Cloud computing framework for a hydro information system

Blagoj Delipetrev[1, 2]✉, Andreja Jonoski[2], Dimitri Solomatine[2]

[1]Facutly of Computer Science, University “Goce Delcev”, Stip, Republic of Macedonia.
[2]Department of Integrated Water Systems and Governance, UNESCO-IHE Institute for water education, Delft, The Netherlands.

✉Email: blagoj.delipetrev@ugd.edu.mk

Abstract

The cloud computing framework of the hydro information system is based on three concepts that are closely related: cloud, service oriented architecture and web geographic information system. The architecture of the prototype hydro information system contains three tiers. The bottom tier is a distributed relational database (PostgreSQL, PostGIS) that store geospatial and other types of data. The middle tier is GeoServer web application that manages and presents geospatial maps and data from the bottom tier. GeoServer is a platform that abstracts relational database and provide data services for the developed web services using OGC (Open Geospatial Consortium) standards. The top tier contains three developed web services: 1) web service for geospatial data management 2) web service for supporting water resources modeling and 3) web service for optimization of water resources. The web services are build using several programming languages (JavaScript, AJAX, PhP, Java), additional applications (Eclipse, ESRI ArcGIS, Adobe Dreamweaver), libraries (OpenLayers) and geospatial standards (OGC). All components and software packages of the hydro information system are open source. The cloud framework for the hydro information system is flexible for adding additional services and components. The system advantages over previous technologies are in accessibility, availability, flexibility, distribution of computational resources, scalable computation power, interoperability, internet based collaboration platform and dissemination of valuable information over the internet, between stakeholders and general public.

Keywords: *Cloud computing; hydroinformatics; water resources management (WRM), geographic information system (GIS).*

Introduction

Water resources are esencial human need and especially important for population growth, higher living standard, industrialization, power generation and agriculture. Climate change is creating additional pressure for quality water resources management. All these factors demand for development of integrated water management solutions. Integrated water resources management is often done by computer designed models. Advancement in ICT (Information Communication Technology) is making possible to develop robust and complex models. Historically usual problem in models creation was lack of data and information. Nowadays data is gathered from many sources including terrestrial measurements, sensor networks and satellites. The prices of sensors and measurement equipment due to ICT are getting lower and lower. The immense increase of data and information demands development of new powerful computer models. The computer models should be able to gather all relevant data, model complex physical, chemical, biological, social and other factors, possess vast processing power and give solution for integrated water management problems. The research presented in this article is in this direction. Starting point for the development of the integrated water management is development of a hydro information system. A hydro information system is a computer based system that gathers all water related data (precipitation, rivers, population, agriculture, etc.) and provide platform for building specialized applications and services.

A hydro information system can be build using different technologies and platforms. Often these systems include relational databases, geographic information system (GIS) applications, water modeling software, etc. Water related data usually have important geospatial component. GIS (1) provide modeling platform for almost all types of geospatial data and that’s why it is a preferred integration platform for water resources modeling. GIS can be used as foundation for a hydro information systems (2-4) based on open source or commercial software components. A hydro information system can be based on distributed and interoperable web services (5). The article review several possible software components and technologies that are suitable for cloud hydro information system.

The cloud computing framework for the hydro information system presented in this article has several advantages over previous solutions. The main objective is to be available and accessible all the time and from everywhere using only a web browser. The hydro information system framework is flexible for adding additional applications and services or upgrading the existing ones. The framework is scalable and depending of the workload additional computational resources can be added or subtracted. The hydro information system is distributed and interoperable e.g. designed to operate in distributed computer environment on several physically independent servers. All these advancements are possible because the cloud computing framework is built on three latest ICT concepts: Cloud, Service Oriented Architecture and web GIS.

The presented cloud computing framework for the hydro information system has three web services:
1) Web service for geospatial data management.
2) Web service for supporting water resources modeling (WRM).
3) Web service for water resources optimization.
The web services are composed of different software components, (GeoServer, PostgreSQL, PostGIS) programming languages, (AJAX, JavaScript, PhP, Java), libraries (OpenLayers) and geospatial standards (OGC).

