http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 13 (5): 948-953, 2014
ISSN 1812-5638 / DOI: 10.3923/itj.2014.948.953

Application of Peer-to-peer Data Transmission Mode in the WebGIS

[1]Hua Yang, [2]Zhou-Fang Li and [3]Wei Wei
[1]School of Information Engineering, Kaifeng University, Kaifeng, Henan, China
[2]School of Information Science and Engineering, Henan University of Technology, ZhengZhou, Henan, China
[3]School of Computer Science and Engineering, Xi'an University of Technology, Xi'an, 710048, China

Abstract: WebGIS is a kind of technology, which employs the Internet as its environment, adopts Web pages as the user interface of GIS software, combines the GIS technology with the Internet and provides all kinds of GIS application with GIS function. It's applied in all kinds of GIS field, However, In the condition of the upsurge in the number of clients, it's easy to form "bottleneck" of computing resources in the server side, causing the server load and network latency. It can affect the user's normal use in severe condition. This study solves the bottleneck of the server-side problem by changing the architecture of the existing WebGIS. The experimental results show that the new transmission mode can solve the bottleneck problem at a certain extent.

Key words: P2P Web, WebGIS, spatial data transmission

INTRODUCTION

The application of GIS is promoted in various fields in a very high speed within the global at the present, it's widely used in earthquake prevention and disaster reduction (Yao, 2002), environmental monitoring (Yi, 2007) E-government (Li and Ning, 2006), agricultural production (Chen and Mao, 2003), Land and Resources Information Services (Xiong and Zhao, 2006), Electricity sector (Xu *et al*., 2005), Water conservancy, transportation, communications, marine and military and many other fields, Desktop GIS has been unable to meet the needs of resource sharing the information age, Currently, WebGIS based on network technology and GIS is an important direction of research, However, the server of WebGIS will receive a lot of requests with the increasing of number in the case of the client, and the server's computing power is limited, so it will cause the server's load is large to deal with these large number of requests.

Literature, Tu (2008) put the P2P web into the WebGIS and use a centralized P2P web architecture, where, the specialized data resource server and specialized directory server are provided in the system, the data resource server is used to store all of the spatial data, address information of each machine node and the related data information that this machine node is requesting at the current time are recorded in the directory server, each code provides services as a server, at the same time, it also gets data from a neighbor node as a role of client node in the network. When a client requests the data, it first makes a request to the directory server, then it looks for whether a machine is downloading the same data, if there is, then it obtains data from this machine, if not, then obtaining data from the resource server and registering in the directory server.

Literature Yu *et al.* (2008) also use the P2P technology, but they use a distributed unstructured peer, where, a node finds resource by using the flooding way, first, it asks whether neighbors have the requested resource, if not, continuing to look from neighbors' neighbors, until it finds success or failure.

The whole system is divided into ordinary node layer, super node layer, the server layer by using hierarchical organizational structure in literature, Pan *et al*. (2009). All nodes within the network are divided into many groups and a super node which has the best performance is chosen in every group, the rest are ordinary nodes. Members in a same group can communicate directly with each other, can exchange their resource lists and can query the data stored by other members of the group through the super node. When a node looks up data, first it finds from the data resources distribution table in local, if the search is successful, it access to the data directly, when the search is unsuccessful, it will send a query request to the super-node. Super-node will query data from other groups to establish a connection if the search is successful and get the data, else it will get data from the server.

Despite a number of methods have been developed to improve the efficiency of WebGIS' data acquisition and achieved good results but there are areas for improvement. Literature, Zeng *et al.* (2008), Xu *et al.* (2007) and Duan and Hu (2008) have changed the architecture of WebGIS but they use a centralized peer network structure, the central node's failures may cause

Corresponding Author: Hua Yang, School of Information Engineering, Kaifeng University, Kaifeng, Henan, China

the system to crash, resulting in the entire system does not work properly and because of the probability that two nodes download a data at the same time in reality is not very high, it may lead to algorithms in the literature can not obtain good results. They query data using the flooding way in literature Lu and Wu (2004) and Li and Zhang (2007), this way will increase the network traffic, take up a lot of bandwidth and increase the burden on the network. In literature, Yu *et al*. (2008) and Pan *et al*. (2009), because of it is very frequent that nodes join and leave the P2P network, so it has led to waste a lot of bandwidth in the exchanging information.

WEBGIS IMPROVED ARCHITECTURE

The improved architecture of WEBGIS is called P2PWebGIS, it takes full advantage of the hosts resources within the P2P, so that each host both as a client can obtain data, but also as a server can share data with other nodes, thus it also can increase the ways that clients obtain data, reduce the number that client access to servers, relieve the pressure on servers and improve the speed that client obtain data.

