e.g. LSOS) or modeling services (e.g. LMMP). All of which may be developing tools within their environments. Bringing these communities in to the PDS v4 discussion may be beneficial to their own development targets as they may leverage PDS V4 capabilities. In turn it may help nurture a more robust developer's community for tool sharing and standards participation.

Dave Linick

1. The proposed architecture for PDS-4 appears to be appropriate. It exploits the state of the practice technologies, positions the PDS to provide better and easier access to the users, including more support for cross-node and multi-disciplinary users, and strikes a good and tunable balance between a centralized and decentralized approach. It allows for continued evolution of the PDS as the customer needs evolve. It does this in a way which allows the transition from PDS-3 to be done in a gradual, considered, and flexible way.

2. The flexibility inherent in this architecture provides hooks for its future evolution and recognizes that not all of the customer needs/desires can be accommodated in the timeframe of PDS-4 development (October, 2011). However, it is not always clear what specific capabilities will be provided on that time frame, and what will be added later as time and resources permit. This can lead to misalignment of expectations. I recommend that the specific requirements that will be satisfied by October, 2011, be crisply defined by the project and agreed to by the sponsor. This provides a measurable way of assessing the success of the project and assures that the project is clear about what is expected.

   (One specific expectation that should be clarified is when no new project will be permitted to use assume PDS-3.)

   Perhaps one approach to this that recognizes the uncertainty in the required implementation effort and schedule is to prioritize the requirements and indicate which of those the project commits to have completed in the required time frame. The remaining requirements can be satisfied if time and resources allow, but are not "guaranteed." They also provide a blueprint for future evolution.

   I further recommend that a way of doing this is to conduct a formal level 4 requirements review for each of the 6 projects. These requirements should be specific enough to provide a clear indication of what is to be done by that project. Review participants should include representatives from the stakeholders – customers, nodes, and sponsors.

3. There is a great deal of interest in international "interoperability." The PDS should implement the architecture so as not to preclude this, but needs to remember that this is a NASA effort. Nothing should be done to enable interoperability that detracts from meeting the requirements to satisfy the NASA community.

4. Difficult for the Board to assess the adequacy of resources and schedule. The result of the requirements reviews, and drawing the circle around which requirements represent commitments, will permit a better analysis of this.

Attachment 2

Requests for Action

(RFAs)

Title:	**PDS 2010 System Review**	**Date:**	**22-24 Mar 2010**
Author:	**Jan Merka**	**Email:**	**jan.merka@nasa.gov**
Topic:	**The core data model and data dictionary are crucial for distributed queries**		

Comment/Concern: The data model and dictionary are critical components of the PDS4 architecture but their development is not much discussed.

The data model is planned to consist of 'core' and 'local/extended' parts where the latter will be created by DNs or specific data domains. The content of the core data model will however determine the type and features of distributed queries possible within the PDS system. The model development and management plans unfortunately lack sufficient details how this will be achieved.

Recommendation:

- **The model development should be fast-paced and firmly controlled to allow its quick employment.**
- **The model should be comprehensive and well tested with strict core rules.**
- **The extended/local models should be directed by templates and/or examples provided by the core model developers.**

Title:	**PDS 2010 System Review**	**Date:**	**22-24 Mar 2010**
Author:	**Jan Merka**	**Email:**	**jan.merka@nasa.gov**
Topic:	**Data node holdings overlap in content**		

Comment/Concern: It is unclear to what extent data node holdings overlap in content and how is this managed.

The PDS data nodes archive products but sometimes a product might be (for convenience or performance reasons) available from several data nodes. The presentations did not touch on this issue, so it is unclear whether this situations in fact occurs at PDS and if yes how is it handled. In particular, how would this affect the archiving process and/or the distributed searches (cross-DN searches).

Recommendation:

- **Please provide a brief discussion addressing my concerns.**

Title:	**PDS 2010 System Review**	**Date:**	**22-24 Mar 2010**
Author:	**Jan Merka**	**Email:**	**jan.merka@nasa.gov**
Topic:	**How much NASA funding goes towards supporting collaboration with international partners?**		

Comment/Concern: PDS2010 is a NASA-funded project that needs to work closely with international partners. The budget however does not specify what portion of funding goes towards this task.

