Spatial Cloud Computing and Performance Evaluation of Cloud Storage Providers

Martin Lnenicka, Jitka Komarkova, and Eva Milkova

***Abstract* -** Performance is an important factor while evaluating cloud services. The technology and architecture that deployment models and services in cloud environments offer, are also important area of research and development of geographic information systems (GIS), especially with costs and performance metrics. This paper deals with cloud computing technologies in nowadays GIS and their role in the presentation and the availability of the spatial data via internet (spatial cloud computing). The first part is focused on the role of cloud storage services in architecture of GIS. It is followed by performance and costs evaluation of selected cloud storage providers (CSPs) while using spatial data, with emphasis on measurements of response time, both for upload and download speeds.

***Keywords* -** Spatial cloud computing, cloud storage, geographic information systems, performance and costs evaluation.

I. Introduction

TECHNOLOGY continuously changes and creates new ways of managing and storing information. Cloud computing changed the way users access the Information and Communications Technologies (ICT) systems, use resources and the way of managing and delivering computing technologies, services and solutions. Users may use distributed resources (infrastructure, storage, databases, and applications) without having to deal with implementation or configuration details. Clouds promise lower costs, faster implementation, and more flexibility using mixtures of technologies, and the associated tools for achieving this.

Cloud environments offer computing resources in a pay-per-use approach, which allows businesses to use their applications, adding capacity on-demand or automating workload elasticity without any software installation and access their personal files at any computer with internet access using only an internet browser. The pay-per-use model of cloud computing is also significantly cheaper for than the pay for everything up front model of internal ICT.

M. Lnenicka is with the Institute of System Engineering and Informatics, Faculty of Economics and Administration, University of Pardubice, Studentska 95, 532 10, Pardubice, Czech Republic as a Ph.D. student (corresponding author, phone: +420-466036075; fax: +420-466036010; e-mail: martin.lnenicka@student.upce.cz).

J. Komarkova is with the Institute of System Engineering and Informatics, Faculty of Economics and Administration, University of Pardubice, Studentska 95, 532 10, Pardubice, Czech Republic (corresponding author, phone: +420-466036070; fax: +420-466036010; e-mail: jitka.komarkova@upce.cz).

E. Milkova is with the Department of Informatics, Faculty of Science, University of Hradec Kralove, Rokitanskeho 62, 500 03 Hradec Kralove, Czech Republic (e-mail: eva.milkova@uhk.cz).

Geographic Information System (GIS) is usually 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. Better performance also means lower costs, and increased margins through less waste. So, cloud computing could be a new computing, storage and delivery model for GIS.

Cloud storage is a young industry for processing and storing very large amounts of data, but it has a great promise for growth. The benefits of network based applications have already led to the change from server-attached storage to distributed storage. These networked storages helped improve storage utilization and data center efficiency. New data are generated locally by the user and then are uploaded to the cloud and synchronized. Data can be retrieved from the cloud when local data was lost.

Cloud Storage Providers (CSPs) such as Amazon, Google and Microsoft have 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. Cloud storage services differ in pricing scheme and performance characteristics, so potential customers need to identify those that can deliver the appropriate price/performance levels to meet their needs.

This case study is focused on performance comparison of CSPs, concretely how quickly CSPs can write and read spatial data stored as files of different sizes, and also on the evaluation of costs incurred. These results are also one of the indicators for decision to deploy GIS architecture into cloud. However this case study focuses more on spatial data.

II. Spatial Cloud Computing and Data Storage Services for Better Performance of Geographic Information Systems

Cloud computing refers to the hardware, systems software, and applications delivered as services over the internet. It enables the user to access computing resources anytime from anywhere, location independent resource pooling, rapid resource elasticity, usage-based pricing, transference of risk, etc.

In general, cloud providers fall into three categories [3]:

- Infrastructure as a Service (IaaS): offering web-based access to storage and computing power. The consumer does not need to manage or control the underlying cloud infrastructure but has control over the operating systems, storage, and deployed applications.
- Platform as a Service (PaaS): provides an application platform, or middleware, as a service which giving developers the tools to build and host web applications.
- Software as a Service (SaaS): applications that are accessible from various client devices through a thin client interface such as a web browser.