DESIGN OF P2PWEBGIS

The logical structure of P2PWebGIS is shown in Fig. 1, which is divided into four levels, namely the presentation layer, resource retrieval layer, WEB layer and GIS application layer.

Presentation layer is the public interface of interaction between the user and the system, it is responsible for the user's request should be sent to the server and receive the data from server. After the data is received, presentation layer organizes the data in logic and displays them on the browser for users to view, it also can response to the user's actions, complete map browsing, panning, zooming and other operations when

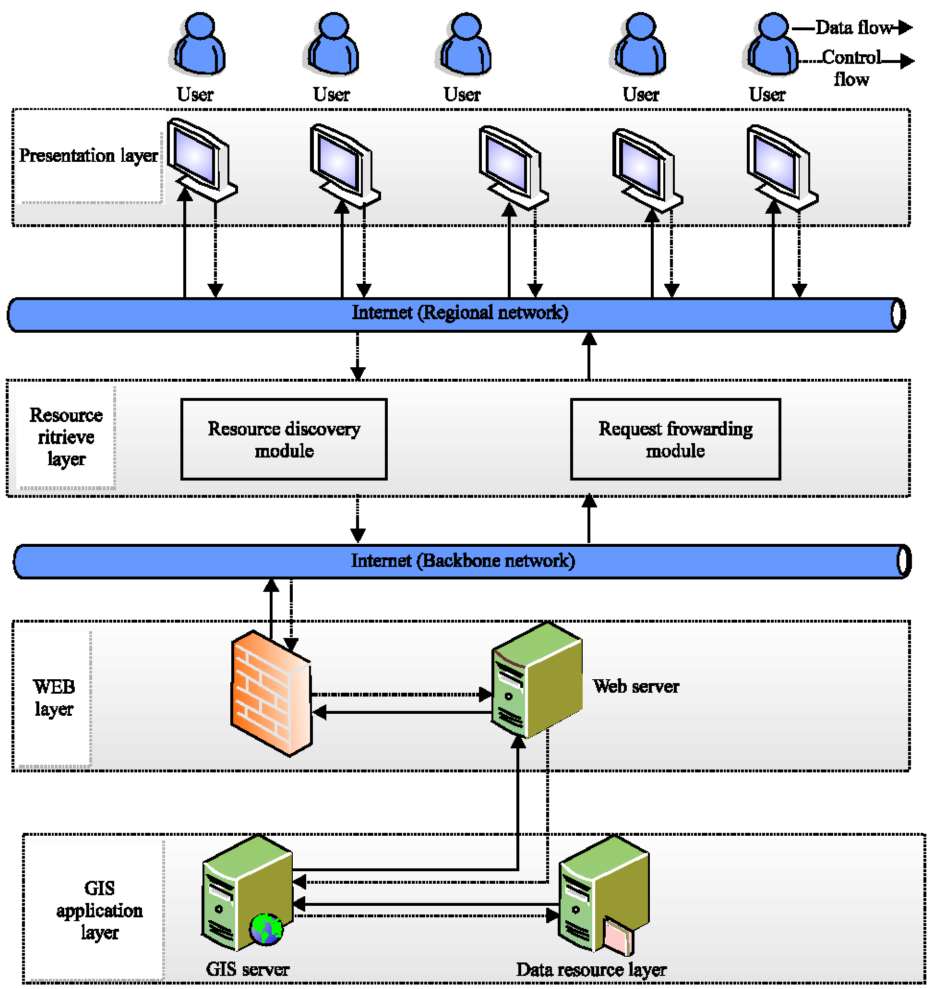

Fig. 1: P2PWebGIS structure design

users browse the map. Firewall to protect the stability and security of the system at run-time. After receiving the request, this layer resolves the user's request and calls GIS application layer function module for processing by using the remote object access technology and then returns the results to the client.

When a user requests data resource to the server, resource retrieval layer is responsible for finding the nodes which stores the data resource from network according to the file name requested. After finishing the research, this layer selects a host which has the minimum load, maximum network bandwidth from these nodes and servers which store the requested data, then it gets data from this host to reducing the pressure on the server and improving the speed of accessing to data.

Web layer is mainly responsible for receiving a service request coming from the client and blocking out the illegal operation of the user with the help of the GIS application layer is mainly to complete the definition storage, integrity constraints, retrieval of data and the relational database management work, this layer also can receive and process user's requests submitted by the Web layer and return the results to the Web layer. GIS application layer is the most important of all layers. All the analysis and processing of GIS data are completed by the GIS server in this layer within the system.

