CyberGIS Gateway for Enabling Data-Rich Geospatial Research and Education

Yan Liu, Anand Padmanabhan, and Shaowen Wang
CyberInfrastructure and Geospatial Information Laboratory (CIGI)
National Center for Supercomputing Applications (NCSA)
University of Illinois at Urbana-Champaign
Urbana, IL, U.S.A.
{yanliu, apadmana, shaowen}@illinois.edu

***Abstract*— This paper describes CyberGIS Gateway, an online problem-solving environment, for multiple science communities to conduct data-rich geospatial research and education.**

Keywords—CyberGIS; spatial analysis; cyberinfrastructure; science gateways

I. INTRODUCTION

CyberGIS [1] represents a new-generation of geographic information systems (GIS) based upon advanced cyberinfrastructure (CI) and consists of three complementary modalities: 1) CyberGIS Gateway (referred to as Gateway hereafter), which provides an online environment for making CyberGIS capabilities accessible to a large number of users for various research and educational purposes; 2) CyberGIS Toolkit, which creates and maintains a collection of loosely coupled scalable geospatial software capabilities empowered by advanced CI environments such as the NSF Extreme Science and Engineering Discovery Environment (XSEDE) (http://xsede.org); and 3) GISolve [2], a middleware framework that bridges the CyberGIS Toolkit and Gateway to manage the complexity of accessing heterogeneous and distributed CI resources and services.

Gateway (http://gateway.cybergis.org) provides an online problem-solving environment for a number of science communities (e.g., biology, geography and spatial sciences, hydrology, public health, and statistics). Driven by various community needs, Gateway is designed to provide a rich set of highly interactive user interface components for interactive cyberGIS analytics and visualization and shield complexity of CI access from users. An open service integration framework and a HTML5-based mashup library are established to achieve seamless integration among user interface components, cyberGIS functions, and CI resources and services. Gateway employs GISolve to manage multiple cutting-edge CI environments (e.g., the NSF XSEDE, the Open Science Grid (OSG) (http://opensciencegrid.org)) and cloud computing resources (e.g., ArcGIS Online from Esri (http://arcgis.com)) to provide interoperable and scalable cyberGIS analytics.

II. DESIGN AND COMPONENTS

Gateway is built on rich-client Web technologies to provide a data- and interaction-rich environment for users to access data-rich analytics. Gateway access to CI is made seamless by using GISolve middleware and the community account model (on XSEDE) or the virtual organization (VO) model (on OSG). Details of Gateway design can be found in [1].

A. User Interfaces

Interactions with GIS user interfaces are common for data manipulation and visualization, configuration of function parameters, result validation and comparison, etc. With the advances in Web technologies, especially HTML5, user interface components and interaction capabilities of Web browser-based Web environment are nowadays comparable to desktop graphic user interface (GUI). We implemented an open mashup library framework to develop and componentize HTML5 user interface components customized for cyberGIS use. The implementation uses Ext JS (http://sencha.com), Sencha Touch, and OpenLayers (http://openlayers.org) for user interface and mapping support. Specifically, the Gateway user interface is designed with the following features:

Model-View-Control (MVC)-based framework. MVC [3] development model becomes mature in rich-client Web application development. Gateway employs Ext JS MVC framework to develop various user interface components such as data selection panels, computation management panels, map panels, etc. Each gateway application is also developed as a standalone MVC application.

Customizable user interfaces. Gateway implements generic user interface templates for user management, access control, and data and analytical functions that can be extended to support the integration of individual gateway applications. These functions are encapsulated as client-side JavaScript libraries and server-side PHP libraries.

Streamlined Web and mobile interface development. Enabled by the interoperable library support between Ext JS and Sencha Touch, minimum effort is required to convert a Web application into mobile Web application, vice versa. In MVC framework, only view and partial controller functions need modification.

B. Open Service Integration Framework

As an online environment, Gateway needs to interact with external data and analytical Web services hosted by various organizations. These Web services include such types as SOAP, REST, and OGC WCS / WFS / WMS / WPS (http://www.opengeospatial.org/). We have developed an open API approach to providing a consistent and highly usable interface for application developers and online environments to interact with backend service implementations, leading to a flexible development framework in the context of service-oriented software integration. The framework includes four types of functional APIs focusing on security, application integration, computation management, and data and visualization, implemented as REST services. Currently, this framework provides multi-language client programming support, including PHP, Perl, Python, and Java.

