

# A Service Pipelining Model for Web GIS Based on Distributed Environment

Weng Yu and Cheng WenYi

*College of Information Engineering, Minzu University of China*  
*dr\_wengyu@126.com*

## Abstract

*Pipeline is one of the classic technologies to enhance efficiency. When the amount of data is becoming more and more big when processing, it becomes a big problem that how to improve the efficiency with existing resources. In the meantime, a large number of older hardware and software feel difficult to deal with new task in new environment with the accumulation of time. In order to ameliorate efficiency of web GIS with low cost, this paper propose a service pipelining model for Web GIS based on distributed environment. The system build a basic service library to reuse plenty of existing resource, and design a segment system to organize pipeline in level of service to improve efficiency. The model can be used in many platforms and occasions, not limited to the application we proposed in this paper. The model we proposed has been proved work well in our experiments and test.*

**Keywords:** *web GIS, pipeline, service composition, software reuse*

## 1. Introduction

When processing large data and attempting to provide flexible service, there is still issues due to different data source and changeable user needs [1, 2]:

- a). Not all types of data processing are suitable for Map Reduce model.
- b). With the years of accumulation, the existing technology to the new platform or system migration often need a lot of previous work repetition.
- c). the user's demands changes frequently and cannot accurately predict.

In order to combine various existing resources with low cost, improve the efficiency of data processing, construct fast and flexible service platform, we proposes a service pipeline system based on distributed environment in this paper. The system is to build a library of basic services, connect distributed environment and running basic services by segment system, and provide a variety of complete services via pipelines.

Although there will be bottleneck and running conflicts in pipeline technology, the proper use can significantly improve the throughput of the system, improve the efficiency of the system.

Construction of the application or machine level pipeline is one of the effective ways to improve system performance and efficiency, and in a distributed environment in the form of service to organize data and computing resources, it can integrate all resources flexible and safe. Build application level or machine level pipeline is one of the exertive means to improve system performance and efficiency, and organize data and compute resource in services can integrate all resources flexible and safely.

We did a detailed analysis and depict four key problems in Chapter 3, and propose our model in Chapter 4. In Chapter 5, we accomplished a instance and an construct system to test and verify our theory.

## 2. Related Works

Few researcher involved in pipeline system of services though pipeline system has been applied in manufacture and micro level of instruction widely [3-5]. The service pipelining model for Web GIS based on distributed environment we proposed relay.

For purpose of code reuse and convenient system integration, large-scale system and open system usually adopts component model. Component model of the system is not only conducive to code reuse, and makes system integration easier and more flexible, the late maintenance more convenient and cost less. Function modules act as components in component model system, and Sometimes referred to as plug-ins .Eclipse, IBM's contribution to open source programming IDE, is a typical component model software. Users install different plug-in according to different requirement, to undertake a variety of programming language development. Most dynamic link library (DLL) of Microsoft's Windows operating system are essentially provide as components. These DLL not only provide services for Windows itself, many third-party software also need to call the interfaces they provided to take advantage of system resources. In addition, Many Hara, h. *et al.*, this paper proposes a proxy-based approach of software component reuse way, make the software components in get used again in another way. [6]

In von Neumann computer's microprocessor system, pipeline development has experienced from single-function line to multifunctional line, linear assembly line to the nonlinear assembly line. Issues, such as the conflict, related problems in Instruction pipelining technique, has a mature solution. In industrial, each processing act as a segment, and segments connect each other, assembling materials and parts automatically. Each process line has one or more lines, compared to the mill production, greatly improved its production quantity per unit time.

There is few researchers involved in software service level of pipeline. When deal with the rapid processing of large amounts of data, researchers tend to use model of Map Reduce. But many types of data processing is not suitable for use Map Reduce to deal with. In these cases, people used to improve the server configuration to improve the efficiency.

Web Service technology widely rely on technology, such as TCP/IP protocol, the HTML standard, and. NET platform, Java, XML specification has been very mature. [7,8] Web Service platform system XSD data types and object information transmission protocol SOAP and Web services description language (WSDL), as well as the electronic commerce standard UDDI.[9,10]there are various forms of implementation and encapsulation. Combination between Web Service services, service is also a hot topic. [11, 12, 13] Maolin Tang and Lifeng Ai proposed a genetic algorithm based on constraints of best choice service composition method, good results have been achieved in the evaluation experiment. Chao Ma and Yanxiang He put forward a kind of service composition visualization and formalized method, greatly simplifies the operational ease of service composition. [14]

Our work not only come down to both software reuse and service combine, but also include hardware reuse and large data processing.