Adding resource retrieval layer in the original system agencies is the improved architecture, it can solve this server computing "bottleneck" and bandwidth "bottleneck". Resource retrieval layer contains two modules, namely, resource discovery module and the client requests forwarding module. The main functions of resource discovery module are to retrieve resources in P2P, find the best host in ability from these hosts and servers which store the resources, retrieve data from the best host and reduce the number of requests to server which are sent by client in order to alleviate the pressure on server. When the resource discovery doesn't find the host which stored the query data within the P2P or when the performance of all the hosts storing the resources is poor than the server, client requests forwarding module forwards the request of user to the server and gets data according to OGC standards.

DESIGN AND IMPLEMENTATION OF RESOURCE DISCOVERY MODULE

Resource discovery module stores resource information by using a distributed hash table and searches for the hosts storing the requested data in the P2P by using the Chord resource lookup mechanism so that the client can obtain data from other clients within the P2P, reduce the number of requesting server and change the data acquisition mode existed. The module works as follows.

Joining node: When a user first makes a request to the server, it begins executing the operation of joining node, adding the node to P2P.

Resource information dissemination: When a node joins in P2P, it needs to perform resource publishing operations, share data in the cache. When a node acquires tiles from another node, it needs to perform resource publishing operations too and share their newly acquired tile data. Resource publishing contains two steps:

- The current node calculates the hash value of the tile name by using the hash function and then it looks up the node S which stores the resource basing on the hash value in the entire network, this process is similar to the process of data lookup
- Adding the key-value pair <key, y> into the distributed hash table which is saved by node S, key is the value which is got by hashing the tile name, y is the IP address of the node which stores the tile data

Client requests data: According to the current operation of user, the client sends the standard request format to the resource retrieval module, the format is got by processing and combining the relevant information of the resolution and the latitude and longitude coordinates range of the data needed to obtain, then it tells the resource retrieval layer which data needs to be found.

Data search: When a local resource retrieval module receives a data request, first it calculates the names of all of the requested tiles according to the standardized naming convention of spatial data, then it hashes the name of each tile into an M-bit binary value by using the compatible hash function, this value is called key value expressed by key, then looking up the host nodes which save this key value according to the Finger tables that saved by all nodes. Because the tile data searched may have many backups within the P2P, therefore, it can find a lot of hosts that store tile data in the hash table and get all the hosts set {S} including remote sensing data.

Access to data: If the number of hosts inside set {S} is greater than zero, then finding host C that has the most computing power and the best network bandwidth from this host set {S}, getting the overall performance index Ci of this host, then getting the performance index Cj of WebGIS server, comparing the size of Ci and Cj, if Ci is the larger, getting data from the host C, else, getting data from the server according to the OGC specifications. If the number of hosts inside set {S} is less than zero, but also directly getting data from the server basing on the OGC specifications and then storing data into the cache in accordance with standardized naming.

Client displaying: The client displays the acquired data in the web browser, users can easily browse, zoom in, zoom out, drag and drop operation.

Exit node: After user completes browsing data in the client, it needs to perform the exit operation and it must do two things in exit node: The first thing is to update the current node's successor pointer field following the previous nodes and the current node's predecessor pointer field in the subsequent nodes, the second thing is to transfer all the keywords stored in the current node to the successor node.

DESIGN AND IMPLEMENTATION OF THE REQUEST FORWARDING MODULE

If it doesn't find the tile data in the P2P or the server's performance is better than the performance of all clients saving tile data, client requests forwarding module is responsible for converting the data request sent by resource discovery module to meet the OGC standards, forwarding the data request to the server and waiting for the server's data and receiving data that is returned by the server, this module's work flow is as follows.

Receiving the data request: Because not all the data is obtained from the P2P, when the client gets data from the server, the client requests forwarding module should receive the data requests coming from resource discovery module and request data needed by client to the server.

Request data: To meet the compatibility criteria, so that the existing WebGIS server can be applied to the new architecture, when the data is requested, the requested conditions are organized into requested data in accordance with the OGC WMS and WCS protocol standards in the paper and then sending the request to the server again.

Access to data: The module receives data returned by the server, names the data according to the standard organization rules and stores data into the cache specified by the client.

Resource publishing: After the data is received, it notifies the resource discovery module for resource publishing and shares the new data obtained from the client to P2P.

