THE SPATIAL DATA SERVER BASED ON OPEN GIS STANDARDS IN HETEROGENEOUS DISTRIBUTED ENVIRONMENT

Mi-Jeong Kim, Eunkyu Lee, Byoung-Woo Oh and Minsoo Kim
ETRI, 161 Gajeong-dong, Yuseong-gu, Daejeon, 305-350, KOREA
{kmj63341, ekyulee, bwoh, minsoo}@etri.re.kr, Tel: +82-42-860-1575, Fax: +82-42-860-4844

Abstract

Recently, due to explosive increase of users' interest of internet and wide spread of the distributed computing environment, users usually want to use spatial data reside in remote servers through network. However, spatial databases usually have their own incompatible data formats and it is almost impossible to integrate and share the all kinds of spatial information and services because of the diversity of spatial information systems. So it results in duplicated investment problems of spatial information.

In this paper, we proposed a spatial data server for interoperability with efficient performance in distributed environment. In order to provide interoperability and reusability for the spatial data, we adopted standards of Open GIS Consortium. So user can retrieve and manage a diverse spatial data in heterogeneous distributed environment and prevent the duplication development investment.

INTRODUCTION

For the purpose of efficient management of spatial data, Geographic Information System (GIS) is emerged. Various GIS that can analyze, process, and manage spatial data have been developed. Because they have their own storage format, integration or conversion of the spatial data consumes a lot of cost and time. Furthermore, interoperability among the spatial data seems infeasible in distributed environment.

Recently, there has been rising concerns to integrate and connect a developed spatial information services without consideration of location in the heterogeneous distributed environment. So loosely-coupled systems are widely used for distributed computing environment. The Open GIS Consortium (OGC) makes efforts to overcome these problems and develops implementation specifications for interoperability of the spatial data.

We developed a spatial data server related to the web service framework for spatial information services using the international standards developed by the OGC, such as the Simple Features specification for OLE/COM (SFO), Web Feature Services (WFS), and Geography Markup Language (GML). The spatial data server provides interoperable access of spatial data in the distributed environment.

The next chapter shows you the overview of web services and web services model of OGC as a related works. The third chapter shows you the details of the developed spatial data server of this paper and the forth chapter presents process of system. And the last chapter concludes this paper and discusses about future works.
RELATED WORKS

Overviews of web services

Recently, the web services concept is rapidly rising as a new solution to solve the integration problem among heterogeneous application systems. The web services concept is a kind of standardized software technology that can integrate and share various computer programs. This web services concept has an advantage of flexibility by perfectly defining not mutual interoperable methods but only standard specifications for mutually sharable data among distributed systems (Kim et al., 2003). So the web services has an advantage that can transparently access any web servers in any places, with any devices and at any time.

To locate, access, and use arbitrary resources is difficult practically in the open, distributed systems. Because client will generally not know in advance where a desired resource is located. The web services provide a set of protocols that allow applications to expose their functionality and data to other applications over the Internet. The web services are implemented using XML to exchange data between clients and application servers over HTTP or sockets (W3CC, 2003b).

The web services architecture has three essential roles: service provider, service registry, and service requestor. The service provider publishes the availability of their resources to repository using WSDL. The service provider means group, enterprise, and individual offering the spatial data and spatial information services. The service registry is acting as a registry or clearinghouse of services using UDDI technology. The service registry is defined as a process to find requested spatial data list by comparing the users request from the published metadata by service provider. It supports the dynamic binding between service provider and service requestor. The service requestor performs service discovery operations on the registry to find the providers. After search of metadata, the service requestor obtains data from service provider using SOAP.

The Web Services Architecture can be depicted in an equivalent manner as the “Publish-Find-Bind”. The “Publish” used to advertise data and services to a registry. The “Find” used by service requestor to locate specific service types or instances. The “Bind” used when a service requestor and a service provider negotiate (OGC, 2003b). The basic process of web services framework as follows (Figure. 1).

![Figure 1: The basic process of web services framework.](image)

The web services standard architectures is composed of XML (extensible Markup Language), UDDI (Universal Discovery Description & Integration), WSDL (Web Services Description Language), SOAP (Simple Object Access Protocol). The purpose of UDDI is to build a distributed global registry that could be accessed through web environment. The WSDL as a kind of language that define usages of web services is used in order to describe the interface name, argument and return value of serviceable programs (W3CC, 2003c).
The SOAP is a protocol that enables users to mutually communicate their services under distributed environment using powerful XML (W3CC, 2003a).

**Web services model of OGC**

The Open GIS Consortium (OGC) provides a system architecture model related to open web service for spatial information. The OGC fundamentally proposed various standard implementation specifications so that they can smoothly serve diverse spatial information. Based on these implementation specifications, OGC announced the ORM (OpenGIS Reference Model) in order to design web services of various spatial information (OGC, 2003b). We adopted international standards for implementation made by the OGC: Web Feature Services (WFS) Implementation Specification, Geography Markup Language (GML) Implementation Specification, and Simple Features Specification for OLE/COM (SFO).