The open-service API framework makes distributed software and services interoperable. Such interoperability is critical for cyberGIS to share digital resources and services while providing a large number of users with data-rich geospatial problem-solving capabilities.

III. Spatial Data and Analytics

Through Gateway, end users have access to a collection of advanced spatial data and analysis services. These services are integrated into gateway applications, each with a highly-usable user interface with details and complexity of CI-based computation hidden from users. Each service is integrated into Gateway by going through rigorous software engineering practice and computational intensity analysis in order to resolve computational bottlenecks associated with data size and computational intensity. A set of Gateway applications representing cutting-edge research in multiple science domains has been developed and made available to general public. Due to the page limit, one application is highlighted in this section.

A. Flumapper

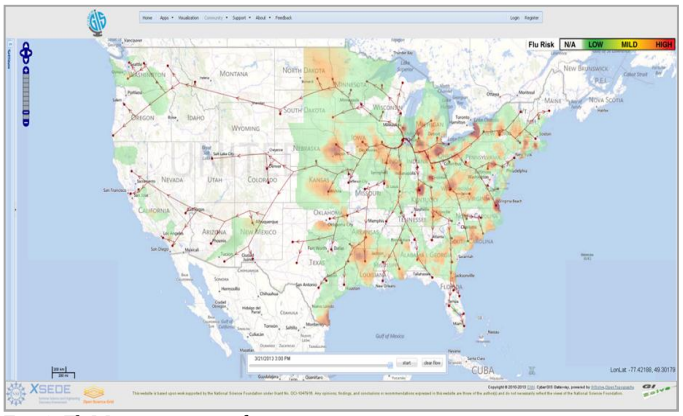

Fig. 1. FluMapper user interface.

Flumapper [4] (Fig. 1) is a gateway application for interactive spatial exploration of massive location-based social media data for providing early detection information on influenza-like illnesses (ILI) risks. FluMapper uses real-time Twitter data streams, computes in near real-time spatial patterns of ILI and spatial and temporal movement patterns (i.e., flow mapping) using a set of spatial analytics. To address the challenges of near real-time collection, processing, storage, query, and analysis of massive Twitter data across multiple spatial and temporal scales, a space-time cube mechanism is constructed based on CI to provide a holistic solution to the effective integration of various data and analysis components. In Gateway, however, the Flumapper application is presented as an integrated map interface with rich information about flu risk and flow patterns. This interface is adaptive to both mobile and Web client platforms. The intensive computation needed to create and maintain the map is completely hidden from end users.

B. Data Collection and Handling

In addition to the support for user data space, Gateway provides access to several commonly used data sources that host big geospatial datasets.

USGS National Elevation Dataset (NED). NED (http://ned.usgs.gov) is a major digital elevation model (DEM) data source. To make it more accessible in cyberGIS related communities, we created a set of data processing tools, including clipping and reprojection, for accessing the entire NED based on interactive visualization.

ArcGIS Online. Gateway provides access to cloud-based ArcGIS Online from Esri. ArcGIS Online data can be listed and used in Gateway by using the developed ArcGIS Online data plugin. This gateway plugin module interacts with ArcGIS REST API to retrieve vector and raster data listed on user's ArcGIS Online account. This interoperability demonstrates the complementarity of two representative cyberGIS environments and promises a sustainable approach to the emerging cyberGIS software ecosystem.

C. Workflows

Gateway services can be orchestrated to create application workflows. For example, the viewshed application is a high-performance GPU-based approach to large-scale visibility analysis. However, a highly usable viewshed application in Gateway must invoke a set of service and user interface components. Specifically, the open service API to the viewshed analysis service includes the invocation of USGS NED and OpenTopography (http://www.opentopography.org/) data access and processing services, GISolve data transfer and job management services, and Gateway visualization services. The componentized Gateway software framework allows the development of sophisticated user interface components in the interactive viewshed application in Gateway.

IV. Education and Outreach

While Gateway is widely used for scientific research, education and training are another important area that Gateway focuses on.

A. Education and training

Gateway has been used by hundreds of undergraduate and graduate students to learn advanced CI, service-oriented computing, GIS, and spatial analysis and modeling on multiple campuses (Fig. 2). In academic year 2012-2013, six courses

conducted during the Summer and Fall of 2012 and the Spring of 2013 at the University of Illinois, University of North Carolina at Charlotte, and Idaho State University, including one online course, used Gateway applications in classrooms. Students involved gained concrete and synergistic understanding about advanced CI, cyberGIS, and related scientific problem solving.

