PAY-PER-USE REVENUE MODELS FOR GEOPROCESSING SERVICES IN THE CLOUD

B. Baranski [a,*], T. Deelmann [a], B. Schäffer [b]

[a] Institute for Geoinformatics, University of Münster, Weseler Straße 253, 48151 Münster, Germany – {baranski, th.deelmann}@uni-muenster.de
[b] 52°North, Martin-Luther-King-Weg 24, 48155 Münster, Germany - schaeffer@52north.org

KEY WORDS: Cloud Computing, Geoprocessing, Web Processing Service, Business Models, Service Level Agreements

ABSTRACT:

Cloud Computing is one of the latest trends in the mainstream IT world and hints at a future in which the storage of data and the hosting of applications are no longer performed on local computers, but on distributed third-party facilities. From a provider perspective, Cloud Computing enables companies to significantly increase their hardware utilization rate and allows external customers to use the company's infrastructure on pay-per-use revenue models. From a client perspective, it enables the on-demand allocation of sufficient resources to solve complex computational problems or to scale all kinds of applications. Geographic Information Systems (GIS) have been under constant change in recent years. Emerging web service technologies stimulated the evolution from desktop- and data-centric GIS to distributed and loosely coupled web services merged into the Spatial Data Infrastructure (SDI) concept. The processing of data is an essential part of daily work of GIS experts. With increasing amount of available data and higher requirements on adopted algorithms, the data processing part in SDIs becomes crucial in terms of practicability and performance. Cloud Computing offers a technical opportunity for the on-demand provisioning of sufficient resources for computing-intensive algorithms and an economic opportunity to support future business models. This paper presents an OGC Web Processing Service (WPS) implementation that uses Cloud Computing resources to perform geoprocessing tasks. The resource usage is monitored and the accumulated usage costs are charged to the service consumer account. The implementation will be demonstrated in a real world use case that is based on governmental forest management.

1. INTRODUCTION

Geographic Information Systems (GIS) have been under constant development during the last decade. Emerging web services technologies – such as the SOA paradigm - promote the development from classical desktop- and data-centric GIS to distributed and loosely-coupled architectures composed of open and interoperable web services merged into the Spatial Data Infrastructure (SDI) concept (McLaughlin et al. 2000). In the past, open standards based SDIs - for instance based on standards developed by the Open Geospatial Consortium (OGC) - focused on the retrieval, portrayal and processing of geospatial data through web services (Kiehle et al. 2006). They have shown a great potential for enabling the market value of geospatial data as for instance presented in (Fornefeld et al. 2004). Current SDI development faces different challenges as for instance an increasing amount of available data due to advanced data acquiring technologies and new emerging laws - as for instance the Infrastructure for Spatial Information in the European Community (INSPIRE) directive – and corresponding Quality of Service (QoS) requirements. With such an increasing amount of available data and higher requirements on adopted algorithms, the data processing part in SDIs becomes crucial in terms of practicability and performance.

Cloud Computing - as one of the latest trends in the mainstream IT world - offers a technical opportunity for the on-demand provisioning of sufficient resources for computing-intensive algorithms. The allocation of external cloud resources (e.g. high capacity storage or computing power) could be done nearly in real-time. That gives service providers the option for handling huge amount of data and peak load very efficiently without operating their own datacenters and also without heavily investing in infrastructure in advance. Furthermore, Cloud Computing offers an economic opportunity by means of pay-per-use revenue models. That gives service providers the option for acting as value added service providers. While the processing part in SDIs has already been tackled (Schäffer et al. 2009) (Friis-Christensen et al. 2007) (Weiser and Zipf 2007), the utilization of Cloud Computing and related methods and technologies has not been regarded in the broader context of geoprocessing yet.

The overall aim of this document is to develop, implement and demonstrate a concept for realizing pay-per-use revenue models for geoprocessing services. The following major research questions will be addressed by this document. How can the WPS specification and corresponding implementations be extended or integrated in a new licensing system in order to settle service consumer's account for resource usage (e.g. storage or CPU cycles)? How can such a licensing and settlement system be realized by means of open and standardized interfaces and encodings? How can such a licensing and settlement system be realized, so that any existing WPS client implementation could use the processing capabilities without being changed?