EXPERIMENT TO TEST

Experimental environment:

- Hardware environment
- Client computer: 12 units
- Configuration: CPU: Intel Core i3 dual-core processor, clocked at 2.2 G; Memory: 2 G; NIC: 100 M
- Server: 1 unit
- Configuration: CPU:Intel Core i5 dual-core processor; clocked at 2.5 G, Memory: 4 G, NIC: 100 M
- Switch: 1 unit
- Software environment
- Server: Operation system: windows 7; Web server: GeoServer 2.1.1 version
- Client: IE browser, P2PWebGIS client

Experimental design: This experiment mainly tested the WMS service of WebGIS, when clients access to the data, it tests the pressure created on the server side expressed with the CPU utilization and obtains the time that the client access to data in seconds (S) as a unit.

Test existing architecture: In the 12 computers, each runs a set of tools Microsoft Web Application Stress Tool which is developed by Microsoft's website testers specifically to test the actual pressure of website. Through this powerful stress testing tool, you can find the possible impact on website's service by using a small amount of computer simulation a large number of users on-line. Each client requests to the server 1000 tile data through this tool to simulate large number of users to access, because of the data obtained by this architecture only serves for the local client, therefore, it obtains this 1000 tile data in order in the twelve clients and doesn't get duplicate data in the middle of obtaining data, the program will automatically end when it finished that getting the 1000 tiles data. Observing the CPU utilization of GeoServer server during the middle of the process and accessing to the calculation pressure of server-side.

Test P2PWebGIS architecture: In P2PWebGIS architecture, each acquired data is random, it may have been obtained before, may not have been accessed to, the data may be in a neighbor client, the data may be not in the neighbors and join in the loop statement to get data, in the experiment, it not only requires a client to obtain 1000 tile data, but also the twelve machines should send

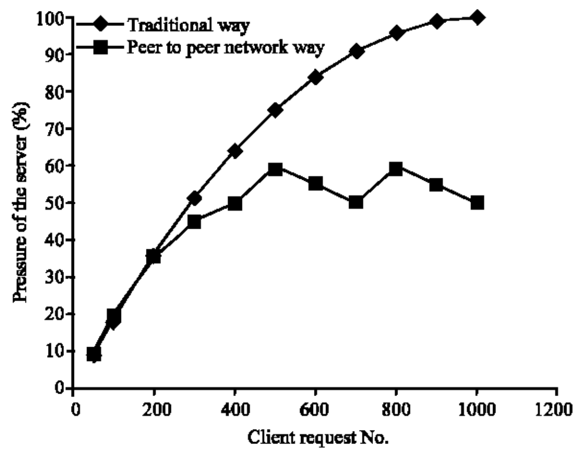

Fig. 2: Server CPU usage

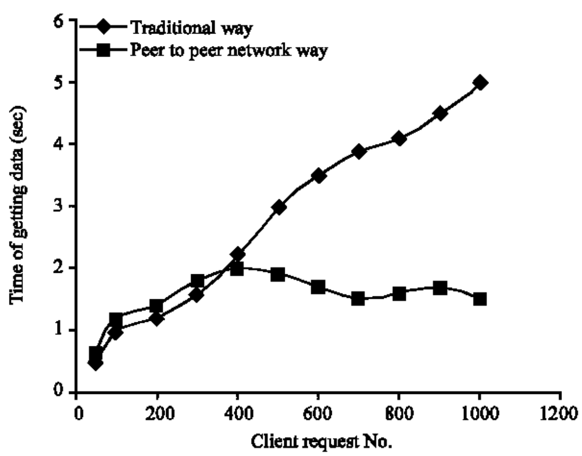

Fig. 3: Client data acquisition time diagram

requests simultaneously, it gets the data of pressure on the server uninterruptedly during the process that clients obtain the data.

Experiment analysis: Through testing the WMS service, the data are shown in Fig. 2.

Through Fig. 3 it can be seen that the two architectures have the same pressure on the server at the beginning but when the number of requests is greater than 200, the pressure on the server which is the WebGIS server using the exiting architecture increases with the increasing of the number of users, but the pressure on the WebGIS server using the P2PWebGIS architecture shows a stable trend with the increasing of the number of users, sometimes there are fluctuations in the situation, but volatility is not obvious. The main reason that occurs this situation is: at the beginning of WebGIS, client node has no data, then the two architectures, like, must get data from the server, so the pressure on servers is the same, but after running for some time, the client node has enough cache data, the advantage of P2PWebGIS can be reflected, a part of the data can be obtained from the neighbors, so the pressure on the server tends to balance.

