High Performance, Federated and Service-Oriented Geographic Information Systems

Ahmet Sayar
asayar@cs.indiana.edu

Introduction

A Geographic Information System (GIS) [1, 2] is a collection of computer hardware and software for capturing, managing, analyzing, and displaying all forms of geographically referenced data.

The general purpose of a GIS is extracting information/knowledge from the raw geo-data. The raw-data is collected through sensors, satellites or any other ways and stored in databases or file systems. The data goes through the filtering and rendering services and, presented to the end-users in human recognizable formats such as images, graphs, charts etc. A well-known example of a GIS is a map viewer which processes layers of geospatial data to create map images. GISs are used in a wide variety of tasks such as urban planning, resource management, emergency response planning in case of disasters, crisis management and rapid response etc.

Over the past decade, GIS has evolved from traditional centralized systems to distributed systems [3]. Centralized systems provide an environment for stand-alone applications in which data sources, rendering and processing services are all tightly coupled and application specific. Therefore, they are not capable of allowing seamless interaction with the other data or processing/rendering services. These deficiencies of the centralized systems and the improvements in the internet technologies encouraged the academy, governments and businesses to start using a distributed system approache. Distributed systems are composed of autonomous hosts that are connected through a computer network. They aim to share data and computation resources, and collaborating on large scale applications.

The main challenge in sharing of data and computation resources and integrating GIS data is the heterogeneity of different sources. It has been studied from the perspective of how to resolve the semantic differences between heterogeneous data sources, mapping different schemas, and providing standard service and query interfaces. The adoption of GIS open standards for online service and data model solve the heterogeneity problems to

some extent. Open Geospatial Consortium (OGC) and ISOTC-211 are the well-known universally standards we use for this respect.

A large variety of different data sets are available in various specialized repositories, and users and geo-science applications would like to access these distributed heterogeneous data sources through uniform service interfaces enabling unified querying from a single access point. In the literature, these requirements are explained as “federation” which is initially used by the database community [4, 5] to federate heterogeneous databases. Federation is established by working out the interoperability and inter-service communication issues among the distributed heterogeneous sources.

Distributed GIS systems typically handle a large volume of datasets. Therefore the transmission, processing and visualization/rendering techniques used need to be responsive to provide quick, interactive feedback. There are some characteristics of GIS services and data that make it difficult to design distributed GIS with satisfactory performance. One of them is that services provided by a GIS typically require heavy CPU usage due to the complex computation involved in the underlying computational geometry. Another one is that GIS services often transmit large resulting datasets such as structured data, images, or large files in tabular-matrix formats.

In this thesis, we will first research architectural design requirements of a federated GIS framework (in accordance with Service Oriented Architecture (SOA)) composed of Web Service components of commonly accepted GIS Open Standards. Secondly, we propose a novel grid-enabled [6, 7] aggregator map services (federator) for optimizing the performance and responsiveness of the federated Service-Oriented GIS.

References

[1] GIS Research at Community Grids Lab, Project Web Site: http://www.crisisgrid.org.

[2] Ahmet Sayar, Marlon Pierce, Geoffrey Fox OGC Compatible Geographical Information Services Technical Report (Mar 2005), Indiana Computer Science Report TR610

[3] Peng, Z.R. and M. Tsou, Internet GIS: Distributed Geographic Information Services for the Internet and Wireless Networks. 2003: Wiley

[4] McLeod and Heimbigner (1985). “A Federated architecture for information management", ACM Transactions on Information Systems Vol 3, Issue 3: 253-278.

[5] Sheth and Larson (1990). "Federated Database Systems for Managing Distributed, Heterogeneous, and Autonomous Databases". ACM Computing Surveys Vol 22, No.3: 183-236.

[6] Fox, G. and M. Pierce. Web Service Grids for iSERVO. in International Workshop http://www.eps.s.u-tokyo.ac.jp/jp/COE21/events/20041014.pdf on Geodynamics: Observation, Modeling and Computer Simulation University of Tokyo Japan October 14 2004. 2004

[7] Fran Berman, Geoffrey C, Fox, Anthony J. G. Hey., Grid Computing: Making the Global Infrastructure a Reality. John Wiley, 2003.