A COMPARISON OF PLATFORM AS A SERVICE (PAAS) CLOUDS WITH A DETAILED REFERENCE TO SECURITY AND GEOPROCESSING SERVICES

Byron Ludwig* and Serena Coetzee

Department of Computer Science, University or Pretoria, Pretoria, 0002, South Africa - byronludwig@gmail.com, scoetzee@cs.up.ac.za

Commission IV, WG IV/5

KEY WORDS: platform as a service, PaaS, geoprocessing, cloud computing, distributed computing, service level agreement, SLA, security

ABSTRACT:

Cloud computing is an emerging computing paradigm aimed at running services over the internet to provide scalability and flexibility. The advantages in using the cloud for start-up and small businesses that lack infrastructure have been shown to far outweigh the disadvantages. Cloud platform services, also known as Platform as a Service (PaaS), provide a computing platform or solution stack on which software can be developed for later deployment in a cloud. However, there are a number of security challenges because users of the cloud have to rely on third party companies to provide confidentiality, integrity and availability. Geoprocessing is the manipulation of geographic information, ranging from simple feature overlays and geocoding to raster processing and advanced climate modelling. The Open Geospatial Consortium's (OGC) Web Processing Service (WPS) defines a standardized interface that facilitates the publishing of geospatial processes. Parallelization and distribution of geoprocessing services have received much attention lately, including running them in a cloud. However, work on the security aspects of geoprocessing in a cloud is limited. In this paper, we compare three PaaS cloud computing solutions, namely Microsoft Azure, Google App Engine and GroundOS, with a detailed reference to cloud security concerns. An analysis of the security mechanisms and Service Level Agreements (SLA) provided by these PaaS clouds is presented. We then look at the implications of these security issues for geoprocessing services and the OGC's WPS specifically, investigating potential security pitfalls when developing a WPS in a PaaS cloud. Finally, recommendations for future work are presented.

* Corresponding author.

1. INTRODUCTION

Cloud computing is largely a combination of existing technologies that have been around since the early 1990's. These technologies include: grid computing; utility computing; and most recently virtualisation. Each of these technologies forms a layer in the cloud computing stack. Cloud computing allows a user to pay only for the resources used instead of paying a fixed cost; this is the concept of utility computing. One of the main drivers that launched cloud computing is virtualisation technology, which allows resources to be dynamically scaled on demand. This has a close relationship to the utility computing model where each additional virtual instance created will have associated with it additional costs because of the more resources provided by that instance. Cloud computing can be seen as another form of distributed high performance computing, which has similarities to cluster, parallel and grid computing. Based on our understanding and a study of various cloud definitions, we have come up with the following definition for a cloud:

> *A cloud is a utility based computing model that provides a service, and allows virtualised resources to be easily and efficiently scaled on demand.*

Geoprocessing is the manipulation of geographic information, ranging from simple feature overlays and geocoding to raster processing and advanced climate modelling. The Open Geospatial Consortium's (OGC) Web Processing Service (WPS) defines a standardized interface that facilitates the publishing of geospatial processes (OGC 2007). Parallelization and distribution of geoprocessing services on grids and clouds have received much attention lately. For example, the research agenda for geoprocessing services proposed by Brauner et al. (2009) recommends further research into the use of cloud and grid computing to overcome the performance obstacle in geoprocessing services that are used in SDIs. In earlier work we analysed the technology choices for data grids in a spatial data infrastructure (SDI), including the use of a number of OGC web services (Coetzee and Bishop 2009). Here we limit our focus to the OGC WPS in relation to security in PaaS clouds. Work on the security aspects of geoprocessing in a cloud, which are investigated in this paper, is limited.

Cloud computing provides three service models that provide different levels of control and security. These levels are, in decreasing order of control and increasing order of security:

1. Infrastructure as a Service (IaaS);
2. Platform as a Service (PaaS); and
3. Software as a Service (SaaS)

Each service model can be seen as a layer with IaaS at the base allowing full control of resources and storage, PaaS in the middle allowing development on an existing platform and finally, SaaS providing limited development opportunities but having appeal to end-users. Each layer provides different development and/or deployment opportunities that can be matched to the resource requirements of individuals and businesses.

Security is one of the greatest concerns currently preventing large-scale adoption of the cloud. This issue is emphasised in numerous recent literature articles, either stating that cloud computing security is still immature or just unreliable. Examples can be found in Everett (2009), Grossman (2009), Hutchinson *et al.* (2009), Kaufman (2009), and Sloan (2009), which all raise the question of security as a concern in the cloud computing environment. Since cloud computing is such a new and talked about topic, numerous blog and web articles are also talking about security-related concerns in a cloud. Some examples are found in the following: Cloud Security Alliance (2009); InfoSecurity (2009); Knights (2009); and Twentyman (2009).