Most cloud service providers, at a broad level, have tariffs for the kind of elastic computing, the elastic storage, or the elastic bandwidth. The data are stored in cloud service provider's devices on multiple machines across the entire virtual layer. The data are also hosted on devices that belong to infrastructure provider. The cloud service provider needs to ensure users that the security of their data is being adequately addressed between the partners, that their virtual environments are isolated with sufficient protection, and that the cleanup of outdated images is being suitably managed at cloud infrastructure provider's storage machines. [9]

A. Benefits and Risks in Cloud Storage

Cloud storage is a service-oriented model of networked online storage where data are stored in virtualized pools of storage which are generally hosted by third parties. The typical structure of cloud storage includes storage resource pool, distributed file system, Service Level Agreements (SLA), and service interfaces, etc. Physically, the resources may be located all over the world. Cloud storage services may be accessed through a web service Application Programming Interface (API), command line tool or through a web-based user interface to create, share, and manage data. [23]

Three different network entities can be identified in representative network architecture for cloud data storage [7]:

- User – who has data to store in the cloud and rely on the cloud for data computation, consist of both individual consumers and organizations.
- CSP – who has resources and expertise in building and managing distributed cloud storage servers, owns and operates cloud computing systems.
- Third Party Auditor (TPA) – an optimal TPA, who has expertise and capabilities that users may not have. It is trusted to assess and expose risk of cloud storage services on behalf of the users upon request. Third party auditing provides a transparent yet cost-effective method for establishing trust between data owner and cloud server.

Cloud storage can be basically classified based on its use cases and software/hardware requirements for: backup, file synchronization, distributed file system and content sharing. In many cases, storage products combine multiple storage types.

Just about everything in CSPs infrastructure except block storage devices is available across all availability zones in a given region. If provider loses the entire availability in zone B, nothing happens. The application continues to operate normally, although performance level can be degraded. If provider loses availability in zone A, he will need to bring up a new load balancer in availability of zone B and promote the slave in that availability zone to the master. The system can return to operation in a few minutes with a little or no data loss. [22]

It is often easier to set up and maintain applications that take advantage of storage in the cloud than deploying an equivalent service using own hardware and software. Also elimination of the costly systems and people required to maintain them typically provides organizations significant costs savings that quite well balance the fees for cloud storage. Other benefits of cloud storage include availability (i.e., being able to access data from anywhere), backup of data, archive, cyclical peak workloads, disaster recovery, content sharing and leveraging subscriber policies across geographic distances, etc. [25]

Despite cloud computing's many benefits, it is important to be aware of the risks and failures in cloud architecture. Security and privacy are two of ICT professionals' top concerns. Typical security and privacy examples include data storage and data transfer protection, vulnerability management and remediation, personnel and physical security, application security; data privacy; and identity management. Applications are remotely deployed in a virtualized runtime environment using shared hardware/software resources, and of course hosted in a third-party infrastructure. Some of these factors change at runtime and thus cannot be fully predicted and controlled. Applications hosted on remote clouds may have also lower controllability and observability, compared with conventional in-house hosted applications. [8], [17]

Even though different CSPs offer nearly identical service commodities, customers can experience vendor lock-in: It can be prohibitively expensive for clients to switch from one provider to another. Storage providers charge clients for inbound and outbound bandwidth and requests as well as for hosting the actual data. A client who moves from one provider to another, pays for bandwidth twice, in addition to the actual costs of online storage. The resulting vendor lock-in gives storage providers leverage over clients with large amounts of data. And finally, as of yet, there are no standards to ensure interoperability or free movement between cloud providers. [1], [17]

B. Spatial Cloud Computing as a New Architecture for Geographic Information Services

Before the cloud technologies were widely available, GIS was the privilege of companies that could afford high up-front investments in powerful hardware, high maintenance costs and expensive software licenses. Now available as a service, GIS is more affordable and much more adaptable to the needs of specific users. Data stored in databases can be retrieved at real-time and processed. The results obtained after computation and analysis can also be stored for future reference. The important factor for the application users is the flexibility it provided them for accessing the GIS application service. The user may be located in different parts of the world and be able to access the service. [6], [17]