International collaboration is extremely important for PDS2010 functionally and, ultimately, for its success. The PDS budget should address, at least estimate, how much effort is dedicated to supporting international collaboration.

Recommendation:

- **Estimate how much effort will go towards supporting the international components and parts of PDS2010**

Title:	**PDS 2010 System Review**	**Date:**	**22-24 Mar 2010**
Author:	**Jan Merka**	**Email:**	**jan.merka@nasa.gov**
Topic:	**What is PDS2010 relationship to the NASA Virtual Observatories?**		

Comment/Concern: The NASA Virtual Observatories (VxOs) have been developing distributed searches driven by SPASE metadata for several years and their work may be relevant to PDS2010.

It was not clear to me whether the PDS team considered technologies and approaches used by the VxOs. They may not be applicable for the PDS2010 environment but a brief discussion would be appropriate.
Several question about the potential PDS-VxO relationship come to mind, for example: Will the VxOs point to or search the PDS data? Will PDS leverage any VxO services? Will there be a translation of the PDS data model to/from the SPASE data model?

Recommendation:

- **Please address the PDS-VxO interactions, if any.**
- **If the VxO are irrelevant for PDS2010, briefly comment on the reasons.**

Title:	**PDS 2010 Review**	**Date:**	**22-24 March 2010**
Author:	**Dave Heather**	**Email:**	**dheather@rssd.esa.int**
Topic:	**Documentation Inconsistencies and comments**		

Comment/Concern:

- Harvest Tool SRD/SDD section 8.0, and Registry Service SRD/SDD section 8.0: '…no planned phasing with regard to the implementation with all planned capabilities available in Build 1.' Does the Registry Service not rely on input from the Harvest Tool? There must be some sort of phasing for these in order to work.
- Report Service SRD/SDD: I don't understand or see the difference between Use cases described in Sections 3.3. and 3.4. Please clarify.
- PDS 2010 Operations Concept: Section 4.3 mentions a 'Registry/Inventory Service' and a 'Harvest/Ingest service'. Should be plain 'Registry Service' and probably just 'Harvest service'.
- PDS2010 Operations Concept: Section 5.2.17. The first point here states that the 'Data Provider assembles a peer-review committee'. Should it not be the PDS Node that does this?

Recommendation: Clarify and update documentation as necessary

Title:	**PDS 2010 Review**	**Date:**	**22-24 Mar 2010**
Author:	**Peter Shames**	**Email:**	**peter.shames@jpl.nasa.gov**
Topic:	**Support for global data searches (#1 intro, #16 architecture, and elsewhere)**		

Comment/Concern: ***The capabilities for global and targeted data searches have yet to be developed.***

The team has quite rightly deferred working on this until some of the seminal data modeling and data standardization work has been done and until the ingest functionality has been defined. Given available resources this is entirely understandable. This work seems to be a next priority.

Recommendation: My recommendations:

- As part of the design process identify next, in detail, how the global and targeted searches work within the architecture
- Ensure that global searches are adequately supported while also supporting the different types of targeted and local searches that have been identified

Title:	**PDS 2010 Review**	**Date:**	**22-24 March 2010**
Author:	**Dave Heather**	**Email:**	**dheather@rssd.esa.int**
Topic:	**Harvest and Registry Tool Comments**		

Comment/Concern: Just a few comments that may need clarification in the documentation.

The Harvest tool requires cfg file input by Discipline Node, but it is not clear to me why this should be needed. The purpose of the tool is to crawl through the data holdings and identify new or updated products. Why not automatically scan for new items / time stamps in the data holdings using an xml image to map the last run of the tool?

It is not clear from the documentation if the possibility will exist for the Harvest and Registry tools to handle the situation where there is a need to step back to an older, previously ingested version of a data product or data set. For example, if a data provider releases a newly calibrated product, and later finds serious issues with the calibration, they may request to remove access to the latest file and to make the previous version available. I assume this could work with the 'Deprecate Artifact' Registry function, but it was not 100% clear from the documentation that this had been considered.

Recommendation: Clarify the points in the documentation.