## 3. Problems Describe and Analyses

Service pipeline is the core of the system. Which faced many questions in design scheme of service pipeline system, four of the most important questions are:

- a). What form service pipelines?
- b). How to constitute a service pipe line?
- c). How to solve the runtime issues?
- d). How to combine services?

In order to solve these questions, we did a lot of assumptions and arguments before Summarized as follows.

### 3.1. The Resource of the Service

Service line system relay on a large amount of basic service. Basic services running on line system in the service is best be some function relatively single, less operation time.

A segment is made of a basis or several basic services .Considering the Efficiency, basic service shouldn't cost too much time.

In order to meet the requirements of the pipeline system, and as far as the integration of all possible reusable resources, we construct a basic service library, to fusion, standardize all kinds of service, to provide basic services to candidates for the pipeline system.

Basic service library is make of numbers of services with function relatively unitary and process relatively simple. The basic services in library mainly has the following sources:

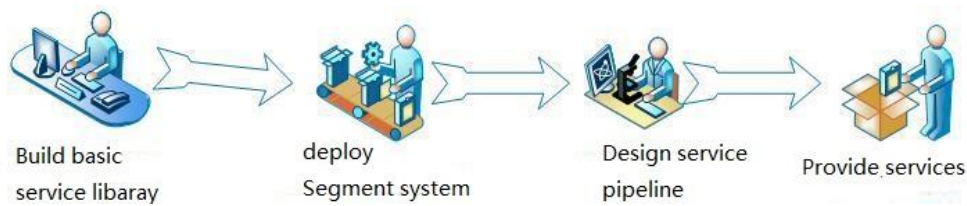
- a). Re package independent existing functional module .Many operational functions can be achieve in different forms, such as the jar format of the Java runtime library under Windows and DLL Dynamic link library, under the environment of Linux so dynamic link library and the.NET runtime library *etc.*. They running in different environment and provides API in different forms. As a result, there is much difficult in using them in a certain environment. We can repackage these resource, expose API in a Unified standard, and provide services by running these repackaged resource on machines equipped necessary facilities
- b). Repackage open web service and authorized web service. Different web service has different way to call, and different interface. A web service could be used more than once. Which means that once the interface of a web service changed, many point of the system should be updated. It is necessary to repackage these web service and expose an unified interface to reduce dependence on external characteristics of the web service system stability [15, 16].
- c). Divide service when adding new service. A service can be divided into some small services. Divide services  $M_1, M_2, M_3, M_4, \dots, M_n$ , into  $C_1, C_2, C_3, C_4, \dots, C_n$  services respectively. The total services we get by dividing these n services  $C_{total} \leq C_1 + C_2 + C_3 + \dots + C_n$ . That's because service  $M_i$  and service  $M_j$  may contains some same functions. If the basic services library is large enough, we just need to add few basic service to library to combine the new service.

### 3.2. Cooperation in Machines

The cluster's communication adopt TCP/TP agreement. Segment system amongst will adopt query mode to conversation forwardly. segment system amongst and pipelined computer system adopt heart beating model to retain communication trouble-free operation, that is to say it will transmission an inquiry signal voluntary and make decision based on feedback information when some system require others resources or need to know present states. Machines in cluster work together like workers in a workshop. Each machine accomplish a part of the task they skilled. We can design a special subsystem to organize the communications between machines to make them work in lines.

### 3.3. Design of Service Pipelines

Pipelined computer system utilize widely in industrial product and WEIZHILING in recent years, however, service pipelined computer system mention rarely.



**Figure 3-1. Steps to Build System**

First of all service function should more simplify and structure uncomplicated to unified standardization service's ends, then before republication simply repackage the opening web service to unified standardization service that structure basic service library. According to user or user's demand to combine service form to structure service pipelined computer system that formed by basal service tin basic service library, which purpose to increase of system efficiency and throughput capacity.

As picture 3-1 described ,we integration software resource to structure basic service library ,then package to pipeline system and design reasonable pipeline to render service for external.

