Thesis Proposal:

High Performance, Federated and Service-Oriented Geographic Information Systems

Ahmet Sayar
asayar@cs.indiana.edu

Introduction

- Definition of GIS
- Explanation of IS requirements in GIS
- History: From central to distributed
- Interoperability and heterogeneity of distributed application specific IS.
- Federation of data+services.
- Performance issues: exchange, transform and process large data sets
- Summary:
    - Architectural issues for Service-Oriented capability-based federated GIS
    - High performance and responsive GIS
        - Aggregator services
        - Pre-fetching, parallel processing and caching techniques

Motivation

- Distributed nature of data and computation resources
    - Heterogeneity, interoperability
    - Data is provided by diff vendors in diff formats with diff service interfaces
- Single point of access, query and display of heterogeneous data sources (*federation*)
    - Through integrated data views (seamless interaction with data sources)
    - Needs for informatory and comprehensible data (hierarchical?)
        - Definitions of services and comprehensible data formats such as maps
- Needs for high performance and responsiveness
    - Emergency early-warning systems such as homeland security and natural disasters
- Motivating examples for federation and performance requirements (see the document)
- Summary: Above needs require us to research on
    - Federated and Service-Oriented (Interoperable) GIS systems

        - Seamless data access and unified querying of heterogeneous distributed data from a single access point
    - Interoperable and responsive Information System architecture for GIS
        - Handling large scale data
        - Common data models
        - Common service interfaces for data sources

Architecture

- Interoperable Service Oriented GIS
    - Definitions of Services (Generalization)
        - WFS – provides data in common data model (GML)
            - Has different adaptors and wrappers depending on the original data source (such as relational database or plain ASCII files)
        - WMS – provides comprehensible data (map images)
    - Services, their capability files and common data model (GML)
    - Extending the service definitions as Web Services
    - Additional components of the architecture
        - Interactive-smart user interface
            - Interactive display and unified query over integrated data- view
        - Wrappers for integrating other WMS
            - For non-web service OGC compatible WMS and,
            - For Google Maps
- Federated GIS
    - Creating federator service based on WMS
        - AWMS (WMS Aggregator): WMS extended with aggregator capabilities and high-performance techniques
    - Service and capabilities extensions for federation
        - WFS-based common service interfaces to heterogeneous data sources
        - Capability-defined data+service chaining and referencing
            - WMS to WFSs
            - WMS to WMSs
            - WFS to WFSs
    - Web-based interactive single point data access and querying architecture
        - Through integrated data view
        - Capability based dynamic GUI update.
            - Capability comes from initially interacted service (AWMS)
        - Hierarchical data definitions in capability files
            - Project->map ->layer ->{vector (GML), raster and binary data}
            - Data and service bindings

- Abstracting the system to the general science domain
    - Formulating the system and defining the requirements
        - Capability (service+data defns), ASL, ASFS, ASVS, ASAS.
    - From WFS to ASFS (Application Specific Feature Service)
    - From WMS to ASVS (Application Specific Visualization Service)
    - From AWMS to ASAS (Application Specific Aggregator Service)
    - Capability requirements
    - From GML common data to ASL (Application Specific Language)
- High Performance design Issues
    - Data transfer (streaming – pub/sub based messaging middleware)
    - Data rendering (pull parsing over common data model)
    - Federator (AWMS) oriented design
        - Pre-fetching (un-frequently changing data)
        - parallel processing and caching (frequently changing data)

Applications of the Proposed Work

- Coupling the federated GIS Data System with Geo-Science Applications
    - Integrating the System with other CGL-Lab projects for Geo-Science Applications.
        - Sci-Plotting (my development), WS-Context HPSearch Web Services
        - Displaying correlation of input and output of scientific applications
            - Three-layer structured display through Sci-Plotting Service
    - Interactive decision making tools
        - AJAX (Asynchronous Java and XML) and Web Service integration with mediator Service
        - Ex use cases: Google Map + WFS, Google Map + WMS
    - Use Case Scenarios on real Geo-Science Applications
        - Pattern Informatics (PI)
        - Virtual California (VC)
        - Interdependent Energy Infrastructure Simulation System (IEISS)
- Animating time-series data (such as earthquake seismic) as streaming map movies
    - Time-series data animation
    - Streaming map movies architecture for GIS
        - Frames (static map image) creation
            - Frame is map created through proposed GIS system
            - Invocation stage -chain of requests based on time-periodicity
            - Data retrieval stage –chain of returns
        - Map streams creation, publishing, and fetching and displaying
    - Applications to Virtual California and Pattern Informatics.

Research Issues

- Developing Interoperable and Service-Oriented GIS
- Developing high-performance federated GIS
    - Developing capability file-based federated GIS system
    - WFS-based federation of data
- Enhancing the system responsiveness by dealing with
    - Data transfer and rendering issues
        - Streaming data transfer through message based middleware
        - Advanced parsing and rendering techniques
    - Performance-enhanced federation design
        - Federator (AWMS) oriented approaches
        - Pre-fetching, caching and parallel processing or combinations
- Abstraction of the federated GIS
    - Defining the requirements for the generalizing the proposed capability-file based federation for general application domains
    - Defining the capability file, service types, service functionalities and message formats.
- Secondary Research Issues:
    - Framework for coupling Geo-applications with the proposed GIS system
    - Framework for map animations and streaming map movies - time-series data
    - User portals enabling single point data access and querying through integrated data views.

Expected Contributions

- Introducing capability file-based federation approach to distributed Application Specific Information Systems (ASIS),
- And define the requirements for generalizing the architecture to the other application science domains.
- Introducing interoperable Service-Oriented GIS with universally accepted standards and Web Service technologies.
- Minor Contributions:
    - Introducing the innovative techniques for high performance and responsive data rendering. Performance hurdles are due to:
        - Data characteristics: variable sized and un-evenly distributed
        - Using semi-structured common data model.
    - Introducing a web-based interactive single point data access and querying architecture through integrated data view
        - Through hierarchical data definitions and service bindings in capability files