

Using Cloud and Python for promoting Jawi Usage

Mohd Zamri Murah, Hamdan Abdul Rahman, Khairuddin Omar

Pattern Recognition Research Group
Center for Artificial Intelligence
Universiti Kebangsaan Malaysia
zamri@ftsm.ukm.my

Abstract. Cloud computing is an innovative computing technology for building web-based applications. The fundamental concept of cloud computing is that applications and data are on a computer somewhere out there in the cloud. A cloud is a huge collection of data centers around the world. Each cloud has large processing power, redundancy, and huge storage capabilities. It is scalable, robust, maintenance free, and eases of application deployment. In the cloud, a user buys resources (software, hardware, storage, memory, processing power) as a service. If you need extra storage, you buy extra storage. There is no need to upgrade or buy new hardware. The application on the cloud can scale up to be used by one hundred users or millions of users. There is no need to change the application code for scalability. Python is an open-source scripting language that is widely used to develop prototype and web-based applications. The current version has extensive support for UNICODE, thus make it easier to code and manipulate non-ASCII documents such as Arabic, Jawi and Chinese. In this project, we have developed a cloud-based application for writing Jawi using python language. It is developed to illustrate the feasibility of using cloud-based computing for distributing web-based application and the ease of development using Python.

Keywords: cloud computing, Python, Jawi

Introduction

Cloud computing is an innovative technology for building web-based application. It involves many advances technology such as virtualization, distributed computing, utility computing and software as a service. This computing model is based on SOA (software as a service) model that offer flexibility, reduced maintenance cost and services on demand (Baun, Kunze et al. 2011).

There exist many definitions for cloud computing. Among them, cloud computing is a form of computing where a user will buy his resources for IT from IT many ser-

vice providers. These computing resources may include storage, computing power, computing platform and software. Another definition is that cloud computing is a computing model based on user requirements where a user can buy or rent computing resources based on his requirements. These computing resources can be anything.

The term ‘cloud’ in cloud computing is often a reference to a cloud-shape diagram that represents Internet or web. In its basic form, cloud computing is a computing model that is capable of offering multiple computing services using shared IT infrastructures which includes processing, storage and networking. This computing model is highly flexible and scalable. It can cater to small a scale usage or a large scale usage. A user can also purchase IT resources based on his need (Taylor 2005; Stanoesvska-Slabeva, Wozniak et al. 2010).

The term ‘cloud computing’ was first used in 1996 to describe a computing model where all the software and storage would be in the Internet. However, due to technology limitations and other factors, this computing model was not widely adopted. In 2006, the model is reintroduced when big Internet companies like *google.com*, *yahoo.com*, *amazon.com* and *ibm.com* began to build large data centers for their business activities. These companies began to offer their computing facilities to other business using many current technologies such as software virtualization and virtual storage. This computing model is called ‘cloud computing’. Since then, various type of services and software have been implemented based on cloud computing model (Ommeren, Duivestein et al. 2009; Rubio 2010).

Jawi usage in writing and printed communication has decreased considerably over the years. There are hardly any major newspapers or magazines or books in Jawi now. This has made it difficult to expose Jawi to the current generation. Although the government has introduced Jawi as a subject in schools, many from the current generation cannot use Jawi accurately. Even today, there are many Jawi spelling errors in Jawi related programs in television and Jawi newspapers or magazines.

In this paper, we discuss a few features of cloud computing and how it can be used for software development using Python. For illustration purpose, we have developed a prototype Jawi Transliteration System based on the cloud computing model.

Our main contributions from this paper are;

- Cloud computing model offer a new paradigm and a new environment for computing and system development. This model is both flexible and scalable. Cloud computing is based on user requirement. A user can buy additional resources as needed. There is no need on the user to buy additional hardware or software.
- Cloud computing offer a different framework for system development. In this framework, the user is required to follow the methodologies used by the cloud service providers. Therefore, a user is restricted in the tools and the methodologies he can use.

- To provide a service or an application to a wide audience of users, we can use cloud computing. For instance, in order to promote Jawi, we can develop Jawi application in the cloud.

Cloud computing model

In cloud computing, all computing infrastructure (computing, storage, networking) is move to the cloud or Internet. A developer will develop a system or software and put the system or software in the cloud. A user will access the system from the cloud as a service. All the computing, storage and networking are manage by the cloud computing provider. The developer doesn't need to worry about regular maintenances of the infrastructure. If the system is successful and attracts a large number of web users, the developer can buy additional resources to cater for the increase usage. If the usage of the system decreases, the developer can reduce his resources requirement. In this way, cloud computing is both flexible and scalable.

