

A horizontal banner image featuring a sequence of celestial bodies from left to right: a blue planet (Earth), a brown planet (Mars), a brown planet (Jupiter), and a brown planet (Saturn). The text "Planetary Data System" is overlaid in white on the right side of the banner.

Planetary Data System

# **System Ingestion Demonstration**

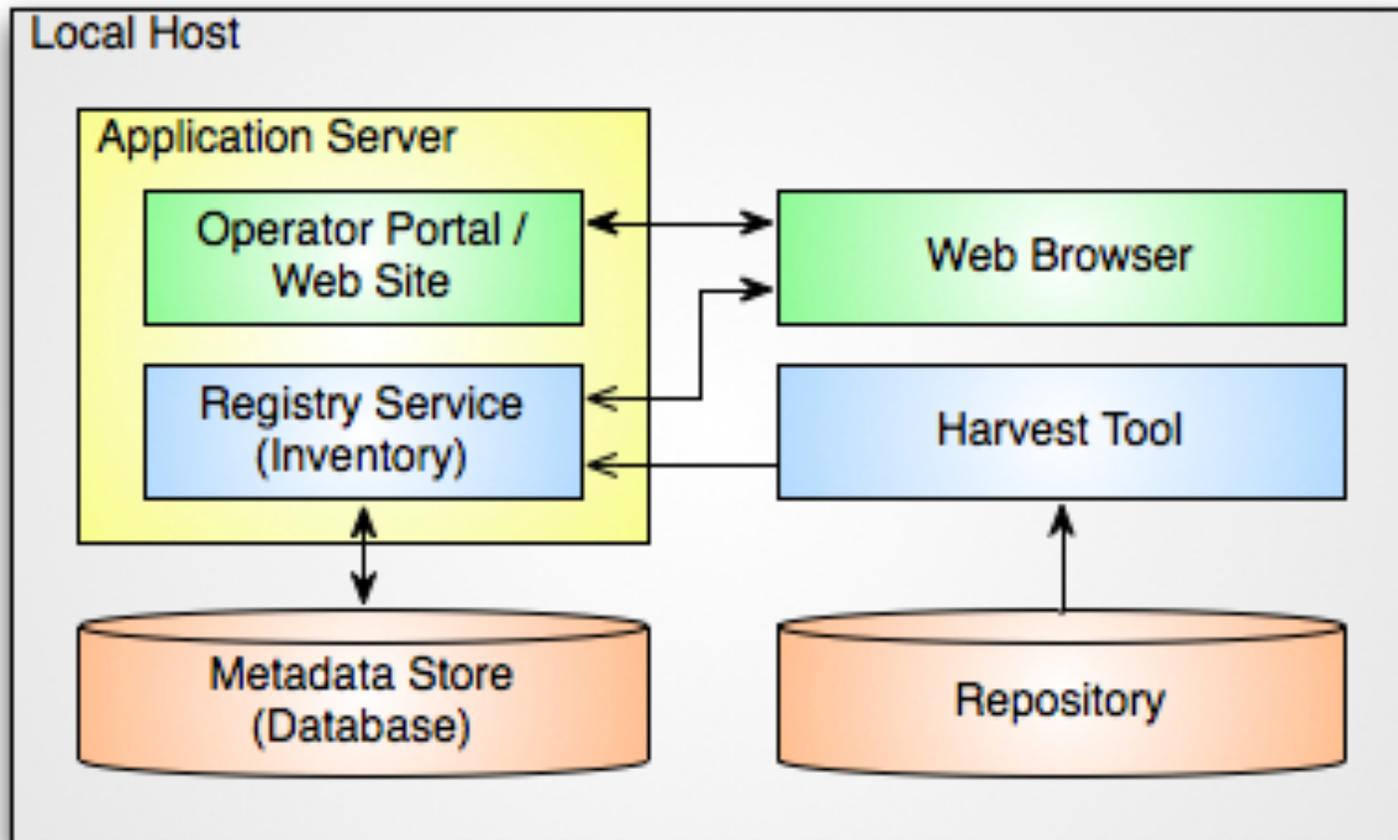
Technical Session  
Pasadena, CA  
February 28 – March 2, 2011

Sean Hardman  
Paul Ramirez

# Overview

- The demo exhibits the progress made in core component development and integration.
- The Harvest and Registry UI components utilize the API interface of the Registry Service.
  - Several examples of the REST-based API are on display as well.