System can take advantage of software resource and hardware resource, insulate user ` service manifestation mode and basement hardware source to render single standard service .In every node arrange pipeline service system to link hardware resource layer and pipeline.

Segment system is a program working on a node machine. It makes the machine work as segment in a pipeline. The principle of segment is as shown in flowing picture.

Segment systems has no difference, while the machines and basic services they rely on are different. The differences between services is born, which is related to their functional characteristics. The difference of machines can be controlled and adjusted, we should deploy services on right machines. Segment system is an important layer isolation hardware and service in the whole system. Some basic services need larger memory, is suitable for deployment in the configuration memory on the machine running. Some services to the floating point arithmetic demanding, had better first deployed in arithmetic processor machines; Also some services need outside the network access (*e.g.*, invoke external Web services), you may need to have the twin wire machine. Service combined with the machine properly can take the best advantages of hardware and software.

### **3.4. Runtime Issues**

A service composition system services divide a service into a number of basic services, according to which we design service pipeline. Responsibilities of pipeline scheduling system is to realize static optimization assembly line, assembly line, assembly line, dynamic adjustment assembly line. Service composition should avoid foreseeable conflicts, to avoid the loop in the design the service pipeline. Pipeline scheduling system monitoring lines running state, dynamically testing bottlenecks, and open the bottleneck to repetition.

## **4. System Design**

According to the four key problem our system has four kernel subsystem and a configure system. The basic services system is the base of whole system, and the segment is the bone of the system. However, the service pipeline dispatch system and service combine system determine the effect and performance. The configure system make the system much more easy to use and operate. The flowing picture shows the relationship between these subsystems briefly.

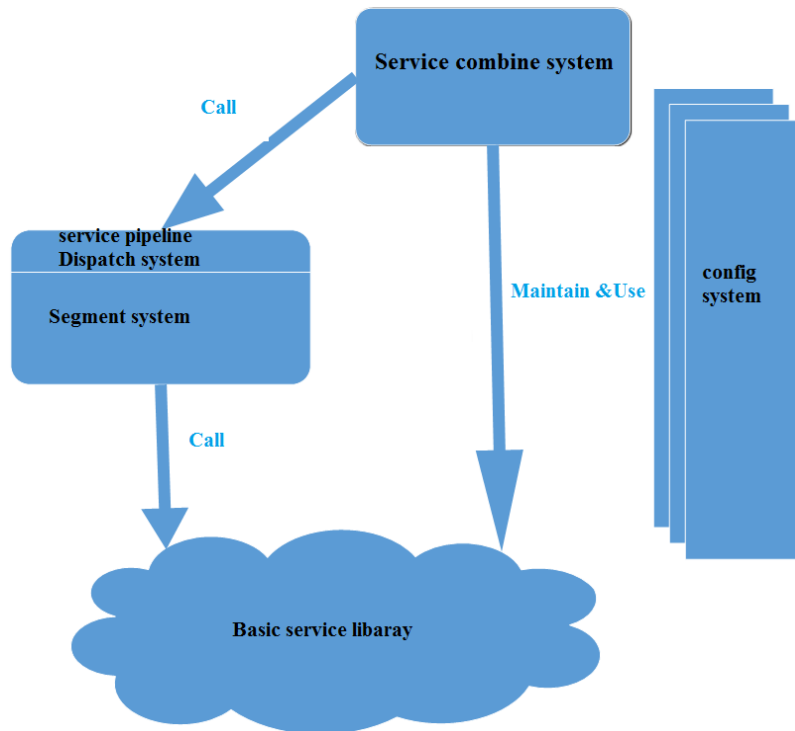


Figure 4-1. Segment System Module

#### 4.1. Basic Service Library

By means of repackaging, basic service library can resume plenty of existing software resource. With help of Hybrid programming, we use JAVA code to call these resource.

Packaged services Comply with the uniform interface rules, and should also be marking service some runtime features, known as the system resource usage and operation need of environment, *etc.*

#### 4.2. Segment System

The segment system of service pipelining system includes input model, process model, and output model.