The reminder of this paper is structured as followed. Chapter 2 gives a brief introduction into the basic concepts and related work. Chapter 3 presents the requirements and the proposed approach for realizing pay-per-use revenue models for geoprocessing services in the cloud. Chapter 4 describes the use case and the proof-of-concept implementation. Chapter 5 concludes the findings of the presented work and gives an outlook about remaining open issues for future research.

2. BACKGROUND

This Chapter gives a brief introduction into the basic concepts and related work.

2.1 Web Processing Service

In the context of web-based geoprocessing, the OGC Web Processing Service (WPS) interface specification (OGC 2007) evolved as the de facto standard. It provides a straight-forward approach to publish and execute geoprocesses over the web. According to the WPS interface specification, a geoprocess is defined as any calculation operating on spatially referenced data. Such a process can range from a simple geometric calculation (e.g. a simple intersect operation) to a complex simulation process (e.g. a climate change model).

In detail, the WPS interface specification describes three operations, which are all handled in a stateless manner. The GetCapabilities operation is common to any type of OGC Web Service and returns the service metadata. In case of WPS, it also returns a brief description of the processes offered by the specific service instance. To get more information about the hosted processes, the WPS provides process metadata through the DescribeProcess operation. This operation describes all parameters, which are required to run a process. Based on this information the client can perform the Execute operation upon the designated process. As every OGC Web Service, the WPS communicates through HTTP- GET and HTTP-POST based on an OGC-specific XML-message encoding. However, the WPS interface specification does not describe any aspect regarding licensing as it is designed in this document.

2.2 Cloud Computing

Cloud Computing is one of the latest trends in the mainstream IT world (Gartner 2008) (Gartner 2009). The used cloud metaphor represents the internet or other large networking infrastructures. The paradigm behind the Cloud Computing buzzword hints at a future in which the storage of data and computations are no longer performed on local computers, but on distributed facilities operated by third-party storage and computational utilities (Foster 2008). The key characteristic of a cloud is the ability to scale and provide computational power and storage dynamically in a cost efficient and secure way over the internet. By outtasking application and data to computational facilities operated by third parties, clients do not need to operate their own large-scale computational infrastructure anymore. Thereby, existing fixed costs can be transformed into variable costs and create a business advantage. Furthermore, cloud resources (e.g. storage or computational power) could be allocated nearly in real-time and applications are able to scale automatically on-demand (e.g. in case of high amounts of requests). This allows cloud users to handle peak loads very efficiently without managing their own infrastructures.

In essence, Cloud Computing is not a completely new concept. It moreover collects a family of well known and established methods and technologies under the umbrella of the term Cloud Computing - for instance Software as a Service (SaaS) as a model for software deployment and virtualization as an efficient hosting platform (Sun Microsystems Inc. 2009). Besides, it describes a paradigm of outsourcing applications and specific tasks to a scalable infrastructure and therefore consequently enabling new pay-per-use business models with less upfront investments.

2.3 Related Work

To speed up the processing of large amounts of data and perform complex calculations, the use of Grid Computing or related methods and technologies are a good choice for achieving high calculation performance. Although the application of Grid Computing is not novel to the mainstream IT-world (Foster et al. 1998), in the context of geospatial applications and OGC Web Services (OWS) only little research has been conducted; see (Baranski 2008), (Lanig et al. 2008), (Padberg et al. 2009) and (OGC 2009a).

Cloud Computing has not been regarded in the broader context of SDIs yet. In (Schäffer et al. 2010a) a merger between SDIs and Cloud Computing infrastructures was evaluated on a conceptual level. In (Baranski et al. 2009) the scalability aspect for a cloud enabled WPS implementation is challenged and proven exemplary in the Google Cloud (Google App Engine). In (Kim et al. 2009) the Amazon Elastic Compute Cloud (EC2) was used to provide the high level of performance for impact assessment studies of climate change that require considerable amount of data.

The licensing aspect of geospatial services is partially covered by the OGC Geospatial Digital Rights Management (GeoDRM) Reference Model (OGC 2006). In (Schaeffer et al. 2010b) a security enabled architecture is presented, in which geoprocessing services can be enhanced in order to support ad-hoc licensing, without any prior offline negotiated licenses being necessary between service consumer and service provider.

3. CONCEPT

A classification of potential license models and an abstract architecture for realizing pay-per-use revenue models for geoprocessing services is introduced in this chapter.