# Deployment (Local Host)



# Installation Steps

1. Download the distribution packages (Registry Service, Registry UI, Harvest) from:  
<http://pds-engineering.jpl.nasa.gov/pds2010/development/1.0.0/>
2. Unpack the packages.
3. Launch an Apache Tomcat server on your local machine.
4. Follow the installation instructions for the Registry Service and Registry UI packages.
5. Deploy the WAR files to the Apache Tomcat server.
6. Verify the installation with the Service and UI.

# Engineering Node Site

## Release 1.0.0 Page

PDS 2010 - About PDS 2010

http://pds-engineering.jpl.nasa.gov/pds2010/development/1.0.0/

NASA NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

+ NASA Homepage  
+ NASA en Español  
+ Contact NASA

Search the Engineering Node:

Planetary Data System  
Engineering Node

Home Standards Tools Contact Us Feedback

PDS 2010

Last Published: 2010-11-01 Dev Home | PDS 2010

**Documentation**  
**About PDS 2010**  
System Architecture (pdf)  
System Requirements (pdf)  
System Releases

**Components**  
Harvest  
Preparation  
Registry  
Report  
Security

**Project Documentation**  
▶ Project Information

Built by: **maven**

**About PDS 2010**

The PDS 2010 effort will overhaul the PDS data architecture (e.g., data model, data structures, data dictionary, etc) and deploy a software system (online data services, distributed data catalog, etc) that fully embraces the PDS federation as an integrated system while leveraging modern information technology.

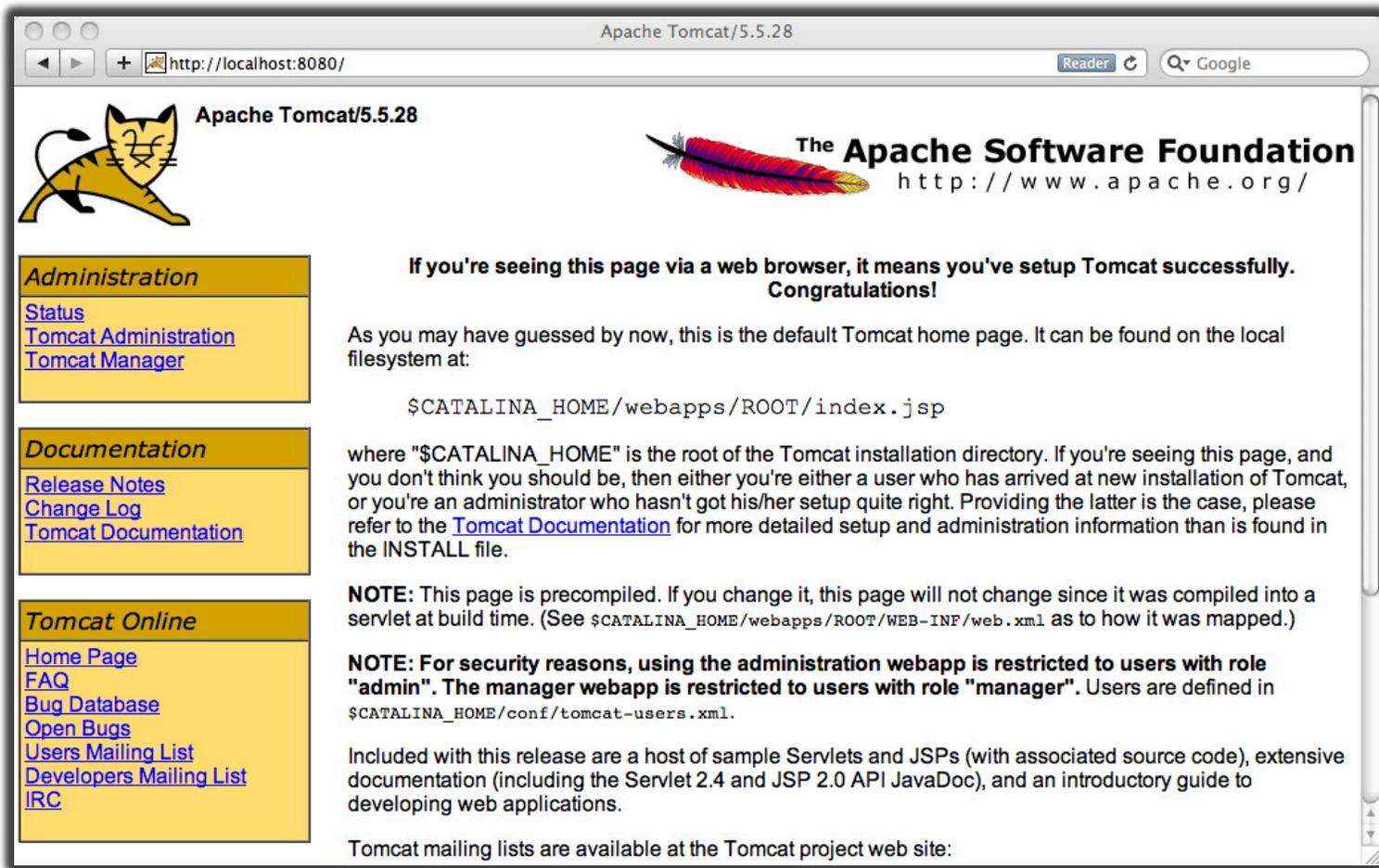
Copyright © 2005-2010, by the California Institute of Technology.  
ALL RIGHTS RESERVED. United States Government Sponsorship acknowledged.