Title:	**PDS 2010 Review**	**Date:**	**22-24 Mar 2010**
Author:	**Peter Shames**	**Email:**	**peter.shames@jpl.nasa.gov**
Topic:	**Improve architecture description and understanding (#17 Service Design)**		

Comment/Concern: ***#17, pg 7 is very confusing, looks like layered chart but it really just a list of standards related to client, presentation, and logic layer clients and servers, needs a better approach such as RASDS or SCA***

Most of the architecture discussion (aside from data architecture) is focused on software architecture, little truly addresses system architecture. This chart is an example. While it looks like a representation of a layered architecture it is little more than some notional representation of layers along with a set of standards that might be relevant to the layer. In order to make clear what is actually being designed and implemented it would be really useful to have some clear diagrams that show each service element, its service interfaces and technical bindings, and how it relates to (and uses) other PDS 4 services in the process of delivering user services.

A separate but related issue is that the physical (deployment) architecture was never really shown. I have no clear idea of what class of processors are required, how functions are allocated to them, how failover is handled, what storage is needed, what network bandwidth is required, nor any comparison to what is available. It appears that this is not an issue since we were told that commodity class servers were being used, but fielding a highly reliable distributed system (a design goal) requires that these issues be considered and documented. Similarly there was no discussion of planned system performance.

Recommendation: My recommendations:

- Update the design materials for the architecture to make the actual service architecture and layering clear, derive these from some existing documented method, such as RASDS, SOA-RM, or SCA
- Clarify the actual service interfaces, technical / protocol bindings for each service element
- Document how these elements interact in the delivery of user services
- Define a physical / deployment view for the architecture showing hardware, software deployment, network and storage capacities for all nodes in the distributed system
- Simulate or at least prototype how the distributed system will operate under something like real conditions and evaluate end to end performance

Title:	**PDS 2010 Review**	**Date:**	**22-24 Mar 2010**
Author:	**Peter Shames**	**Email:**	**peter.shames@jpl.nasa.gov**
Topic:	**Improve monitor data (#17 Service Design)**		

Comment/Concern: ***Suggest that the PDS should include monitor data reflecting performance of all nodes as a measure of how well they handle requests and a signal of loading impacts and need to upgrade nodes and or links***

Related to the RFA asking that the a priori understanding of the physical / deployment architecture be improved, this RFA suggest that you may wish to evaluate and improve the sorts of monitor data that are to be provided from all of the distributed system physical and service elements. This would include gathering and analyzing a variety of performance information,

Recommendation: My recommendations:

- Analyze the expected deployment and allocation architecture of the PDS 4 system
- Identify resources that need to be monitored
- Identify key performance metrics to be gathered to allow assessment of the health and delivered performance of the system
- Develop any necessary performance requirements (which seem to be missing) and levy them on the system elements as needed
- Develop the necessary monitoring requirements and design the capabilities
- Leverage commercial capabilities as you have elsewhere

Title:	**PDS 2010 Review**	**Date:**	**22-24 March 2010**
Author:	**Dave Heather**	**Email:**	**dheather@rssd.esa.int**
Topic:	**Management of Information Model (Operation Concept Doc)**		

Comment/Concern: Section 5.2.1 of the Operations Concept document outlines the management of the information model. One of the scenarios listed is the 'Delete Operation' in which an existing class can be removed from the model.

I have a big concern here (and likewise with the removal of keywords / permissible values in the dictionary), in that it becomes extremely difficult for data providers to remain compliant if values or classes are completely removed from the model/dictionary. It is especially difficult for long missions where a version of the PDS Standards is selected and 'frozen' at the beginning of the mission. If keywords or classes these data providers use are then removed from the model during the mission, it becomes far more complex to remain compliant and to manage the information model / dictionary versions originally chosen.

No mechanism appears to be in place to manage this sort of situation in the current setup.

Recommendation: Provide a robust versioning for both the dictionary and information model to allow for full validation/ingestion/use of data compliant to older versions in cases where the model evolves during a mission. All PDS4 tools should also be developed to allow for this scenario.

Title:	**PDS 2010 Review**	**Date:**	**22-24 Mar 2010**
Author:	**Peter Shames**	**Email:**	**peter.shames@jpl.nasa.gov**
Topic:	**Tension between integrated system goals and node autonomy**		

Comment/Concern: ***There is a tension apparent between the nodes, who have and want to maintain autonomy, and the PDS users some of whom may want a more integrated system than they have had in the past. Not clear that the right balance has been struck.***

This issue never appeared in any particular set of slides, but emerged during discussions among the board, the PDS EN team, and the DN members. There is no question that the programmatic structure, with the DNs funded separately from the PDS project itself, represents an organizational challenge. It is also seems to be clear that the DNs want to maintain their autonomy, to make their own technical choices, and to meet the needs to their discipline users in the way that seems best to them.