The different service models offered by cloud providers determine the security mechanisms needed to provide adequate privacy and data protection in the cloud. Current published research lacks detailed explanations and more importantly practical experience of security measures provided by cloud computing providers. Many questions have been raised about security in cloud computing but few answers exists at this stage, as cloud computing is still in an adaptation or peak of inflated expectations phase.

In this paper we present the results of a comparison of three PaaS clouds, with specific reference to the three security goals:

1. Confidentiality;
2. Integrity; and
3. Availability

An analysis of the security mechanisms and Service Level Agreements (SLAs) provided by these PaaS clouds are presented, as well as results from experiments that were run in the three PaaS clouds. Finally, the implications of these results for writing a WPS in a PaaS cloud are discussed.

2. PAAS CLOUD COMPARISON

In this section we present the results of our comparison of three PaaS cloud computing solutions, namely Microsoft Azure (MWA), Google App Engine (GAE) and GroundOS (GOS). First we provide some background on the comparison and then we describe the results of our comparison.

2.1 Background

Figure 1 shows the different cloud service models and how they relate to security and user control over resources. Typically, as you move up through the layers from IaaS to SaaS, there are fewer security risks for the user and provider but at the cost of less control by the user. Each layer serves a different purpose to serve both users who are just regular internet users, as well as developers.

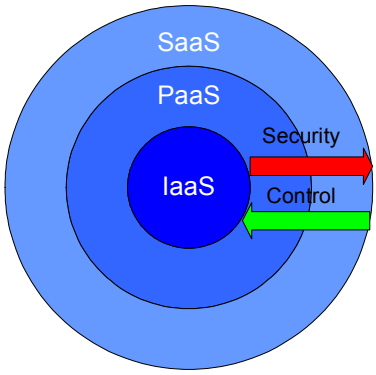

Figure 1. Cloud service models related to security and user control

While the IaaS model provides the most control over development, network and storage, it is also the most susceptible to risk. A user can upload anything they want onto the machine, or even develop malicious code straight on the machine. This model can be further attacked by taking advantage of the privacy and confidentiality agreements and violating the terms of service agreement. Further up the model layers, less is at risk for the provider and user. At the top, SaaS does not provide the chance to create internal attacks by users (obviously still susceptible to insider employee attacks); only external attacks, such as denial of service attacks, affecting availability are possible.

Our comparison focuses on the PaaS layer, which provides the user with the ability to develop and deploy custom applications onto the cloud infrastructure. Most cloud providers largely restrict development to a few specific development languages at this level. This model provides a greater level of abstraction of resources where the user is not given control of lower level resources.

The clouds in our comparison, namely Microsoft Windows Azure (MWA), Google App Engine (GAE) and GroundOS (GOS) are all examples of PaaS clouds. MWA and GAE are proprietary clouds whereas GOS is an open source cloud. These three clouds were selected to provide insight into what current cloud providers have to offer in terms of general cloud offerings and security in their clouds. In order to be able to compare apples with apples, the clouds were all chosen to be PaaS clouds. Also, PaaS clouds are interesting for developing a WPS that takes advantage of the scalability capabilities of a cloud.

2.2 Results

Table 2 shows the five most important aspects of security that a cloud should address. Availability is usually stipulated in an SLA. This is considered by many providers as the place to address availability and some SLAs consist of issues on availability only. Availability, which is often called uptime, can be defined as how long a service (i.e. the cloud) will be available or online. Availability is largely provided by reliable software and reliable scalability under pressure. For example, MWA promises an annual uptime of 99.95% for computational operations. Such a percentage of uptime is very promising for any type of business.

Confidentiality and integrity of data stored in the cloud is mainly provided by encryption and password security, ensuring that only authorised users get access to data.

MWA provides authentication to the entrance of its developer web portal (interface) and also standard SQL authentication and authorisation practices to their database through logins and GRANT/DENY/REVOKE commands (Microsoft Windows Azure 2009).

GAE provides authentication though its web portal and single-sign on authentication for first time registration. Single-sign on allows a generated password to be used only once. Upon sign-up to use GAE and request for a domain name a password is generated and sent to the user's cellular phone. This generated password can then be used to activate the user's account, after which the password cannot be used again.

