PDS-2010 Project Planning

PDS Planning and Assessment WG

http://pds-engineering.jpl.nasa.gov/index.cfm?pid=100&cid=118

April 3, 2008

PDS Agenda

- Background
- Status
- PDS-2010 Vision Statement
- Characteristics
- Projects
- Timeline
- Transition
- Management Plan
- Node Discussions
- MC Requests

Background

- August 2007 - PDS formed study working groups (WGs) to assess PDS3 and make recommendations for the design and implementation of PDS-2010
- December 2007- Formed a Project Planning WG
  - Members: Beebe, Crichton, Joy, LaVoie, Martin, Stein
- February 2008 & March 2008 - Met with nodes to discuss PDS-2010 project vision, structure and schedule
- March 2008 - Changed title of project to “PDS-2010”
  - Note: We are proposing that new version of the standards be PDS4, but the overall project be “PDS-2010”

Status

- Draft Level 1,2,3 Requirements
  - http://pds-engineering.jpl.nasa.gov/projects/PDS4/pds4-level123-requirements-20080123dc.pdf
- Project Plan*
  - Characteristics
  - PDS3 Issues
  - Projects
  - Timeline
  - Management Plan
- Input from Nodes

* The WG is compiling the project plan document from the node input

Preview -- MC Exec Session

- Affirmation of Vision
- Affirmation of Schedule
- Agreement on Project Structure
- Begin process of moving to next phase of project formulation (High Level Architecture)
  - Wrap up assessments and recommendations from study teams and disband (by June 1, 2008)
  - Setup design team(s) to begin system definition (June 2008)
  - Finalize Level 1,2,3 Requirements (by July 2008)
- Agree on plan to freeze PDS3

Vision

Vision Statement: With PDS-2010, PDS will provide the community with planetary science archiving standards that are international, consistent and simple to adopt and use. It will provide online services for using its data archives, allowing users to quickly access and transform data from across the federation of PDS nodes. Its data providers will be given adaptable tools that enable them to design, prepare and deliver data efficiently to PDS for archiving and distribution. Its data and services will be managed and delivered from a highly reliable and scalable computing infrastructure that is designed to protect the integrity of the data and virtually link PDS nodes into an integrated data system.

PDS4/PDS-2010 Planning Discussions with DNs*, **

PPI	February 15, 2008
GEO	February 26, 2008
Imaging	February 28, 2008
Atmospheres	March 7, 2008
Small Bodies	March 10, 2008
PSI	March 25, 2008
NAIF	March 26, 2008
Radio Science	March 26, 2008
Rings	April 1, 2008

* Summary notes posted at PDS-2010 Project Page

** Plan is to have several discussions throughout development

Characteristics of PDS-2010

- Unambiguous Standards
  - Logical, consistent, concise
- International
- Online Federation of Services
  - Data management, transformation, geometry, etc
  - Common and discipline-specific
- Archive/Data Integrity across System and all Interfaces
- Integration with Data Producers As Early as Possible
- Efficient (e.g., automation)
- High Speed Data Exchange and Data Delivery Standards
- Tracking and Reporting
  - Service usage, deliveries, etc
- Coordination and Management of the System
  - Effective governance
- Flexible Search Facilities
  - Discipline-specific
- Highly Scalable, Reliable Computing Infrastructure
- Capture extended information and documentation for users
  - Allow users to build a knowledge-base