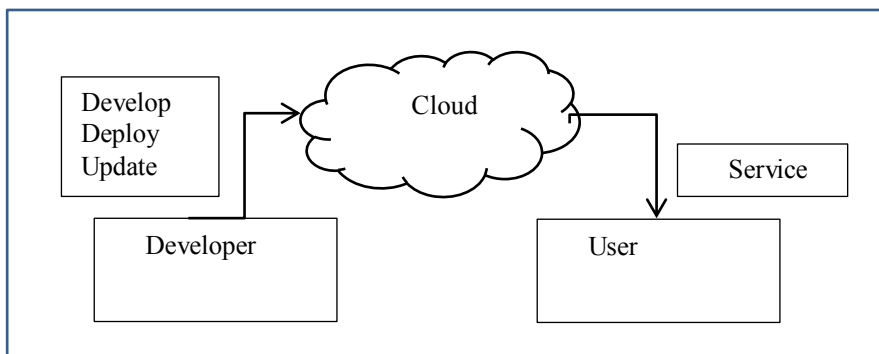


Fig. 1. A basic model for cloud computing. A developer will develop, deploy and update his application on the web. The user will access the application as a service from the web.

Some of basic features of cloud computing are;

- A user will access the application on the cloud as a service
- The service that can offered by the cloud include data processing, storage, software, application, virtualization and networking
- A developer who uses the cloud to deploy application will be charged based on his actual usage of the computing resources. This is similar in concept to electricity.
- A developer can mash up his application with other applications from the web. This mash up will create another application. The developer can create a complex application based on this idea.

MyGOSSCON 2011

- In the cloud, program and data are on a neither computer somewhere out there and neither you know nor care where that computer is.

Cloud offers a new way of looking about applications development. A developer can write almost any application in the cloud. However, there are three types of applications that are suitable for cloud;

- (a) Collaborative applications that allow people to work together, share data, communicate or collaborate. An example is online meeting or online storage.
- (b) Service applications that allow the users to do activities on the web. For instance, a subscription-based online music store. This application allows the user to browse a large collection of music, to search for a particular music, to read reviews from other users, and finally to buy a selected music.
- (c) Application that use large computations for a specific purpose such as genetic research or disease modeling. The cloud model allows you to purchase the required computing power for your applications.

Some of the example of cloud computing are;

- (a) Amazon EC2 (Elastic Computing Cloud). EC2 provides many different platforms for development such as Linux, Solaris or Windows Servers. The data can be stored in DB2, Informix, MySQL or Oracle. The applications can be implemented in Perl, Python, Ruby, Java, and C++ or C#. The applications can be deployed using Web sphere, Web logic or IIS.
- (b) Amazon S3 (Simple Storage Service). S3 is a storage system. A user can uses S3 to store database system, backup or other items.
- (c) IBM Computing on Demand. IBM provides a cloud service based on IBM's suite of web tools that uses Web sphere and DB2.
- (d) Microsoft Azure. Azure is a Windows based cloud platform. It uses a combination of web services technologies such as SOAP, REST, Servlets and Silverlight that allow developers to create applications that are similar to desktop applications. However, it is closely tied to the Windows platform.
- (e) Google apps engine. A cloud service offered by `google.com`.

Python

Python is a general-purpose, open source computer programming language. It is optimized for software quality, program portability, and component integration. Python is used in areas such as Internet scripting, systems programming, user interfaces, product customization, numeric programming, and more. Python is deployed in a wide variety of products and roles. Its current user base are Google, YouTube, Industrial Light & Magic, ESRI, the BitTorrent file sharing system, NASA's Jet Propulsion Lab, the game Eve Online, and the National Weather Service.

Python's application domains range from system administration, website development (django), smartphone scripting (android), numerical analysis (Numpy), and computer games (Alice), and spacecraft control.

Python has many good features such as a readable and maintainable syntax; easy integration with external components in other languages; a multi-paradigm design (OOP, procedural, functional), modular structures; and a vast collection of libraries. Its features make it a flexible and agile language, ideal for both simple tasks and large application development.

Although it is a general-purpose language, Python is often called a *scripting language* because it makes it easy to utilize and direct other software components. Perhaps Python's best asset, though, is simply that it makes software development more rapid. In many cases, Python serves as *enabling technology*.