Spatial Cloud Computing (SC2) adds geography to the cloud computing paradigm and also provides dynamically scalable geo-information technology, spatial data, and geo-applications as a web service. It can be represented with a framework including physical computing infrastructure, computing resources distributed at multiple locations, and a spatial cloud computing virtual server that manages the resources to support cloud services for end users. Because the geo-technology infrastructure, the services and the data are provided; there is no large initial investment in time and costs, or ongoing maintenance. This is important because the costs of an enterprise GIS can be quite significant from a variety of factors including software licensing, applications development, data management, and ICT infrastructure. [24], [25]

When a company offers a GIS service or an access to spatial data – whether in the cloud or in a traditional data center – that company generally provides its customers with SLA that identifies key metrics (service levels) that the customer can reasonably expect from the service. The ability to understand and to fully trust the availability, reliability, and performance of the cloud is the key conceptual block for many technologists interested in moving into the cloud. For GIS applications the cloud can prove to be an approach to provide computing or storage capacity as a service, provisioned from a parallel, on-demand processing platform that leverages economies of scale to different types of users and organizations requiring GIS application services. Implementing a traditional enterprise GIS in an organization requires people with specialized skills. By providing the GIS functionality and data as a web service, SC2 eliminates the need for in-house GIS capabilities [6], [22], [24]

Fig. 1 shows a proposed model of GIS architecture in the cloud, in which spatial cloud computing middleware is the most important part. This bottom layer cloud computing infrastructure includes computing resources, storage, and network, which are managed by the middleware layer. It is a connection between cloud computing and spatial applications developed for the new cloud platform.

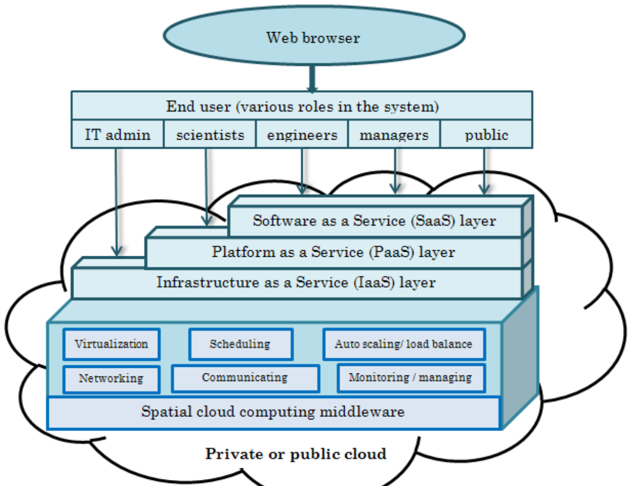

Fig. 1 Cloud based GIS architecture

C. *Performance and Costs Evaluation*

Developing, testing, and studying the performance of cloud systems is quite complex. 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 and how many users are using it at the same time).

The use of the cloud is cost-associative; a customer pays only for the computing time and resources which are equivalent to the total lease time of virtual machines. Thus, it is important to evaluate it not only from the performance perspective but also from costs to performance efficiency view, which was devised to evaluate the gain in response time in relation to the costs. [19]

The following metrics are used for the costs and performance evaluation [19]:

Response time r_j, of a job j is the time interval between the arrival and the departure of the job (in this study it means the time of transfer from default location to the cloud storage of selected CSP). Its Average Response Time (ART) is defined as [19]:

$$ART = \frac{\sum_{j=1}^{N} r_j}{N} \quad (1)$$

Where N is the total number of jobs (transfers).

The other performance metric is weighted response time, which takes into account the size of file for each parallel job (response time of selected CSP is measured simultaneously). The response time of each job r_j is weighted to its number of tasks $p(x_j)$ so Average Weighted Response Time (AWRT) is defined as follows [19]:

$$AWRT = \frac{\sum_{j=1}^{N} p(x_j) r_j}{\sum_{j=1}^{N} p(x_j)} \quad (2)$$

The costs of cloud storage usage derive from the Lease Time (LT) of storage space (usually monthly) and inbound / outbound data transfer (usually per GB based on monthly usage of cloud storage).

$$LT_{monthly} = P_U * X + P_I * Y + P_o * Z \quad (3)$$

Where P_U is price for 1 GB of data in cloud storage per month, P_I is price for inbound data transfer, P_o is price for outbound data transfer (both for GB per month); X is size of files (GB) in cloud storage, Y is size of transferred data (GB) to cloud storage and Z is size of transferred data (GB) from cloud storage.