All of the compared clouds allow SSL encryption. SSL is one of the most widely used communication protocols on the internet. It provides encrypted communication between a client and server, as well as optional client and server authentication (Pfleeger & Pfleeger 2007). While SSL provides network communication encryption, data encryption is provided by programming libraries and database management systems as discussed earlier.

	GAE	MWA	GroundOS
Availability	No SLA and no mention of guaranteed uptime.	Provided by SLA	Problem of user
Integrity	Encryption Authentication	Encryption Authentication	Problem of user Encryption
Confidentiality	Privacy policy Encryption Authentication	Privacy policy Encryption Authentication	Problem of user Encryption
Authentication	Single-sign on Username & password	Username & password	Username & password
SLA	No	Yes	No

Table 2. Comparison of security in the three PaaS clouds

Encryption in GAE is available through programming language libraries. MWA and GAE require the developer to handle encryption and decryption when storing sensitive data; data is not encrypted by default. MWA, GAE and GOS provide network communication encryption through SSL, which has to be set up by the developer. Protection of user data is stipulated in privacy policies, which detail the conditions under which someone is considered to be violating privacy.

MWA and GAE both provide an automated failover system, which will relocate a user's data to another data center if the current data center were to fail by some disaster (Google Apps 2007, Microsoft Windows Azure 2009). While cloud providers generally provide redundant backups it is recommended that users make personal backups regularly.

A number of experiments were run in the GAE and MWA clouds to test the security aspects that the cloud providers promised to provide. These experiments were run only in the GAE and MWA clouds because they are the proprietary solutions and therefore promise a secure environment in their terms of service and SLAs. GOS on the other hand, does not require SLAs and terms of service agreements because applications are run in a private cloud, for example, inside a company and thus not available to outside users. While these experiments are by no means complete and exhaustive, neither very complex, they are simply a means to further the understanding of security concerns in clouds. There is no better way to compare clouds than to run one's own experiments and see if they are doing what they say they should be doing – the proof is in the pudding! The Python programming language was used in the GAE cloud and C# was used in the MWA cloud, as these were the only languages available at the time the experiments were run in August 2009. The same experiments were run in both clouds. The main focus was to find potential points of attacks that are available to outside users from inside the cloud or against the service provider itself. The four experiments are described below.

Write to file system. The first experiment consisted of creating and writing out a file into the cloud. If allowed, this functionality can be used maliciously to create and upload Trojan programs or some other malicious files into the cloud. It was expected that this would not work because it is a major security threat to the cloud providers' internal storage systems.

Ping request. The second experiment attempted to create a socket to a host that would simply ping the user to see if a connection can be made. The security implications of this are obvious: various types of denial of service attack, such as smurf attacks and ping of death attacks can be launched through a socket.

Spawn thread. A third experiment attempted to create one or more threads, which can be used to run multiple instances of something, valuable to any type of attack. To protect the cloud, threads should work as specific worker roles in MWA and should be blocked by GAE's sandboxing.

System call. The fourth experiment attempted to execute a system call that launches a command prompt console. If a user can gain access to a console, this inevitably means the user has full control over the target computer. Therefore, this type of code in the cloud should be strictly forbidden and prohibited for the obvious reason stated above.

Table 3 and 4 show the results of the experiments. They show that security measures are in place and provide adequate protection against attacks by users' applications running in the cloud. A more detailed report of the experiments can be found in Ludwig (2009).

Cloud providers are very careful when it comes to security in the cloud because it is currently the main concern preventing large-scale adoption, specifically by large businesses.

	GAE	MWA
Write to file system	Denied	Denied
Ping request	Denied	Denied
Spawn thread	Denied	Granted
System call	Denied	Denied

Table 3. Experiment results

	GAE	MWA
Write to file system	No support for I/O	Permission error raised
Ping request	Block opening of sockets by customised Python API	I/O error raised
Spawn thread	Limited to one thread (main)	Allowed but limited to two worker roles
System call	No API support	Permission error raised

Table 4. Preventative measures by the clouds in the experiments

3. IMPLEMENTING A WPS ON A PAAS CLOUD

There are many advantages of deploying a geoprocessing service in a cloud. For example, a compute intensive WPS, such as one that does climate modelling, could benefit from the scalability of the resources in a cloud; or a WPS that suddenly becomes popular (e.g. one that does polygon intersections), could benefit from the user scalability provided by the cloud. While geoprocessing services probably stand to benefit most from processing (resource) scalability, implementations of OGC's Web Feature Service (WFS) stand to benefit from the scalability of user requests that a cloud provides.