Through Fig. 3 it can be seen that the two architectures have the same time to obtain data at the beginning, but when the number of requests is greater than 400, the time of obtaining data by server which is the WebGIS server using the exiting architecture increases with the increasing of the number of users, but the time of obtaining data by the WebGIS server using the P2PWebGIS architecture shows a stable trend with the increasing of the number of users, sometimes there are fluctuations in the situation, but volatility is not obvious. The main reason that occurs this situation is: at the beginning of WebGIS, client node has no data, then the two architectures, like, must get data from the server, so the time of obtaining data is the same, but after running for some time, the client node has enough cache data, the advantage of P2PWebGIS can be reflected, a part of the data can be obtained from the neighbors, so it can achieve the effect that the speed of obtaining data becomes faster with the increasing of users.

CONCLUSION

The main contents and conclusions of this study are as follows:

- Through study and research about the related technology of peer and the analysis about the architecture of WebGIS, the study designed the architecture of WebGIS based on peer
- Through the use of peer technology, so that each client can contribute their own resources, provide services to other clients and reduce the pressure on server when the number of users increased dramatically
- Through the development of a WebGIS basing on P2PWebGIS, it is verified that this architecture can solve the bottleneck problems between in the server-sides and can reduce the pressure on the server

The future research directions of this study are as follows: Data consistency issues still need to be considered in this study, because the spatial data needs to be updated, how to make each WebGIS client's data are up to date and ensure that the shared data is valid is the direction of future research.

ACKNOWLEDGMENTS

We would like to thank the anonymous reviewers for their valuable comments. This work is supported by National Natural Science Foundation of China (No. 61203094, No. 11226173 and 61102163), the project of Henan Province (No. 12A520009) and the Science and Technology Development Project of Henan Province (No. 122102210230). This work is also supported by Scientific Research Program Funded by Shaanxi Provincial Education Department (No. 2013JK1139) and Supported by China Postdoctoral Science Foundation (No. 2013M542370) and supported by the Specialized Research Fund for the Doctoral Program of Higher Eduacation of China (Grant No. 20136118120010). This work is also supported by Scientific Research Program Funded by Xi'an University of Science and Technology (Program No. 201139). This job is also supported by Science and Technology Project of Xi'an (CX1262⑨).

REFERENCES

Chen, S.H. and D.M. Mao, 2003. The establishment of xinjiang nutrient partition management system basing on WebGIS. J. Modern Agric., 11: 27-28.

Duan, L. and B.Q. Hu, 2008. Research on the efficiency of the spatial information service in the P2P network. Proceedings of the 16th International Conference on Geographic Information Science and Technical, November 5-7, 2008, Irvine, CA, pp: 147-149.

Li, Z. and S.N. Ning, 2006. Research and design of electronic document flow basing on WebGIS. J. Comput. Eng. Appl., 35: 191-194.

Li, C.H. and W. Zhang, 2007. One kind of WebGIS architecture based on P2P network. Mapping Geograph. Inform. Space, 5: 22-24.

Lu, R. and J.P. Wu, 2004. Discussion about the applications of P2P technology in WebGIS. Mapping Geograph. Inform. Space, 27: 7-10.

Pan, S.M., Z.W. Yu and H. Wang, 2009. Massive terrain roaming search model based on P2P. Appl. Res. Comput., 26: 1659-1662.

Tu, Z.F., 2008. Design and Development of WebGIS Basing on Ajax. Central China Normal University, HUBei.

Xiong, D.L. and J.S. Zhao, 2006. Design and implementation of land and resource network information system basing on WebGIS. World Geographical Inform., 5: 17-20.

Xu, T.Y., X.P. Wang and Y.L. Cao, 2005. Research about network query system of rural power resource basing on WebGIS. J. Shenyang Agric. Univ., 6: 718-720.

Xu, J., C.H. Li and T.Z. Chen, 2007. Research about the data transfer method of GIS basing on P2P technology. Comput. Eng. Design, 28: 2851-2853.

Yao, B.H., 2002. Research and implementation of shock mitigation system basing on WebGIS. Institute of Engineering Mechanics, China Earthquake Administration (CEA), China.

Yi, Z.H., 2007. Design and Implementation about the Environmental monitoring and analysis of distributed WebGIS. J. Mapping Scient., 1: 138-139.

Yu, Z.W., S. Zheng and Z.M. Li, 2008. One kind of large scale terrain data transfer mechanism based on the hybrid of P2P. Mapping J., 37: 243-249.

Zeng, M., Y.X. Yang and Y.P. Feng, 2008. The application of P2P technology in the campus GIS planning and management. Comput. Modernization, 4: 88-90.