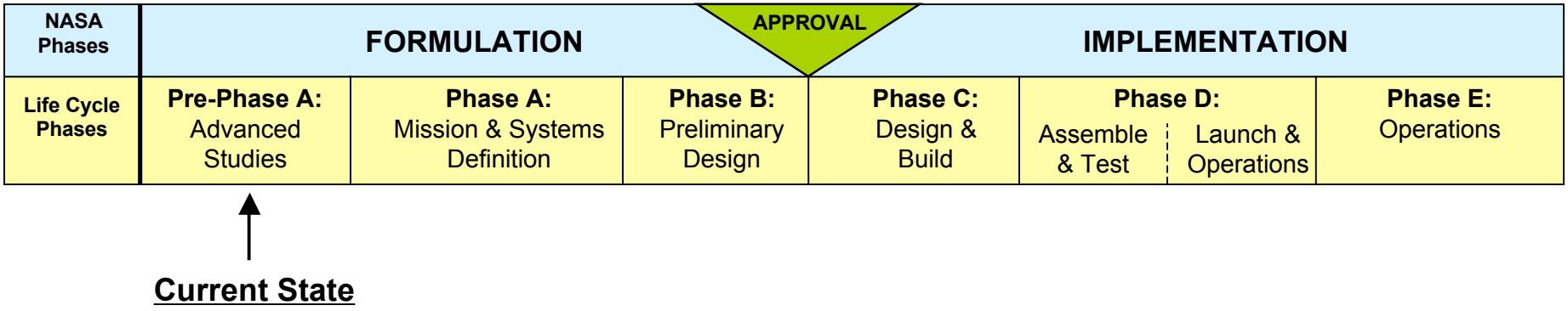
Proposed PDS Project Structure**

- Study Phase
- High Level Architecture (i.e. System Definition)
- Implementation Phases*
  - Phase I: Data Standards, Data Integrity and Core Infrastructure
  - Phase II: Catalog System, Archive Tools and Data Migration and Core Infrastructure Integration
  - Phase III: Search, Distribution, High-speed Data Exchange
