A Comparative Study of Geographical Information Services in Public and Private Clouds

Peichao Gao¹, Zhao Liu¹, Kun Tian¹ and Meihui Xie¹

¹Institute of Geomatics, Department of Civil Engineering, Tsinghua University, Beijing 100084, China

Abstract

In recent years, three trends have been developed in the field of Geographical Information System (GIS): 3D GIS, big data, and especially cloud GIS. With the rapid progress of cloud computing in the IT industry, more and more cloud GISs emerge and are now changing the traditional geographical information service mode step by step. Compared to large-scale software must be installed before use in the past, cloud GIS enables anyone to utilize geographical information services (GIServices) anywhere and anytime. According to different types of deployments, cloud GIServices can be divided into public cloud GIServices and private cloud GIServices. GIServices are generally deployed on public clouds when local infrastructures are inadequate or applications are relatively simple, but with the increase in the requirement of data security and the number of customized applications, GIServices deployed on private clouds will become the norm. This paper strives to compare and contrast public cloud GIServices and private cloud GIServices and give insights into the essential characteristics of both.

Key words: Geographical information service, Public cloud, Private cloud, Cloud GIS.

1. Introduction

Information is wealth nowadays, but with the advancement of technology, information tends to be complex and diverse. As a significant part of the information stream, spatial information is playing an increasingly important role. To manage and make effective use of those spatial information, geographical information systems are often utilized.

Geographic information system (GIS), also known as geospatial information system or geo-information system, was created to store, manage and analyze data relating to space and geography (Li, 1994). Doctor Roger Tomlinson originated and developed the Canada Geographic Information System (CGIS), the first GIS in the world, in the 1960s (Azzam and Robinson, 2013; Macdonald and Crian, 1985). GIS gradually came to light from then on, and was applied to a wide variety of fields.

The 2006 launch of Amazon Web Services (Orna et al., 2013), the first public cloud computing service in the world, contributed to the emergence of cloud GIS. Providing geographic information services (GIServices) by deploying GIS on cloud computing has gradually become a trend. This paper studies GIServices both in public cloud computing and private cloud computing, and analyzes the differences between them from several aspects.

2. GIServices on the Cloud

In the last 20 years, GIS has gone through different stages of development, such as standalone GIS software, web GIS and
distributed GIS. A typical representative of standalone GIS software is ArcGIS Desktop, which is released by the Environmental Systems Research Institute (ESRI) (Orna et al., 2013) and needs to be installed on a high-performance computer. Web GIS is a combination of GIS with the Internet (Karnatak et al., 2007). People can directly access the functions of GIS through the Internet, and utilize GIS without any software installation on their computers. Distributed GIS is GIS in a distributed environment (Tait, 2005), and can be employed to deal with compute-intensive tasks.

Cloud GIS is the cloud-based GIS software and services (Bhat et al., 2011). Users can easily access cloud GIS through the network, and employ cloud-powered GIS to analyze and process their own spatial data. The emergence of cloud GIS changed the traditional mode of GIServices. GIServices in the cloud environment can be divided into four categories: GIServices infrastructure as a service, GIServices data as a service, GIServices platform as a service and GIServices application as a service (Peng et al., 2013; Wu and Wu, 2011). GIServices infrastructure as a service means the infrastructure, which includes the resources of computation, storage and network as well as GIS software, will be provided to customers. In the mode of GIServices data as a service, location information, remotely sensed imagery from satellites, and other frequently-used data will be categorized, stored, and provided to users. GIServices platform as a service is like the platform as a service, which is represented by Google App Engine. In the mode of GIServices platform as a service, a developing environment which is useful to GIS developers will be deployed on the cloud. GIServices application as a service will supply clients with basic GIS-based applications, as well as applications for specific use (Ji et al., 2012), through the Internet.

Cloud computing is mainly divided into public cloud computing and private cloud computing (Bo and Wang, 2011). Public cloud computing is a kind of cloud computing which is provided by third-party vendors, while private cloud computing is usually on-premises and built by users themselves. Cloud GIServices, based on cloud computing, can also be divided into two types: public cloud GIServices and private cloud GIServices. Public cloud GIServices are GIServices deployed on the data centers of public cloud computing vendors, whereas private cloud GIServices are GIServices based on cloud computing data centers established by individuals or organizations.

