

CI System & Service Description Overview

LEAD services are mainly used to execute the workflows launched using LEAD portal. All these services that are part of LEAD infrastructure are implemented as Web services or wrapped as Web services. These services mainly communicate with each other using SOAP messages. In addition to the basic Web services specifications like SOAP and WSDL, these Web services also support WS-Addressing and WS-Messaging. WS-Addressing is mostly used to provide an asynchronous communication mechanism for the services to communicate. WS-Messaging, on the other hand, provides a means to track things happening within various distributed services. For example, the workflow execution engine sends out events in various stages, like start, end, failure, successful invocation, of its workflow execution. All the communication channels are secured using https as the transport layer. (For more information on Web services technologies used within LEAD project, please refer [resources](#))

Services

Services hosted within LEAD environment can be broadly categorized in to two classes: namely; transient and persistent services. Transient services will start and perform any tasks they have and dies. Persistent services on the other hand will always be running throughout the life time of the system.

Since all these services are part of a distributed system, each of them employ various security mechanisms for their communications. Some of the services are using transport level security to secure the wire level communications between them. In addition to that all the entities including users and services should authenticate them selves to the services that they are interacting. LEAD system uses security certificates issued by a certificate authority (CA) for most of the authentications.

Let's now understand briefly about the different services hosted within LEAD infrastructure. More information about these services can be found within service manuals for each services and information model, event model documents.

Auditing Service - It collects auditing notifications sent by application services and contacts tera-grid accounting databases to update accounting information. [Manual](#)

Calder - The Calder stream processing system is a continuous query grid service that enables applications to submit long running continuously executing queries on data streams. In LEAD, we use the Calder system to mine the real-time weather data and trigger forecast workflows. Event detection is carried out over Level II data produced by the NEXRAD radars. Filtering, aggregation, self-joins are supported by the Calder events processing system. The data mining algorithms are part of the ADaM data mining toolkit developed at University of Alabama Huntsville. [Manual](#)

Data Catalog -

Data Movement and Naming (DaMN) - DaMN is responsible for all the data movements within LEAD project. If GFac wants to stage for an application service to execute, [GFac](#) will provide the source and destination URLs into DaMN and DaMN will do the transfer. DaMN can handle numerous URL schemes. [Manual](#)

DSC (Dynamic Service Creator) - The Dynamic Service Creator (DSC) is a proxy service that accepts an input message to a Web service from a workflow engine on behalf of an actual Web service instance, finds an appropriate Web service instance, and invokes the Web service instance. DSC uses XRegistry to look for available Web services instances, and when no Web service instances are available, DSC contacts GFac to create a new Web service instance. [Manual](#)

GFac - GFac wraps any command line application and provides a Web services interface to it. When an application service is invoked with a given set of input parameters, it runs the application with those input parameters (possibly on a cluster of resources), monitors the application and returns the results to the user. This toolkit can be used by application providers (a.k.a application service providers or service providers) to wrap any command-line application as an application service without writing any

program code or modifying their applications in any way. The toolkit also provides a generic web service client that allows users to securely access any application service created by the toolkit from the convenience of a Grid Portal. [Manual](#)

GPEL - This is the main workflow execution engine within LEAD. GPEL implements WS-BPEL, the workflow composition language and provides an infrastructure to deploy and execute those workflows. [Manual](#)

GEOGUI - GeoGUI provides geographical query user interface to the lead users and anonymous user. Users can search data product by it's temporal, spatial, product type, and other terminologies over the public and private data catalogs. This interface is integrated into the lead portal interface under the tag of "data search". [Manual](#)

Host Scheduler Service - There might be multiple hosts to run applications. The host scheduler service ranks these hosts and help to pick the best host to run a particular application. [Manual](#)

WS-Messenger - WS-Messenger is a Web service-based publish/subscribe XML message brokering service for service-oriented computing. It supports both WS-Notification (3/5/2004) and WS-Eventing (8/2004) specifications and provides mediation between them. It has expressive subscription options for consumers to specify their interest, including topic-based subscriptions, XPath-based subscriptions and hybrid subscriptions. WS-Messenger is the central "glue" that holds the LEAD architecture together. [Manual](#)

Message Box - Even though some services are subscribed to various events in the system, they might not be ready to process those events as and when they occur. Services can use message box to accumulate the event notifications and later process them, providing a pull based notification processing mechanism. [Manual](#)

MyLEAD Agent - The myLEAD Agent service is a service running above the myLEAD server. It actively keeps track of user experiments and workflows, and intelligently manages users' workspace. The Agent is a stateful service that saves its clients from a heavy dependency stack if they were to interact with the myLEAD server directly. [Manual](#)

MyLEAD Server - MyLead is a service that records metadata regarding the data and derived products from the meteorological experiments performed by scientists using the LEAD grid. Metadata is recorded for binary products using the LEAD Metadata Schema (LMS) that has been adopted by the LEAD grid for describing data products. [Manual](#)

MyLEAD Publisher - The myLEAD Publisher service allows its users to push data objects from the personal workspace to a community repository, and thus makes them available to others. [Manual](#)

Query Service - Query Mediator mediates the query issued by users between different data catalogs running on the LEAD system so that users can query over the heterogeneous data catalogs through unified interface. Query Mediator accesses Data catalog and MyLead personal catalog. [Manual](#)

XRegistry - Xregistry is a Document registry that supports users, recursive groups and sharing. It is designed for LEAD atmospheric discovery project and supports four specific document types (Host Descriptions, Application Descriptions, Service Descriptions and Concrete WSDLs) and a any document type. It consist of a one persistent server, that provides a Web Service interface to manage documents. [Manual](#)

SCMS (Session Credential Management Service) - It provides a session oriented management of the user's credentials. And also system (drlead) credentials. [Manual](#)

WCS (Workflow Configuration Service) - A Web service that can configure and invoke a workflow. It is invoked by GPEL for dynamic workflow cases.

Workflow Monitor - This is a utility services which can be used to monitor already executed workflows. You can search for workflows, using various criteria, and look for the status, SOAP messages exchanged. [Manual](#)

XBaya - XBaya is a graphical client program for composing, launching, and monitoring workflows. XBaya is designed to be easy and intuitive to use for domain scientists. [Manual](#)

In addition to the above service manuals, LEAD documentation also contain following documents.

Information Model - This document captures the information placement of the LEAD workflow system. The information includes static configurations, input data files, as well as runtime states of a workflow. [Document](#)

Event Model - Event notifications are critical part of LEAD project. This document captures the events fired and captured within LEAD environment. [Document](#)

Resources

[An Introduction to the Web Services Architecture and its Specifications.](#), Curbera, Kurt and Box. 2004.

[The Next Step in Web Services](#) Curbera, Khalaf, Mukhi, Tai and Weerawarana, CACM Oct 2003.

[An Introduction to WS-Addressing.](#) Beth Linker

[WS-Eventing for Dummies.](#) A simple overview of ws-eventing. [part II](#).