3.1 License Model

In the presented scenario for pay-per-use revenue models, the service provider offers specific geoprocessing functionality to the service consumer by hosting WPS instances (on-demand) in the cloud. The geoprocessing service consumer has to pay the service provider for the geoprocessing functionality usage. Furthermore, the geoprocessing service provider has to pay the cloud infrastructure provider for the resource usage. Therefore, the geoprocessing service provider acts as a value added service provider and forwards the resource usage costs (plus additional business costs) to the service consumer.

Two different *payment models* could be identified. Firstly, the geoprocessing service consumer has to pay the geoprocessing service provider for the complete hosting of a WPS instance with a specific set of offered processes (flatrate). The usage costs for instance depend in that scenario on the booked up-time of the WPS instance. Secondly, the geoprocessing service consumer has to pay the geoprocessing provider for each single execution of a specific process (pay-per-use). The usage costs for instance depend in that scenario on the executed process and the amount of input data to be processed; or for instance on the plain used CPU time for the single process execution.

To ensure that service providers fulfil promised Quality of Service (QoS) guarantees (such as geoprocessing service availability or computational resource availability), normally a formal contract between service consumers and service providers - a Service Level Agreement (SLA) - is concluded prior to the service consumption (Lee 2002). Two different *contract models* could be identified for such a license negotiation process. Firstly, so-called “click through” licenses are concluded prior to each single service consumption (short term contract). In that scenario, the geoprocessing service consumer has to read the terms of the license prior to each single process execution. After accepting them, he could execute a specific process. Secondly, the service consumer concludes a usage license for a specific period of time in advance (long term contract). In that scenario, the geoprocessing service consumer is able to execute processes und the terms of a previously negotiated license without having to accept the terms of license prior each single process execution.

Furthermore, two different *accounting models* could be identified independent of the payment models and independent of the contract models. Firstly, the service consumer account is charged after the geoprocessing functionality consumption is finished (postpaid). The service consumer account is charged for instance directly after a process execution is finished or for instance on a monthly basis. Secondly, the service consumers’ account is charged in advanced (prepaid). A differentiation can be made between an unlimited (for instance no usage costs limit per month) or limited (for instance a monthly based upper limit of usage costs) prepaid-scenario.

Generally, the geoprocessing service provider is free to design the *pricing models* for his service offering. He could offer different QoS level to different users for different prices. On the one hand he could for instance create a kind of a ‘gold license’ (asking a high price) in which he guarantees high computational power and immediate process execution. On the other hand he could for instance create a kind of a ‘silver license’ (asking a lower price compared to the ‘gold license’) in which he guarantees medium computational power and only a queued process execution (e.g. with a guaranteed process execution start time within a specific time frame).

3.2 Publish-Find-Bind

The overall activity in SDIs is the retrieval, portrayal and processing of geospatial data through interoperable web services above organizational boundaries. Most SDIs implement the popular publish-find-bind pattern that identifies three roles that are involved in the process of service consumption: the service provider, the service broker and the service consumer (Massuthe, 2005). From this simple point of view, the SDI service provider hosts the geospatial data in his own organizational boundaries and operates the (geoprocessing) services on his own computing facilities. But if we consider that the data is mostly collected by other parties than the service provider, we can identify the data provider as an additional role in SDIs. Furthermore, the general trend in the mainstream IT (emerged in the Cloud Computing paradigm) hints into a future where services and data are hosted on third party infrastructures. Therefore, we can identify a resource provider as an additional important role in SDIs.

Since the service consumption now should be performed under the terms of a previously negotiated license, the basic publish-find-bind pattern now must be enhanced to with an 'agree' phase in which both parties – the (geoprocessing) service consumer and the (geoprocessing) service provider – agree to certain guaranteed service qualities and corresponding service usage costs. The publish-find-agree-bind pattern presented in (INSPIRE 2008) and (OGC 2009b) covers rights management in SDIs with a focus on restricting the access to SDI services. Aspects like authentication, authorization and pricing are covered, but service quality aspects are missing and a role differentiation on the service consumer’s side is not available.