3. Comparing Public Cloud GIServices and Private Cloud GIServices

3.1 Main Features

Public cloud GIServices are usually established on remote public cloud infrastructures, which means that all the layers of GIServices, including infrastructure layer, data layer and application layer, are realized in the remote public cloud. Only through the Internet can the local users access public cloud GIServices. However, private cloud GIServices can rely only on a local area network. Compared to public cloud GIServices, the transmission rate of private cloud GIServices is less affected by the wide area network. Since private cloud GIServices rely on the local area network, which is usually much faster than the wide area network, organization-oriented GIServices are usually deployed on a private cloud.

The computing resources rented by public cloud GIServices are only a part of the public cloud computing resource pool. Therefore, the creator of public cloud GIServices only possess the ownership of the data layer and application layer, while the ownership of the infrastructure still belongs to the vendor. This mode has both advantages and disadvantages. On the one hand, the creator of public cloud GIServices can directly utilize the infrastructure layer provided by public cloud vendors without maintenance. This will help the creator concentrate on GIServices instead of cloud computing. On the other hand, the security of public cloud GIServices depends to a large extent on public cloud vendors instead of the creators themselves.

Since the overall architecture is realized and possessed by individuals or organizations themselves, the data center is usually located in their personal residence or the place where the organization is located. The benefit of an on-premises data center is that administrators can optimize the architecture. In addition, users will achieve faster access speeds, higher quality of service and better security when using private cloud GIServices.

Public cloud GIServices and private cloud GIServices have different service objectives. Public cloud GIServices are intended to provide the general public with geo-information services or location-based services for general use. For example, ArcGIS Online (Kouyoumjian, 2010) based on the Amazon Web Services, provides Internet users with methods to display, inquire and analyze maps. The service objects of private cloud GIServices are generally agency internal users with specific purposes. For example, Hochschule Furtwangen University in Germany created a private cloud named CloudIA (Doelitzscher et al., 2011) for the purpose of teaching, and researchers (Fan et al., 2012) from the Chinese Academy of Science established private cloud GIServices for scientific research.

In terms of scalability, public cloud GIServices have the ability to rapidly extend. Since the resources rented by public cloud GIServices are only a small partition of the public cloud resource pool, the performance of public cloud GIServices can be rapidly improved by renting more resources when needed. However, to improve the performance of private cloud GIServices entails buying more physical hardware. Obviously, for private cloud GIServices, more time will be consumed.

The main features of public cloud GIServices and private cloud GIServices are shown in Table 1.
3.2 Applicable Scenarios

GIServices deployed on public clouds will be an ideal choice when local resources of infrastructure are insufficient or initial capital is limited. Public cloud vendors provide renters with an elastic bottom layer, which can replace the local infrastructure layer to a large extent. Another applicable scenario is that the existing local infrastructure cannot be effectively virtualized, and there is not enough money to add new infrastructures at the same time. For small-scale applications or large-scale applications in the preliminary stage, choosing the public cloud as a deployment platform can reduce the cost.

Though all the GIServices can be deployed on public clouds, private clouds are still worthwhile because of security and independence. GIS is an industry that requires a high level of security for data, especially high-precision data. Leaking such high-precision data, whether intentionally or unintentionally, may even violate state law. Deploying high-precision GIS data on private clouds is relatively reliable. In addition, unlike public cloud GIServices, GIServices based on private clouds can work independent of the external network. Private cloud GIServices rely on the internal local area network instead of a wide area network, which means the GIServices will not stop running even if the wide area network is cut off.

One of the aims of cloud computing is to provide low-cost but high-performance computing ability. Large companies or universities often have a lot of basic computing facilities, such as Server clusters and PC clusters. If those basic computing facilities can be utilized to build a cloud, a lot of costs will be cut, which is more in line with the original purpose of cloud computing. The fact that public cloud GIServices unable to make full use of local facilities has become a driving factor of private cloud GIServices.