A WPS can be deployed on any of the three layers of a cloud, namely IaaS, PaaS or SaaS. In an IaaS cloud, the user has full control over resources but virtualisation happens at a low level and the user therefore has the responsibility to address many security risks. Deploying geoprocessing on an IaaS cloud makes sense if an organization needs to deploy a variety of WPSs and has the required human resources to administer the IaaS cloud. In a PaaS cloud virtualisation happens at a higher level of abstraction and therefore some of the development challenges of building scalable applications are handled by the cloud platform. These include the security measures discussed in the previous section. In a SaaS cloud, an existing WPS is deployed. Developing geoprocessing services such as a WPS on a PaaS platform is easier and safer than on an IaaS cloud, while giving the user more control than in a SaaS cloud.

Typically geographic data is required as input for geoprocessing and a geoprocessing service could also produce geographic output data. It is important that copyright and other rights of the data are protected. A project by the International Organization for Standardization's (ISO) technical committee for geographic information (ISO/TC 211, www.isotc211.org) is currently developing a reference model for digital rights management (DRM) functionality for geospatial resources (GeoDRM). Examples of geospatial resources are geographic data files and geoprocessing services. The ISO work follows on earlier work in OGC by the Geo Rights Management Working Group (http://www.opengeospatial.org/projects/groups/geormwg). The reference model aims to be as close as possible to other resources, such as music, text, or services. This reference model should be implemented for a WPS in a PaaS cloud to ensure proper protection of data and intellectual property rights.

In a WPS input data is supplied either embedded in the Execute request, or referenced as a web accessible resource. Both ways are possible in a PaaS cloud because they do not rely on accessing files on a local file system. However, the WPS source code cannot create any temporary files that could be required during processing. Instead, such data has to be stored temporarily in cloud database tables.

The WPS specification was written to allow existing software interfaces to be wrapped up and presented to the network as web services (OGC 2007). However, in a PaaS cloud, existing software cannot necessarily be used. Instead, deployment in a PaaS cloud requires the whole WPS to be ported to the PaaS cloud and if required, conversion to a programming language supported by the PaaS. This conversion includes the removal or rewrite of any code with potential cloud security violations, such as ping requests, system calls and threads. Also, the code should be implemented to incorporate the cloud features that are automatically provided by the PaaS. Depending on how the existing geoprocessing is implemented, this could be a considerable effort, which is only worth the effort if cloud features such as availability, scalability, unlimited storage and security measures that are automatically provided in the PaaS cloud, are exploited.

As can be seen from our experiments, PaaS was available in only a few programming languages in August 2009 and since then, only a few more have been added. Further, the security

measures, as well as other cloud features, are provided in different ways in different PaaS. Developing a geoprocessing service on a specific PaaS has the risk of locking the developer into a specific vendor's platform. Initiatives such as the Open Cloud Manifesto (2009) and the Open Cloud Consortium (2009) aim to counter vendor lock-in by developing standards for interoperability between clouds and benchmarks for cloud computing.

4. CONCLUSION

Security is very important when the internet is involved as the internet creates anonymity and removes boundaries. Cloud computing exists on the internet backbone and gives users the ability to connect from anywhere. Hence, security in cloud computing is a vital consideration. The security goals of confidentiality, integrity and availability are commonly used in the field of security and we applied them here to security in cloud computing, as well as to geoprocessing.

In order to analyse the various PaaS clouds, it was necessary to support the theoretical validation of the study with practical validation. While the experiments run in the cloud were neither exhaustive nor complex they provided a step towards understanding the cloud better, especially in terms of the security restrictions. The comparison of PaaS clouds provides valuable insight into how security in cloud computing is addressed. The practical experience of running applications in the cloud provides valuable knowledge into how security measures in the cloud are implemented and what the benefits of its use are. Finally, we discussed the implications of the security measures for the development of geoprocessing services, such as OGC's WPS, in a PaaS cloud. These results are valuable to all WPS developers.

Further work in our research group is investigating the implementation of a WPS for geocoding on a cloud. A specific research question is how to access the address reference data that is required for the geocoding.

With cloud computing being a fairly new technology, there is a huge avenue for further research in security related to geoprocessing services on a cloud platform. New protocols to aid in interoperability between providers can be designed and established. Security models could be developed for cloud computing that could exist at one or multiple layers of the cloud conceptual stack. Trust plays a big part in large companies adopting and utilising the cloud and could be overcome with the solution of an independent trust body to certify a cloud provider's status.

