Performance Testing of Cloud Storage while Using Spatial Data

MARTIN LNENICKA[1], JITKA KOMARKOVA[1], EVA MILKOVA[2]
1) Faculty of Economics and Administration
University of Pardubice
Studentska 95, 532 10 Pardubice
2) Faculty of Science
University of Hradec Kralove
Rokitanskeho 62, 500 03 Hradec Kralove
CZECH REPUBLIC
martin.lnenicka@student.upce.cz, jitka.komarkova@upce.cz, eva.milkova@uhk.cz

Abstract: This paper deals with the meaning of the cloud computing technologies in nowadays GIS and their role in the presentation and the availability of the spatial data via internet. The first part is focused on the role of cloud storage, GIS and their intersection. It is followed by performance testing of selected cloud storage services while using spatial data, with emphasis on measurements of upload and download speeds.

Key-Words: Cloud computing, Cloud storage, Spatial data, Performance testing

1 Introduction

Cloud computing changes the way users access the ICT systems, which means resources (computers, infrastructures, data storage and applications) and the way of managing and delivering computing technologies, services and solutions. It is a new model of delivering computing resources in a pay-per-use approach, which allows businesses to use their applications without any software installation and access their personal files at any computer with internet access using only an internet browser.

Cloud storage is a young industry for processing and storing very large amounts of data, but it has a great promise for growth. Cloud storage providers (CSPs) such as Amazon, Google and Microsoft have already been dependent on the cloud storage to manage and run their businesses, and they have just begun to offer varying degrees of performance and reliability to smaller enterprises. Since there are no two CSPs which are equal, potential customers need to identify those that can deliver the appropriate price / performance levels to meet their needs.

Geographic information systems (GIS) is being implemented not only as a standalone system but integrated into other systems. It is a collection of wide range of tools that captures, stores, analyses, manages and utilizes large volumes of geospatial data and delivers geodata and services for massive concurrent users. Therefore, one of the requirements is the ability to handle the huge volume of spatial data and ensure the required performance with operational flexibility. This case study is focused on cloud storage systems and their performance testing, concretely how quickly CSPs can write and read spatial data stored as files of different sizes.

2 Problems of Cloud Computing, GIS and Storage in the Related Works

Cloud computing refers to both the applications delivered as secure services over the internet and the hardware and software systems in the data centres that provide those services. The services offered in the cloud can be divided into: cloud infrastructure – distributed multi-site physical infrastructure (enabled by server virtualization); cloud platform (e.g. web server); cloud storage, which describes the ability to safely store a file in the cloud (formerly utility computing); cloud applications (applications delivered as a service), and cloud clients – presentation layer (e.g. browsers or mobile devices).

Cloud storage is a service-oriented model of networked online storage where data is stored in virtualized pools of storage which are generally hosted by third parties. Physically, the resources may be located all over the world. Cloud storage services may be accessed through a web service API, command line tool or through a web-based user interface to create, share, and manage data [6].

Cloud computing is an emerging technology for processing and storing very large amounts of different data types and sizes. Sometimes anomalies and defects affect parts of the cloud infrastructure, resulting in a performance degradation of the cloud.

ISBN: 978-1-61804-130-2

Authors in [2] propose a performance measurement framework for Cloud Computing systems, which integrates software quality concepts from ISO 25010. The importance of performance and scaling in cloud environments is discussed in [3]. Authors in [9] analyze the performance of cloud computing services for scientific computing workloads. There is performed an empirical evaluation of the performance of four commercial cloud computing services including Amazon, and there is compared the performance characteristics and cost models of clouds and other scientific computing platforms through trace-based simulation. Suitable platforms and standards to develop applications and store data are also solved in [13].

In addition, GIS functions and services that use spatial data are geographically and logically distributed according to the source of data, location of computing facilities and organizations. The spatial analyses on large amount of data, such as [10], are complex and computationally intensive. In order to share and work with geodata and the computation results among geographically dispersed users, a scalable and low cost cloud computing platform is a good solution for GIS application. It also enables many users to interact together, exchanging and collaborating with data pertaining to multiple disciplines. The spatial data can be stored in the cloud without paying attention to details of huge volume data storage and spatial data security. The possible applications of cloud computing as a new storage and delivery GIS model are discussed in [4], [11], [14] or [15].

