**Integrating Map Services into Science Grids**

**Ahmet Sayar**

***ABSTRACT***

***Statement of the research problems:***

A Geographic Information System (GIS) is a system for creating and managing spatial data and associated attributes. Maps are created by overlaying various geo-spatial features.

Traditional GIS mapping approaches mostly depend on the desktop applications and require expertise and high amount of resources. They have centralized server-client models for Web-based environments which do not provide cross-vendor or cross-product interoperability. They WMShave also some other problems due to the nature of the geo-spatial data (distributed and heterogeneous) and lacking of proprietary data formats and service methodologies. OGC [3] and ISO/TC211 are the standard bodies developing standards for making geographic information and services neutral and available across any network, application, or platform. Even if we solve the interoperability problem through the standard bodies we still need some other requirements in order to integrate geo-spatial data into the Science Grids.

OGC’s service definitions are HTTP GET/POST based. HTTP has limited data transfer capabilities and used in tightly coupled, point-to-point and synchronous centralized applications. In contrast, creating Map Services and integrating them into Science Grids require asynchronous communication models to cope with the high number of participants and long-running codes, transfer of large data between services, coupling data sources and high performance tools.

Besides being interoperable through Web Services and standard specifications, Map Services should also have some other quality of services to fulfill Science Grids requirements. These might be summarized as (1) they should be dynamic and up-to-date, (2) they should be able to federated when it is needed and (3) they should be scalable.

We basically investigate the problems of distributed map processing and rendering of scientific data and information from the interoperability, scalability and performance points of view.

***Proposed solutions:***

The Map Services are Web Services. This enables Map Services to be loosely coupled and, registered, located and invoked remotely. Map Services for Science Grids are described by both WSDL and “capability” separately for different purposes. “Capability” term is inherited from OGC, defines service metadata which is different from Web Service. We federate Map Services to get more advanced and detailed map through the federation of their capabilities. In our approach, each service is described by generic and domain specific metadata descriptions that can be queried through Web Service invocations. Going beyond the enablement of service discovery, this approach enables at least three important things. First, services of the same type that provide a subset of the request can be combined into a “super-service” that spans the query space and has the aggregate functionality of its member services. Second, the capability metadata can be used to determine how to combine services into filter chains with interconnected input-output ports. Third (and building on the previous two), capabilities of “super-services” can be broken into smaller, self-contained capabilities that can be associated with specific services.

We also introduce advanced workload management architecture to eliminate the burden of parsing and rendering of large XML based geo-spatial data with the help of capability federation of Map Services. Workload distribution can be done in at least two ways.  First, the image itself can be split into smaller bounding box tiles and each tile can be farmed out to a slave Map Service. Second, the layers themselves can be generated by different Map Service slaves. One could also combine these two approaches.

Map Services are capable of exchanging and federating their capabilities dynamically. This enables service chaining and workload management. Service chaining composed of Map Services and geo-spatial data services (WFS-Web Feature Service) [5]. WFS provide vector data (in XML based GML [6] format) enhancing the map images by overlaying as layers. Data transfers between the services in the chain are accomplished by publish-subscribe based messaging system. It eliminates broken service connection due to long-running services.

We also wrap Google Map and Dislin [1] Plotting libraries as Web Services and integrate them to the Science Grids.

***Test and experiments:***

We will implement the Open Geospatial Consortium’s Web Map Service (WMS) [4] and will integrate it with GIS and other services for data access, information management, and scientific plotting. This will involve cooperation with other lab research projects in GIS and will be integrated with geophysical applications from the ServoGrid [2] project. The GIS services will be described with capability metadata that will test our proposed framework. We will demonstrate WMS federation and Map-Plotting service integration. Based on this work, we will design experiments that will illustrate capability “meta-querying” and service aggregation and federation into “super-services” described above. We will compare and contrast the performance results obtained from the basic Map Services and the one enhanced with workload management and fault tolerance. We will make our tests via our implementation of smart map tools. Smart map tools allow users to create interactive queries and, analyze, display and animate (such as movies for time series data) the geo-spatial information and the data.

***Expected contributions:***

Our contributions are summarized as below:

(1)We merge two important software worlds: GIS and Web Service Architectures, (2) we extend OGC capability specifications with the workload management, (3) we create an architectural framework for dynamic capability federation of Map Services enabling workload management and service chaining, (4) we create Web Service-based scientific plotting services and Map Services and, integrate them into Science Grids, (5) We create “smart map” tools as portlets for the science portals, (6) we design high performance Web Service architecture for distributed Map Services to support archived geospatial data.

***Usage Scenarios:***

Ex: Integrating Distributed Map Services and Plotting Services into

* Home-Land security project at Los Alamos National Labs – (IEISS Interdependent Energy Infrastructure Simulation System): Overlaying outage area (calculated through the IEISS Science Grid) over the Google Map or other Map Services.
* Earthquake forecasting project at NASA-JPL Labs - (Pattern Informatics): Plotting probability of the earthquake happenings (based on the data calculated through the Science Grid) over the maps from our Map Services.

***REFERENCES***

*[1] Dislin Scientific* Data Plotting Libraries, available at <http://www.mps.mpg.de/dislin/>

[2] Aktas, M., Aydin, G., Donnellan, A., Fox G., Granat, R., Grant, L., Lyzenga, G., McLeod, D., Pallickara, S.,Parker, J., Pierce, M., Rundle, J., Sayar, A. and Tullis, T., “iSERVO: Implementing the International Solid Earth Research Virtual Observatory by Integrating Computational Grid and Geographical Information Web Services” Technical Report December 2004, to be published in Special Issue for Beijing ACES Meeting July 2004.

[3] The Open Geospatial Consortium, Inc. web site: <http://www.opengeospatial.org/>

[4] de La Beaujardiere, Jeff, Web Map Service, OGC project document reference number OGC 04-024.

[5] Vretanos, P (ed.) (2002), Web Feature Service Implementation Specification, OpenGIS project document: OGC 02-058, version 1.0.0.

[6] Cox, S., Daisey, P., Lake, R., Portele, C., and Whiteside, A. (eds) (2003), OpenGIS Geography Markup Language (GML) Implementation Specification. OpenGIS project document reference number OGC 02-023r4, Version 3.0.