The hydro information system and the web services are successfully tested on the hydro system Zletovica. The hydro system Zletovica is located in the east-north part of the Republic of Macedonia. Basic model of the hydro system Zletovica was constructed, containing the reservoir Knezevo, river network, main water users, irrigation canals, and agriculture lands usign the web service for supporting WRM. With the web service for water resources optimization, Krezevo optimal reservoir operation was calculated. The testing of the hydro information system demonstrated that cloud framework can be used by jointly by several users working in the same environment. The presented system is a great internet collaboration platform for distributed users. Distributed users can access the system over internet, jointly view, enter data, model and view results. The cloud framework for the hydro information system is the first step towards development of a truly integrated water modeling and management solutions.

Motivation

The motivation for the cloud hydro information system is in three main ideas: Cloud computing, Service Oriented Architecture (SOA) and web GIS. Common in these three ideas is that are all internet based.

Cloud computing is the latest buzz word in the ICT. The general idea is relatively old and can go back to UNIX servers and terminals, before beginning of personal computer (PC) era. In those times UNIX servers had processing, storage power and provided services for terminal that only had an input device like monitor and keyboard. Now industry is going back to a similar concept where all data and processing power will be in the Cloud while our devices (laptops, mobile phones) will access and use cloud services via internet. Analogous to this concept is the electric grid where we use power as service without understanding of system components (6). The main advantages of cloud computing compared to previous technologies are accessibility and availability. Cloud computing is available and accessible, anytime and from everywhere. Additional advantages are scalability, interoperability, data portability, applications versioning, backup, etc. Many concerns of average users like operating systems, drivers, data, etc. are solved on cloud level. The major drawback is that a system will not work without internet. The following characteristics define the cloud (7):

- Illusion of infinite computing resources.
- Internet-based data processing where devices have on-demand access to data, resources and applications.

- Platform of web based application and services accessed with web browser same as application installed on a local computer.

All major IT companies like Google, Facebook, Apple, Microsoft and Amazon are building their cloud infrastructure and services. The most popular web application like Facebook, Gmail and other are already cloud based.

The second idea is Service Oriented Architecture. SOA is a design methodology for development, integration and implementation of computer system. SOA (8) defines how to integrate various heterogeneous application and platforms into one solution. Previously applications had defined API (Application Programming Interface) to communicate with other components. SOA design defines interface and protocols for communication between different components and applications based on XML (eXtended Markup Language) messages. Communication messages are transferred between applications “on demand”. Implementing SOA provides the following advantages:

- Integration independent of programming language.

Different system components can be coded in various programming language. SOA defines interfaces and communication protocols between components and can integrate various programming language components.

- Improving existing systems.

SOA can be used to include legacy systems components and applications into newly developing solutions. Legacy components will be upgraded to create interface and communication protocol with newly developed components. Integration of legacy components is important advantage of SOA.

- Components reuse and organization agility.

Current web service technology gives opportunity to create software components and offer them as a service to other companies. SOA defines components and services as building applications blocks. Proper design of components and services offers possibility for their combinations and integrations into new software solutions.

Historically information systems evolved from monolith architecture called 1-tier that contained application interface, processing and data. This architecture was replaced by 2-tier that separated data from presentation and application tier. With the separation of presentation and application as separate tiers, 3-tier architecture was created. SOA is n-tier architecture that support composition of several complex distributed tiers. SOA can combine several data, presentations and applications tiers into integrated solution.

The third idea is web Geographic Information Systems (GIS). GIS are successfully used in modeling almost everything with geographical information. There are numerous examples of systems based on GIS that are used for WRM. GIS software have tremendous growth parallel with the ICT progress. Opening of the GPS signal to general public and embedding GPS receives into many devices (mobile phones, tablets, etc.) created new market for geospatial data and applications. Today web based geospatial mobile applications and distributed web GIS (9) are one of the most exploited technologies.