**Web feature services (WFS) implementation specification**

The Web Feature Services (WFS) Implementation Specification provides operations to obtain spatial data and metadata about them. The WFS specification defines the method that transmits the vector map of GML format (OGC, 2002). Sometimes, the WFS proposes the functions of spatial operator and transaction services besides the vector-based map service.

**Geography markup language (GML) implementation specification**

The Geography Markup Language (GML) is an XML encoding for the modeling, transport and storage of geographic information. Using GML, you can deliver geographic information as distinct features, and then control how they are displayed in a Web browser (OGC, 2003a). Just as XML is helping the Web to clearly separate content from presentation, so GML needs the specific viewer in client, for example GML viewer, to display vector map. The GML approach is a great improvement over the historical reliance on simple GIF/JPG image maps for the following reasons: better quality maps, works on a browser, custom map styling, editable maps, more sophisticated linking capabilities, better query capability, control over content, service chaining, and so on (OGC, 2003a). The version of the GML that we implemented is GML 3.0 released in 2003. Figure 2 shows an example of GML used in response.

![Figure 2: The example of GML.](image-url)
Simple feature specification for OLE/COM (SFO)
The Simple Features Specification for OLE/COM is used for uniformed access of spatial database from disk without consideration of its specific storage format (OGC, 1999).

The data provider conforms to the Simple Features specification for OLE/COM. A data provider is prepared to a vendor specific spatial database or file, such as ARC/SDE, GE Small World, Shape file, etc. For example, if an administrator wants to serve spatial data stored in ARC/SDE or DXF file, then both SDE data provider and DXF file data provider have to be installed on the server computer respectively.

ARCHITECTURE
The web services concept is adaptable in the field of application systems for integration of spatial information. We developed the spatial data server that could serve various spatial data. It has some advantages of being flexibly and powerfully owing to adapting the web service standard specifications of OGC. The spatial data server provides the mechanism to offer a rapid spatial data management, client connectivity, and communications, and support the international standards. It consist of Spatial Data Manager (SDM), Spatial Query Manager (SQM) and Web Services Manager (WSM) as shown in Figure 3.

Figure 3: Overview of spatial data server.

Spatial data manager (SDM)
The Spatial Data Manger (SDM) provides main-memory based spatial data manager. It provides XML based main-memory management, GML encoding, spatial data processing, compression to reduce data size, and encryption to protect the data. It uses spatial index and main-memory manager for rapid response. Also it supports access spatial data without consideration of data formats because use data provider.
The spatial data server based on open GIS standards in heterogeneous distributed environment

**Spatial data management component (SMC)**
The Spatial data Management Component (SMC) manages spatial data in the main memory and fetches spatial data corresponding to user’s query. Because the cost of the main memory has been decreased gradually, content providers who want to support rapid response may willingly increase the main memory for better services. We support up to 64 Gigabytes of physical main memory if they are installed in the server computer. The physical main memory supports by the MS-Windows’ kernel guarantees that it is never swapped to the disk by operating system for context switching (Oh et al., 2003). We exploited main memory based spatial data server to manage spatial data for efficient performance by reducing time to read data from disk for each request.

The spatial data server supports access spatial data without consideration of data formats in the distributed environment. An administrator can select data provider that conforms to the Simple Feature specification for OLE/COM.

The SMC manages a target data format in the main memory to eliminate conversion time from intermediate format on service. The format of the spatial data in main memory is GML. It reduces time-consuming operations such as reading database from disk, conversion from a specific storage format of spatial data to well-known-binary (WKB) in data provider component for unified access among different formats, and conversion from the WKB to the GML for each request. Also it is expected to provide interoperability by implementing interfaces in compliance with OGC specifications. The spatial data to be accessed by the client and service provider reside in main-memory and disk.

The SMC is implemented as an out-of-process COM server to be shared with other processes, such as the SCT and the SQM.

**Spatial data index component (SIC)**
The Spatial data Index Component (SIC) creates and manages spatial index structure while loading the features. The SIC supports efficient access for spatial data using feature pointers that refer the offsets of the individual features (Oh et al., 2003). The index is used for rapid response time when the SDM processes user’s request that includes BBOX parameter. We currently implemented a hash-based approach for the spatial index method and R*-tree method using same interface to select spatial index with reflecting the properties of each spatial dataset. The spatial indexes always reside in main-memory for performance.

**Spatial data compression component (SCC)**
The spatial data server supports compression for efficiency of transmission time in sending GML data via network. The compression is developed using zlib which can be obtained from http://www.gzip.org/zlib.

**Spatial data encryption component (SEC)**
The spatial data server supports encryption for security. The encryption uses Crypto API which is provided by Microsoft.

