

SOA for Web GIS Application in Cloud

¹Harsh Bhor

¹K J Somaiya Institute of Engineering & IT, Sion, Mumbai, Maharashtra, India.

¹hbhor@somaiya.edu

Abstract : Cloud computing is rapidly emerging as a technology trend that almost every industry that provides or consumes software, hardware, and infrastructure can leverage. The service Oriented Architecture is a very promising architecture for practical implementation of the next generation geographical information systems. This paper explain the Service Oriented Architecture for Web GIS application distributed platform. The spatial cloud computing applies geographical information systems to the cloud computing paradigm. It provides dynamically scalable geographic information technology, spatial data, and spatial applications as a web service. GIS web services were designed to provide the hosted spatial data and GIS functionality to integrate the customized GIS applications to perform basic geo-processing tasks, such as address matching, map image display, and routing, without maintaining GIS tools or the associated geographical data. It has an important application prospect in the Geographical Information System.

Keywords — Cloud Computing, Geographical Information System, Service Oriented Architecture, Technology, GIS Platform..

I. INTRODUCTION

Cloud computing is rapidly emerging as a technology trend that almost every industry that provides or consumes software, hardware, and infrastructure can leverage. The technology and architecture that cloud service and deployment models offer are key areas of research and development for geographic information system (GIS) technology[a]. With cloud computing, the cloud computing oriented applications presents on-demand self-service, ubiquitous network access, location-independent resource pooling, rapid elasticity, and pay-per-use. It has three delivery models, SaaS (software as a service), PaaS(platform as a service), IaaS(infrastructure as a service)[4,5]. The Google's Google Reader and Google Docs is the SaaS examples for cloud computing [6]. Amazon's Elastic Compute Cloud provides compute capacity in the cloud and is a leading IaaS. Cloud Computing offers cost effective solutions that are dependable, and reliable. The value of the Cloud Computing industry is expected to reach hundreds Billion in five years.

Spatial Cloud Computing researches to apply geography to the cloud computing paradigm [6, 7]. It provides dynamically scalable geographic information technology, spatial data, and spatial applications as a web service. The applications do not need the massive investment in specialized spatial data, and software typically required in an enterprise GIS implementation. It has the benefits of all the cloud computing. With spatial cloud computing, The Spatial Cloud Computing provides the foundation data as part of the core service. Web Services are the software components that can be accessed through the Internet in other applications. GIS Web services are the software components that can provide the hosted spatial data and GIS functionalities and can be accessed and integrated to the practical customized GIS applications. Developers can apply the GIS Web services to perform geographical information processing and return the results to the customized applications without maintaining the basic GIS system or the geographical data. The service and cloud computing oriented architecture is a very promising architecture for practical

implementation of the next generation geographical information systems.

The GIS web services have been thoroughly investigated in recent years [8]. There are a number of researches and developing projects of applying the web services to GIS system by several different companies and organizations. The Open GIS Consortium (OGC) developed the web-based geographic information and services accessible via the Internet through the OGC consensus process [9]. Each service supports a certain GIS operation function that returns metadata describing the nature of the service, the methods to register Web Service, the addresses to access the service or to contact its provider, the type of data provided by the service, and so forth.

II. OVERVIEW OF CLOUD IN GIS

A. Cloud Computing

Cloud computing is a way of computing shares computer resources on Internet instead of using software or storage on a local computer [1-3]. Cloud computing is an outgrowth of the ease-of-access to remote computing sites provided by the Internet. In concept, it is a paradigm shift whereby details are abstracted from the users who no longer have need of, expertise in, or control over the technology infrastructure in the cloud that supports them. Cloud computing describes a new supplement, consumption and delivery model for IT services based on the Internet, and it typically involves the provision of dynamically scalable and often virtualized resources as a service over the Internet. The term cloud is used as a metaphor for the Internet, based on the cloud drawing used to depict the Internet in computer network diagrams as an abstraction of the underlying infrastructure it represents. Typical cloud computing providers deliver common business applications online which are accessed from a web browser, while the software and data are stored on servers.

Geographic Information System includes a variety of advanced features in information processing, and its basic functionality is data collection, management, processing, analysis and output. Relying on these basic functions, GIS implements a variety of applications through the use of spatial analysis, modeling, network technology, database and data integration technology, and further development environment to meet the broad needs of users[10]. The emergence of cloud computing, no doubt brings new opportunities for GIS.