Two organizations are important for web GIS. The first is Open Geospatial Consortium (OGC) where 417 commercial, government, non-profit and research organization worldwide collaborate in development and implementation of standards for geospatial contest, data and services. The second is Open Source Geospatial Foundation (OSGeo) which mission is to support and promote open source geospatial software accessible under OSI-certified license. OSGeo promotes software packages and component for web mapping, desktop applications, geospatial libraries, etc.

These three main ideas together with water as essential hyman resource are main motivation in realization of the cloud computing framework for the hydro information system.

Software components and technologies

The cloud computing framework for the hydro information system can be build using different system components and technologies. The requirements for components and technologies are following: 1) support of cloud computing concept, 2) operate in SOA, 3) flexible for adding new or upgrading the existing components and services, 4) interoperable, 5) work in distributed environment, 6) scalable, 7) support latest OGC standards 8) and preferably open source. Most of the selected software components and technologies comply with the listed requirements. The cloud computing framework for the hydro information system has three main tiers: database tier, middle tier for managing geospatial data and front end web services tiers. Several software packages and components were reviewed and the most appropriate selected and included into tiers development.

Three relational databases were investigated for the database tier: Microsoft SQL Server, MySQL and PostgreSQL. Microsoft SQL Server is a commercial database that support almost all previously mentioned requirements and widely used in geospatial solutions. MySQL is an open source database used in many applications but not so often in geospatial data management. PostgreSQL together with PostGIS is a preferred solution for storing and managing geospatial data (10). PostgreSQL and PostGIS are open source and can work in distributed heterogeneous environments.

The middle tier for geospatial data management was researched between OSGeo open source software packages. There are two popular web applications: GeoServer and MapServer. GeoServer is build in JAVA and support OGC standards like WMS (Web Map Service), WFS (Web Feature Service), WFS-T (Web Feature Service – Transactional), WCS (Web Coverage Service) etc. GeoServer can be configured to manage geospatial data from relational databases and provide services for other applications. MapServer have similar characteristic, but final decision was to select GeoServer as a middle tier application.

The front end web services will be connected to middle tier application GeoServer and indirectly to the geospatial data from relational database. OSGeo support several open source components suitable for front end web services: Geomajas, MapFish, and OpenLayers. Geomajas is a complete web GIS application with client-server integration for presenting and editing geospatial data using OGC standards. MapFish is package for building geospatial application based on Pylon Python web framework. OpenLayers is JavaScript library for presenting and processing of geospatial data (11). Geomajas and MapFish are already mature products and their customization is relatively difficult and complex. On the other hand OpenLayers is just a library and application development will start from scratch. After longer testing and investigation of the three components OpenLayer was selected because it was much easier to build the solution starting from the beginning than customizing the source code from MapFish and Geomajas.

Except these three main software components, additional software and applications were used in development of the system. MapInfo and ESRI ArcGIS and open source software uDIG were used for creation of geospatial maps and data. The front end web services were developed in Adobe Dreamweaver using several programming language like JavaScript, AJAX and PhP. Application for optimization of water resources is build in Java using Eclipse. The main tiers are flexible for adding additional components and tiers and in the future system development it is possible to include more components.

Architecture of the hydro information system

The cloud computing framework of the hydro information system is composed of two web interfaces on which three web services are running. The architecture of the hydro information system and data communications links between services are shown on Figure 1. The three web services are:

1) Web service for geospatial data management.
2) Web service for supporting WRM.
3) Web service for optimization of water resources.

The web service for geospatial data management consist of two main components: relational database HMak created in PosgreSQL and GeoServer web application. Hmak database stores relevant geospatial and other types of data. HMak also stores data for the web service for supporting (WRM) and the web service for optimization of water resources. GeoServer is connected to HMak and provides web interface for viewing, presenting and downloading geospatial data. GeoServer provides framework for web services development using OGC standards.