Feb 28 - Mar 2, 2011

System Ingestion Demonstration

5

# Apache Tomcat



The screenshot shows a web browser window with the title "Apache Tomcat/5.5.28". The address bar shows "http://localhost:8080/". The page content includes the Apache Tomcat logo (a yellow cat), the Apache Software Foundation logo (a red feather), and the text "The Apache Software Foundation http://www.apache.org/". The main content area features a congratulatory message: "If you're seeing this page via a web browser, it means you've setup Tomcat successfully. Congratulations!". Below this, it explains that the page is the default Tomcat home page and provides the file path: `$_CATALINA_HOME/webapps/ROOT/index.jsp`. It also includes two notes: one about the page being precompiled and another about security restrictions for the administration webapp. A sidebar on the left contains navigation links under three categories: "Administration" (Status, Tomcat Administration, Tomcat Manager), "Documentation" (Release Notes, Change Log, Tomcat Documentation), and "Tomcat Online" (Home Page, FAQ, Bug Database, Open Bugs, Users Mailing List, Developers Mailing List, IRC).

Apache Tomcat/5.5.28

The Apache Software Foundation  
<http://www.apache.org/>

**If you're seeing this page via a web browser, it means you've setup Tomcat successfully. Congratulations!**

As you may have guessed by now, this is the default Tomcat home page. It can be found on the local filesystem at:

```
$_CATALINA_HOME/webapps/ROOT/index.jsp
```

where "\$CATALINA\_HOME" is the root of the Tomcat installation directory. If you're seeing this page, and you don't think you should be, then either you're either a user who has arrived at new installation of Tomcat, or you're an administrator who hasn't got his/her setup quite right. Providing the latter is the case, please refer to the [Tomcat Documentation](#) for more detailed setup and administration information than is found in the INSTALL file.

**NOTE:** This page is precompiled. If you change it, this page will not change since it was compiled into a servlet at build time. (See `$_CATALINA_HOME/webapps/ROOT/WEB-INF/web.xml` as to how it was mapped.)

**NOTE:** For security reasons, using the administration webapp is restricted to users with role "admin". The manager webapp is restricted to users with role "manager". Users are defined in `$_CATALINA_HOME/conf/tomcat-users.xml`.