  - Phase IV: Value Added Services

** Each phase has a set of projects and is aligned w/ fiscal years*

** More to come on the schedule in the next few slides

Following the Mission Lifecycle Context, where is PDS-2010?



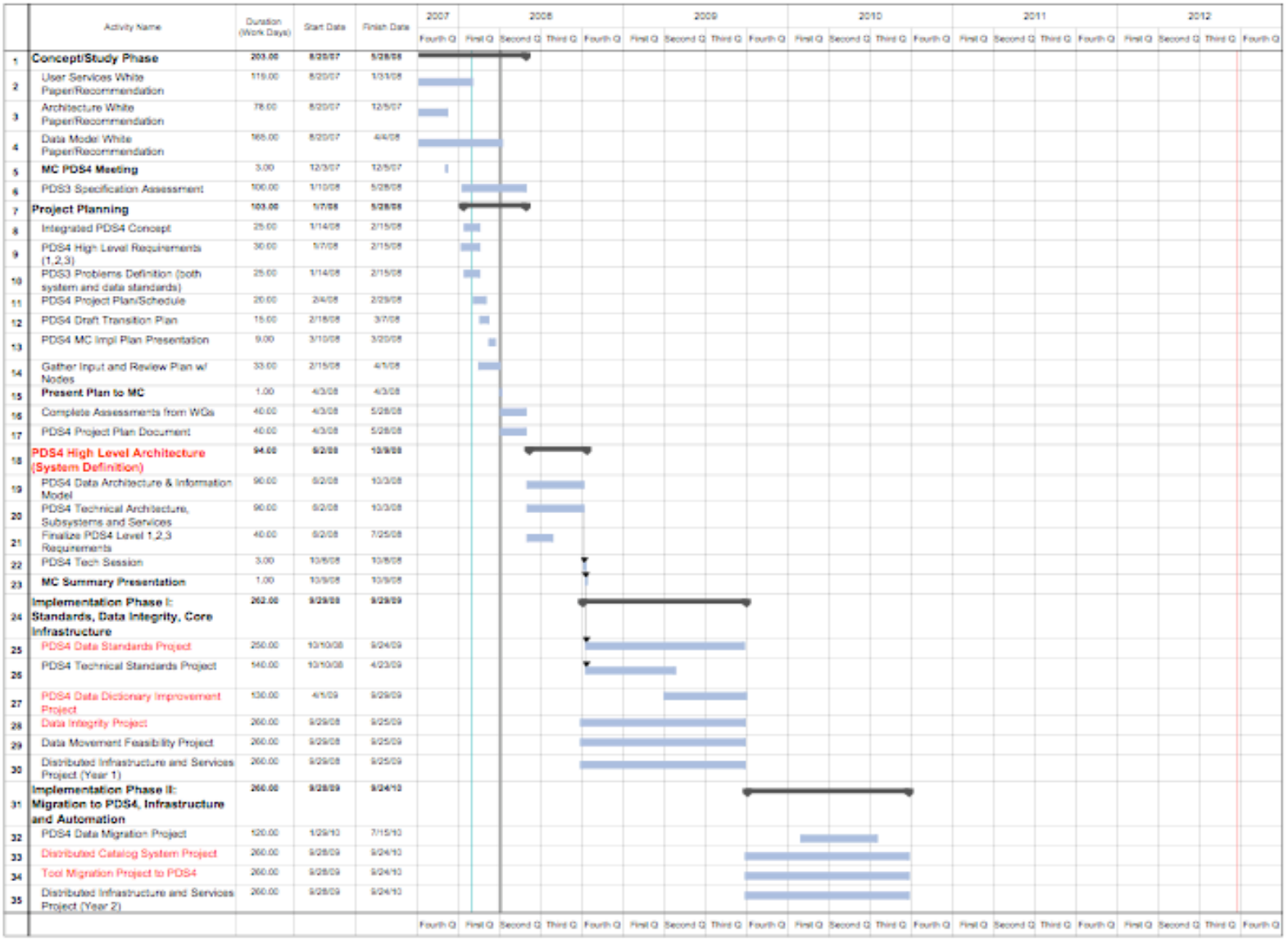
- To date, we're in the early formulation stage (i.e., Pre-Phase A)
  - PDS-2010 Study Phase