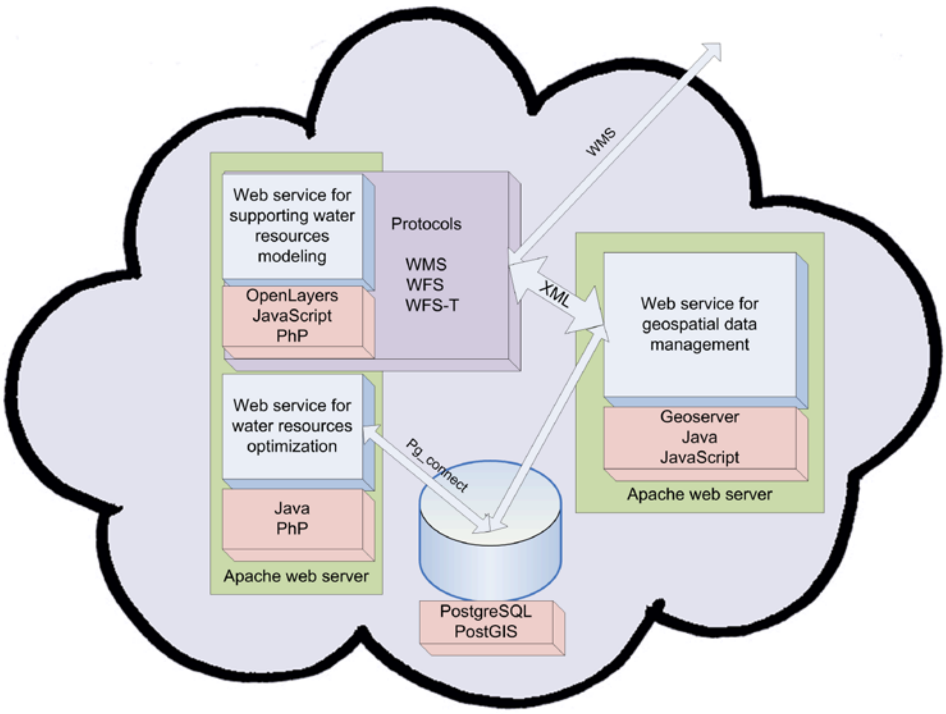

Figure 1. Architecture of the cloud hydro information system

The web service for supporting WRM is build using OpenLayer library and own developed code in JavaScript, AJAX and PhP. The web service provide user interface for WRM. The water resources are modeled using six geospatial vector layers: rivers, canals, users, reservoirs, land agriculture and inflow stored in HMak. The geospatial vector layer except geospatial objects contains additional attributes informations. The attributes informations for each vector layer are simple and it is just a demonstration of the possibility to assign additional informations for every geospatial object e.g river attributes are ID, Name, Category and Goes_In. Actually the web service for supporting WRM is a customized web GIS application. The web service for supporting WRM is connected with GeoServer using WFS-T protocol and via GeoServer to geospatial data from HMak. The communication between the web service for supporting WRM and GeoServer is asynchronous or "on demand" with XML messages. The web service has two WMS connections for the background maps from Google Earth and OpenLayers WMS. The web service offers toolbar for editing, creating, modifying and saving of geospatial object and form for entering attributes. The web service user interface is shown on figure 3 in the section of testing of the hydro information system and the web services.