The last performance metric is Cost-Performance Efficiency (CPE), which is evaluated by combining LT with the ART and is defined as [19]:

$$CPE = (-D_{LT}) + (-D_{ART}) \quad (4)$$

Where D_{LT} is the relative difference (%) in LT between two of CSPs and D_{RT} is the relative difference (%) between their ART.

In conclusion: In clouds, performance counts two times. Low performance means not only long waiting times, but also high costs. The advanced user, on the other hand, would also know if there is a way to optimize its application so as to reduce the costs of its run without sacrificing performance. The high-end user, who cares more for performance than for the costs, would like to know how to choose the best configuration to maximize the performance of his application. In the cloud computing environment configurations can be easily changed in order to fit the user needs. Whereas performance is crucial for content delivery, the costs structures of the different CSPs have to be examined too. [9]

III. Related Works

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. Authors in [4] 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 [5]. Authors in [14] 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 costs 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 [20]. The questions of how the standards of IT service management might change with the standardization of cloud computing is discussed in [15].

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 [16], 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 costs 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. Spatial cloud computing as a new paradigm for geographic information services is discussed in [24]. The possible applications of cloud computing as a new storage and delivery GIS model are shown e.g. in [6], [17], [21] or [26]. Telemetry data transmission, storage and processing system [13] can be given as an example of an application which provides large volumes of various data (including spatial data). During the race the application should be permanently available and it should provide enough fast response.

IV. Performance and Costs Evaluation of Cloud Storage Providers

The study is focused on evaluation of a response time, which is represented by upload/download times and speeds. The other goal is costs evaluation, which takes into account actual prices for cloud storage services. All performance and costs factors have to be considered in the initial conditions of the following case study.

A. *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

Service name	Storage and transfer	Costs	For free
Amazon S3	First 1 TB / month…Over 5000 TB / month	$0.125 per GB…$0.055 per GB	5 GB of 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 costs.

Service name	Storage and transfer	Costs	For free
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 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	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.

This is not a complete list, however it does provide a good representative overview of CSPs. Customers can usually choose between the geographic redundancy service which replicates data between two geographically distant sites, so 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 storage (it is cheaper).

B. 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 in this study is free access), it is followed by the performance evaluation of cloud storage while using spatial data, costs evaluation 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 costs and performance evaluation

	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 the performance evaluation 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) [11]. From Table 1 there were chosen only Amazon S3 [2], Windows Azure [18] and Google Cloud Storage [12] for the evaluation because OpSource Cloud Files and GoGrid are not supported by Gladinet.

The collected data were then statistically analyzed using Microsoft Excel 2010.

C. Comparison of Response Times

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 [10] was won by Windows Azure Storage – see Table 3. These results will be later compared with our own performance evaluation.

Table 3 Results of download speed test by CloudHarmony

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 speeds were measured with various spatial data files of different sizes for all three 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 (parallel jobs) 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. There are also highlighted times with the slowest upload and download speed for each CSP.

Four performance metrics from these measurements were collected – average upload speed, upload time, average download speed and download time.

Table 4 Speed results and times of upload and download – file size of 3,57 MB

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 and time	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 and time	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

Table 5 Speed results and times of upload and download – file size of 120 kB

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 and time	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 and time	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. 2 and 3 then show the progress of upload and download speeds (both with the spatial data file of 3,57 MB) between 10:00 – 16:00. Whereas 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 CSP. Another finding is that the upload speed from location 2 to Amazon cloud storage is significantly slower compared to Google and Windows – see Fig. 2.