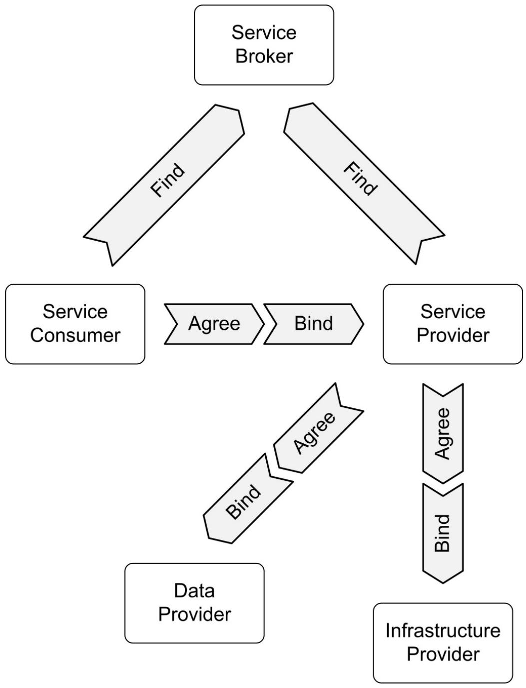

Figure 1. The popular publish-find-bind pattern is extended with an additional agreement phases

3.3 Architecture

The presented architecture for realizing pay-per-use revenue models for geoprocessing services is based on a common policy-based XACML security architecture (OASIS 2000) and it incorporates findings resulting from research conducted in last OGC testbeds and a modified version of the OGC DRM reference model (OGC 2007b) for the incorporation of geoprocessing services. In (Schaeffer et al. 2010b) it is studied how standard geoprocessing services can be enhanced in order to support ad-hoc license agreements directly in-process, without any prior offline negotiated agreements being necessary between geoprocessing provider and geoprocessing user. The presented abstract architecture extends the security-focused policy-based architecture proofed during last OGC testbeds in such a way to support pay-per-use revenue models for geoprocessing services.

Three different domains could be identified in the context of geoprocessing services – the service consumer domain, the service provider domain and the infrastructure provider.

The service provider domain consists of a *Service Provider* who offers specific geoprocessing functionality via the WPS interface. A specific WPS instance is hidden behind *License*

Proxy, which acts as a Policy Enforcement Point (PEP) in the proposed policy-based architecture. The license proxy is an OGC-compliant proxy component which enriches the original geoprocessing service metadata with precondition elements. In general, preconditions publicly announce a potential service consumer, which conditions (in our scenario licenses) must be fulfilled before service consumption. This concept ensures interoperability by allowing services to fulfil all required preconditions prior to the service invocation. Furthermore, the license proxy decides - by utilizing a Policy Decision Point (PDP) - whether the service consumer is allowed to access the geoprocessing service (or specific processes) under the terms of the previously concluded license.

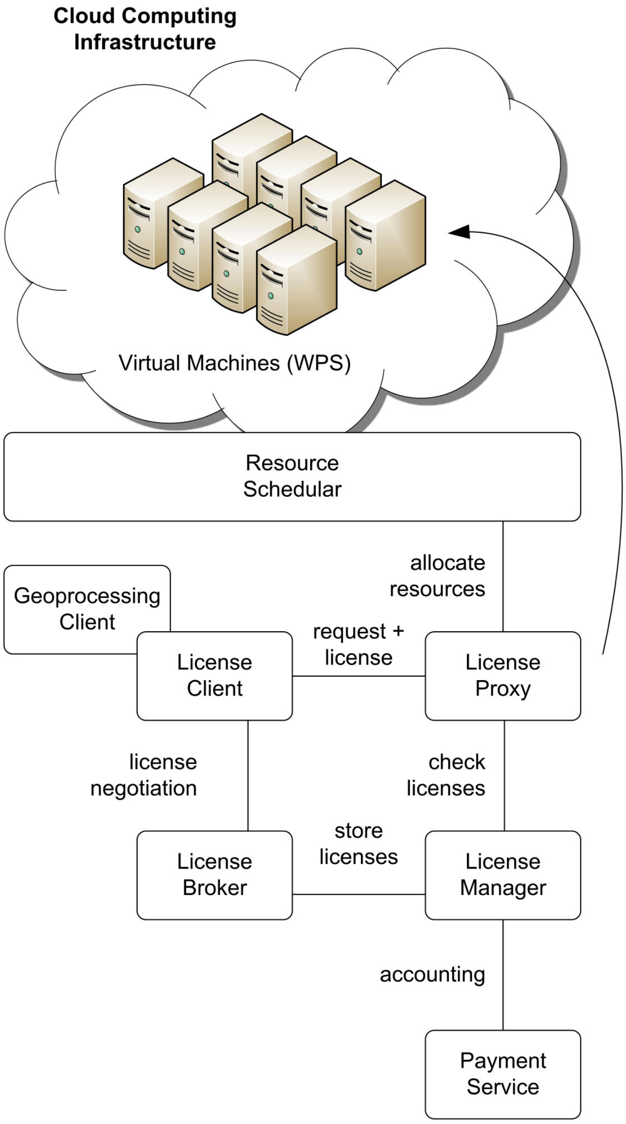

Figure 2. The required components for realizing pay-per-use revenue models for geoprocessing services