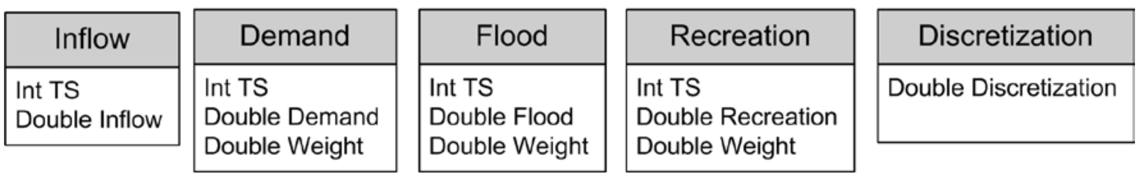

Figure 2. Input tables for the web service for optimization of water resources

The web service for optimization of water resources is composed of Java application DP for optimal reservoir operation and web interface. DP application is based on the dynamic programming algorithm for optimal reservoir operation from the book (12). DP use 5 input tables shown of figure 2. "Inflow" table contains timeseries values of water inflow into the reservoir. "Demand" table is total demand requested by users from the reservoir. The "Flood" table gives upper limit of the reservoir level while "Recreation" table is lower limit of reservoir lever. The "Weight" in the tables of "Demand", "Recreation" and "Flood" describes importance of satisfying the objective. "Discretization" table store values of reservoir storage discretization. The dynamic programming algorithm calculates optimal reservoir operation. Results of the optimization are visualized and shown on separate web page. Special web interface is developed for including DP application into the system. In the process of upload of the five input tables into HMak, specially developed PhP script verifies data. After upload of input data, DP application can be started. Results of the optimization are presented on separate web page. Screen prints of the optimization results are shown in the next section on figure 4.

Testing of the hydro information system and the web services

The testing of the hydro information system and the web services is performed on Zletovica river basin. Zletovica river basin is located in north-eastern part of the Republic of Macedonia. The modeling of the river basin Zletovica starts with the web service for supporting WRM. At the beginning, layer is selected in which geospatial objects will be drawn. Using the toolbar geospatial objects are drawn and attributes are entered. After all geospatial objects are entered, the model of Zletovica river basin is created that contains the river network, the reservoir Knezevo, the agriculture land and the irrigation canals presented on figure 3.

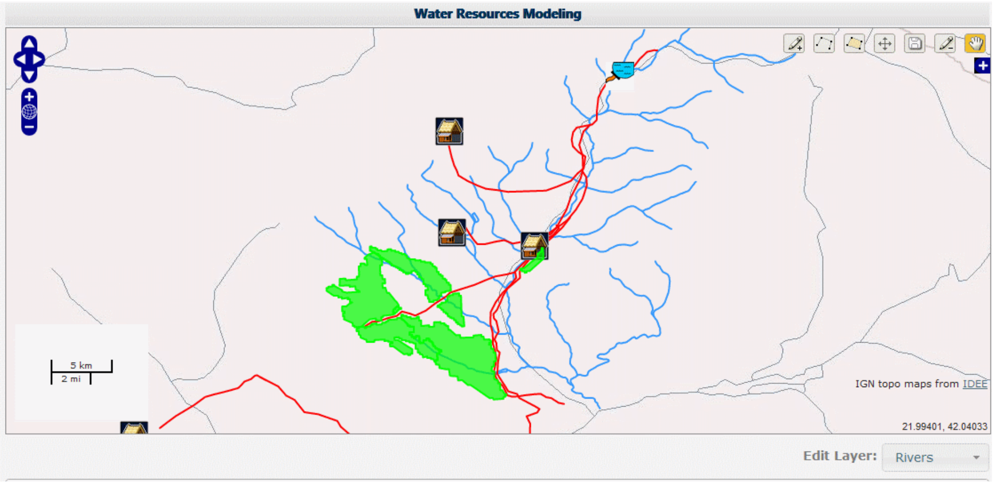

Figure 3. Model of the hydro system Zletovica

The web service for optimization of water resources is tested on the reservoir Knezevo. Timeseries data about river basin Zletovica where analyzed and sintetized in five arrays to fit developed application DP. After that reservoir Knezevo was selected and data was uploaded. The optimal reservoir operation results are presented on graphs containing reservoir inflow, total demand and optimal reservoir operation curve figure 4. Tabelar results are also available for additional processing.