- Need to transition to start definition of the system
  - PDS-2010 High Level Architecture

Projects

- High Level Architecture
  - PDS System Architecture
    - Definition of elements of the system
    - Definition of core services
  - PDS Data Architecture
- Phase I
  - Data Standards
  - Data Dictionary
  - Distributed Services Infrastructure
  - Data Movement Feasibility
- Phase II
  - Data Migration
  - Distributed Service Infrastructure
  - Distributed Catalog System
  - PDS Portal Feasibility
  - PDS Archive Tool Development
- Phase III
  - PDS Portal, Search and Distribution
  - Provider Tools
- Phase IV
  - Node Discipline Services
  - Knowledge-base

Project Schedule (Page 1)

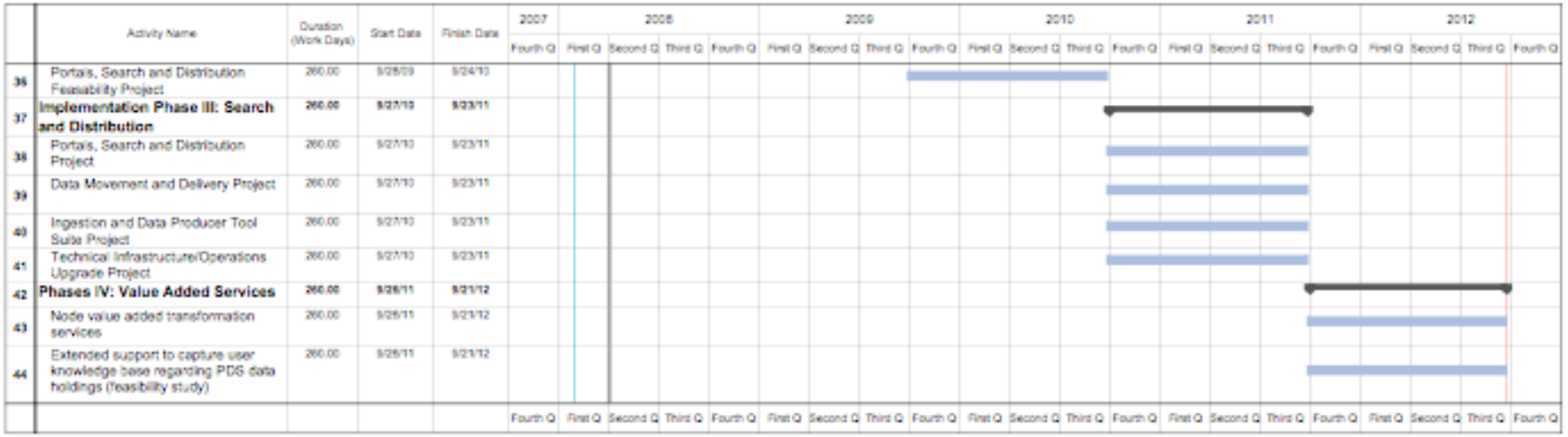


	Activity Name	Duration (Work Days)	Start Date	Finish Date
1	**Concept/Study Phase**	**203.00**	**5/25/07**	**5/26/08**
2	User Services White Paper/Recommendation	119.00	5/25/07	1/31/08
3	Architecture White Paper/Recommendation	78.00	5/25/07	12/5/07
4	Data Model White Paper/Recommendation	165.00	5/25/07	4/4/08
5	**MC PDS4 Meeting**	3.00	12/3/07	12/5/07
6	PDS3 Specification Assessment	100.00	1/10/08	5/26/08
7	**Project Planning**	**103.00**	**1/7/08**	**5/26/08**
8	Integrated PDS4 Concept	25.00	1/14/08	2/15/08
9	PDS4 High Level Requirements (1,2,3)	30.00	1/7/08	2/15/08
10	PDS3 Problems Definition (both system and data standards)	25.00	1/14/08	2/15/08
11	PDS4 Project Plan/Schedule	20.00	2/4/08	2/29/08
12	PDS4 Draft Transition Plan	15.00	2/18/08	3/7/08
13	PDS4 MC Impl Plan Presentation	9.00	3/10/08	3/20/08
14	Gather Input and Review Plan w/ Nodes	33.00	2/15/08	4/1/08
15	**Present Plan to MC**	1.00	4/3/08	4/3/08
16	Complete Assessments from WGs	40.00	4/3/08	5/26/08
17	PDS4 Project Plan Document	40.00	4/3/08	5/26/08
18	**PDS4 High Level Architecture (System Definition)**	**94.00**	**6/2/08**	**10/9/08**
19	PDS4 Data Architecture & Information Model	90.00	6/2/08	10/3/08
20	PDS4 Technical Architecture, Subsystems and Services	90.00	6/2/08	10/3/08
21	Finalize PDS4 Level 1,2,3 Requirements	40.00	6/2/08	7/25/08
22	PDS4 Tech Session	3.00	10/6/08	10/8/08
23	**MC Summary Presentation**	1.00	10/9/08	10/9/08
24	**Implementation Phase I: Standards, Data Integrity, Core Infrastructure**	**262.00**	**9/29/08**	**9/25/09**
25	PDS4 Data Standards Project	250.00	10/10/08	9/24/09
26	PDS4 Technical Standards Project	140.00	10/10/08	4/23/09
27	PDS4 Data Dictionary Improvement Project	130.00	4/1/09	9/29/09
28	Data Integrity Project	260.00	9/29/08	9/25/09
29	Data Movement Feasibility Project	260.00	9/29/08	9/25/09
30	Distributed Infrastructure and Services Project (Year 1)	260.00	9/29/08	9/25/09
31	**Implementation Phase II: Migration to PDS4, Infrastructure and Automation**	**260.00**	**9/28/09**	**9/24/10**
32	PDS4 Data Migration Project	120.00	1/29/10	7/15/10
33	Distributed Catalog System Project	260.00	9/28/09	9/24/10
34	Tool Migration Project to PDS4	260.00	9/28/09	9/24/10
35	Distributed Infrastructure and Services Project (Year 2)	260.00	9/28/09	9/24/10

2007 2008 2009 2010 2011 2012

Fourth Q First Q Second Q Third Q Fourth Q

Project Schedule (Page 2)