In contrast, the private cloud is more customized to meet the needs of GIServices. Although there are many public cloud computing services, they are not designed for GIServices. Therefore, they cannot fully meet the needs of GIServices in data structures, algorithms, etc. Private clouds designed for GIServices can be customized to meet the user's needs.

3.3 Building Strategies

Cloud computing has three main service modes, namely Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS) (Luo et al., 2011). In IaaS, the resources of computation, storage and network are provided in the form of virtual machines (Doelitzscher et al., 2011) to users. In PaaS, what users get is a development platform with runtime and toolkits. In SaaS, consumers can utilize software provided by cloud computing vendors through the Internet. GIServices, whether based on a public cloud or a private cloud, can be integrated with all the service modes of IaaS, PaaS and SaaS.

The general process of establishing public cloud GIServices: (1) select a cloud computing vendor; (2) pay for a resource pool; (3) deploy software, data and service; (4) achieve IaaS, PaaS or SaaS objectives. On the other hand, the general process of establishing private cloud GIServices: (1) use existing hardware or purchase new hardware; (2) virtualize the hardware; (3) create a resource pool; (4) deploy software, data and service; (5) achieve IaaS, PaaS or SaaS objectives. Figure 1 shows the contrast between the two establishing processes.

Obviously, building strategies of public cloud GIServices and private cloud GIServices are different. For public cloud GIServices, emphasis should be placed on performance and scalability of the cloud in the initial stage of building, while for private cloud GIServices, emphasis should be placed on functions of the cloud in the beginning. After obtaining a basic cloud computing platform, creators of public cloud GIServices should customize the public cloud to meet the needs of GIServices, while private cloud GIServices creators should customize the public cloud to meet the needs of GIServices.

Table 1. Main features of public cloud GIServices and private cloud GIServices.

<table>
<thead>
<tr>
<th>Maintenance for bottom layers</th>
<th>Public cloud GIServices</th>
<th>Private cloud GIServices</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Security degree</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Expansibility</td>
<td>Excellent</td>
<td>Good</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data centers</th>
<th>Off-premises</th>
<th>On-premises</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service objects</td>
<td>A wide range of users</td>
<td>A narrow range of users</td>
</tr>
<tr>
<td>Usage</td>
<td>Ordinary</td>
<td>Specified</td>
</tr>
<tr>
<td>Network</td>
<td>Wide area network</td>
<td>Local area network</td>
</tr>
<tr>
<td>Maintenance for bottom layers</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Main features of public cloud GIServices and private cloud GIServices.
3.4 Technical Difficulties

For public cloud GIServices, the main technical difficulty is integrating GIServices and public cloud computing. Public clouds should be customized or even improved according to the characteristics and needs of GIServices. Not all of the infrastructures and platforms provided by public cloud computing vendors can meet the requirements of GIServices to the system environment. For instance, the data query language of Google App Engine is Google query language (GQL), which is similar to but somewhat less functional than the structured query language (SQL) (Karimi et al., 2011). So researchers from the University of Pittsburgh (Karimi et al., 2011) specifically improved their methods before deploying services on Google App Engine for storing and operating triangulated irregular networks (TIN). In addition, Google App Engine only supports Java/Python languages and non-relational databases (Karimi et al., 2011), both of which are not convenient in terms of GIServices staff. Researchers from the Chinese Academy of Science (Karimi et al., 2011) found that Google App Engine cannot meet their requirements for storing, retrieving and operating data, so they redesigned the Q-Tree and R-Tree for GIServices on Google App Engine.

Since public cloud GIServices are accessed through the wide area network, the way to ensure a high speed of data transmission has become another key point for public cloud GIServices. Academics from Wuhan University (Yue et al., 2013) tested the transmission time of GIServices both on Google App Engine and Microsoft Azure. They found the transmission time is much different with Google App Engine and Microsoft Azure, even with the different data centers of Microsoft Azure. They also found the transmission time increases with the amount of data growing. Data transmission in a big data era will be a very difficult issue (Huang et al., 2004). In the era of big data, for public cloud GIServices, methods should be developed to automatically select the nearest data center from the customer and to transfer large amounts of data efficiently.