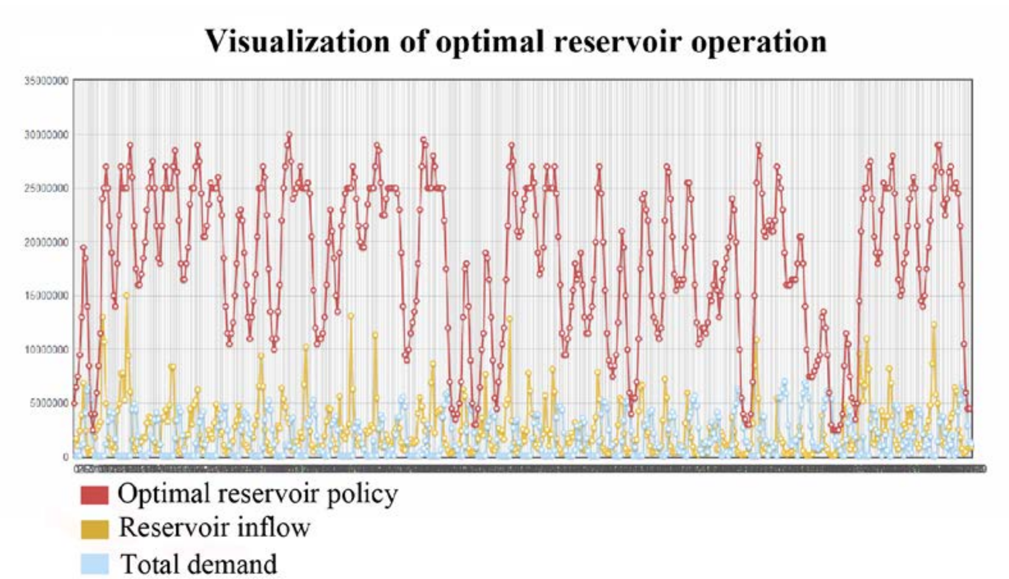

Figure 4. Optimal reservoir operation results

The web service for geospatial data management provide platform for building additional services. The web service for geospatial data management have web interface for viewing, searching and downloading geospatial data from HMak figure 5. The web service manage geospatial data stored in HMak containing Macedonian river network, lakes, reservoirs, meteorological stations, towns, ect.

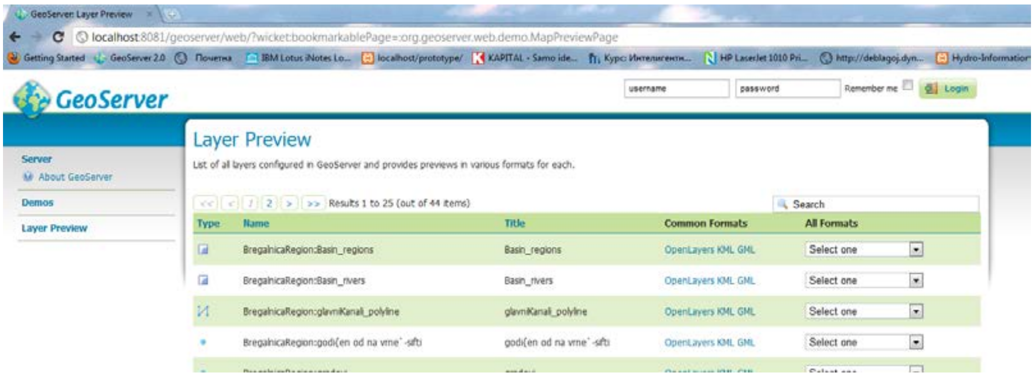

Figure 5. The web service geospatial data management user interface

Discussion

The presented cloud computing framework for the hydro information system is a new concept in building WRM applications. The system advantages over previous technological solutions are in several key points. The system is accessible and available everywhere and all the time. The demonstration shows that using only a web browser and internet connection it is possible to work with the system. All data, models, results, processing is performed in the cloud. The system framework is designed in SOA. The system components are flexible for adding additional or upgrading existing ones which is demonstrated with the including the web service for optimization of water resources. The overall system is scalable and depending on the workload additional hardware and software resources can be added or subtracted. The source code and most of the components are interoperable and can be installed on various platform, servers and operating systems. Data portability and application versioning are no longer an issue. Data is stored in the cloud in standardized formats and application is the same for all users accessing the services.

Maybe one of the most important advantages over previous solutions is the collaboration platform. The presented system enable collaboration platform where geographically dispersed users can access the same working environment. Users can jointly draw geospatial objects, enter attributes, model the river basin and run optimizations. Currently all users have the same privileges on the system e.g. all can perform the previously mentioned activities. In the future development of the system different types of users and different working environments will be created. Group of users will work in a separate working environment. Users will have different privileges for working with the web services. The next step in system development is the creation of different types of users and environments.