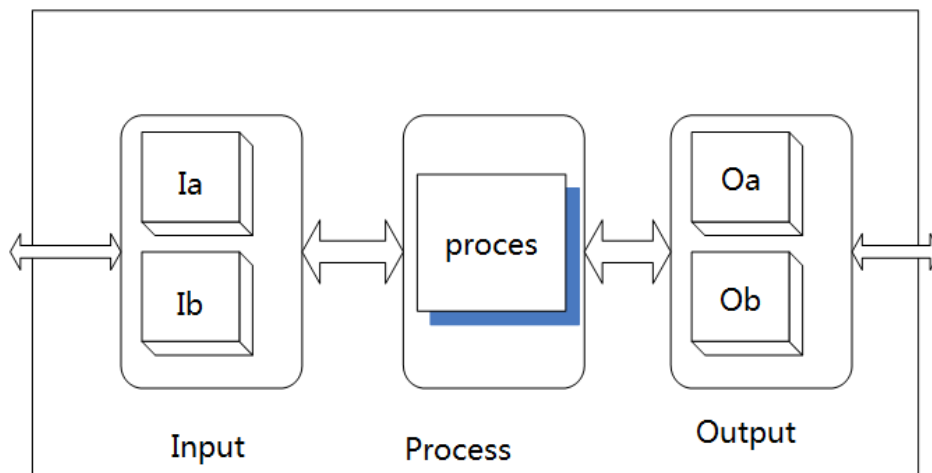


Figure 4-1. Segment System Module

Input model has two input buffer, and a thread for input data receive. Input receive thread respond the previous segment's request, and accept task when one or more input buffer is accessible.

When the process model fetch input data , it take the data in buffer who is received first and remark it as free after get the whole data in this buffer.

Output module includes two out buffer and network port. The network port is used to communicate with next segment.

When both two output buffer of output module is mark as busy, the output module will refuse the process module to write. When output module send data, it chose the buffer who filled first.

Service pipeline system divide large data process into some simple process, and deploy these basic services on different machines. As a result, we can modify the machine to specialization to be good at process a certain kind of data to improve the data throughput and reduce the response time.

### **4.3. Service Pipeline Dispatch System**

Assembly line scheduling system is responsible for the coordination between the periods of system communication, determine the basic service deployment, connection system become the assembly line, and solve the production line running problem.

Service line scheduling system itself in the form of services deployed on the platform, as a key service system to allocate the basis of other services.

### **4.4. Service Combine System**

Service composition system has two main functions. First, connect some small service portfolio based service in the library, making it a direct response to the upper demand of service. Second, the number of need to improve performance or efficiency of the service chain to pipeline scheduling system, to form the service line.

Isolated from the service composition system services and the implementation of the form, the system on a unified service, use of the service line below to efficiently complete the service function.

### **4.5. Configure System**

Configuration management system involves the configuration of the whole system and monitor, basically has the following features: firstly, provide a graphical interface real-time display system of the current situation. Secondly, Manual fine tuning system parameters. What's more, Record log analysis system. Besides, Provide maintenance based service library interface and System security monitoring. Visibility, controllable configuration management system make the whole system became much easier to use.

## **5. Experiments and Analyses**

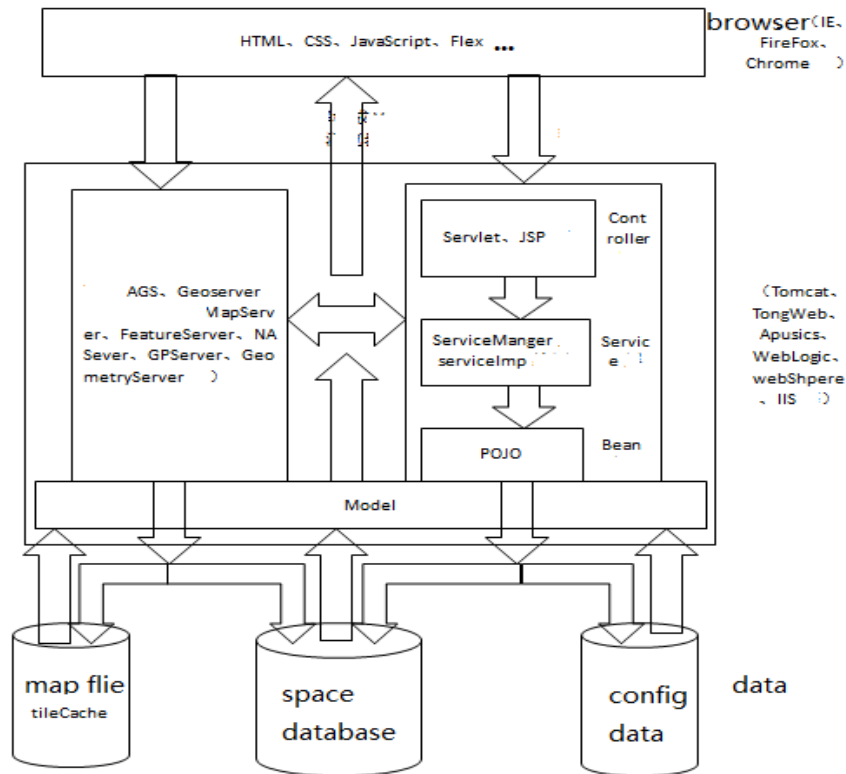
### **5.1. Web GIS based on Service Pipelining Model in Distributed Environment**

