

Registry Service Summary

Distributed Infrastructure Design Team

May 31, 2009

Overview

In the Service Oriented Architecture (SOA) environment the service registry provides a "yellow pages" directory of all web services. Web services are a class of resources that can be discovered and accessed by users. A web service will have an access protocol and a response format for the results. The SOA architecture is linked to the use of some standard protocols. The Simple Object Access Protocol (SOAP) is like an envelope for communications between clients and servers. The Web Services Description Language (WSDL) is used to describe what the service provides and how and where to get it. Several other standardized capabilities have been added to this scheme to provide features like security, reliable messaging and notification. There is a solid foundation of Java-based utility software that supports web services (JAX-WS) however there are some compatibility issues between different versions of the components. Some people think the JAX-WS scheme has gotten too complicated. Also, there are some issues with how large response files are delivered with SOAP. Over the past couple years there has been a strong push to provide REST-ful services. REST is an architectural style, not a standard. However REST is implemented on standards that include HTTP, URL, XML and mime types. WSDL 2.0 has extended support for describing REST services. The payoff of using a standard definition for service protocols is that it allows resources to be publicized and opens them up to new uses. There are utilities to take WSDL files and generate code in several popular languages. On the other hand, if every service uses a completely different protocol, this capability will not be so useful. So there is also a need to standardize protocols for categories of services.

Current Practice within PDS

At the present time many PDS resources are only available via human interaction with a custom interface for each different resource. These include several search systems, volume browsers, and hundreds of individual html pages which encapsulate a lot of planetary science knowledge. The Engineering Node high-level catalog includes links to the search systems and volume browsers for some data sets. There are a few PDS-D profile and product servers, which allow data set and data product inventories to be searched and data products to be retrieved using OODT protocols. The Map-a-Planet system and the OnMars and OnMoon servers provide a Web Map Service (WMS) interface to for automated generation of map products. DITDOS is a REST-based service for exploring PDS volumes. There are also a few FTP servers at the nodes. There is no convenient mechanism for identifying and using these services.

Current Practice outside PDS

Earth Science has built ECHO, a centralized metadata database (including browse data), which is accessible using web services protocols. A handful of specialized search applications have been developed to access the database search APIs. ECHO has also established an extended service registry (using a protocol called UDDI) for data manipulation services, but it is not being actively used. The OPENGIS community uses the Catalog Service for Web (CSW) as a distributed registry of web map and web coverage servers and of catalog metadata as well, but this service has not been widely implemented. The astrophysics community with NVO and IVOA have developed a distributed registry system which identifies thousands of servers (often one per data set) which use a small set of simple service protocols to provide access to specific data types (tables, images, spectra). Astrophysics search and display systems (Aladin, Datascope) utilize the registries to determine which servers to access. The space and solar physics community has developed a uniform metadata model for describing data, documents and services (SPASE data model). It has a registry and repository architecture that can be queried with a specialized query language (SPASE-QL) and accessed with REST-like URLs. In the commercial world the idea of global service registries supporting b2b interactions seems to have waned. The focus is on enterprise service registries where all services and service artifacts (wsdl files, schemas) are accessible for exploitation and reuse. There is also a strong element of governance, with the service registry/repository acting as a specialized content management system. A business standards group named OASIS has developed the ebXML registry standard, which is used in many commercial service registries and in "freebXML" an open source implementation.

Observations

There are many PDS resources that are not visible except by manually traversing web sites and learning several different access protocols. One of the key recommendations of the User Support working group was to make PDS resources more transparent to users and provide a single point of entry. A PDS service registry will serve this purpose. It is important to note that a service registry will not be so useful if every service has a different protocol. Thus it is important to come up with standard service types and standard protocols for accessing those services. The ebXML registry standard seems to have strong support and is well aligned with PDS registry needs.