The service consumer domain consists of the *Service Consumer* who interacts with a *Geoprocessing Client* to access the *Geoprocessing Service*. Since the geoprocessing service is hidden behind a proxy component (which restricts the service access to service consumers who previously concluded a license), the service consumer domain contains also a (web- or desktop-based) *License Client* which enriches the original geoprocessing client with license negotiation capabilities. A *License Broker* provides license templates for each geoprocessing service and accepts (or rejects) licenses from the service consumer (the negotiation process is realized through the proposed license client). If a license offer is accepted, the license broker stores the license at a *License Manager*. The license manager allows the service consumer and the service provider to check the status of concluded licenses.

The *Infrastructure Provider* (in our case a Cloud Computing provider such as Amazon) offers potentially unlimited computational resources (e.g. virtual machines). These computational resources could be allocated on-demand by the geoprocessing service provider to handle the geoprocessing tasks and typically the service provider has to pay only for real resource usage (no fixed costs). The service provider needs a *Resource Scheduler*, which manages the required computational resources. The resource scheduler automatically allocates new resources (e.g. in times of high request rates) and frees previously allocated resources (e.g in idle times) at the infrastructure provider. On the one hand, the resource scheduler has to allocate sufficient resources at the infrastructure provider to realize an adequate Quality of Service (QoS) for the service consumer (for example a service availability or a maximum computation time). On the other hand, the resource scheduler has to minimize the accumulated infrastructure usage costs for the service provider.

In addition to the license enforcement functionality, the license proxy component (which acts as a PEP in the policy-based architecture) now utilizes the resource scheduler in order to allocate sufficient computational resources at the infrastructure provider according to the previously concluded license of the geoprocessing service consumer.

Furthermore, a *Payment Service* is required to accumulate the usage costs of the geoprocessing service consumers according to the concluded licenses and according to the delivered QoS. In some cases, the geoprocessing service provider has to pay some license penalty fees if he failed to deliver the promised service quality goals (if such penalty fees are part of the previously negotiated contract). The payment service also realizes or initiates the accounting (the charging of the geoprocessing service consumer account).

4. DEMONSTRATION

Geographic Information Systems (GIS) and their functionalities already play an important role in the context of forest management. The focus here is typically on inventory and planning, as well as in the production of detailed forest maps. For the forest managers various issues arise. Which areas of the forest belong to which forestry category? What is the shortest path from an arbitrary place in the forest to the next main road? What about the ownership of specific areas? What about the protection of specific areas against harvesting? The following example describes a real-world scenario which benefits largely from Cloud Computing. It illustrates the presented ideas and acts as a proof-of-concept.

To ensure the efficient transport of harvested wood, a forest manager wants to calculate the shortest and safest route from the harvesting area through the forest to the nearest main street. A fully functional desktop-based GIS seems to be too oversized (as well as too expensive) for this single purpose. Furthermore, specifically developed expert software for the forestry economy typically contains too much functionality and such software also seems to be too expensive in acquisition and maintenance since it is not in daily use. An OGC-compliant WPS offering route planning processes for forested terrain that could be booked and used on-demand based on pay-per-use payment model seems to be an appropriate and cost-efficient tool for forest managers in such situations.