Included with this release are a host of sample Servlets and JSPs (with associated source code), extensive documentation (including the Servlet 2.4 and JSP 2.0 API JavaDoc), and an introductory guide to developing web applications.

Tomcat mailing lists are available at the Tomcat project web site:

**Administration**  
[Status](#)  
[Tomcat Administration](#)  
[Tomcat Manager](#)

**Documentation**  
[Release Notes](#)  
[Change Log](#)  
[Tomcat Documentation](#)

**Tomcat Online**  
[Home Page](#)  
[FAQ](#)  
[Bug Database](#)  
[Open Bugs](#)  
[Users Mailing List](#)  
[Developers Mailing List](#)  
[IRC](#)

# Apache Tomcat Manager

The Apache Software Foundation  
http://www.apache.org/

Tomcat Web Application Manager

Message: OK

Manager

[List Applications](#) [HTML Manager Help](#) [Manager Help](#) [Server Status](#)

| Path                              | Display Name                            | Running | Sessions | Commands   |
|-----------------------------------|---|---------|----------|--|
| /                                 | Welcome to Tomcat                       | true    | 0        | <a href="#">Start</a> <a href="#">Stop</a> <a href="#">Reload</a> <a href="#">Undeploy</a> |
| /svn                              |   | true    | 0        | <a href="#">Start</a> <a href="#">Stop</a> <a href="#">Reload</a> <a href="#">Undeploy</a> |
| /balancer                         | Tomcat Simple Load Balancer Example App | true    | 0        | <a href="#">Start</a> <a href="#">Stop</a> <a href="#">Reload</a> <a href="#">Undeploy</a> |
| /host-manager                     | Tomcat Manager Application              | true    | 0        | <a href="#">Start</a> <a href="#">Stop</a> <a href="#">Reload</a> <a href="#">Undeploy</a> |
| /jsp-examples                     | JSP 2.0 Examples                        | true    | 0        | <a href="#">Start</a> <a href="#">Stop</a> <a href="#">Reload</a> <a href="#">Undeploy</a> |
| /manager                          | Tomcat Manager Application              | true    | 0        | <a href="#">Start</a> <a href="#">Stop</a> <a href="#">Reload</a> <a href="#">Undeploy</a> |
| <a href="#">/registry-service</a> |   | true    | 0        | <a href="#">Start</a> <a href="#">Stop</a> <a href="#">Reload</a> <a href="#">Undeploy</a> |
| /registry-ui                      | Registry User Interface                 | true    | 0        | <a href="#">Start</a> <a href="#">Stop</a> <a href="#">Reload</a> <a href="#">Undeploy</a> |
| /schema                           |   | true    | 0        | <a href="#">Start</a> <a href="#">Stop</a> <a href="#">Reload</a> <a href="#">Undeploy</a> |
| /servlets-examples                | Servlet 2.4 Examples                    | true    | 0        | <a href="#">Start</a> <a href="#">Stop</a> <a href="#">Reload</a> <a href="#">Undeploy</a> |
| /tomcat-docs                      | Tomcat Documentation                    | true    | 0        | <a href="#">Start</a> <a href="#">Stop</a> <a href="#">Reload</a> <a href="#">Undeploy</a> |
| /webdav                           | Webdav Content Management               | true    | 0        | <a href="#">Start</a> <a href="#">Stop</a> <a href="#">Reload</a> <a href="#">Undeploy</a> |

# Registry Status

## Initial Installation

- Navigate to the Registry Service status interface:
  - <http://localhost:8080/registry-service/registry/status>

- The following should appear in your browser:

```
<statusInformation xmlns='http://  
registry.pds.nasa.gov' classificationNodes='0'  
classificationSchemes='0' services='0'  
products='0' associations='0'  
serverStarted='2011-02-28T11:37:44.992-08:00'  
status='OK' />
```

# Product Listing

Registry Browser

http://localhost:8080/registry-ui/

GUID LID Name Submitter Object Type Status

Any Object Type Any Status Update

**Product Registry**

| Name                        | LID | Version ID | Object Type | Status |
|-----------------------------|-----|------------|-------------|--------|
| There is no data to display |     |            |             |        |

1 of 0

Num Records: 0

# Demo Steps

1. Configure the Registry Service by loading the Data Object scheme and nodes.
2. Configure Harvest Tool for the personnel collection.
3. Register the personnel collection.
4. Verify the registration with the Service and UI.
5. View product details by selecting a product.
6. View association details by selecting an association.
7. Configure Harvest Tool and register the Atmospheres collection.