III. SOME TECHNOLOGIES OF CLOUD COMPUTING FOR WEB GIS APPLICATION

Cloud computing technology architecture consists of 4 levels: virtualization, infrastructure, platform and application. Every level is regarded as a service, i.e. Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a Service (SaaS). Many technologies are integrated and extended in cloud computing platform, such as virtual machine, on-demand and pay-by-use model, distributed file system, parallel computing, grid computing, huge data storing and searching, distributed resources management, and so on.

Google is one of the most famous cloud computing providers. Many people are using the Google's cloud services, such as Gmail, Docs, Maps and so on. All these services are built on its powerful cloud computing platform. Some key technologies used by Google are briefly mentioned below.

A. Hadoop Distributed File System (HDFS)

Apache Hadoop is a Java software framework that supports data-intensive distributed applications under a free license. It enables applications to work with thousands of nodes and petabytes of data. Hadoop was inspired by Google's Map-Reduce and Google File System (GFS) technologies.

The HDFS is a distributed file system designed to run on commodity hardware[5]. HDFS is highly fault-tolerant and is designed to be deployed on low-cost hardware. HDFS provides high throughput access to application data and is suitable for applications that have large data sets.

A HDFS cluster consists of a single NameNode and a number of DataNodes. The NameNode is a master server that stores the metadata of file system, provides mapping between HDFS and the users' file system, and is responsible to manage the file storing service. The DataNode manages the storage of users' data on the node, creates and deletes data blocks.

B. Map-Reduce

Map-Reduce is a programming model for data-intensive computing, which consists of two functions: map function and reduce function [6]. The former applies a map operation to a set of items, and produces a set of intermediate key/value pairs on a group of distributed computer nodes. The latter merges all intermediate values associated with the same key on distributed computer nodes. A Map/Reduce job usually splits the input data-

set into independent chunks which are processed by the map tasks in a completely parallel manner. The framework sorts the outputs of the maps, which are then input to the reduce tasks. Typically both the input and the output of the job are stored in a file-system. The framework takes care of scheduling tasks, monitoring them and re-executes the failed tasks.

C. Bigtable

Bigtable is a distributed storage system for structured data, and is a very effective mechanism for storing very large amounts of data in a distributed environment[7]. A Bigtable is a sparse, distributed, persistent multidimensional sorted map. The map is indexed by a row key, column key, and a timestamp, each value in the map is an uninterpreted array of bytes.

Bigtable is built on several other pieces of Google infrastructure. Bigtable uses the distributed Google File System (GFS) to store log and data files. A Bigtable cluster typically operates in a shared pool of machines that run a wide variety of other distributed applications, and Bigtable processes often share the same machines with processes from other applications. Bigtable depends on a cluster management system for scheduling jobs, managing resources on shared machines, dealing with machine failures, and monitoring machine status.

IV. SERVICE ORIENTED ARCHITECTURE FOR GIS PLATFORM

The goal of the Service Oriented Distributed GIS platform is to provide both the dynamic GIS service components for publication vector and raster maps to Internet and interactive viewing on web browsers. It supports the complex rendering, navigation and allows working with multiple layers, thematic maps, hyper linked features and attribute data. The fundamental GIS platform is based on the Client/Server architecture. The server spreads vector maps to client. The map features are stored on a SQL database according OGC SFS and WMS that can be used as raster maps source. The GIS data interface supports the ArcView shape files and MapInfo MIF files. The server side is constructed based on J2EE architecture.

Web Services are the software components that are well defined, self-contained, and does not depend on the context or state of other services [10]. Web services essentially use XML to create a robust connection. The web services architecture has

three roles: a provider, a requestor, and a broker. The provider creates the web service and makes it available to clients who want to use it. A requestor is a client application that consumes the web service. The broker, such as a service registry, provides a way for the provider and the requestor of a web service to interact. The provider, requestor, and broker interact with each other through the operations of publish, find, and bind. A provider informs the broker about the existence of the web service by using the broker's publish interface to make the service accessible to clients. The information published describes the service and specifies where the service is located. The requestor consults the broker to locate a published web service. With the information it gained from the broker about the web service, the requestor is able to bind, or invoke, the web service. The following figure illustrates a basic service and cloud computing oriented architecture.