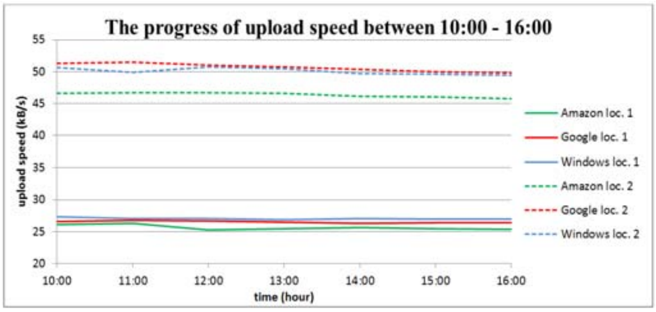

Fig. 2 The progress of upload speed file size of 3,57 MB between 10:00 – 16:00

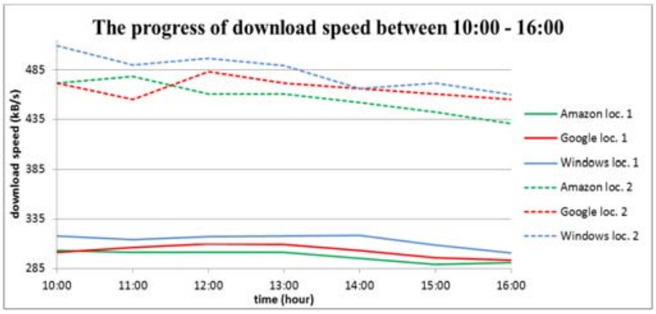

Fig. 3 The progress of download speed file size of 3,57 MB between 10:00 – 16:00

Fig. 4 and 5 then show the progress of upload and download speed (both with the spatial data file of 120 kB) between 10:00 – 16:00. The best upload speed, which is comparatively more consistent, can be expected from Windows Azure.

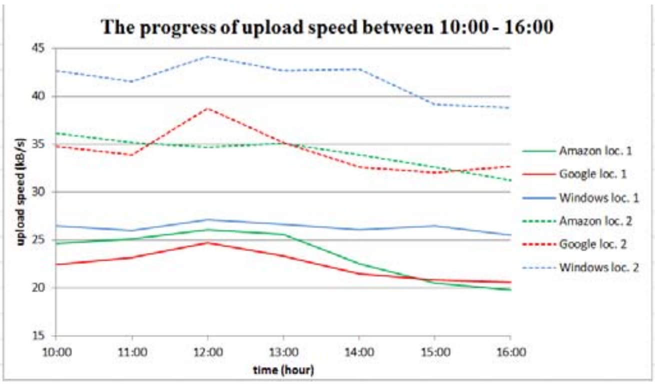

Fig. 4 The progress of upload speed file size of 120 kB between 10:00 – 16:00

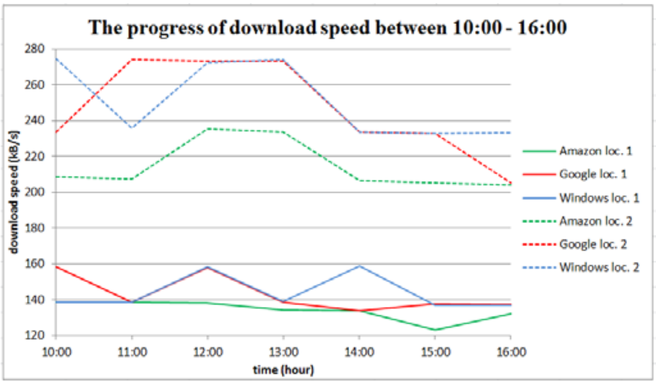

Fig. 5 The progress of download speed file size of 120 kB between 10:00 – 16:00

The following Fig. 6 shows results of ART for upload and download times of spatial data files. 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.