At the same time the PDS 2010 project has goals of:

- ***On-line services allowing users to access and transform data quickly from anywhere in the system***
- ***A highly reliable, scalable computing infrastructure that*** protects the integrity of data, ***links the nodes into an integrated data system***, and provides the best service to both data providers and users

Accomplishing these is going to require a level of integration and standardization, of processes, naming, searching, user interfaces that exceeds what is presently done in PDS 3.

Recommendation: While it is clear that the whole team, EN and DNs has worked together really effectively, it appears that some further levels of integration will be required if these goals are truly to be achieved. My recommendations:

- Review the requirements to make sure that they are sufficient to lend the needed guidance to the architecture and design effort
- Ask the SMD management responsible for the PDS 2010 project and all the nodes to ensure that appropriate guidance is provided to all the elements involved in this project
- Ensure that the integrating technical elements are adequate to the task. This must include: global search terms and language, common GUIs, agreed target names and name disambiguation, integrated search mechanisms, and common data access mechanisms
- Continue the excellent work among all the nodes to ensure that the global goals are met and that local discipline user and provider needs are still satisfied

Title:	**PDS 2010 Review**	**Date:**	**22-24 Mar 2010**
Author:	**Peter Shames**	**Email:**	**peter.shames@jpl.nasa.gov**
Topic:	**Maintaining consistency among global and specialized schema (#5 Ops Concept, and elsewhere)**		

Comment/Concern: ***There is no strict means for managing consistency between user (and provider) derived schemas and new info model, but they expect to have a schema validation capability and may also define a "schema for schemas"***

This topic was discussed twice during the presentations, but the answers still sounded to me like "the hooks are in to support this, but we do not yet have a concrete plan for how to do it". Since there are a lot of opportunities for "XML hackers" to corrupt even the best defined schemas it would be appropriate to design in and agree on some processes for managing this. Similarly, there are lots of opportunities for new names for all sorts of different elements to proliferate. There are mechanisms for managing hierarchical namespaces, but there may need to be a process for identifying and resolving conflicts and promoting key elements to global status.

Recommendation: My recommendations:

- As part of the design develop agreed processes for managing the evolution of the schemas
- As part of the design develop mechanisms for identifying collisions and multiple definitions throughout the namespaces
- Develop processes for resolving ambiguities and collisions and for promoting key terms to global status

Title:	**PDS 2010 Review**	**Date:**	**22-24 Mar 2010**
Author:	**Peter Shames**	**Email:**	**peter.shames@jpl.nasa.gov**
Topic:	**Support for global name resolver (intro and elsewhere)**		

Comment/Concern: ***Not clear that there meta-data and target name consistency that would permit cross node queries to be successful. Not clear how this same problem is going to be handled in the even broader international IPDA case? Will there be some SIMBAD like name resolver (http://simbad.u-strasbg.fr/simbad/sim-fid), which seems particularly useful for all of the small bodies objects?***

This issue also surfaced during discussions among the board, the PDS EN team, and the DN members. Within the astronomy community many different names may be used for the same physical object, for that matter, the same thing happens with spacecraft, where there are MER1 & 2, MER A & B, and Spirit and Opportunity. I suspect the issue may be the most challenging for small bodies where there are a lot of them, they may have different names, and, it appears, some may even change type from comet to asteroid.

Recommendation: My recommendations:

- As part of the design process identify if some sort of global name resolver has a place in the architecture
- Determine if there needs to be one for the whole PDS 2010 or ones for specific DNs
- Define a global mechanism that makes this service available, ala Simbad, for use in queries
- Do the work to populate it (them)

Title:	**PDS 2010 Review**	**Date:**	**22-24 March 2010**
Author:	**Dave Heather**	**Email:**	**dheather@rssd.esa.int**
Topic:	**Validation layer for Node-level requirements in the standards**		

Comment/Concern: There was a concern raised during the review board discussions that there is very little structure/architecture provided for validation of the changes made to the Dictionary and Information Model by the individual Nodes.