For private cloud GIServices, hardware virtualization is one of the technical difficulties. Virtualization is the most important technical foundation for cloud computing. The virtualization layer, the logical abstraction of physical resources in cloud computing (Foster et al., 2008), is the basis for a highly centralized and automated IT service management. After virtualization, resource utilization can be improved by dynamic resource allocation based on the performance of each application. According to different objects to be virtualized, virtualization can be divided into Server virtualization, application virtualization, desktop virtualization, network virtualization, storage virtualization, etc. (Lunsford, 2009; Ruest and Ruest, 2009) In the process of establishing private cloud GIServices, server virtualization is utilized most frequently. Server virtualization can "split" a single high-performance server into multiple independent virtual servers. There are two main architectures of server virtualization (Bugnion et al., 2012; Chang et al., 2013), namely hosted architecture and bare metal architecture. Hosted architecture, as shown in Figure 2, means the hypervisor is installed, like an ordinary application, on the operating system of the host machine. Bare metal architecture, as shown in Figure 3, means the hypervisor is directly installed on the host machine that without any operating system. Creators of private cloud GIServices can choose to

![Figure 1. Building strategies of public cloud GIServices and private cloud GIServices](image1.jpg)

![Figure 2. The hosted architecture](image2.jpg)
utilize a single one or a mixture of the two architectures based on the specific situation of the local infrastructure.

A distributed storage system for data should be taken in account both public cloud GIServices and private cloud GIServices. Since a single data storage node will be ineffective when dealing with data that exceeds a certain amount, multiple storage nodes are recommended. The format, projection and time of data should be unified before distributing to different storage nodes.

### 3.5 Security Arrangements

A survey in 2009 from Gartner, a market research firm in the US, showed that the major reason that more than half of the companies are reluctant to cloud computing in the short term is safety (Lin et al., 2011). It is important to keep data safe both in public cloud GIServices and private cloud GIServices. The safety of data refers to not only the confidentiality of geographical information, but also the integrity of geographic information and the availability of GIServices (Liu et al., 2007).

Since established at a local place, private cloud GIServices have greater autonomy in security protection. There are many available methods (Liu et al., 2007) such as set a password, access control, install a firewall, install anti-virus software, etc. However, for public cloud GIServices, data security relies heavily on the maintenance of public cloud vendors and the proper use of public cloud security products by administrators.

Amazon Web Services (AWS) launched the AWS CloudHSM service to protect users' sensitive data. AWS CloudHSM is a professional hardware security module that can help users generate, store and manage data encryption keys, which can only be obtained by users themselves. Ali cloud, a leading cloud computing vendor in China, launched a product named "cloud shield" to ensure data safety.

For private cloud GIServices, since the virtualization of physical facilities is completed by creators, virtualization security cannot be ignored. But in fact, virtualization security in private cloud GIServices is often overlooked. This is because people tend to manage the virtualized environment as a real physical environment, and use the same security arrangement as in virtualized environments. However, the traditional security arrangement is ineffective in virtualized environments (Jing et al., 2012). There are three main approaches (Van Cleeff et al., 2009) to ensure the safety of the virtualization layer: (1) virtualization technology design; (2) application design; (3) virtualization deployment and management. If the virtualization is based on commercial software, such as VMware, Citrix, Microsoft, Red Hat, software packages provided by virtualization vendors can be used at the same time to ensure the security of virtualization. For example, VMware vShield can be used to ensure the security of virtualization in a VMware-based virtualized environment (Basak et al., 2010).

Apart from the level of cloud computing, both the safety of public cloud GIServices and private cloud GIServices can be strengthened on the level of GIServices. As shown in Figure 4, the security of GIServices is divided into three parts by Wang Wei et al. (Wang, 2013): infrastructure security, system security and management security. The security of infrastructure relies mainly on firewall software and anti-virus software. The system security refers to security logs, authentication, digital signature, cryptography system, etc. The management security is mainly the establishment of a management system.

In summary, if dividing the security of cloud GIServices into six parts: physical security, network security, host security, system security, application security and data security, creators of public cloud GIServices should place emphasis on the latter three, whereas creators of private cloud GIServices need to pay attention to all the six parts of security. Table 3 showed the comparison of key points of safety for public cloud GIServices and private cloud GIServices.