Cloud computing using google apps engine

Google offers a cloud computing service called google apps engine (GAE). This service offers a free account for any user. With the free account, a user is given the ability to run up to 10 applications along with these features;

- 6.5 hours of CPU time
- 10GB/day for outgoing/incoming bandwidth
- 1GB of storage
- 2,000 e-mails/day

If a user needs more, he can buy additional resources for each category. With this free account, a developer can experiment and build new applications for the cloud.

In order to develop for the cloud, you need to have the proper SDK (software development kits) either for Python or Java. Currently, the SDK supports Python 2.7, Java and Go. Support for Python and Java is in a stable state, while support for Go is still in experimental stage. Go is a new language specifically developed for the web development by Google.

The process of developing the application can be divided into three stages;

- (a) A developer will develop the application using the proper SDK on his own computer. He can test the application locally using real data or real situation. During this stage, the application is run in a simulated environment.
- (b) When the application has passed all the necessary tests, the developer can deploy the application on the cloud at web site appspot.com. The application will be given an id by the developer, and users can access the application using

MyGOSSCON 2011

this id. For instance, if an application is given an id called `rumi-to-jawi`, the users can access this application at `rumi-to-jawi.appspot.com`.

- (c) A developer will continue to update his application by adding new features or fixing bug. When the developer is satisfied with the changes, he can deploy the new version in the cloud and the cloud will automatically serve the new version to the users. This process of adding new features, fixing bugs, and deploying will continue repeatedly during the lifetime of the application.

Jawi transliteration

Jawi transliteration is the process of converting Malay word in Rumi script into Malay word in Jawi script. This process can be done using two different methods. In the first method, the process is based on rules where the rules are derived from analyzing the syntaxes and the rules of Rumi spelling and Jawi Spelling. Based on our research, this method has an accuracy of about 70% because it is very difficult to accurately extract the rules for transliterating Rumi to Jawi. However the advantage of using this method is that a user can transliterate any words, including what is called out-of-the-bag words.

The second method is based on corpus. A corpus is prepared based on established sources such as dictionaries and word lists. Using this method, the accuracy is 100% if the Rumi word is in the corpus. If the Rumi word is not in the corpus, the accuracy is 0%. This method cannot transliterate out-of-the-bag words. The main problem using the corpus method is the corpus itself. The process of preparing and checking a corpus is both time consuming and tedious.

A prototype of cloud computing application for transliterating Rumi to Jawi using Python is deployed at web site `rumi-to-jawi.appspot.com`. This prototype is based on the corpus method, and we hope to enhance to further enhance the prototype using rules.

The first prototype is written using C# and .NET platform as a standalone application for Windows platform. This prototype is based on rules. The accuracy of the application is about 70%. The application is migrated into a web-based application and deployed at `www.jawi.ukm.my`. This prototype is deployed using the traditional client-server model. The latest prototype is rewritten in Python and deployed at `rumi-to-jawi.appspot.com` using cloud computing model.

The benefits of deploying this prototype on the cloud are;

- We do not need to worry about infrastructure maintenance or upgrading. We can focus on coding, adding new features and improving our application.

- The application is scalable. If, by a chance, the application become popular and is access by 1 million users at one time, the cloud infrastructure can handle the requests and the sudden surge of interest.
- The application is on the web. Thus, it is accessible by anyone using whatever tools they use such as browser, smartphone or tablets. It is available anytime from anywhere.

Result and Discussion

In this paper, we has discussed about cloud computing and its features. Cloud computing model offer a new approach in applications development. It is well suited for web-based applications that involve collaborations and services. We have shown the usage of this model by developing and deploying a Jawi transliteration system. Using cloud, a developer can develop and deploy all types of innovative applications on the web.

References

1. Baun, C., M. Kunze, et al. 2011. *Cloud Computing – Web-Based dynamic IT services*. Springer.
2. Lutz, Mark. 2010. *Programming Python*. O'Reilly.
3. Lutz, Mark. 2008. *Learning Python*. O'Reilly.
4. Ommeren, E., S. Duivestein, et al. 2009. *Collaboration in the Cloud*. Microsoft.
5. Sanderson, Dan. 2010. *Programming Google App Engine*. O'Reilly.
6. Stanoesvska-Slabeva, K., T. Wozniak, et al. 2010. *Grid and Cloud Computing*. Springer.
7. Taylor, I. J. 2005. *From P2P to Web Services and Grids*. Springer.
8. Vouk, M. A. 2008. *Cloud Computing – Issues, Research and Implementations*. Journal of Computing and Information Technology.

Pusat Pengajian Sains Komputer
Fakulti Teknologi dan Sains Maklumat
Universiti Kebangsaan Malaysia

Email: zamri@ftsm.ukm.my