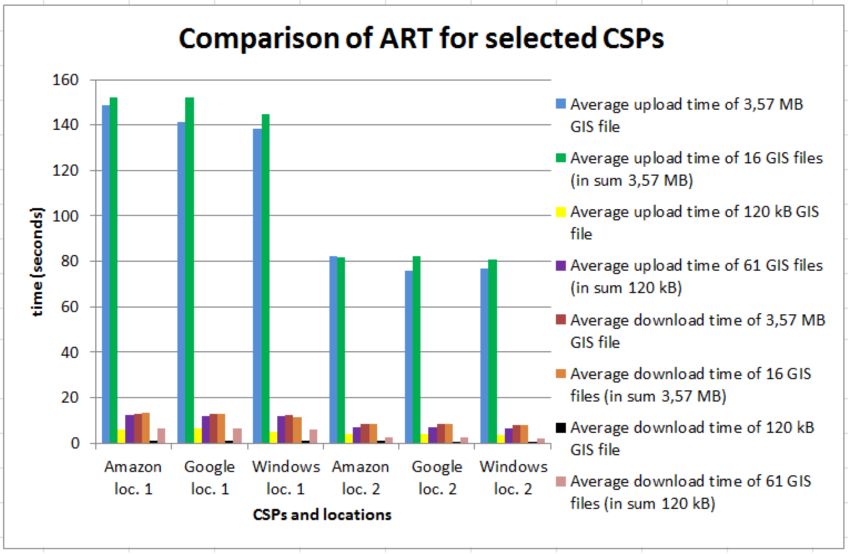

Fig. 6 Comparison of average upload and download times

A. *Costs Evaluation*

Since the use of cloud is costs associative, the system has to maintain a good analogy between response time (in this study it is represented by upload/download times and speeds) and costs (for the use of cloud storage). Transparent measurement and billing will increase the trust level of users towards cloud services. Pay-as-you-go subscription or pay-as-you-consume model of measuring and billing are popular for cloud services. The service gets the status of the SLA, and invokes the credit service, which debits the user credit card or account and informs the user.

There are many pricing strategies such as Random-Access Memory (RAM) hours, Central Processing Unit (CPU) hours and capacity, bandwidth (inbound/outbound data transfer), storage space (GB of data), software license and fees, and subscription-based pricing. [3]

This costs evaluation is focused only on capacity, inbound and outbound data transfer. Costs are calculated for model (monthly) with:

- Size of data in cloud storage – 50 GB,
- Inbound data transfer – 10 GB,
- Outbound data transfer – 30 GB.

Table 6 Actual prices for selected CSPs

Service name	Capacity (GB)	In (GB)	Out (GB)
Amazon S3	From 0 to 1 TB for $0.125 monthly.	free	First 1 GB per month free, otherwise from $0.12 per GB.
Google Storage	From 0 to 1 TB for $0.12 monthly.	free	From 0 to 1 TB for $0.12 per GB.
Windows Azure	From 0 to 1 TB for $0.125 monthly.	free	First 5 GB per month free, otherwise from $0.12 per GB.

The CPE was calculated by using (4) for both of upload and download times with file size of 3,57 MB, measured from both locations together. ART was calculated using (1). LT was calculated by using (3) and was same for upload and download. The results are shown in Table 7 and Table 8. Based on CPE results is Windows Azure 13,6 % (16,8 % in download time) ahead of Amazon S3 and 9,1 % (13,5 % in download time) ahead of Google Storage.

Table 7 ART and LT for selected CSPs

		Amazon S3	Google Storage	Windows Azure
Up-load	ART	116,2 s	112,8 s	110,1 s
	LT	$9,73	$9,6	$9
Down-load	ART	10,7 s	10,5 s	9,9 s
	LT	$9,73	$9,6	$9

Table 8 CPE for selected CSPs

CPE (upload / download)	Amazon S3	Google Storage	Windows Azure
Amazon S3	1	- 4,4 %	- 13,6 %
		- 3,2 %	- 16,8 %
Google Storage	4,4 %	1	- 9,1 %
	3,2 %		- 13,5 %
Windows Azure	13,6 %	9,1 %	1
	16,8 %	13,5 %	

V. Conclusion

Spatial cloud computing could help increase performance and cost effectiveness of GIS, as well as utilize available distributed computing resources, especially data storage. Users then pay only a small amount of money for GIS services in cloud systems they access. For example Esri (producer of ArcGIS) offers a variety of cloud-based applications and services in partnership with Amazon.

This paper describes only one of the steps of the way of solving the problem of spatial cloud computing and also the role of performance and costs metrics in cloud storage. In order to obtain comparable results, it is necessary to test cloud storages from multiple locations across the world, also with different initial conditions.

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

Users do not like to wait, time is money. Thus, speed and system response time are very important for them, as well as costs. 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 and also customers, who may have access to cloud storage. According to the case study, the upload speed is comparatively low for all tested solutions. For the spatial data, there is one more promising alternative – to download spatial data from multiple locations.

References

[1] H. Abu-Libdeh, L. Princehouse, H. Weatherspoon. RACS: A Case for Cloud Storage Diversity, *SoCC '10 Proceedings of the 1st ACM symposium on Cloud computing*, 2010, pp. 229-240.