**Spatial data configuration tool (SCT)**
The Spatial data Configuration Tool (SCT) is an executable file which invokes the spatial data server at the beginning. The configuration information that is set by the administrator
can be stored as XML file for reuse in the future. The SCT can be used to show status and logging information of the server.

**Spatial query manager (SQM)**

The Spatial Query Manager (SQM) implements the Web Feature Services (WFS) specification of OGC to fetch spatial data. It is used to request to access spatial data in web environment. The WFS provide a web interface for GML data transmission. The GetCapabilities, DescribeFeatureType, and GetFeature operations was implemented (OGC, 2002). We implemented basic WFS that did not support transactional operations because the spatial data was managed in the main memory regardless of write-operations. The SQM consist of the GML Server to return the GML document and the WKB Server to return the WKB data for user's request through HTTP. And SQM uses the Microsoft Internet Information Server (IIS) as a web server.

It was implemented as an Internet Server Application Programming Interface (ISAPI) extension which guaranteed more efficient service than CGI approach though many clients request the service at the same time. The ISAPI extension is loaded only once by web server when the first service is requested, but it is shared when another request is invoked. It minimizes overload of the server and enables rapid service.

Also we extend it to support progressive transmission to reduce latency time by representing only received spatial data (Oh et al., 2003). The remainder can be continuously received during representation with multi-threaded process. Once, the client draw already transmitted geometry data, the client should merge remainder part of geometry to the previous data and redraws periodically. More time goes; more detailed spatial data can be displayed.

**Web services manager (WSM)**

This spatial data server provided the framework of web service for spatial information. The Web Services Manager (WSM) publishes to metadata describing its capabilities and network address. We implemented a wrapper which uses the spatial data server as back-end for the web services using SOAP. The WSM is composed of GML Server for web services to return GML document and WKB Server for web services to return the WKB data for user's request through the SOAP.

It is advantageous to allow fast binding between service requesters and service providers in dynamic, heterogeneous environments. It is integrating the business processes of different organizations through the interoperability of spatial information system. It is applicable to the spatial information to share with a general user, business fields, public sector.

**PROCESS**

**Preparation of service**

First of all, an administrator executes the SCT. The SCT should not be terminated until the service for stable response. The SCT enumerates the data providers that is installed on the server machine for the administrator to select a data provider for connection. The administrator can select data providers and datasets to load into main memory. The configuration that includes the data provider and the list of feature tables selected by the administrator can be stored in XML file for the next use.
And the administrator can make to start a service with size of main memory to be allocated. The SCT tries to allocate the size of physical memory and to load the selected spatial dataset to the allocated memory using the SDM.

**Loading the spatial data**
The SMC gets the parameters such as number of features a list of type names, and MBR (Maximum Boundary Rectangle) information of the whole dataset to make the memory arranged for preparation of data loading. The SDM reads to be serviced spatial data through data provider from databases or files, and loads them to the allocated main memory. Also the spatial index is created.

**Publishing the services**
The Spatial Data Server advertises the availability of their spatial data and spatial information service to UDDI server, and Web Client can then query a UDDI server to discover resources of interest and determine how to access them.

**Use of data**
The user easily retrieves a diverse spatial data through web browser. User can submit a request to the SQM by the web server (i.e., Microsoft Internet Information Server) to obtain spatial data via internet. The SQM parses the request that is transferred from the web server and calls corresponding method. And they transfers the result to the user via internet in GML format for spatial data. If the request contains BBOX parameter, the SDM uses spatial index for filtering feature data.

Figure 4 shows examples of application using the spatial data server using main memory. The spatial data displayed in the figure is obtained from the spatial data server after search of metadata. The display area is developed as ActiveX control. It provides zoom in/out, panning, 3D-like representation, and layer control.

![Figure 4: An application of spatial data server.](image-url)
CONCLUSIONS

We developed new solution that could easily integrate and to share spatial information and services among different organizations under distributed environment. Especially, this system provided the extensibility and flexibility as a solution to integrate and share spatial information and services under by adapting the web services standard of OGC.

The developed spatial data server could be practically used and planed to be used in national projects such as nationwide spatial clearinghouse system, national marine GIS, etc. (Kim et al., 2003). However, this spatial data server has some potential problem in the performance when spatial data or services are served in the GML format through internet because of large volume of GML data. Therefore, we will focus our next research on the speedup of the spatial data server.

ACKNOWLEDGEMENT

This work was performed as a part of the project named “Development of Integration Technology for Spatial Information Systems” and supported by Korean Ministry of Information and Communication.

REFERENCES

Kim, M.S., Oh, B.W., Kim, M J. and Jang, B.T., 2003: Design of 4S web services model for the infrastructures of national spatial clearinghouse of spatial data and services. International Symposium on Remote Sensing (ISRS) 2003