This paper was limited to the comparative study of one type of service model, namely the PaaS model, and we discussed implications of the security measures on a geoprocessing service. Further research can be conducted at the IaaS layer, which will provide much more insight into the inner working of the cloud, its security restrictions and how these apply to geoprocessing services such as a WPS. The IaaS layer can be further investigated for building applications that are non-web based and require a specific operating system. Implications for implementations of other OGC web services, such as the WFS, on both PaaS and IaaS cloud platforms would enhance the understanding of implementing OGC web services on clouds. Also to be investigated is if and how data transfer between different geoprocessing services can be reduced in an IaaS cloud, thus improving performance of not only geoprocessing services, but also geoprocessing service chaining.

Furthermore, the advantages and disadvantages of open source and proprietary clouds can be compared to each other in order to determine any additional benefits. Open source initiatives may have a head start on interoperability and standards as seen with the Open Cloud Consortium.

Cloud computing as it stands today, is a viable option for businesses and individuals. The benefits for start-up companies are far outweighed by the disadvantages and can only improve over time. However, any potential company or individual wanting to utilise the resources that the cloud computing environment has to offer, needs to weigh the advantages against the disadvantages and compare various providers for all concerned issues. Some providers offer varying amount of detail in their terms of service and SLAs and provide different platforms, development languages and service models. All these issues need to be addressed to find the optimum solution for deployment of geoprocessing services.

REFERENCES

Brauner J, Foerster T, Schaeffer B and Baranski B, 2009, Towards a Research Agenda for Geoprocessing Services, *12th AGILE International Conference on Geographic Information Science*, Leibniz Universität Hannover, Germany, 2009.

Cloud Security Alliance, 2009, Security Guidance for Critical Areas of Focus in Cloud Computing, http://www.cloudsecurityalliance.org (accessed 9 June 2010)

Coetzee S and Bishop J, 2009, An analysis of technology choices for data grids in a spatial data infrastructure. *Spatial Data Infrastructure Convergence: Research, Emerging Trends, and Critical Assessment* edited by B van Loenen, J Zevenbergen and J Besemer, Nederlandse Commissie voor Geodesie/Netherlands Geodetic Commission, 48, 2009.

Everett C, 2009, Cloud computing - A question of trust, *Computer Fraud & Security*, 2009(6), pp. 5-7

Google Apps, 2007, *Comprehensive review of security and vulnerability protections for Google Apps*, www.google.com/a/help/intl/en/admins/pdf/ds_gsa_apps_whitepaper_0207.pdf (accessed 9 June 2010)

Grossman RL, 2009, The Case for Cloud Computing, *IT Professional*, 11(2), pp. 23-27

Hutchinson C, Ward J & Castilon K, 2009, Navigating the Next-Generation Application Architecture, *IT Professional*, 11(2), pp. 18-22

InfoSecurity, 2009, Three Quarters Of Organisations To Increase Security For Cloud Computing, http://www.infosec.co.uk/page.cfm/Action=Press/PressID=1242 (accessed 22 Feb. 2009)

Kaufman LM, 2009, Data Security in the World of Cloud Computing, *IEEE Security and Privacy*, 7(4), pp. 61-64

Knights M, 2009, Can security kill cloud computing?, http://www.itpro.co.uk/610299/can-security-concerns-kill-cloud-computing (accessed 9 June 2010)

Ludwig B, 2009, *A comparison of PaaS clouds with a detailed reference to security*, BSc Hons research report, University of Pretoria, South Africa, November 2009.

Microsoft Windows Azure, 2009, *Azure Services Platform*, http://www.microsoft.com/azure/data.mspx (accessed 9 June 2010)

Open Cloud Manifesto, 2009, Open Cloud Manifesto, http://www.opencloudmanifesto.org (accessed 9 June 2010)

Open Cloud Consortium, 2009, Open Cloud Consortium, http://www.opencloudconsortium.org (accessed 9 June 2010)

Open Geospatial Consortium (OGC), 2007, *OpenGIS Web Processing Service*, Reference number OGC 05-007r7, Versions 1.0.0, Open Geospatial Consortium, Inc.

Pfleeger CP & Pfleeger SL, 2007, *Security in Computing*, Fourth edition, Prentice Hall, Upper Saddle River, New Jersey, USA.

Sloan K, 2009, Security in a virtualised world, *Network Security*, 2009(8), pp. 15-18

Twentyman J, 2009, Security concerns for cloud computing, http://www.computerweekly.com/Articles/2009/03/27/235439/security-concerns-for-cloud-computing.htm (accessed 9 June 2010)