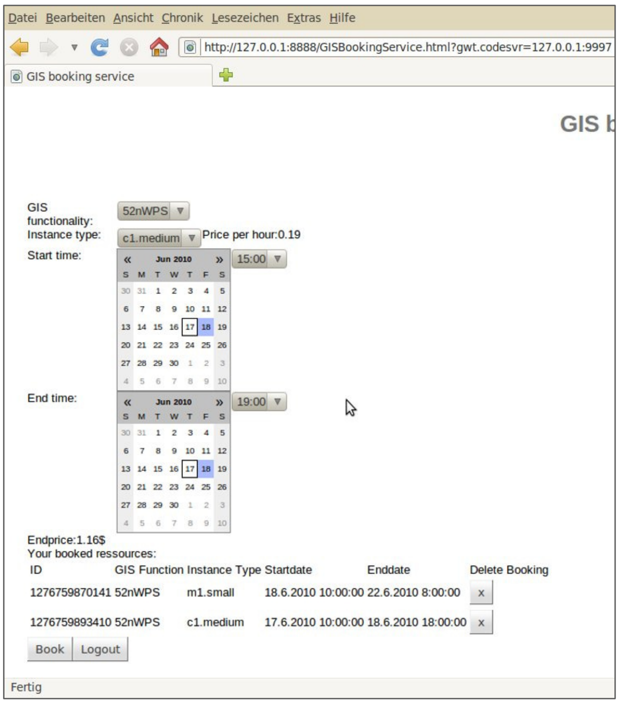

Figure 3. The web-based license client enables users to book a geoprocessing service instances for a specific period of time

The proof-of-concept implementation is based on the open source 52°North WPS implementation and the security and licensing components developed at the 52°North Security & GeoRM Community. The security components generally enable business and access control process for geospatial services and spatial data infrastructures. They have been extended in order to support the allocation of resources in the Amazon Cloud (Amazon EC2) and to initiate the accounting process. The proof-of-concept implementation utilized multiple mainstream IT standards. The OASIS eXtensible Access Control Markup Language (XACML) describes a basic security architecture, protocol and license encoding, which serves as the foundation for the presented concepts. This was combined with OASIS' Security Assertion Markup Language as a common encoding for security tokens. On the message level, the OASIS WS-Security (WSS) describes a secure message exchange that is used on a technical level to transport security and license tokens, encrypt and sign message. This specification is accompanied by WS-Trust for managing trust and WS-Policy for the encoding of preconditions.

5. CONCLUSION

This Chapter concludes the findings of the presented work and gives an outlook about still remaining open issues for future research.

This document describes an abstract architecture for the seamless realization of pay-per-use revenue models for geoprocessing services. Different licensing and accounting models are described and could be realized with the proposed abstract architecture. The presented abstract architecture covers the complete license negotiation process between geoprocessing service consumers and geoprocessing service providers. Geoprocessing service consumers are now able to find adequate geoprocessing service providers according to functional (for instance offered processes) or non-functional (for instance offered QoS guarantees and corresponding pricing models) service requirements. With the presented architecture, geoprocessing service providers are now able to manage the underlying infrastructure in an efficient manner by utilizing external Cloud Computing resources. In addition, Geoprocessing service providers are now easily able to act as value added service providers by forwarding their infrastructure costs directly to the service consumers. Even though the presented architecture is mostly technology independent (an abstract architecture containing a description of required components and communication), an implementation could be realized with common web services technologies (for instance XACML as a license encoding and WS-Policy for promoting the license preconditions). The presented architecture is not only limited to geoprocessing services but could also be extended (by modifying the license encoding and the PEP mechanisms) in such a way to support other geospatial services (for instance for data delivery).

The commercial dimension of the presented approach presents a basis for sustainable SDI business models. The billion Euro market of classical GIS desktop software can thereby be merged with SDI concepts. By incorporating Cloud Computing technologies, we move towards a world where geoprocessing functionality is ubiquitously available like electrical power from the power outlet. Further research needs to be conducted on this topic. Especially the merger of public and private clouds towards hybrid clouds is interesting in this context. Questions like "When does it make sense to go into a public cloud" needs to be answered. Selective outsourcing in such hybrid clouds, where only less sensitive georesources is transferred to public clouds are also of high interest.

REFERENCES

Baranski, B. (2008). Grid Computing Enabled Web Processing Service. In E. Pebesma, M. Bishr, & T. Bartoschek (Eds.), Proceedings of the 6th Geographic Information Days, IfGI prints (Vol. 32, pp. 243-256). Presented at the GI-days 2008, Muenster, Germany: Institute for Geoinformatics.

Baranski, B., Schäffer, B. and Redweik, R. (2009). Geoprocessing in the Clouds. Presented at Free and Open Source Software for Geospatial Conference. Sydney, Australia

Fornefeld, M., Oefinger, P. and Rausch, U. (2003). The market for geospatial information: potentials for employment, innovation and value addes. MICUS Management Consulting GmbH. Düsseldorf, Germany

Foster, I, Kesselman,C (1998) The Grid: A Blueprint for a New Computing Infrastructure. Morgan Kaufmann Publishers.