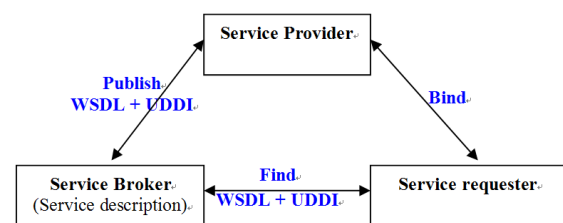


Figure 2. Web services oriented architecture

The web services use SOAP (Simple Object Access Protocol), WSDL (Web Services Description Language), UDDI (Universal Description, Discovery and Integration), and WSIL (Web Services Inspection Language) to communicate. Web services are published on the UDDI registry. A WSDL document defines the target service so the client knows what the service does. SOAP standardizes the way a Web service communicates with a client and allows programs written in different languages and on different platforms to communicate. SOAP works with standard Web protocols including XML, HTTP, and TCP/IP, as well as WSDL.

GIS Web Services conceal complexity. Web Services embed its complex data processing within itself in server side. Client side application developer make use of Web Services through standard interface which described in Web Services WSDL, therefore the application developer don't need dealing with Web Services internal process.

The following principles are considered in the design for the service oriented distributed GIS

platform. Web GIS platform should provide a variety of interfaces for integration and interoperability with business and GIS applications.

GIS Web services: Web GIS platform should support the standard enterprise application server technologies to build web services with embedded spatial services, software components and APIs to easily embed spatial capabilities in business applications built using standard technologies and architectures.

Web Mapping Services (WMS): WMS can display map layers from other GIS applications, and allow other applications to display map layers from Web GIS platform .

OpenGIS Consortium interoperability services and standards: The Web GIS platform should be implemented with the OpenGIS interoperability standards:

Geography Markup Language (GML): GML is a common XML-based file format for exchanging spatial information and attributes. The Web GIS platform should support the GM standards.

Common database environment: Specifications on spatial entities and the space/time reference systems have been provided by the OpenGIS Consortium.

V. CONCLUSIONS

The cloud computing technology brings an era for the geographic information services, greatly changes the concepts and ways of thinking of geographic information services. This paper presented a service oriented architecture for web service enabled geographical information platform. It is a distributed, platform independent system architecture that be accessed by lots of different computers in network with different kinds of operating system. The web service framework was applied into the GIS system architecture design. The GIS Web service was used to get and show the geographical information. The service and cloud computing oriented architecture is a very promising architecture for practical implementation of the next generation geographical information systems. It could have an important application prospect in the GIS tools development and application. GIS services to the cloud computing environment and build up a kind of geoCloud.

REFERENCES

- [1] Xiaolin Lu, " An Approach to Service and Cloud Computing Oriented Web GIS Application," Vol. 978-1-4244-5143-2/10, 2010 IEEE
- [2] Yang Xiaoqiang, Deng Yuejin, "Exploration of Cloud Computing Technologies for Geographic Information Services clusters," Sponsored by the project of National 863 Plan, program number: 2007AA120501.
- [3] YANG Jinnan, WU Sheng, "Studies on Application of Cloud Computing Techniques in GIS," *Proceedings of Second IITA International Conference on Geoscience and Remote Sensing*, 978-1-4244-8515-4/10, 2010.
- [4] G. Zhou and G. He, "One Program Model for Cloud Computing," *Cloud Computing*, pp. 589-594..
- [5] M. Vouk, "Cloud computing Cissues, research and implementations," *Journal of Computing and Information Technology*, vol. 16, pp. 235- 246, 2008.
- [6] D. Sui, "GIS, cartography, and the third culture: Geographic imaginations in the computer age," *The Professional Geographer*, vol. 56, pp. 62-72, 2004.
- [7] Shengru Tu, Maik Flanagan, Ying Wu, Mahdi Abdelguerfi, Eric Normand, Venkata Mahadevan, Jay J. Ratcliff and Kevin Shaw, "Design Strategies to Improve Performance of GIS Web Services", *International Conference on Information Technology*, Volume 2, 2004, pp. 444~449.
- [8] Jong Woo Kim, Seong-Seok Park, Chang-Soo Kim and Yugyung Lee, "The Efficient Web-Based Mobile GIS Service System through Reduction of Digital Map", *Lecture Notes in Computer Science 3043*, Springer, 2004, pp. 410~417.
- [9] Borthakur D .The Hadoop Distributed File System: Architecture and Design .2007.
- [10] Dean J.Ghemawat S. Map/Reduce:Simplified Data Processing on Large Clusters. 2004.