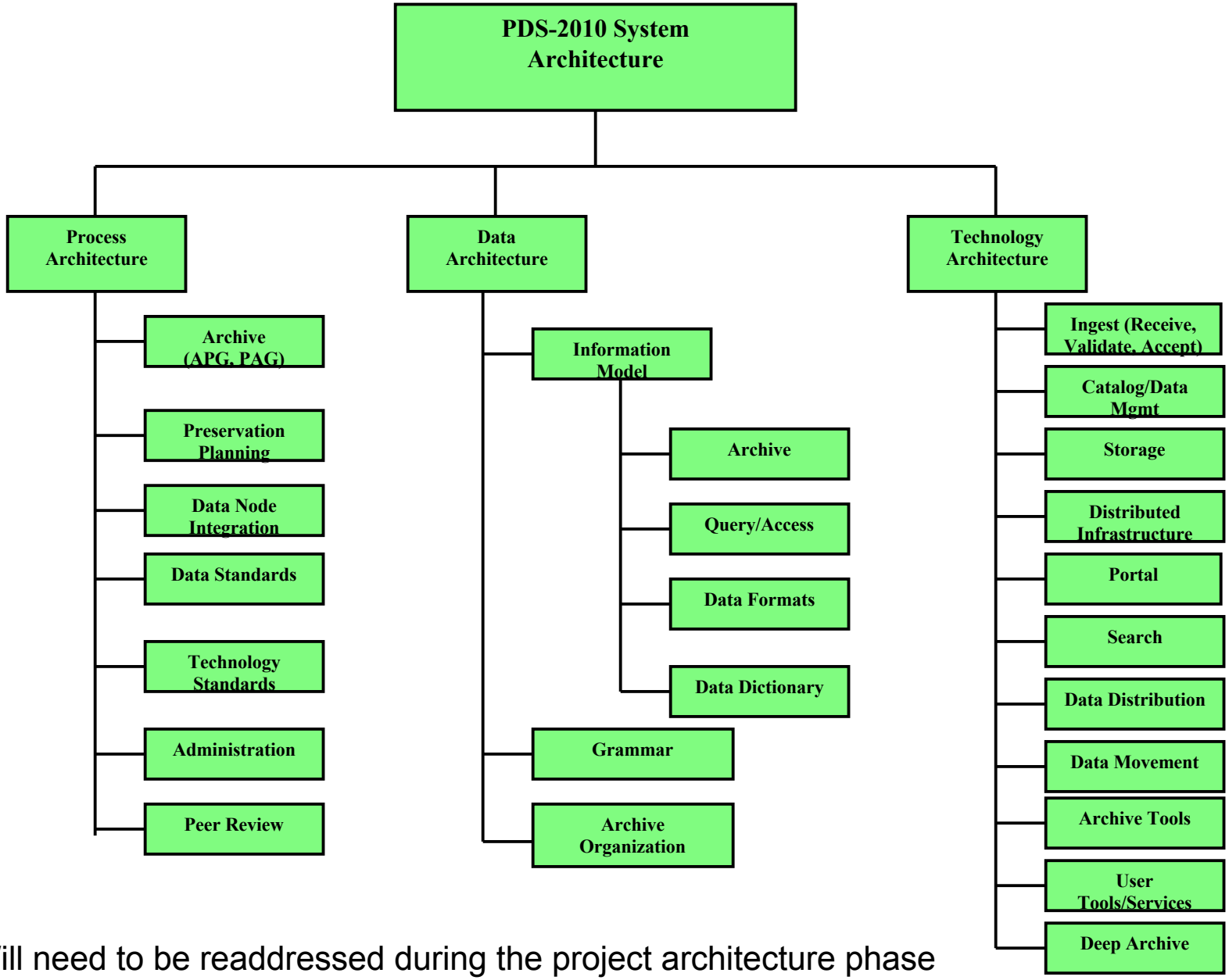


	Activity Name	Duration (Work Days)	Start Date	Finish Date
36	Portals, Search and Distribution Feasability Project	260.00	9/28/09	9/24/10
37	**Implementation Phase III: Search and Distribution**	**260.00**	**9/27/10**	**9/23/11**
38	Portals, Search and Distribution Project	260.00	9/27/10	9/23/11
39	Data Movement and Delivery Project	260.00	9/27/10	9/23/11
40	Ingestion and Data Producer Tool Suite Project	260.00	9/27/10	9/23/11
41	Technical Infrastructure/Operations Upgrade Project	260.00	9/27/10	9/23/11
42	**Phases IV: Value Added Services**	**260.00**	**9/26/11**	**9/21/12**
43	Node value added transformation services	260.00	9/26/11	9/21/12
44	Extended support to capture user knowledge base regarding PDS data holdings (feasibility study)	260.00	9/26/11	9/21/12