GIS (Geographic Information System) is a system supported by computer hardware and software, for the whole or part of the earth surface (including the atmosphere) space about the geographic distribution of data acquisition, storage, management, operation, analysis, display and description of technical system [17].

Since the last century in eighty, a large number of GIS software emerged. They are basically C/S architecture, and professional GIS software should be equipped before users use the GIS software based on their requirements. The desktop to the GIS software has its natural advantages, such as the development of relatively simple, does not need the Internet support, and can do more complex functions, relatively rapid response, *etc.* But at

the same time, its disadvantage should not be ignored. Be the first to bear the brunt of is not convenient to popularize, and inconvenient popularization reason lies in the natural disadvantage of C/S system itself: not easy to update, not easy to cross platform, not easy to use user specified in not to install the client's case [18, 19].

As the Internet into explosive growth stage, the network of laying and speed of ascension have dramatically increased, which provides a solid environment for the development of Web GIS. So GIS system based on B/S is more and more began to provide services, and in the RIA (rich client) technology, AJAX (dynamic web pages) and mature, the emergence of Web GIS also can show the same effect as C/S and function [20]. And the big development of Web GIS, but also the needs of users.



**Figure 5-1. Construction of Web GIS**

Users of personalized customization needs more and more high, so the open-source community, the emergence of a large number of free and open Web GIS technology, Open Layers is an Web GIS technology born to meet the need of user customization.

The fundamental framework of Web GIS is shown in the figure below:

The framework of Web GIS framework with other Web projects don't have a lot of difference in nature. The only difference is that the Web GIS function services need to provide some map services.

As all of the Web project, the front-end consists of HTML, CSS, JS, Flex language composition and so on. According to the requirements and language features, select the appropriate language development is of vital importance.

Compared with other general Web project, Web GIS has an obvious difference reflected on map server. General map server are provided by the existing software solutions, such as AGS, GeoServer, MapServer, SuperServer, etc. Use the map server we can realize the map publishing and good support for some spatial query and spatial operations.

Structural features of GIS data is not obvious, and have a variety of forms. We take Hbase as main storage system to strengthen the extension and Keep the original structure data

HBase is a distributed open source database that and stored in the column, the technology comes from Fay Chang wrote Google paper "Big table: a structured data distributed storage system". Just like Big table using the Google File System (File System) provided by the distributed data storage, HBase on Hadoop provides a similar to the ability of Big table. HBase is Apache Hadoop's subprojects. HBase is different from general relational database, it is a suitable for unstructured data storage database. Another different is that HBase based on column model rather than the row model.

HBase can organize the minority areas of data rapidly and effectively. Through the API we can quickly locate the position of the data.

We import and southwestern minority ethnic characteristics of ancient villages in Hbase data, as the two data services on GIS.

Web GIS can be divided into several basic services, such as GIS management services, map WMS (web map service) service, image file manage service, *etc...*

GIS data service can be divided into basic services like data storage service, data query service, service read data point to the file, data screening services, and data clustering services.

The services described above will be added into basic service library service, and Web GIS service and GIS data exist in the Foreign Service list, but no independent in system implementation, by the foundation to implement service composition, connection. Minority ethnic village will use web GIS and GIS data services.

In this instance system, we can design two pipeline respectively according to the web GIS service and data service. When designing the pipeline, we follow the principles, of "simple line, balance, avoid conflict, predict bottlenecks", and minimize runtime issues. The two service pipeline has no obvious conflict because we have enough equipment as redundancy.