One of the issues with PDS3 has been that the Nodes have interpreted the Standards differently from one another, and each has entered their own keywords to the Dictionary or placed restrictions on values / keywords used. Over the years this has produced a system in which data providers are given different requirements for the same data product depending on the Node to which they are delivering. Similarly, data users are presented with products that can vary in content quite significantly depending on the Node (or international partner) from which data are retrieved.

The PDS2010 system presented seemed to provide a good deal of flexibility in the dictionary and information model for use by the Nodes, which is good. However, no mechanisms seemed to be in place to validate these changes and avoid running into the situation we now have in PDS3. It is possible that, to a certain extent, this can be avoided by having a very robust PDS4 Standards Reference document, but this was not available at the time of the review. It is also not clear how the Standards Reference will provide both a clear list of requirements for PDS compliance and the specialized (and evolving) requirements of a Node. It is therefore also worth considering the possibility to add a more automated validation layer at the top-level for the changes requested by Nodes, ensuring core PDS-compliance requirements are not compromised.

Recommendation: Three recommendations arise from this concern

- Ensure that this is considered as the design process continues
- Provide a very robust Standards Reference document that will ensure that the specialized requirements of the individual nodes can be met without compromising the core requirements.
- If at all possible, put in place an automated validation layer that will ensure changes made by Nodes will not compromise the core standards

Title:	**PDS 2010 Review**	**Date:**	**22-24 Mar 2010**
Author:	**Peter Shames**	**Email:**	**peter.shames@jpl.nasa.gov**
Topic:	**Primacy of Information Model (#9 PDS Data Architecture)**		

Comment/Concern: ***#9, pg 7, there is an issue in that the info model is actually defined in an ontology, but the slide says "implemented as XML schema", this should be made clear***

From discussion it became clear that the authoritative information model for the data architecture is actually an ontology maintained using an open source tool, Protégé. This slide obscures that fact and it is really never made clear in these materials even though it is alluded to. This approach, of using a formal ontology and deriving schema, documents, and data structures from it is a real strength of the design and it should be acknowledged and taken advantage of wherever it can be.

This approach can be somewhat daunting to those unfamiliar with it, but it is very powerful and the team is to be applauded for using it and for finding ways of making derived materials accessible.

Recommendation: My recommendations:

- Update the design materials for the architecture to make this approach clear
- Leverage the approach wherever it makes sense to do so, since it is a real strength for maintaining conceptual and expressive clarity
- Sustain the processes for translating these core design materials into more easily understood and accessible forms for others to use

Title:	**PDS 2010 Review**	**Date:**	**22-24 March 2010**
Author:	**Dave Heather**	**Email:**	**dheather@rssd.esa.int**
Topic:	**Tools planned for PDS4**		

Comment/Concern: Tools presented during the review did not include the following, which should be considered in order to provide the best service to data providers and end-users:

- Conversion tools from PDS3 to PDS4 *and* from PDS4 to PDS3
- Porting of PDS4 data to the most commonly used scientific packages. Providing a replacement for NASAView style visualization is good, but scientists would want codes that would allow for software such as IDL, Envi and ISIS to open PDS4 data.

During the review, Slide 29 of the System Architecture presentation listed a large number of 'Science Related Services' including:

- Coordinate System Transformation
- Calibration on the fly
- Map overlays
- All-Purpose geometry engine

I saw no associated plans for this sort of service in the documentation and the nature of these services would make it extremely difficult to incorporate into a general PDS tool-suite.

Recommendation: Consider provision of tools for PDS3<->PDS4 transformation and for porting to the most used scientific packages.

The science related services seem to me to be primarily node-related, and should probably not be considered as part of the top-level PDS tool suite.

Title:	**PDS 2010 Review**	**Date:**	**22-24 March 2010**
Author:	**Dave Heather**	**Email:**	**dheather@rssd.esa.int**
Topic:	**Tool Distribution - comment**		

Comment/Concern: The reasoning behind distribution of tools to Nodes is unclear in many instances.

Registry Service SRD/SDD Section 6.1: It is not clear to me why the preferred deployment for the Registry Tool should be locally at Node level. If the objectives for PDS2010 include providing access to any PDS data from anywhere within the system, and moving the nodes into an integrated data system, it may be more logical to have a single Registry for the entire PDS holdings, automatically updated each time the Harvest Tool is run at each Node. I see no advantage to running the tool at the Nodes above having a centralized instance that will aid in the objectives noted above.

