PDS-D – THE PLANETARY DATA SYSTEM DISTRIBUTION SUBSYSTEM. J. S. Hughes, S. Lavoie, J. Wilf, R. Joyner, and D. Crichton, NASA Jet Propulsion Laboratory (4800 Oak Grove Drive, Pasadena CA 99109-8001).

Introduction: The Planetary Data System (PDS) Distribution Subsystem (PDS-D) provides on-demand, web-based search, retrieval, and distribution of science data products from a loosely coupled collection of distributed heterogeneous data repositories that comprise the PDS archive. The development of PDS-D was initiated when it was realized that it would be cost prohibitive to use CD/DVD media to distribute the large volumes of data expected from future missions. Requirements gathering for PDS-D began at the end of 2001 and by October 2002, the subsystem supported the first release of Mars Odyssey data. [3] The development of PDS-D caused minimal impact to existing PDS resources while it maintained the geographically distributed and discipline oriented nature of the PDS.

PDS-D Architecture: The PDS-D subsystem was implemented using a multi-tiered architecture. The top tier includes user interfaces and applications that provide data search, retrieval, and transformation functionality for science users. The second tier consists of middleware that supports message-driven processing. A middleware request broker manages service request messages from top tier client applications to server applications. Two server applications, the profile and product servers, provide search and retrieval functionalities and interface to catalogs and data repositories in the bottom tier of the architecture.

PDS-D Components: *Metadata.* Since its inception in the mid 1980's the PDS has focused on the development and capture of metadata to describe the science data in the archive. This metadata, captured in data product labels and catalog databases, provides the means to search, correlate, and use the science data. It also allows the creation of self-describing science data packages. This wealth of metadata is key to the success of PDS-D since it supports effective communications between the distributed PDS-D components.

Middleware. Message-driven processing software or middleware uses a request broker to handle service request messages from clients to server applications. The message-driven paradigm addresses both interface as well as scalability issues since the number of component interconnections increases linearly as new components are added. The Object Oriented Data Technology (OODT) framework [2] is the foundation for the PDS-D middleware and provides the messaging mechanism, product and profile servers, distributed server management, and plug-in capabilities for user tools. The PDS-D middleware is configured as a single downloadable package and was installed at every node participating in PDS-D.

Product Server. Product servers provide a common system interface to differing data repositories for data product retrieval. For the Odyssey delivery of PDS-D, a standard product server was installed at each of the PDS nodes distributing Odyssey data. A product server accepts one or more data product identifiers and returns the requested results. Any user application can request service from a product server through a standard HTTP or JAVA interface.

Profile Server. Profile servers provide a common system interface to differing data catalogs for data product search. For the PDS-D delivery supporting Mars Odyssey, a single profile server was installed for searching the master data set catalog. The profile server accepts a query for data sets and web-based resources and returns the description of each match in a standardized XML document. A data set view interface integrates the results so that users can review the results and subsequently link to all web resources associated with the data set.

Data Set Browsers. Data set browsers are top tier user applications that provide science users with search, browse, retrieval, and transformation functionality. The PDS-D architecture allowed existing PDS user applications such as the Planetary Atlas, to be integrated into the architecture as "custom" data set browsers. Such custom browsers access data products from both new and existing data repositories using the common product server interface. "Default" data set browsers were quickly developed and installed for data sets that had no existing browsers.

Data Repository. A PDS-D data repository is essentially one or more PDS archive volumes copied to system disk under a standardized directory layout. Since a PDS-D requirement is that all distributed data be PDS standards compliant, the repository contains all the necessary information to search and correlate data products. For the Odyssey delivery of PDS-D, each of the participating discipline nodes configured a data repository. An additional repository at the Arizona State University provided access to the large 4TB data set. In addition, a centralized repository was configured for backup and was used in several instances when Internet connections to remote repositories were slow.

Data Notification and Subscription. A new requirement to provide access to data as soon as it was available necessitated the release of subsets of data sets. PDS-D allows users to subscribe for notification

of such releases. The notification e-mail includes direct links to the resources providing access to the data.

Conclusions: The PDS-D subsystem provides seamless access to distributed heterogeneous science data repositories and catalogs. It is a data driven architecture and supports location independence and information hiding. It is also scalable and extensible and provides client APIs for product search and retrieval. Most importantly it maintains geographically distributed data repositories and thereby conforms to CODMAC recommendations for discipline data systems by keeping data in the hands of the scientific experts and promoting closer ties with mission instrument teams. By reconfiguring existing PDS resources using a multi-tiered, message-driven architecture, PDS-D leveraged and had minimum impact on existing PDS system resources and is now successfully distributing Mars Odyssey data.

References: [1] Slavney, S., Arvidson, E., Guinness, E.A. (2002) LPSC XXXIII 1303. [2] Crichton, D. et al., (2000) CODATA Science Search and Retrieval using XML and http://oodt.jpl.nasa.gov/ [3] LaVoie, S., (2002) PDS-D01 Quarterly Status Report.

Acknowledgements: The authors wish to acknowledge the following individuals for their efforts in the successful development of PDS-D. OODT Team:Sean Kelly, Thuy Tran, Shu Lui, Jerry Crichton ;PDS Central Node:Cynthia Hall-Atkinson, Chris Hawley; PDS Discipline Nodes:Pam Woncik, Elizabeth Duxbury-Rye, Ed Guinness, Susie Slavney, Tom Stein, Noel Gorelick, Lyle Huber, Todd King, Patty Garcia, Lee Elson, Steve Levoe, Dick Simpson.