Foster, I., Zhao, Y., Raicu, I., and Lu, S. (2008). Cloud computing and grid computing 360-degree compared. In 2008 Grid Computing Environments Workshop. IEEE

Friis-Christensen A, Ostlander N, Lutz M, Bernard L (2007) Designing Service Architectures for Distributed Geoprocessing: Challenges and Future Directions Transactions in GIS, Blackwell Publishing, 11: 799-818

Gartner (2008) Gartner Says Cloud Computing Will Be As Influential As E-business. Gartner Press Release. [Online] Available http://www.gartner.com/it/page.jsp?id=707508

Gartner (2009a) Gartner Says Cloud Application Infrastructure Technologies Need Seven Years To Mature. Gartner Press Release. [Online]. Available: http://www.gartner.com/it/page.jsp?id=871113

Hafner, M., and Breu, R. (2009). Security engineering for service-oriented architectures. Springer. Berlin, Germany

INSPIRE (2008). Network Services Architecture, Version 3.0. Network Services Drafting Team

Kanneganti, R., and Chodavarapu, P. (2008). SOA security. Greenwich, Conn.: Manning

Kiehle, C., Greve, K. and Heier, C. (2006). Standardized Geoprocessing – Taking spatial data infrastructures one step further. Proceedings of 9th AGILE International Conference on Geographic Information Science

Kwang Soo Kim, Doug MacKenzie (2009). Download files for Use of Cloud computing in impact assessment of climate change. Geoprocessing in the Clouds. Presented at Free and Open Source Software for Geospatial Conference. Sydney, Australia

Lanig, S., Schilling, A., Stollberg, B., & Zipf, A. (2008). Towards Standards-based Processing of Digital Elevation Models for Grid Computing through Web Processing Service (WPS). In ICCSA, Lecture Notes in Computer Science (Vol. 5073, pp. 191-203). Presented at the Computational Science and Its Applications - ICCSA 2008, Perugia, Italy: Springer Verlag.

Lee, J. and Ben-Natan, R. (2002). Integrating Service Level Agreements: Optimizing Your OSS for SLA Delivery. John Wiley & Sons, Inc., New York, NY, USA

Massuthe, P., Reisig, W. and Schmidt, K. (2005). An Operating Guideline Approach to the SOA. Annals of Mathematics, Computing & Teleinformatics

McLaughlin, J. and Groot, R. (2000). Geospatial data infrastructure: concepts, cases and good practice. Oxford: University Press

OASIS 2000. eXtensible Access Control Markup Language (XACML) Available: http://docs.oasis-open.org/xacml/2.0/XACML-2.0-OS-NORMATIVE.zip

OGC (2006). GeoRM ReferenceModel. OGC 06-004r3

OGC (2007). OpenGIS Web Processing Service (1.0.0). OGC 05-007r7.

OGC (2009a). OWS-6 WPS Grid Processing Profile Engineering Report. OGC 09-041r3

OGC (2009b). OpenGIS GeoRM Role Model. OGC 09-123

Padberg, A. and Kiehle, C. (2009): Spatial Data Infrastructures and Grid Computing: the GDI-Grid project. Geophysical Research Abstracts 11 (EGU2009-4242)

Sun Microsystems Inc. (2009) Cloud Computing at a higher level. [Online] Available: https://slx.sun.com/files/Cloud_Computing_Brochure_2009.pdf

Schäffer B, Baranski B, Foerster T, Brauner J (2009) A Service-Oriented Framework for Realtime and Distributed Geoprocessing, International Opensource Geospatial Research Symposium, 2009

Schaeffer, B., Baranski, B. and Foerster, T. (2010a). Towards Spatial Data Infrastructures in the Clouds. In Proceedings of 13th AGILE International Conference on Geographic Information Science, Guimarães, Portugal

Schaeffer, B., Baranski, B. and Foerster, T. (2010b). Licensing OGC Geoprocessing Services as a foundation for commercial use in SDIs. Presented at Second International Conference on Advanced Geographic Information Systems, Applications, and Services (GEOProcessing), St. Maarten, Netherlands Antilles

Weiser A, Zipf A (2007) Web Service Orchestration (WSO) of OGC Web Services (OWS) for Disaster Management. In: Li J, Zlatanova S, Fabbri A (ed) Lecture Notes in Geoinformation and Cartography:239–254 Springer, New York