3 Case study of the Performance Testing of Cloud Storage Services

The performance of cloud storage services depends on many influences. Main factors are: the network that transmits the data between storage system and the end user; and the performance of the storage service itself (how scalable is the service provider and how many users are using it at the same time). Thus, these factors have to be considered in the initial conditions.

3.1 List of Cloud Storage Providers

The comparison given in the Table 1 focuses only on public CSPs which are appropriate for business use and which are able to store large amounts of spatial data.

Table 1: - Some characteristics of selected public CSPs (source: authors).

Service name	Storage	Cost	For free
Amazon S3	First 1 TB / month…Over 5000 TB / month	$0.125 per GB…$0.055 per GB	5 GB of Amazon S3 storage and 15 GB of data transfer out each month for one year.
AT&T Synaptic Storage as a Service	1 GB month; inbound; outbound	$0.175 per GB; $0.10 per GB / hour; $0.10 per GB / hour	Nothing for cloud storage.
GoGrid	1 GB / month	$0.15 per GB	Store up to 10 GB / month at no cost.
Google Cloud Storage	First 0 – 1 TB / month…Next 400 TB	$0.12 per GB…$0.085 per GB	5 GB of storage, 25 GB of download and 25 GB of upload data.
HP Cloud Object Storage	0-50 TB / month… Next 950 TB; uploads	$0.12 per GB…$0.10 per GB; free	Nothing for cloud storage.
Microsoft Windows Azure	First 1 TB / month…Next 4000 TB / month	$0.125 per GB…$0.055 per GB	90-day free trial, 35 GB with 50 GB storage transactions.
Nirvanix Cloud Storage	1 GB month; uploads; downloads	$0.25 per GB; $0.10 per GB; $0.15 per GB	Nothing for cloud storage.
OpSource Cloud Files	1 GB; incoming; outgoing	$0.0072 / day; free; first 10 TB for $0.15/GB, over 50 TB for $0.09/GB	14-day free trial.
Peer1 Hosting	1 GB / month	$0.15 per GB (basic) $0.37 per GB (premium)	Nothing for cloud storage.
Rackspace Cloud Files	1 GB / month	$0.10 per GB; free; $0.18 per GB	Nothing for cloud storage.

ISBN: 978-1-61804-130-2

This is not a complete list, however it does provide a good representative list of CSPs. Customers can usually choose between the geographic redundancy service which replicates data between two geographically distant sites so that applications can switch from one site to another (for example, in case of the catastrophic failure of one) and still have all the configuration data, and locally redundant (it is cheaper).

3.2 Proposed Way of Solving the Case Study

The first step have to be the definition of initial conditions – see Table 2 (both locations are situated in the Czech Republic), then choose CSPs according to requirements (the only condition is free access), it is followed by the performance testing of cloud storage while using spatial data, and selection of the best cloud storage provider for spatial data at the end of the case study.

Table 2: - Initial conditions and prerequisites, which were used for performance testing (source: authors).

	Location 1	Location 2
Average download speed	375,1 kB/s	583,7 kB/s
Average upload speed	34,8 kB/s	57,2 kB/s
Average ping	73,5 ms	53,7 ms
HW + SW configuration	Windows 7 Home premium 64bit; AMD A6-3400M; 1,4 GHz; 4 GB DDR3 1333 MHz	Windows XP Home SP3; Celeron Dual-Core T3000 1,80GHz; 2 GB DDR2 667 MHz

For performance measurement and comparison Gladinet Cloud Desktop Professional Edition licensed per-user was used. It can be installed on multiple computers for a single user (a 30-day free trial version can attach only 2 cloud storages) [7]. From Table 1 there were chosen only Amazon S3 [1], Windows Azure [12] and Google Cloud Storage [8] for the evaluation because OpSource Cloud Files and GoGrid are not supported by Gladinet.

The collected data were statistically analysed using Microsoft Excel 2010.

3.3 A Comparison of Upload and Download Speeds

At first, download speed of these three storage services was tested (by the same day as the following testing). Cloud speed test according to CloudHarmony [5] was won by Windows Azure Storage – see Table 3. These results will be later compared with our own performance testing.

Table 3: - Results of download speed test by CloudHarmony (source: authors).