The system will be deployed on the 10 PCS at first, s a reference system to determine the configuration of contrast system. Then, cluster scale will be dynamically extended to 20 PCS, 30 PCS, and comparing with the size of the initial 10 sets respectively.

System is running in a distributed network environment. A distributed environment for classic master-slave structure. The Master node is responsible for monitoring the operation of the whole system, and scheduling system resources. Slave nodes deployed on a variety of basic services, or pick up virtual Foreign Service.

The principle of system deployment diagram as shown in the figure below:

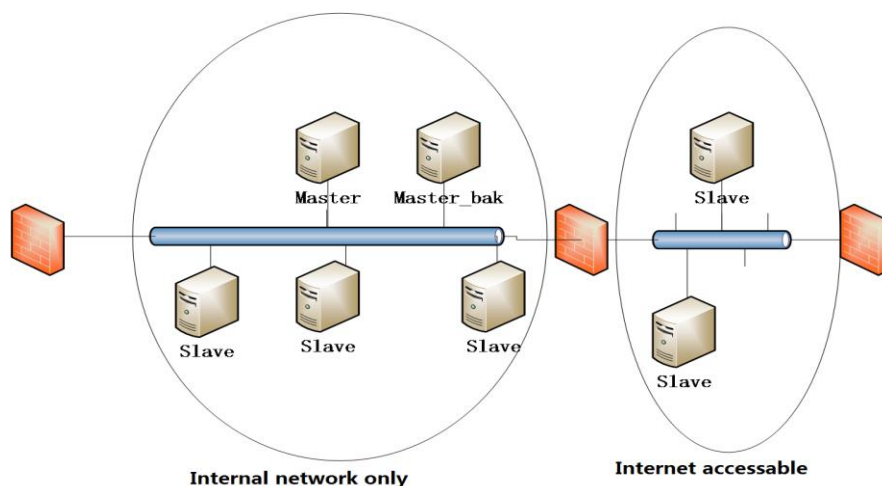


Figure 5-2. Network Deployment Diagram



The Environment and external network system operation is independent of each other, and the internal cluster refused the directly access of external network. Because the basic Service repository contains a part of web services come from external network, and a part of the Service requester feedback data, machines carry these services need to have the ability to access external network. We isolate a part of machines with a firewall to access the external network considering the security of internal cluster. These machine can access external network and exchange a certain kinds of information.

### 5.2. Constract System

Single server edition web GIS share the same principle with web GIS described above. The only difference is that Web GIS service will not be split, and only deployed on a single server When deploying.

The data of contrast system will also be provided by Hbase cluster, as same as our system. The modules Call the Hbase data query, filter, such as clustering function are deployed on the deployment of Web GIS server.

In order to guarantee the comparability of the system, we follow the principle of "cost" to select the server to deploy constast system.

### 5.3. Results and Analyses

On the basis of ensuring the "cost" in principle, we verify the performance of the system by comparing the concurrency, scalability, response time, stability and other indicators.

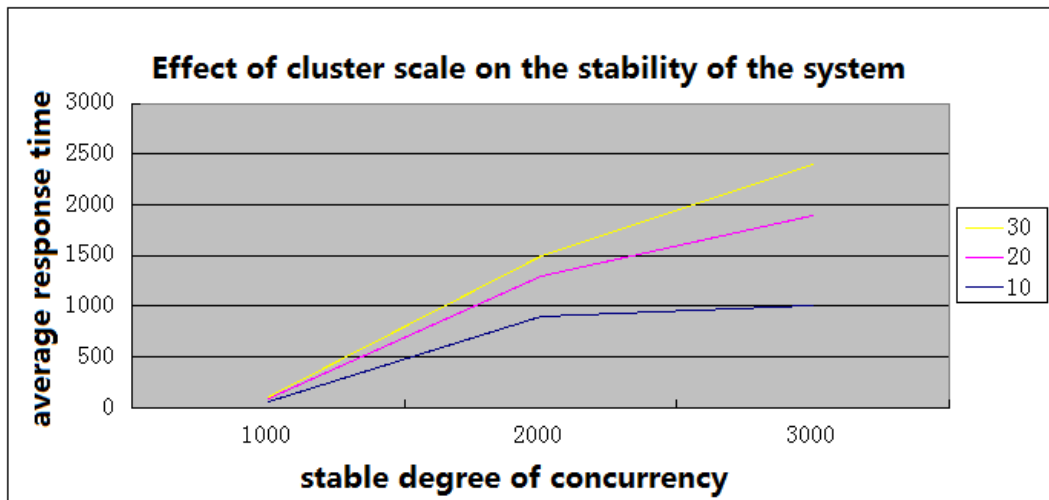
We simulate data requests to call data service of the systems, statistics obtained the following form.