Report Service SRD/SDD Section 8: Why is there only a single EN instance of the Report Service? The paragraph says '…This will allow each Node to submit their metrics to a centralized location.' Does this mean the Nodes will run the tool remotely from the EN to obtain and submit their metrics?

Recommendation: Clarify the reasoning behind these choices in the documentation and/or re-think the distribution policy as needed.

Title:	**PDS 2010 Review**	**Date:**	**22-24 March 2010**
Author:	**Dave Heather**	**Email:**	**dheather@rssd.esa.int**
Topic:	**Standardization of data access / web pages at nodes**		

Comment/Concern: It was mentioned during the presentations that there are significant efforts underway to improve the look and feel of websites as part of the phasing in of PDS 2010.

It would benefit end-users and newcomers to the PDS if the look and feel of the web pages at each Node, especially the data access pages, could be standardized as much as possible. It is understood that all Nodes will have their own requirements to streamline searches against their own specialized data holdings. However, one of the aims presented during the review was to provide

- ***On-line services allowing users to access and transform data quickly from anywhere in the system***
- ***A highly reliable, scalable computing infrastructure that*** protects the integrity of data, ***links the nodes into an integrated data system***, and provides the best service to both data providers and users

A standard look and feel across all PDS sites would be very important in this case.

Recommendation: As much as possible without compromising the needs of the individual Nodes, standardize the look and feel of the websites across the PDS.

Title:	**PDS System Review**	**Date:**	**22-24 Mar 2010**
Author:	**Dave Linick**	**Email:**	**Terry.D.Linick@jpl.nasa.gov**
Topic:	**Requirements Specification**		

Comment/Concern: The flexibility inherent in this architecture provides hooks for its future evolution and recognizes that not all of the customer needs/desires can be accommodated in the timeframe of PDS-4 development (October, 2011). However, it is not always clear what specific capabilities will be provided on that time frame, and what will be added later as time and resources permit. This can lead to misalignment of expectations.

5. **Recommendation:** I recommend that the specific requirements that will be satisfied by October, 2011, be crisply defined by the project and agreed to by the sponsor. This provides a measurable way of assessing the success of the project and assures that the project is clear about what is expected.

   Perhaps one approach to this that recognizes the uncertainty in the required implementation effort and schedule is to prioritize the requirements and indicate which of those the project commits to have completed in the required time frame. The remaining requirements can be satisfied if time and resources allow, but are not "guaranteed." They also provide a blueprint for future evolution.

Title:	**PDS System Review**	**Date:**	**22-24 Mar 2010**
Author:	**Dave Linick**	**Email:**	**Terry.D.Linick@jpl.nasa.gov**
Topic:	**Transition Requirement**		

Comment/Concern:

It is unclear when project and data providers must use the PDS-4 format.

6. Recommendation:

Define and communicate a crisp statement of the requirement to comply with the PDS-4 format.

Attachment 3

Summary Debriefing to the PDS management Council

PDS 2010 System Review

Summary Review Board Debriefing (Preliminary)

March 24, 2010

Board Membership

- David Heather, ESA, ESTEC
- David Korsemeyer, ARC
- David Linick, JPL, Board Chair
- Jan Merka, GSFC
- Andy Schain, NASA HQ
- Peter Shames, JPL

Review Objectives

1. Assess technical architecture and ensure that it is responsive to the needs, drivers and requirements for the PDS over the next decade
2. Ensure PDS has a design that is responsive to the architecture and PDS drivers
3. Assess the implementation plan (schedule, resources, phasing)
4. Ensure that PDS has a deployment and infusion plan for PDS 2010 that includes PDS nodes and missions
5. Assess the transition plan from PDS3 to PDS4 operations
6. Provide overall technical and project management recommendations

Review Agenda (Day 1)