Service name	Transferred	Average speed	Time
Amazon S3	1,91 MB	172,5 kB/s	11,1 s
	650 kB	72,5 kB/s	9,13 s
Google Storage	1,91 MB	168,8 kB/s	11,3 s
	650 kB	69,2 kB/s	9,57 s
Windows Azure	1,91 MB	217,5 kB/s	9,1 s
	650 kB	94,8 kB/s	7,04 s

Upload and download speed was measured with various spatial data files of different sizes for all 3 CSPs. The sizes of files: 1) file size: 3,57 MB; 2) file size: 120 kB; 3) 16 files with size between 100 kB and 500 kB: in sum 3,57 MB and 4) 61 files with size to 5 kB: in sum 120 kB. Transmission speeds were measured simultaneously from two locations hourly from 10:00 to 16:00 (during the working day). The results of the first two measurements are shown in Table 4 and Table 5.

Table 4: - Speed results and time of upload and download – file size of 3,57 MB (source: authors).

Testing CSP from location / time of test			Test 1 – 10:00	Test 2 – 11:00	Test 3 – 12:00	Test 4 – 13:00	Test 5 – 14:00	Test 6 – 15:00	Test 7 – 16:00
Upload speed	Location 1	Amazon	146,5 s; Ø 26,1 kB/s	145,2 s; Ø 26,3 kB/s	150,6 s; Ø 25,3 kB/s	149,3 s; Ø 25,5 kB/s	148,9 s; Ø 25,6 kB/s	149,1 s; Ø 25,5 kB/s	149,8 s; Ø 25,4 kB/s
		Google	140,5 s; Ø 26,6 kB/s	139,8 s; Ø 26,8 kB/s	140,6 s; Ø 26,7 kB/s	141,3 s; Ø 26,5 kB/s	142 s; Ø 26,3 kB/s	141,8 s; Ø 26,4 kB/s	141,9 s; Ø 26,4 kB/s
		Windows	137,9 s; Ø 27,3 kB/s	138,2 s; Ø 27,1 kB/s	138,1 s; Ø 27,1 kB/s	139 s; Ø 26,9 kB/s	138,1 s; Ø 27,1 kB/s	138,4 s; Ø 27 kB/s	138,5 s; Ø 27 kB/s
	Location 2	Amazon	82 s; Ø 46,6 kB/s	81,6 s; Ø 46,7 kB/s	81,7 s; Ø 46,7 kB/s	82 s; Ø 46,6 kB/s	82,7 s; Ø 46,1 kB/s	82,9 s; Ø 46 kB/s	83,2 s; Ø 45,8 kB/s
		Google	75,1 s; Ø 51,3 kB/s	74,9 s; Ø 51,5 kB/s	75,6 s; Ø 51 kB/s	76 s; Ø 50,7 kB/s	76,3 s; Ø 50,4 kB/s	76,9 s; Ø 50 kB/s	77,3 s; Ø 49,8 kB/s
		Windows	75,6 s; Ø 50,6 kB/s	76,5 s; Ø 49,9 kB/s	75,9 s; Ø 50,7 kB/s	76,1 s; Ø 50,5 kB/s	77,4 s; Ø 49,7 kB/s	77,6 s; Ø 49,6 kB/s	77,8 s; Ø 49,4 kB/s
Download speed	Location 1	Amazon	12,7 s; Ø 302,9 kB/s	12,9 s; Ø 301 kB/s	12,8 s; Ø 301,5 kB/s	12,8 s; Ø 301,4 kB/s	13 s; Ø 295,4 kB/s	13,3 s; Ø 289 kB/s	13,2 s; Ø 291,1 kB/s
		Google	12,8 s; Ø 301,1 kB/s	12,6 s; Ø 306,2 kB/s	12,4 s; Ø 310 kB/s	12,5 s; Ø 308,9 kB/s	12,7 s; Ø 302,8 kB/s	13 s; Ø 295,6 kB/s	13,1 s; Ø 293,5 kB/s
		Windows	12 s; Ø 317,7 kB/s	12,2 s; Ø 313,8 kB/s	12,1 s; Ø 316,9 kB/s	12 s; Ø 317,8 kB/s	12,1 s; Ø 318,6 kB/s	12,5 s; Ø 308,3 kB/s	12,9 s; Ø 300,9 kB/s
	Location 2	Amazon	8,1 s; Ø 471,6 kB/s	8 s; Ø 478,1 kB/s	8,3 s; Ø 460,6 kB/s	8,3 s; Ø 460,7 kB/s	8,5 s; Ø 451,8 kB/s	8,7 s; Ø 442,4 kB/s	8,9 s; Ø 430,5 kB/s
		Google	8,1 s; Ø 471,5 kB/s	8,4 s; Ø 455 kB/s	7,9 s; Ø 483,3 kB/s	8,1 s; Ø 471,3 kB/s	8,2 s; Ø 466,1 kB/s	8,3 s; Ø 460,5 kB/s	8,4 s; Ø 454,7 kB/s
		Windows	7,5 s; Ø 509,4 kB/s	7,8 s; Ø 489,8 kB/s	7,7 s; Ø 496,2 kB/s	7,8 s; Ø 489,4 kB/s	8,2 s; Ø 465,9 kB/s	8,1 s; Ø 471,3 kB/s	8,3 s; Ø 460 kB/s