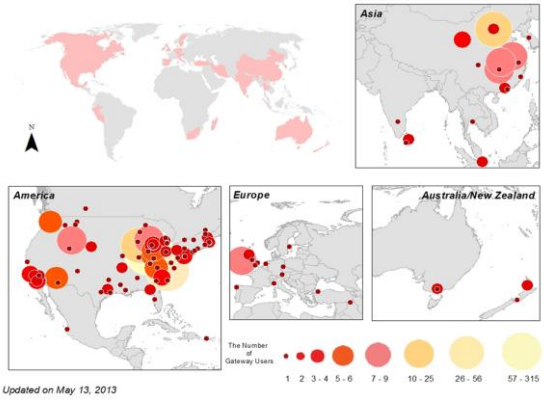

Fig. 2. CyberGIS Gateway user map.

Gateway has been actively supporting the education of XSEDE science gateways and engaging geospatial communities through on-site tutorials and demonstrations. To support education and training of CI-based scientific problem-solving, the open-source SimpleGrid Toolkit [5] includes basic components of GISolve middleware and Gateway. This toolkit has been used in hands-on tutorials conducted at annual XSEDE/TeraGrid conferences from 2007 to 2012 and contributed to startup development efforts of multiple science gateways, e.g., CIPRES gateway (http://phylo.org).

B. Collaboration within Science Gateway Community

Gateway is listed as one of the XSEDE science gateways. XSEDE provides a venue for us to share gateway technologies and experience with other gateway developers while allowing us to learn from and work with other gateways to advance the state of the art. In addition, we participate in the NSF Open Gateway Computing Environment (OGCE) project (http://ogce.org), which targets to create a rich set of gateway software components.

For example, we referred to the iPlant Foundation API (https://foundation.iplantcollaborative.org) when we designed the token service in our open service APIs in order to provide an authentication service among cyberGIS portals and web services. In turn, we contributed our experience to the development of a prototype identify service in Apache Rave, a software project in OGCE. The security requirements gathered from cyberGIS communities were documented and shared with the NSF Science Gateway Security project (http://www.sciencegatewaysecurity.org) as a use case for the development of common gateway security components.

With SimpleGrid Toolkit established as a software tool for learning CI and science gateways, we also use this toolkit to incorporate new Gateway technologies and use cases interoperable with other gateways or gateway tools. For example, at the XSEDE'12 conference, we demonstrated the interoperability between Apache Rave portal environment and cyberGIS application gadgets through an extension to SimpleGrid.

V. Concluding Discussions

The Gateway, together with the other two cyberGIS modalities – CyberGIS Toolkit and GISolve middleware – fosters sustainable development of the cyberGIS software environment for geospatial discovery and innovation. Gateway provides a wide spectrum of advanced spatial data and analytics for several science communities and assures the results of dynamic collaborative work will continue to enrich the cyberGIS software environment for effectively enabling data-rich geospatial research and education.

Acknowledgments

This material is based in part upon work supported by NSF under Grant Numbers: 0846655 and 1047916. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number OCI-1053575. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation

References

[1] S. Wang, L. Anselin, B. Bhaduri, C. Crosby, M. F. Goodchild, Y. Liu, and T. L. Nyerges, "CyberGIS software: a synthetic review and integration roadmap," *International Journal of Geographical Information Science*, pp. 1-24, 2013.

[2] S. Wang and Y. Liu, "TeraGrid GIScience Gateway: Bridging cyberinfrastructure and GIScience," *International Journal of Geographical Information Science*, vol. 23, pp. 631 - 656, 2009 2009.

[3] A. Leff and J. T. Rayfield, "Web-application development using the Model/View/Controller design pattern," in *Enterprise Distributed Object Computing Conference, 2001. EDOC '01. Proceedings. Fifth IEEE International*, 2001, pp. 118-127.

[4] A. Padmanabhan, Wang, S., Cao, G., Hwang, M., Zhao, Y., Zhang, Z., and Gao, Y., "FluMapper: An Interactive CyberGIS Environment for Massive Location-based Social Media Data Analysis," in *XSEDE 2013: Extreme Science and Engineering Discovery Environment: Gateway to Discovery*, San Diego, CA, USA, 2013.

[5] S. Wang, Y. Liu, N. Wilkins-Diehr, and S. Martin, "SimpleGrid Toolkit: Enabling Geosciences Gateways to Cyberinfrastructure," *Computers and Geosciences*, vol. 35, pp. 2283-2294, 2009.