**Table 1. Statistics of the Overall System Performance**

	machines	average response time	maximum degree of concurrency	Cost	comments
Contrast system	1	1.34	1200	¥ 20,000	
Instance system	10	0.56	6000	¥ 20,000	
Extended system1	20	0.47	9000	¥ 40,000	
Extend system2	30	0.38	1300	¥ 60,000	

As is shown in Table 1, the improved system is generally better than traditional system. The larger the cluster scale, the greater the performance of the system. However, with the increase of the number of machines, the performance improvement each machine brought is on the decrease. So we need to find a balance point between improvement and cost by controlling our cluster at a appropriate scale.

Further statistical of effect of cluster scale on the stability of system obtain the chart below.



**Figure 6-1. Effect of Cluster Scale on the Stability of the System**

Different size of cluster, stability is not the same. Through the comparison of different size of cluster, we found that the stability of the system enhanced with the growing of the cluster's scale. What's more, with the increase of the number of machines, the performance improvement each machine brought is on the decrease. This regular is similar matches the regular we found in last analyses.

In the test of other services, we get similar result .Our web GIS based on service pipeline system performance. Have the great advantage of the traditional web GIS. The test proved our theory to be effective.

## 6. Conclusions and Next Work

In order to ameliorate efficiency of web GIS with low costs, we propose a service pipelining model for Web GIS based on distributed environment. We build a basic service library to reuse software and unified standard of these API, design a segment to connect machine to construct the service pipeline and reuse low profile computers, and change the way to combine service. The comparison between our improved system and traditional system proved that our theory is very effective. While the comparison between our improved systems we found that cluster in appropriate scale will benefit most, which point out a way to ensure effect and control cost.

According to test data, the next step we will combine the actual situation of the number of users to give the best cluster scale advice, to save energy, get the most cost-effective.

At the same time, we will load more applications on the platform, such as southwest minority ancient books, such as GIS, traditional minority national sports GIS and Hbase data storage service, data service, ancient building panoramic data such as database applications, to extend the platform.

## Acknowledgments

I would like to extend my heartfelt thanks to a host of people, without whose assistance the accomplishment of this thesis would have been impossible. First and foremost, I am deeply indebted to my juniors, Qiao fenghao, Li chengjie, Guo peng, and Jia xiaodong for their patient guidance and constant cooperate during the deployment of this system as well as in the good advice during design of this system. I am also grateful to our test teamer, especially Yang meijia, Wang ruiqi and Zhong ling, whose hard working and valuable advice has benefited me a great deal. This paper is supported by" Construction of a comprehensive database of Development and a information service platform for

Minority Areas”, ”State Ethnic Affairs Commission-Leading talents supported projects”, ”Ministry of Education of the People's Republic of China-Supported by Program for New Century Excellent Talents in University”.