[2] Amazon Simple Storage Service (Amazon S3). *Amazon Web Services* [online]. ©2012 [cit. 2012-11-07]. Available: http://aws.amazon.com/s3/

[3] N. Antonopoulos, L. Gillam. *Cloud Computing: Principles, Systems and Applications*. London: Springer, 1 ed., 2010. 379 p. ISBN 978-1-84996-240-7.

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

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

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

[7] S. Biruntha, V. Venkatesa Kumar, S. Palaniswami. Enabling Data Storage Security in Cloud Computing for Banking Enterprise. *ICNVS'10: Proceedings of the 12th international conference on Networking, VLSI and signal processing*, Cambridge, 2010, pp. 217-224.

[8] J. D. Blower. GIS in the cloud: Implementing a web map service on Google App Engine. *Proceedings of the 1st International Conference and Exhibition on Computing for Geospatial Research and Application*, New York, NY, USA, ACM, Article 34.

[9] R. Buyya, J. Broberg, A. Gościński. *Cloud computing: Principles and Paradigms*. Hoboken, N.J.: Wiley, 1 ed., 2011, 637 p. ISBN 978-0-470-88799-8.

[10] Cloud Speed Test. *CloudHarmony: transparency for the cloud* [online]. ©2011 [cit. 2012-11-07]. Available: http://cloudharmony.com/speedtest

[11] *GLADINET: Cloud Storage Access, Sync, Team Collaboration Solutions* [online]. ©2008-2012 [cit. 2012-11-07]. Available: http://gladinet.com/

[12] Google Cloud Storage. *Google Cloud Platform* [online]. ©2012 [cit. 2012-11-07]. Available: http://cloud.google.com/products/cloud-storage.html

[13] O. Horak. Telemetry Data Transmission, Storage and Processing System Supporting Ski-mountaineering Race. WSEAS TRANSACTIONS on COMMUNICATIONS, 2008, Vol 7., No. 9, pp. 964-973.

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

[15] M. Jansen. Will Cloud Computing Change Standards in IT-Service Management. *International Journal of Computers and Communications*. 2012, Vol. 6, No. 1, pp. 9-16.

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

[17] V. Kouyoumjian. GIS in the Cloud: The New Age of Cloud Computing and Geographic Information Systems. *Esri* [online]. June 2011. [cit. 2012-11-07]. Available: http://www.Esri.com/library/ebooks/gis-in-the-cloud.pdf

[18] MICROSOFT. *Windows Azure: Microsoft's Cloud Platform* [online]. ©2012 [cit. 2012-11-07]. Available: http://www.windowsazure.com/

[19] I. A. Moschakis, H. D. Karatza. Evaluation of gang scheduling performance and cost in a cloud computing system, *The Journal of Supercomputing*, 2012, Vol. 59, No. 2, pp. 975-992.

[20] S. Pastore. 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.
[21] S. Pandey. 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.
[22] G. Reese. *Cloud Application Architectures: Building Applications and Infrastructure in the Cloud*. Sebastopol, CA: O'Reilly Media, 1 ed., 2009. 189 p. ISBN 978-0-596-15636-7.
[23] C. Wang, K. Ren, W. Lou, J. Li. Toward Publicly Auditable Secure Cloud Data Storage Services. *IEEE Network Magazine*, Vol. 24, No. 4, 2010, pp. 19-24.
[24] H. Williams. Spatial Cloud Computing (SC2) - White Paper: A New Paradigm for Geographic Information Services. *SKE: Home of GeoPortal and Spatial Cloud Computing* [online]. August 2012. [cit. 2012-11-07]. Available from: http://www.skeinc.com/pages/Downloads/SC2_White_Paper_August_2012.pdf
[25] J. Wu, L. Ping, X. Ge, Y. Wang, J. Fu. Cloud Storage as the Infrastructure of Cloud Computing. *ICICCI '10: Proceedings of the 2010 International Conference on Intelligent Computing and Cognitive Informatics*, 2010, pp. 380-383.
[26] C. Yang, M. Goodchild, Q. Huang, D. Nebert, R. Raskin, Y. Xu, M. Bambacus, D. Fay. Spatial Cloud Computing: How geospatial sciences could use and help to shape cloud computing. *International Journal on Digital Earth,* 2011, Vol. 4, No. 4, pp. 305-329.