Transition

- Transition planning will be on-going throughout PDS-2010 development (e.g., multiple transition points)
- Dependent on new data standards
- Plan for system components
  - Retire aging system components
  - Ensure each has a transition plan
- Obvious major change will be accepting data in PDS4 standard
  - Need adequate support (documents, tools, training) for our data suppliers
  - Must accept data in either PD3/PDS4 standards for several years (e.g., Existing data providers delivering PDS3 vs New data providers delivering PDS4)
  - *Target Acceptance of PDS4 Data: End of Phase II*
- Need to decide what data to migrate

PDS 2010 System Decomposition*



* Will need to be readdressed during the project architecture phase

Component Transition Plan (1)

PDS4 Component	PDS4 Plan	PDS4 Project	Development Phase	Transition Phase
Data Architecture (Archive/Query Model, Data Formats)	Formalize and update the PDS Information Model as PDS4 based on the PDS3 data standards, issues and problems and recommendations of the PDS Information Model WG.	PDS4 Information Model Project	Architecture Design	N/A
Technology Architecture	Formalize the system architecture of PDS identifying the decomposition of the system, the core design patterns (e.g., ingest, search, distribution, deep archive), and initial standards.	PDS4 System Architecture Project	Architecture Design	N/A
Data Standards (Standards Reference, Grammar, Archive Organization)	Develop a new PDS standards reference based on the updated data architecture. Release final PDS3 standards reference.	Data Standards	Phase I	End of Phase I
Data Dictionary	Develop a new data dictionary structure and set attributes. Clean up of the existing keywords and values.	Data Dictionary	Phase I	End of Phase II
Technology Standards	Identify a core set of technology standards including platform, operating system and interface standards.	Technology Standards	Phase I	End of Phase I
Distributed Infrastructure	Develop, implement and deploy a distributed service architecture for PDS including the ability to search, access, download, and transform PDS data holdings. Replace PDS-D product, profile services with the new service architecture. Ensure the service architecture can support PDS-D (REST-based) and IPDA services.	Distributed Infrastructure and Services	Phase I, Phase II	End of Phase II
Preservation Planning	Address long term preservation planning including QQC and data usage requirements.	Data Integrity	Phase I	Phase I

Component Transition Plan (2)

PDS4 Component	PDS4 Plan	PDS4 Project	Development Phase	Transition Phase
Storage	Ensure that PDS data holdings are managed in both a primary and secondary storage with geographic separation.	Archive Integrity	Phase I	End of Phase I
Deep Archive	Deliver all PDS data holdings to the NSSDC.	Archive Integrity	Phase I	End of Phase I
Catalog/Data Management	Develop the catalog infrastructure including a centralized data set catalog as well as distributed catalog at all PDS nodes. Support multiple design patterns for populating the distributed catalog system including mechanisms for distributed access and updates.	Distributed Catalog System	Phase II	End of Phase II
Ingest	Deliver tools to support preparation, delivery and ingest of data into PDS	Ingestion and Data Producer Tool Suite	Phase III	End of Phase II
Archive Tools	Migrate VTOOL and LTDTOOL to PDS4 data standards. Ensure the tools can continue to support PDS3 and PDS4.	Tool Migration to PDS4	Phase II	End of Phase II
Portal Search Data Distribution	Replace the current portal structure, search and distribution architecture so it is service based. Upgrade portal software so it is served by a content management system for publishing information on the PDS website. Decommission old PDS portals.	Portal, Search and Distribution	Phase III	End of Phase III
Data Movement	Deliver tools and a package structure for moving PDS data holdings.	Data Movement	Phase III	End of Phase III
Data Management (Repository)	Migrate PDS data to PDS4 structure	Data Migration	Phase III	End of Phase III
User Tools (Services)	Deploy new domain-specific services that are built using the PDS distributed service software and standards.	Node-specific Services	Phase IV	Phase IV +

Other Transition Items

- Archive Process Documents (PAG, APG)
  - Revisit to ensure they are up to date for PDS-2010 (End of Phase II as PDS prepares for accepting new data products)
- Data Node Integration
  - Revisit data node white paper to ensure it is up to date relative to PDS4 system architecture
- Administration
  - Revisit PDS system processes
- Peer Review
  - Revisit to ensure it is up to date relative to PDS-2010