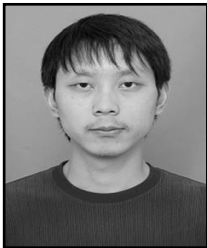
## References

- [1] O. Terzo, P. Ruiu, E. Bucci, F. Xhafa, "Data as a Service (DaaS) for Sharing and Processing of Large Data Collections in the Cloud," Complex, Intelligent, and Software Intensive Systems (CISIS), 2013 Seventh International Conference on , vol., no., pp.475- 480, 3-5 July (2013).
- [2] A. Chaudhuri, T-Y Lee, B. Zhou, C. Wang, T. Xu, H-W Shen, T. Peterka, Y-J Chiang, "Scalable computation of distributions from large scale data sets," Large Data Analysis and Visualization (LDAV), 2012 IEEE Symposium on , vol., no., pp.113-120, 14-15 Oct. (2012)
- [3] R. Canal, A. Gonzalez, J.E Smith, "Very low power pipelines using significance compression," Microarchitecture, 2000. MICRO-33. Proceedings. 33rd Annual IEEE/ACM International Symposium on , vol., no., pp.181,190,( 2000).
- [4] B. Liu, Z. Peng, "Research on the Pipeline Security Protection Technology of Pipeline Integrity Based on Gis," Knowledge Acquisition and Modeling (KAM), 2011 Fourth International Symposium on , vol., no., pp.454,457, 8-9 Oct. (2011).
- [5] P. Lei; K. Yang, "Development of Submarine Pipeline 3D GIS Platform," Image and Graphics (ICIG), 2013 Seventh International Conference on , vol., no., pp.784-788, 26-28 July (2013).
- [6] H. Hara, S. Fujita, K. Sugawara, "Reusable software components based on an agent model," Parallel and Distributed Systems: Workshops, Seventh International Conference on, 2000 , vol., no., pp.447-452, Oct (2000).
- [7] T. Maolin, T. and A. Lifeng, "A hybrid genetic algorithm for the optimal constrained web service selection problem in web service composition", Evolutionary Computation (CEC), 2010 IEEE Congress on, (2010).
- [8] S. Mokarizadeh, *et al.*, Utilizing Web Services Networks for Web Service Innovation. Web Services (ICWS), 2014 IEEE International Conference on, (2014).
- [9] W. Yong, W., *et al.* (2012). An actor-based language to unifying web service orchestration and web service choreography. Computer Science and Information Processing (CSIP), 2012 International Conference on.
- [10] L. Yan, *et al.* Folksonomy-Based In-Depth Annotation of Web Services. Service Oriented System Engineering (SOSE), 2014 IEEE 8th International Symposium on., (2014).
- [11] Z. Zibin, *et al.* "Collaborative Web Service QoS Prediction via Neighborhood Integrated Matrix Factorization." Services Computing, IEEE Transactions on, vol. 6, no. 3, pp. 289-299, (2013).
- [12] R. J. R Raj, and T. Sasipraba, "Web service selection based on QoS Constraints", Trendz in Information Sciences & Computing (TISC), (2010).
- [13] S. Mokarizadeh, *et al.*, "Utilizing Web Services Networks for Web Service Innovation. Web Services (ICWS), IEEE International Conference on, (2014).
- [14] C. Ma, Y. He, "An Approach for Visualization and Formalization of Web Service Composition," Web Information Systems and Mining, 2009. WISM 2009. International Conference on , vol., no., pp.342,346, 7-8 Nov. (2009).
- [15] Z. Yilei, *et al.* WSPred: A Time-Aware Personalized QoS Prediction Framework for Web Services. Software Reliability Engineering (ISSRE), 2011 IEEE 22nd International Symposium on, (2011).
- [16] M. Mohagheghian, R. Sindhgattay, A. Ghose, "Combining Agent Based Modeling with Distributed Constraint Optimization for Service Delivery Optimization," Enterprise Distributed Object Computing Conference Workshops and Demonstrations (EDOCW), 2014 IEEE 18th International , vol., no., pp.296,305, 1-2 Sept. (2014).
- [17] A. Oliveira, M. Urbietta, J. Araujo, A. Rodrigues, A. Moreira, S. Gordillo, R. Gustavo, "Improving the Quality of Web-GIS Modularity Using Aspects," Quality of Information and Communications Technology (QUATIC), 2010 Seventh International Conference on the , vol., no., pp.132-141, Sept. 29 2010-Oct. 2 (2010).
- [18] L. Chuanjin, "On building public service-oriented G-WEB GIS," Environmental Science and Information Application Technology (ESIAT), 2010 International Conference on , vol.2, no., pp.719,722, 17-18 July (2010).
- [19] W. Kehe, B. Xuwei, W. Xiaohui, "Power Web GIS Based on Oracle MapViewer," Internet Technology and Applications, 2010 International Conference on , vol., no., pp.1,4, 20-22 Aug. (2010).
- [20] S. Kubota, T. Sugawara, T. Kosawada, A. Abe, "Web GIS-based information portal system for road maintenance," Geoinformatics, 2011 19th International Conference on , vol., no., pp.1,5, 24-26 June (2011).

## Authors



**Weng Yu** is associate professor of College of Information Engineering, Minzu University of China. His current research interests are distributed computing and service computing.



**Cheng Wenyi** is currently working toward the Master degree in Software engineering at the School of Minzu University of China. His current research interests are cloud computing and Service composition.