# Populate Data Object Scheme and Nodes

- Execute the following from the Registry Service distribution:

```
[node:~/registry-service-0.5.0/bin] % ./load_registry
```

```
* About to connect() to localhost port 8080 (#0)  
*   Trying ::1... connected  
* Connected to localhost (::1) port 8080 (#0)  
> POST /registry-service/registry/schemes HTTP/1.1  
...
```

# Registry Status

## Post Scheme and Node Population

- Navigate to the Registry Service status interface:
  - <http://localhost:8080/registry-service/registry/status>

- The following should appear in your browser:

```
<statusInformation xmlns='http://  
registry.pds.nasa.gov' classificationNodes='50'  
classificationSchemes='1' services='0'  
products='0' associations='0'  
serverStarted='2011-02-28T11:37:44.992-08:00'  
status='OK' />
```

# Scheme Listing

- Navigate to the Registry Service scheme interface:
  - <http://localhost:8080/registry-service/registry/schemes/urn:registry:ObjectType>
- The following should appear in your browser:

```
<classificationScheme xmlns='http://registry.pds.nasa.gov'  
nodeType='UniqueCode' isInternal='true' versionId='1.0'  
versionName='1.0' description='This is the canonical object type  
classification that is one of the core registry objects'  
status='Submitted' objectType='ClassificationScheme'  
name='Canonical Object Type Classification Scheme' lid='urn:uuid:  
305f448a-7924-40a4-9757-9e495f9bf530' home='http://localhost:8080/  
registry-service' guid='urn:registry:ObjectType' />
```

# Node Listing

- Navigate to the Registry Service node interface:
  - <http://localhost:8080/registry-service/registry/schemes/urn:registry:ObjectType/nodes>
- The following should appear in your browser:

```
<response xmlns='http://registry.pds.nasa.gov'>
<results>
  <classificationNode path='/urn:registry:ObjectType/Product/Collection_Browse'
code='Collection_Browse'
parent='urn:registry:ObjectType:RegistryObject:ExtrinsicObject' versionId='1.0'
versionName='1.0' description='A Browse collection is a product that has a table of
references to one or more browse products.' status='Submitted'
objectType='ClassificationNode' name='Collection Browse'
lid='urn:nasa:pds:profile:regrep:ObjectType:Collection_Browse' home='http://localhost:
8080/registry-service'
guid='urn:nasa:pds:profile:regrep:ObjectType:Collection_Browse'/> status='Submitted'
objectType='ClassificationScheme' name='Canonical Object Type Classification Scheme'
lid='urn:uuid:305f448a-7924-40a4-9757-9e495f9bf530' home='http://localhost:8080/
registry-service' guid='urn:registry:ObjectType'/>
```

...

# Configure Harvest

## Launch Script

- Modify the Harvest script to access the appropriate Registry Service instance
- The UNIX or Windows batch script is located in the bin directory of the Harvest distribution.
- Modify the line that looks something like the following for the appropriate URL:

```
java -Dpds.registry="http://localhost:8080/registry-service"
```

# Configure Harvest Personnel Collection Policy File

- Create a configuration file for harvesting the Personnel collection:

```
<policy>
  <collections>
    <file>/Users/shardman/Demo/context_personnel/
collection_context_affil_20101206.xml</file>
  </collections>
  <candidates>
    <productMetadata objectType="Product_PDS_Affiliate">
      <xPath>//PDS_Affiliate/institution_name</XPath>
      <xPath>//PDS_Affiliate/node_name</XPath>
      <xPath>//PDS_Affiliate/affiliation_type</XPath>
      <xPath>//PDS_Affiliate/electronic_mail_address</XPath>
    </productMetadata>
  </candidates>
</policy>
```