Duration	Title	Presenter	Start Time
15	Introductions; Goals	Dan Crichton	9:00 AM
30	PDS Overview, Roadmap and science drivers; mission drivers	Reta Beebe	9:15 AM
45	PDS Engineering Overview and PDS 2010 project and Technical Assessment	Dan Crichton	9:45 AM
15	Break		10:30 AM
30	Architecture Plans (System and Data)	Dan Crichton	10:45 AM
60	Concept of Operations (Planning, Design, Ingestion, Distribution)	Sean Hardman	11:15 AM
60	Lunch		12:15 PM
30	PDS3 Assessment	TBD	1:15 PM
30	PDS4 Data Standards Implementation Plan	Steve Hughes	1:45 PM
45	PDS4 Data Architecture Design (Information Model and Core Data Structures)	Steve Hughes	2:15 PM
15	Break		3:00 PM
45	PDS4 Data Architecture Design (Data Dictionary, Grammar, etc)	Steve Hughes	3:15 PM
30	PDS4 Case Studies from the Discipline Nodes	TBD	4:00 PM
30	PDS4/Mission Alignment (LADEE, MAVEN, others)	Reta Beebe	4:30 PM
480	Adjourn		5:00 PM

Review Agenda (Day 2)

Duration	Title	Presenter	Start Time
30	International Partners and IPDA	Dan Crichton	9:00 AM
30	Interoperability Case Study: Venus Express	Reta Beebe	9:30 AM
30	International Partners and China	TBD	10:00 AM
30	Break		10:30 AM
30	PDS 2010 System Implementation Plan (Development, CM, etc)	Dan Crichton	10:45 AM
60	PDS 2010 System Architecture and Design (Ingestion, Archive, Distribution)	Sean Hardman	11:15 AM
60	Lunch		12:15 PM
90	PDS 2010 IT Services (Security, Registry, …); Support for Science Services	Sean Hardman	1:15 PM
15	Break		2:15 PM
30	PDS 2010 Tools	Sean Hardman	3:00 PM
30	PDS 2010 Test Plans	Emily Law	3:30 PM
30	PDS 2010 Deployment Plan (Builds, Hardware Infra, …)	Emily Law	4:00 PM
465	Adjourn		4:30 PM

Review Agenda (Day 3)

Duration	Title	Presenter	Start Time
45	PDS 2010 Migration & Transition Planning	Emily Law/Dan	9:00 AM
30	Resource Plans and Schedule	Dan Crichton	9:45 AM
30	Wrap up	Dan Crichton	10:15 AM
60	Board Caucus and Recommendations	Board Chair	10:45 AM
165	Adjourn		11:45 AM

Major Observations

1. Proposed architecture appears to be very well-done.
   - It exploits the state of the practice technologies
   - Positions the PDS to provide better and easier access to the users, including more support for cross-node and multi-disciplinary users
   - Strikes a good and tunable balance between a centralized and decentralized approach
   - Allows for continued evolution of the PDS as the customer needs evolve.
   - Allows the transition from PDS-3 to be done in a gradual, considered, and flexible way

Major Observations - 2

2. Balance between the centralized and decentralized approach should be better defined
   - Centralization helps maximize consistency across the PDS, minimizes ambiguity, supports cross-discipline queries, etc
   - Decentralization permits discipline expertise to be applied to the specialized requirements for each node, optimize the needs of for each discipline, etc.
   - Board feels that some greater migration toward consider a stronger system-level (centralized) approach is warranted. This could be accommodated via:
     - Governance
     - Standards (including addressing some level of common "look and feel")
     - Providing tools to incentivize a greater system-level approach at the nodes

Major Observations - 3

3. Engineering Rigor: The Board recommends assuring that an appropriately tailored NPR-7120.8 be applied to the implementation of PDS-4
   - Could be addressed by applying this standard to each of the 6 implementation projects
   - One key is to assure that the requirements are specified at sufficient detail and are reviewed with all stakeholders (users, nodes, sponsor)
     - Need to assure that stakeholders and the PDS are aligned concerning what is to be provided with PDS-4 and that progress can be crisply measured
   - Assure adequate specification of the requirements and design of schemas, data dictionaries, metadata management, etc., relative to providing the needed architecture for system-wide query and search
   - Particular emphasis should be applied to:
     - Requirements on integrated search and query (which drive much of the system design/implementation)
     - Performance requirements

Assessment Against Review Objectives:

1. Achieved.
2. Achieved
3. Partly achieved. The requirements detail presented was not sufficient for the Board to evaluate the resource adequacy.
4. Mostly achieved. Not yet clearly defined when new missions/data providers will be required to use PDS-4
5. Achieved.