ISBN: 978-1-61804-130-2

Table 5: - Speed results and time of upload and download – file size of 120 kB (source: authors).

Testing CSP from location / time of test			Test 1 – 10:00	Test 2 – 11:00	Test 3 – 12:00	Test 4 – 13:00	Test 5 – 14:00	Test 6 – 15:00	Test 7 – 16:00
Upload speed	Location 1	Amazon	5,6 s; Ø 24,6 kB/s	5,4 s; Ø 25,1 kB/s	5,2 s; Ø 26,1 kB/s	5,3 s; Ø 25,6 kB/s	6,1 s; Ø 22,5 kB/s	6,7 s; Ø 20,5 kB/s	6,9 s; Ø 19,8 kB/s
		Google	6,1 s; Ø 22,4 kB/s	5,9 s; Ø 23,2 kB/s	5,6 s; Ø 24,7 kB/s	5,9 s; Ø 23,3 kB/s	6,4 s; Ø 21,5 kB/s	6,6 s; Ø 20,8 kB/s	6,7 s; Ø 20,6 kB/s
		Windows	5,1 s; Ø 26,5 kB/s	5,2 s; Ø 26 kB/s	5 s; Ø 27,1 kB/s	5,1 s; Ø 26,6 kB/s	5,2 s; Ø 26,1 kB/s	5,1 s; Ø 26,5 kB/s	5,3 s; Ø 25,5 kB/s
	Location 2	Amazon	3,8 s; Ø 36,1 kB/s	3,9 s; Ø 35,2 kB/s	4 s; Ø 34,7 kB/s	3,9 s; Ø 35,1 kB/s	4,1 s; Ø 33,9 kB/s	4,2 s; Ø 32,6 kB/s	4,4 s; Ø 31,2 kB/s
		Google	4 s; Ø 34,8 kB/s	4,1 s; Ø 33,9 kB/s	3,7 s; Ø 38,7 kB/s	3,9 s; Ø 35,2 kB/s	4,2 s; Ø 32,6 kB/s	4,3 s; Ø 32 kB/s	4,2 s; Ø 32,7 kB/s
		Windows	3,2 s; Ø 42,7 kB/s	3,3 s; Ø 41,5 kB/s	3,1 s; Ø 44,1 kB/s	3,2 s; Ø 42,7 kB/s	3,2 s; Ø 42,8 kB/s	3,5 s; Ø 39,1 kB/s	3,7 s; Ø 38,8 kB/s
Download speed	Location 1	Amazon	1 s; Ø 138,6 kB/s	1 s; Ø 138,5 kB/s	1 s; Ø 138 kB/s	1,1 s; Ø 134,5 kB/s	1,1 s; Ø 133,9 kB/s	1,2 s; Ø 123,2 kB/s	1,1 s; Ø 132,3 kB/s
		Google	0,9 s; Ø 158,5 kB/s	1 s; Ø 138,8 kB/s	0,9 s; Ø 157,9 kB/s	1 s; Ø 138,5 kB/s	1,1 s; Ø 133,8 kB/s	1 s; Ø 137,9 kB/s	1 s; Ø 137,2 kB/s
		Windows	1 s; Ø 138,7 kB/s	1 s; Ø 138,6 kB/s	0,9 s; Ø 158,5 kB/s	1 s; Ø 139 kB/s	0,9 s; Ø 158,7 kB/s	1 s; Ø 137 kB/s	1 s; Ø 136,9 kB/s
	Location 2	Amazon	0,8 s; Ø 208,7 kB/s	0,8 s; Ø 207,4 kB/s	0,7 s; Ø 235,3 kB/s	0,7 s; Ø 233,4 kB/s	0,8 s; Ø 206,7 kB/s	0,8 s; Ø 205,4 kB/s	0,8 s; Ø 203,9 kB/s
		Google	0,7 s; Ø 233,8 kB/s	0,6 s; Ø 273,9 kB/s	0,6 s; Ø 273 kB/s	0,6 s; Ø 273,1 kB/s	0,7 s; Ø 233,6 kB/s	0,7 s; Ø 232,7 kB/s	0,8 s; Ø 205,3 kB/s
		Windows	0,6 s; Ø 274,5 kB/s	0,7 s; Ø 235,7 kB/s	0,6 s; Ø 272,3 kB/s	0,6 s; Ø 273,9 kB/s	0,7 s; Ø 233,4 kB/s	0,7 s; Ø 232,9 kB/s	0,7 s; Ø 233 kB/s