# Register Collection

## Personnel Collection

- Execute the Harvest launch script with the policy file for the personnel collection.

```
[node:~/harvest-0.2.0/bin] % ./Harvest ../examples/  
demo-policy-personnel.xml -l personnel.txt
```

- Review the resulting log file for successful registration.

```
[avatar:~/harvest-0.2.0/bin] shardman% cat  
personnel.txt  
...  
230 of 232 files are candidate products, 0 skipped  
230 of 230 candidate products registered.  
460 of 460 associations registered.  
...
```

# Registry Status

## Personnel Collection Population

- Navigate to the Registry Service status interface:
  - <http://localhost:8080/registry-service/registry/status>

- The following should appear in your browser:

```
<statusInformation xmlns='http://  
registry.pds.nasa.gov' classificationNodes='50'  
classificationSchemes='1' services='0'  
products='230' associations='460'  
serverStarted='2011-02-28T11:37:44.992-08:00'  
status='OK' />
```

# Product Listing

## Personnel Collection Population

Registry Browser

http://localhost:8080/registry-ui/

GUID  LID  Name  Submitter  Object Type  Status

| Product Registry   |                                     |            |                    |           |
|--------------------|-------------------------------------|------------|--------------------|-----------|
| Name               | LID                                 | Version ID | Object Type        | Status    |
| THOMAS C. STEIN    | URN:NASA:PDS:personnel.TSTEIN       | v1.0       | Product_PDS_Affili | Submitted |
| PHILIPPE L. MASSON | URN:NASA:PDS:personnel.PMASSON      | v1.0       | Product_PDS_Affili | Submitted |
| NOEL S. GORELICK   | URN:NASA:PDS:personnel.NGORELICK    | v1.0       | Product_PDS_Affili | Submitted |
| GENE L. GOLTZ      | URN:NASA:PDS:personnel.GGOLTZ       | v1.0       | Product_PDS_Affili | Submitted |
| TOM MCCARTHY       | URN:NASA:PDS:personnel.TMCARTHY     | v1.0       | Product_PDS_Affili | Submitted |
| TERRY FORRESTER    | URN:NASA:PDS:personnel.TFORRESTER   | v1.0       | Product_PDS_Affili | Submitted |
| JIM S. ALEXOPOULOS | URN:NASA:PDS:personnel.JALEXOPOULOS | v1.0       | Product_PDS_Affili | Submitted |
| LARRY J. GRANROTH  | URN:NASA:PDS:personnel.LJGRANROTH   | v1.0       | Product_PDS_Affili | Submitted |
| DANIEL M. OLSON    | URN:NASA:PDS:personnel.DMOLSON      | v1.0       | Product_PDS_Affili | Submitted |
| DAVID A. KRING     | URN:NASA:PDS:personnel.DKRING       | v1.0       | Product_PDS_Affili | Submitted |
| ROGER PHILLIPS     | URN:NASA:PDS:personnel.RPHILLIPS    | v1.0       | Product_PDS_Affili | Submitted |
| LYLE F. HUBER      | URN:NASA:PDS:personnel.LHUBER       | v1.0       | Product_PDS_Affili | Submitted |
| DAVID C. BLACK     | URN:NASA:PDS:personnel.DBLACK       | v1.0       | Product_PDS_Affili | Submitted |
| ...                | ...                                 | ...        | ...                | ...       |
| Name               | LID                                 | Version ID | Object Type        | Status    |

1 of 5

**Num Records: 230**

# Product Detail Popup

(Query for \*HARDMAN\* then selected product)

The screenshot shows a web browser window titled "Registry Browser" with the URL "http://localhost:8080/registry-ui/". The main interface has search fields for GUID, LID, and Name, with the Name field containing "\*HARDMAN\*". Below the search fields is a table titled "Product Registry" with columns "Name" and "LID". The table contains one record: "SEAN HARDMAN" with LID "URN:NASA:PDS:personnel.SHARDMAN". A "Product Details" popup window is open, displaying the following information:

**Product Details**

- Name: SEAN HARDMAN
- Object Type: Product\_PDS\_Affiliate
- Status: Submitted
- Version Name: 1.0
- Version ID: v1.0
- GUID: urn:uuid:391cd103-0aed-43d7-8a19-8bda96cc8f16
- LID: URN:NASA:PDS:personnel.SHARDMAN
- Home: http://localhost:8080/registry-service
- Node Name: Engineering
- Electronic Mail Address: SEAN.HARDMAN@JPL.NASA.GOV
- Affiliation Type: NODE STAFF
- Institution Name: JET PROPULSION LABORATORY

Below the product details is an "Associations" table:

| Source LID                      | Association Type | Target LID           |
|---------------------------------|------------------|----------------------|
| URN:NASA:PDS:personnel.SHARDMAN | has node         | URN:NASA:PDS:node.EN |

The popup also includes a "Close" button at the bottom left.

# Association Detail Popup (Select the Node association)

The screenshot shows a web browser window titled "Registry Browser" at the URL "http://localhost:8080/registry-ui/". The main content area displays a table with the following data:

| GUID | LID | Name      |
|------|-----|-----------|
|      |     | *HARDMAN* |

An "Association Details" popup window is open, showing the following information:

- Association Type:** has node
- Description:**
- Logical ID:** urn:uuid:3314b5a3-ecb4-4af8-acba-780026df6138
- Name:** cd103-0aed-43d7-8a19-8bda96cc8f16
- Object Type:** Association
- Source GUID:**
- Source Home:** http://localhost:8080/registry-service
- Source Logical ID:** URN:NASA:PDS:personnel.SHARDMAN
- Source Version:** v1.0
- Status:** Submitted
- Target GUID:**
- Target Home:** http://localhost:8080/registry-service
- Target Logical ID:** URN:NASA:PDS:node.EN
- Target Version:** 1.0
- Version ID:** 1.0
- Version Name:** 1.0

The popup also includes a "Close" button at the bottom left. In the background, a table shows the record for "SEAN HARDMAN" with a "Name" field and a "Num Records: 1" indicator at the bottom.

# Product Search

## (Query for \*HARDMAN\* via HTTP)

- Navigate to the Registry Service product interface:
  - [http://localhost:8080/registry-service/registry/products?name=\\*HARDMAN\\*](http://localhost:8080/registry-service/registry/products?name=*HARDMAN*)
- The following should appear in your browser:

```
<response xmlns='http://registry.pds.nasa.gov' numFound='1' start='1'>
  <results>
    <product versionId='v1.0' versionName='1.0' status='Submitted'
objectType='Product_PDS_Affiliate' name='SEAN HARDMAN'
lid='URN:NASA:PDS:personnel.SHARDMAN' home='http://localhost:8080/registry-
service' guid='urn:uuid:391cd103-0aed-43d7-8a19-8bda96cc8f16'>
      <slot name='institution_name' id='800'>
        <value>JET PROPULSION LABORATORY</value>
      </slot>
    ...
```

# Configure Harvest

## Atmospheres Collection Policy File

- Create a configuration file for harvesting an Atmospheres collection:

```
<policy>
  <directories>
    <path>/Users/shardman/Desktop/Demo/pds4_phmt_EDR</path>
    <filePattern>*.xml</filePattern>
  </directories>
  <candidates>
    <productMetadata objectType="Product_Table_Character">
      <xPath>//Identification_Area_Product/Subject_Area/instrument_host_name
      </XPath>
      <xPath>//Identification_Area_Product/Subject_Area/instrument_name</XPath>
      <xPath>//Identification_Area_Product/Subject_Area/target_name</XPath>
      <xPath>//Observation_Area/start_date_time</XPath>
      <xPath>//Observation_Area/stop_date_time</XPath>
    </productMetadata>
  </candidates>
</policy>
```

# Register Collection

## Atmospheres Collection

- Execute the Harvest launch script with the policy file for the Atmospheres collection.

```
[node:~/harvest-0.2.0/bin] % ./Harvest ../examples/demo-policy-atm.xml -l atm.txt
```