Management Plan

- Structure
  - Multiple Projects
    - Each with an implementation plan, requirements, deliverables and schedule
  - Engage discipline nodes to ensure critical discipline input
    - Continue nodes visits and discussions
  - We can't define everything in the formulation phase up front
- Implementation Team
  - Implementation team consists of members involved in development and integration. Nodes are critical to the integration aspects of the implementation
- Operations
  - Develop an integrated view of our system operations and provide periodic presentations to PDS MC.
- Reporting
  - Reporting to the MC will be cross-cutting covering development, integration and operations which will span all nodes

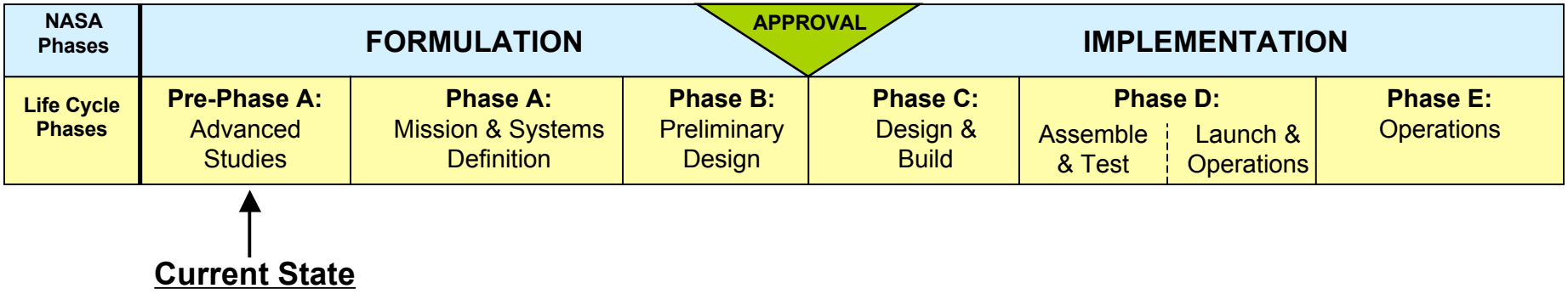
Summary of Node Discussions

- Scope of change of data standards is the largest difference among the nodes
  - A solid assessment of PDS3 is critical to understanding how to improve PDS4 version of standards
  - Almost all nodes suggest the need for a smaller number of data formats, actually the descriptions of data formats. A few nodes have suggested the need to define a few PDS data formats, the data structures being described.
- Agreement that archiving is still a primary responsibility for PDS and should be a prominent part of the PDS-2010 plan
- Support for transformation and processing of data differs greatly across nodes (e.g., node-specific services)
- PDS should divest itself, where possible, of legacy software in PDS-2010

Critical Points

- Budget/schedule is largely dependent on scope of change in our standards
  - Major changes to structure, content, and data dictionary will require that all software is re-written (PDS-wide, nodes, data providers) and that PDS suppliers and users are re-trained
  - PDS will need to decide whether it migrates data forward
- Current plan addresses PDS-wide system and tools, but it will impact node software
  - Need to identify development plan for upgrading node tools and services

When do we transition to Phase A/B?



- WG Recommendation
  - Move to "System Definition" to define High Level Architecture
    - System Architecture (Services, Component Definitions)
    - Data Architecture (Information Model to Product Level)
  - Plan tech session at end of FY to discuss
  - Discuss transition to Phase I projects at Fall MC

For MC Exec Session

- Affirmation of Vision
- Affirmation of Schedule*
- Agreement on Project Structure
- Begin process of moving to next phase of project formulation (High Level Architecture)
  - Wrap up assessments and recommendations from study teams (by June 1, 2008)
  - Setup design team to define system definition (June 2008)
  - Finalize 1,2,3 Requirements (July 2008)
- Define when to freeze PDS3

* This doesn't mean the schedule and/or projects can't change, but that we have a schedule that can be used to drive us forward

Backup

PDS3 Issues and Challenges that Need to be Addressed

- Ambiguous Standards
- Ad hoc storage infrastructure
- Growth in data sizes
- Legacy software that is difficult to maintain
- Varying degrees of search services
- Not well integrated
- International missions and data sharing
- Changes to our data standards are frequent and inefficient
- Mission support challenges; we want to get in as early as possible