Fig. 1 and 2 show the progress of upload and download speed (both with the spatial data file of 3,57 MB) between 10:00 – 16:00. While the upload speed is approximately constant, the download speed decreases slowly in the afternoon. It is apparently affected by the end of working hours in Europe, when for example employees do data backup in cloud storage. It shows how scalable are services of the cloud storage provider. Another finding is that the upload speed from location 2 to Amazon cloud storage is significantly slower compared to Google and Windows – see Fig. 1.

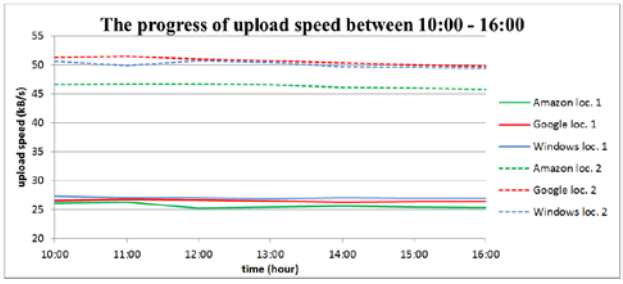

Fig. 1: - The progress of upload speed between 10:00 – 16:00 (source: authors).

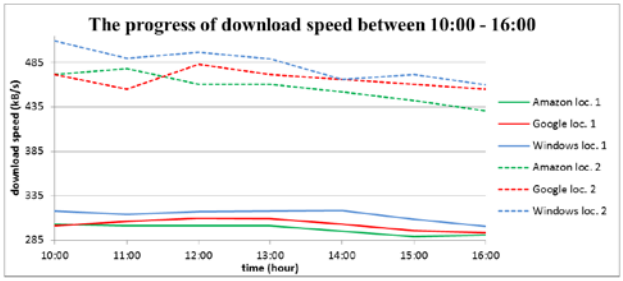

Fig. 2: - The progress of download speed between 10:00 – 16:00 (source: authors).

Table 6 shows results of average upload and download time while working with spatial data. Average upload time of the single file (size of 3,57 MB) is shorter when compared with the folder of the same size containing 16 files. It is even more obvious when using the second file (size of 120 kB). Upload time of the same size folder of 61 spatial data files is twice as long as the single file. From this comparison it is clear, that Microsoft's Cloud Platform – Windows Azure is for spatial data files more suitable than Google and Amazon.

Table 6: - Comparison of average upload and download times (source: authors).

	Amazon loc. 1	Google loc. 1	Windows loc. 1	Amazon loc. 2	Google loc. 2	Windows loc. 2
Average upload time of 3,57 MB GIS file	148,49 s	141,13 s	138,31 s	82,30 s	76,01 s	76,70 s
Average download time of 3,57 MB GIS file	12,96 s	12,73 s	12,26 s	8,40 s	8,20 s	7,91 s
Average upload time of 120 kB GIS file	5,89 s	6,17 s	5,14 s	4,04 s	4,06 s	3,31 s
Average download time of 120 kB GIS file	1,07 s	0,99 s	0,97 s	0,77 s	0,67 s	0,66 s
Average upload time of 16 GIS files (in sum 3,57 MB)	152,20 s	151,93 s	144,93 s	81,74 s	82,04 s	80,57 s
Average download time of 16 GIS files (in sum 3,57 MB)	13,11 s	12,97 s	11,40 s	8,40 s	8,20 s	7,84 s
Average upload time of 61 GIS files (in sum 120 kB)	12,21 s	11,77 s	11,60 s	6,87 s	6,70 s	6,64 s
Average download time of 61 GIS files (in sum 120 kB)	6,47 s	6,26 s	5,79 s	2,31 s	2,24 s	2,21 s

4 Conclusion

This paper is only one of the steps on the way of solving the issue of spatial data in the cloud. In order to obtain comparable results, it is necessary to test cloud storage from multiple locations across the world, also with other initial conditions. Return of investment and implementation costs must not be forgotten as well.