- Review the resulting log file for successful registration.

```
[avatar:~/harvest-0.2.0/bin] shardman% cat atm.txt  
...  
280 of 282 files are candidate products, 0 skipped  
280 of 280 candidate products registered.  
280 of 280 associations registered.  
...
```

# Product Listing

(Select the Product\_Table\_Character type)

The screenshot shows a web browser window titled 'Registry Browser' with the URL 'http://localhost:8080/registry-ui/'. The page features a search and filter interface at the top with fields for GUID, LID, Name, and Submitter. The 'Object Type' dropdown is set to 'Product\_Table\_Character' and the 'Status' dropdown is set to 'Any Status'. An 'Update' button is located to the right of these filters.

Below the filters is a table titled 'Product Registry' with the following columns: Name, LID, Version ID, Object Type, and Status. The table contains 12 rows of data, all with 'Submitted' status and 'Product\_Table\_C' object type. A pagination control at the bottom shows '1 of 6' records, and a 'Num Records: 280' label is displayed below the table.

| Name                        | LID  | Version ID | Object Type     | Status    |
|-----------------------------|--|------------|-----------------|-----------|
| PHOENIX Mars MET Experiment | URN:NASA:PDS:PHX_MET-RAW:MS051EML_00900744690_15CDM1 | 1.0        | Product_Table_C | Submitted |
| PHOENIX Mars MET Experiment | URN:NASA:PDS:PHX_MET-RAW:MS050EML_00900650592_15AFM1 | 1.0        | Product_Table_C | Submitted |
| PHOENIX Mars MET Experiment | URN:NASA:PDS:PHX_MET-RAW:MS074EML_00902819822_1886M1 | 1.0        | Product_Table_C | Submitted |
| PHOENIX Mars MET Experiment | URN:NASA:PDS:PHX_MET-RAW:MS147EML_00909269503_204DM1 | 1.0        | Product_Table_C | Submitted |
| PHOENIX Mars MET Experiment | URN:NASA:PDS:PHX_MET-RAW:MS100EML_00905086235_1BB3M1 | 1.0        | Product_Table_C | Submitted |
| PHOENIX Mars MET Experiment | URN:NASA:PDS:PHX_MET-RAW:MS147EMH_00909269503_204DM1 | 1.0        | Product_Table_C | Submitted |
| PHOENIX Mars MET Experiment | URN:NASA:PDS:PHX_MET-RAW:MS007EMH_00896827901_1145M1 | 1.0        | Product_Table_C | Submitted |
| PHOENIX Mars MET Experiment | URN:NASA:PDS:PHX_MET-RAW:MS125EML_00907303840_1E75M1 | 1.0        | Product_Table_C | Submitted |
| PHOENIX Mars MET Experiment | URN:NASA:PDS:PHX_MET-RAW:MS075EML_00902833089_1887M1 | 1.0        | Product_Table_C | Submitted |
| PHOENIX Mars MET Experiment | URN:NASA:PDS:PHX_MET-RAW:MS064EML_00901908454_1745M1 | 1.0        | Product_Table_C | Submitted |
| PHOENIX Mars MET Experiment | URN:NASA:PDS:PHX_MET-RAW:MS102EML_00905260816_1BF3M1 | 1.0        | Product_Table_C | Submitted |
| PHOENIX Mars MET Experiment | URN:NASA:PDS:PHX_MET-RAW:MS070EML_00902428899_17FFM1 | 1.0        | Product_Table_C | Submitted |
| PHOENIX Mars MET Experiment | URN:NASA:PDS:PHX_MET-RAW:MS131EML_00907835235_1EF9M1 | 1.0        | Product_Table_C | Submitted |

# Wrap Up

- Development has continued for both Harvest Tool and Registry Service.
- Harvest now includes validation and crawler persistence. Currently working on support for PDS3 products.
- Registry now includes expanded support for service registration. Utilized as the registry for the IPDA service registry that is currently under development.
  - This effort will result in a Python API for accessing the Registry Service.

# Questions / Comments