The web services can be upgraded and improved. The web service for supporting WRM can be enhanced with including more layers. In the current implementation 6 vector layers and types of geospatial objects are supported. Easily additional layers and objects can be included e.g. layer of electric power generators, which will improve WRM. The layers attributes in current system implementation are simple but in future this is crucial question. Every layer attributes should be carefully selected. The web service for optimization of water resources can be upgraded with improving the existing algorithm. The existing algorithm can be upgraded with including additional constrains and objectives. Another possibility is to include other algorithms for optimization of water resources based on other techniques like stochastic dynamic programming, reinforcement learning, neural networks etc. The most important part in further system development is the computational framework. The computational framework will define dependancies between objects in the system e.g. river discharge going into the reservoir will be added to reservoir level. Right now geospatial objects are not connected between each other. This will be beginning of a “real” WRM software solution.

The presented system was successfully tested on the hydro system Zletovica. The system is demonstration of the current technologies and standards in development of WRM software. The cloud computing framework for the hydro information system is not the best software for WRM or optimization of water resources but is beginning in development of cloud based WRM solutions.

Conclusion

Desktop applications and software are moving towards the cloud. The cloud computing framework for the hydro information system is research in this direction, or present demonstration of current technologies, standards and software for development of water resources modeling solutions. The presented system advantages are accessibility, availability, scalability, flexibility, interoperability that are possible due to incorporating latest ICT concepts like cloud computing, service oriented architecture and web geographic information systems. Further system development can include additional water related data, urbanization, population growth, infrastructure, etc. and creation of additional web services. The cloud computing framework is build, tested and ready for improving the existing or developing new web services.

References:

(1) D.C. McKinney and X. Cai, *Linking GIS and water resources management models: an object-oriented method.* Environmental Modelling & Software., 17(5), 413-425 (2002)

(2) M. Rao, G. Fan, J. Thomas, G. Cherian, V. Chudiwale and M.Awawdeh, *A web-based GIS Decision Support System for managing and planning USDA's Conservation Reserve Program (CRP).* Environmental Modelling & Software., 22(9), 1270-1280 (2007)

(3) J.Y. Choi, B.A. Engel, and R.L. Farnsworth, *Web-based GIS and spatial decision support system for watershed management.* Journal of Hydroinformatics., 7(3), 165-174 (2005)

(4) A. Leone, S. Shams, and D. Chen, *An object oriented and OpenGIS supported hydro information system for upper Mersey river basin management.* International Journal of River Basin Management., 4(2), 99-107 (2006)

(5) J. Horak, A. Orlik, and J. Stromsky, *Web services for distributed and interoperable hydro-information systems.* Hydrology and Earth System Sciences., 12(2), 635-644 (2008)

(6) N.G. Carr, *The big switch: Rewiring the world, from Edison to Google*, WW Norton & Company., (2008)

(7) M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R.H. Katz, A. Konwinski, G. Lee,D.A. Patterson, A.Rabkin and I. Stoica, *Above the clouds: A berkeley view of cloud computing.* EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2009-28 (2009).

(8) T. Erl, *Service-oriented architecture: concepts, technology, and design*, Prentice Hall PTR, (2005)

(9) M.G. Tait, *Implementing geoportals: applications of distributed GIS.* Computers, Environment and Urban Systems., 29(1),33-47 (2005).

(10) F. Cagnacci and F. Urbano, *Managing wildlife: a spatial information system for GPS collars data.* Environmental Modelling & Software,. 23(7), 957-959, (2008)

(11) N. Kussul, A. Shelestov, and S. Skakun, *Grid and sensor web technologies for environmental monitoring.* Earth Science Informatics, 2009., 2(1),37-51 (2009)

(12) D.P. Loucks and E. Van Beek, *Water resources systems planning and management: an introduction to methods, models and applications,.* Paris UNESCO (2005)