An upload/download test was done to see how quickly it is possible to upload and download data to/from the cloud storage. In this case study Microsoft with Windows Azure was evaluated as the best solution. The results also agree with the download speed test by CloudHarmony.

Users do not like to wait, time is money, thus speed and response time are important. However, the storage in cloud is still not safe to backup sensitive (strategic) business data. But it can be used for information sharing with business partners. The tested cloud storages are still not suitable for heavy transaction processing applications. According to the case study, the upload speed is comparatively low for all tested solutions. However, the promising alternative is to download spatial data from multiple locations.

ISBN: 978-1-61804-130-2

References:

[1] Amazon Simple Storage Service (Amazon S3). *Amazon Web Services* [online]. ©2012 [cit. 2012-08-26]. Available from: http://aws.amazon.com/s3/

[2] Bautista, L.; Abran, A.; April, A. Design of a Performance Measurement Framework for Cloud Computing, *Journal of Software Engineering and Applications*, Vol. 5 No. 2, 2012, pp. 69-75.

[3] Bellenger, D.; Bertram, J.; Budina, A.; Koschel, A.; Pfänder, B.; Serowy, C.; Astrova, I.; Grivas, S. G.; Schaaf, M. Scaling in Cloud Environments, *Proceedings of the 15th WSEAS international conference on Computers*, Corfu, 2011, pp. 145-150.

[4] Bhat, M. A.; Shah, R. M.; Ahmad, B. Cloud Computing: A solution to Geographical Information Systems (GIS), *International Journal on Computer Science and Engineering (IJCSE)*, Vol. 3, No. 2, 2011. pp. 594-600.

[5] Cloud Speed Test. *CloudHarmony: transparency for the cloud* [online]. ©2011 [cit. 2012-08-24]. Available from: http://cloudharmony.com/speedtest

[6] Furht, B.; Escalante, A. *Handbook of Cloud Computing*. 1. ed. London: Springer, 2010. 634 p. ISBN 978-1-4419-6523-3.

[7] *GLADINET: Cloud Storage Access, Sync, Team Collaboration Solutions* [online]. ©2008-2012 [cit. 2012-08-24]. Available from: http://gladinet.com/

[8] Google Cloud Storage. *Google Cloud Platform* [online]. ©2012 [cit. 2012-08-26]. Available from: http://cloud.google.com/products/cloud-storage.html

[9] Iosup, A.; Ostermann, S.; Yigitbasi, M.N.; Prodan, R.; Fahringer, T.; Epema, D.H.J. Performance Analysis of Cloud Computing Services for Many-Tasks Scientific Computing, *IEEE Transactions on Parallel and Distributed Systems*, Vol. 22, No. 6, 2011, pp. 931-945.

[10] Ježek, B.; Vaněk, J.; Procházka, M. Estimation of Response Time for Ground Ambulance Transport. *Journal of System and Management Sciences*, 2011, Vol. 1, No. 5, pp. 69-78.

[11] Kouyoumjian, V. GIS in the Cloud: The New Age of Cloud Computing and Geographic Information Systems. *Esri* [online]. June 2011. [cit. 2012-08-24]. Available from: http://www.Esri.com/library/ebooks/gis-in-the-cloud.pdf

[12] MICROSOFT. *Windows Azure: Microsoft's Cloud Platform* [online]. ©2012 [cit. 2012-08-26]. Available from: http://www.windowsazure.com/

[13] Pastore, S. Distributed computing platforms like clouds and web standards: what could be the solution in an open environment?, *Proceedings of the 10th WSEAS international conference on Applied computer and applied computational science*, Venice, 2011, pp. 195-200.

[14] Pandey, S. Cloud Computing Technology & GIS Applications. *The 8th Asian Symposium on Geographic Information Systems From Computer & Engineering View (ASGIS 2010)*, ChongQing, China, 2010, pp. 1-2.

[15] Yang, C.; Goodchild, M.; Huang, Q.; Nebert, D.; Raskin, R.; Xu, Y.; Bambacus, M.; Fay, D. Spatial Cloud Computing: How geospatial sciences could use and help to shape cloud computing. *International Journal on Digital Earth*. Vol. 4, Issue 4, 2011, pp. 305-329.