### 3.6 Expenditures

Expenditure of private cloud GIServices consists of hardware costs and software costs. Hardware refers to minimal machines, blade servers, ordinary rack servers, high-end...
A Comparative Study of Geographical Information Services in Public and Private Clouds

Table 3. Security arrangements of public cloud GIServices and private cloud GIServices.

<table>
<thead>
<tr>
<th>Private Cloud GIServices</th>
<th>Public Cloud GIServices</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Physical security</td>
<td>(1) System security</td>
</tr>
<tr>
<td>(2) Network security</td>
<td>(2) Application security</td>
</tr>
<tr>
<td>(3) Host security</td>
<td>(3) Data security</td>
</tr>
<tr>
<td>(4) System security</td>
<td></td>
</tr>
<tr>
<td>(5) Application security</td>
<td></td>
</tr>
<tr>
<td>(6) Data security</td>
<td></td>
</tr>
</tbody>
</table>

rack servers, switches, storage, cabinets, etc. Software may include virtualization software, databases, GIS professional software and operating systems. For public cloud GIServices, there is hardly any need to purchase hardware, except to rent a resource pool from cloud computing vendors. In addition, there are no extra costs related to virtualization software and even operating systems as well as databases.

The comparison of hardware/software costs between public cloud GIServices and private cloud GIServices is shown in Table 4. In addition, for private cloud GIServices, there are extra costs for operating and maintenance for data centers.

4. Concluding Remarks

Along with the advancement of science and technology, cloud computing is becoming an essential infrastructure in people’s lives like the resources of water and electricity. GIServices based on cloud computing are changing the traditional mode of GIServices. According to different types of deployments, cloud GIServices can be divided into public cloud GIServices and private cloud GIServices. This paper compared the two kinds of cloud GIServices in terms of their main features, applicable scenarios, building strategies, technical difficulties, security arrangements and expenditures.

According to the differences between public cloud GIServices and private cloud GIServices, it is not difficult to decide which one to use. In terms of their main features, public cloud GIServices rely on a wide area network and off-premises data centers provided by cloud computing vendors, whereas private cloud GIServices depend on a local area network and on-premises data centers which are usually built by users. The maintenance cost of on-premises data centers are relatively high, but their safety performance remains comparatively good. Public cloud GIServices are suitable for a wide range of people. The main applicable scenarios for establishing public cloud GIServices are when local infrastructures are insufficient or initial capital is limited. If there are surplus infrastructures at local place or sensitive data involved, then private cloud GIServices will be a better choice. In terms of their building strategies, public cloud GIServices building is built from scratch, whereas private cloud GIServices building is a process from what users already have. As the overall framework of private cloud GIServices needs to be implemented by the users themselves, there are more technical difficulties when establishing private cloud GIServices than public cloud GIServices. However, the private cloud GIServices are relatively flexible and provide a high level of security. In addition, the investment of private cloud GIServices at the early stage is usually higher than public cloud GIServices.

We hold the idea that cloud GIServices will become the major mode of spatial information service in the future. Since public cloud GIServices and private cloud GIServices have different characteristics and are suitable for different situations, we believe both of them have tremendous

Table 4. Main expenditures of public cloud GIServices and private cloud GIServices.

<table>
<thead>
<tr>
<th></th>
<th>Public cloud GIServices</th>
<th>Private cloud GIServices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hardware</td>
<td>Pay for a resource pool</td>
<td>(1) Minimal machines</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2) Blade servers</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3) Ordinary rack servers</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4) High-end rack servers</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(5) Switches</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(6) Storage</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(7) Cabinets</td>
</tr>
<tr>
<td>Software</td>
<td>GIS professional software</td>
<td>(1) Virtualization software</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2) Databases</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3) GIS professional software</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4) Operating systems</td>
</tr>
</tbody>
</table>

22
development potential. Besides, we believe hybrid cloud GIServices will draw wide attention. It is foreseeable that these different forms of cloud GIServices will play more significant roles in the future.

Acknowledgments

The authors thank the anonymous reviewers for their valuable comments and suggestions.

References


