Spatial cloud computing: how can the geospatial sciences use and help shape cloud computing?

Chaowei Yang[a]; Michael Goodchild[b]; Qunying Huang[a]; Doug Nebert[c]; Robert Raskin[d]; Yan Xu[e]; Myra Bambacus[f]; Daniel Fay[e]

[a] Department of Geography and GeoInformation Science and Center for Intelligent Spatial Computing, George Mason University, Fairfax, VA, USA [b] Department of Geography, University of California, Santa Barbara, CA, USA [c] Federal Geographic Data Committee, 590 National Center, Reston, VA, USA [d] NASA Jet Propulsion Laboratory, CA, USA [e] Microsoft Research Connections, Microsoft, Redmond, WA, USA [f] NASA Goddard Space Flight Center, Greenbelt, MD, USA

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis
Taylor & Francis Group

INVITED PAPER

Spatial cloud computing: how can the geospatial sciences use and help shape cloud computing?

Chaowei Yang[a]*, Michael Goodchild[b], Qunying Huang[a], Doug Nebert[c],
Robert Raskin[d], Yan Xu[e], Myra Bambacus[f] and Daniel Fay[e]

[a]*Department of Geography and GeoInformation Science and Center for Intelligent Spatial Computing, George Mason University, Fairfax VA 22030-4444, USA;* [b]*Department of Geography, University of California, 5707 Ellison Hall, Santa Barbara CA 93106-4060, USA;* [c]*Federal Geographic Data Committee, 590 National Center, Reston VA 20192, USA;* [d]*NASA Jet Propulsion Laboratory, 4800 Oak Grove Drive Pasadena, CA 91109, USA;* [e]*Microsoft Research Connections, Microsoft, Redmond, WA, USA;* [f]*NASA Goddard Space Flight Center, Code 700, Greenbelt, MD 20771, USA*

The geospatial sciences face grand information technology (IT) challenges in the twenty-first century: data intensity, computing intensity, concurrent access intensity and spatiotemporal intensity. These challenges require the readiness of a computing infrastructure that can: (1) better support discovery, access and utilization of data and data processing so as to relieve scientists and engineers of IT tasks and focus on scientific discoveries; (2) provide real-time IT resources to enable real-time applications, such as emergency response; (3) deal with access spikes; and (4) provide more reliable and scalable service for massive numbers of concurrent users to advance public knowledge. The emergence of cloud computing provides a potential solution with an elastic, on-demand computing platform to integrate – observation systems, parameter extracting algorithms, phenomena simulations, analytical visualization and decision support, and to provide social impact and user feedback – the essential elements of the geospatial sciences. We discuss the utilization of cloud computing to support the intensities of geospatial sciences by reporting from our investigations on *how cloud computing could enable the geospatial sciences and how spatiotemporal principles, the kernel of the geospatial sciences, could be utilized to ensure the benefits of cloud computing*. Four research examples are presented to analyze how to: (1) search, access and utilize geospatial data; (2) configure computing infrastructure to enable the computability of intensive simulation models; (3) disseminate and utilize research results for massive numbers of concurrent users; and (4) adopt spatiotemporal principles to support spatiotemporal intensive applications. The paper concludes with a discussion of opportunities and challenges for spatial cloud computing (SCC).

Keywords: digital earth; Cyber GIS; geodynamics; space-time; high-performance computing; geospatial cyberinfrastructure

1. Introduction

'Everything changes but change itself' (Kennedy). Understanding changes becomes increasingly important in the twenty-first century with globalization and geographic

---

*Corresponding author. Email: cyang3@gmu.edu

expansion of human activities (Brenner 1999, NRC 2009b). These changes happen within relevant spatial scope and range from as small as the individual or neighborhood to as large as the entire Earth (Brenner 1999). We use space-time dimensions to better record spatial related changes (Goodchild 1992). To understand, protect and improve our living environment, humans have been accumulating valuable records about the changes occurring for thousands of years or longer. The records are obtained through various sensing technologies, including our human eyes, touch and feel, and more recently, satellites, telescopes, *in situ* sensors and sensor webs (Montgomery and Mundt 2010). The advancements of sensing technologies have dramatically improved the accuracy and spatiotemporal scope of the records. Collectively, we have accumulated exabytes of records as data, and these datasets are increasing at a rate of petabytes daily (Hey *et al.* 2009). Scientists developed numerous algorithms and models to test our hypotheses about the changes to improve our capability to understand history and to better predict the future (Yang *et al.* 2011a). Starting from the simple understanding and predictions of geospatial phenomena from our ancestors thousands of years ago, we can now understand and predict more complex Earth events, such as earthquakes and tsunamis (NRC 2003, NRC 2011), environmental issues (NRC 2009a), and global changes (NRC 2009b), with greater accuracy and better time and space coverage. This process helped generate more geospatial information, processing technologies, and geospatial knowledge (Su *et al.* 2010) that form the geospatial sciences. Even with twenty-first century computing technologies, geospatial sciences still have grand challenges for information technology (Plaza and Chang 2008, NRC 2010), especially with regard to data intensity, computing intensity, concurrent intensity and spatiotemporal intensity (Yang *et al.* 2011b):

- **Data intensity** (Hey *et al.* 2009): Support of massive data storage, processing, and system expansion is a long-term bottleneck in geospatial sciences (Liu *et al.* 2009, Cui *et al.* 2010). The globalization and advancements of data sensing technologies helps us increasingly accumulate massive amounts of data. For example, satellites collect petabytes of geospatial data from space every day, while *in situ* sensors and citizen sensing activities are accumulating data at a comparable or faster pace (Goodchild 2007). These datasets are collected and archived at various locations and record multiple phenomena of multiple regions at multiple scales. Besides these characteristics, the datasets have other heterogeneity problems, including diverse encoding and meaning of datasets, the time scale of the phenomena and service styles that range from off-line ordering to real-time, on-demand downloading. Data sharing practices, which are required to study Earth phenomena, pose grand challenges in organizing and administering data content, data format, data service, data structure and algorithms, data dissemination, and data discovery, access and utilization (Gonzalez *et al.* 2010).
- **Computing intensity**: The algorithms and models developed based on our understanding of the datasets and Earth phenomena are generally complex and are becoming even more complex with the advancement of improved understanding of the spatiotemporal principles driving the phenomena. The execution of these processes is time consuming, and often beyond our computing capacity (NRC 2010). These computing intensive methods extend

across a broad spectrum of spatial and temporal scales, and are now gaining widespread acceptance (Armstrong *et al.* 2005). The computing speed of the traditional sequential computing model and single machine cannot keep up with the increased computing demands. In addition, it is not possible for every organization or end user to acquire high-performance computing infrastructure. This resource deficiency has hampered the advancements of geospatial science and applications. The advancement of computing technology and best use of the spatiotemporal principles would help us to eliminate the barriers and better position us to reveal scientific secrets. On the other hand, problem solutions can be enabled by optimizing the configurations, arrangements and selections of hardware and software by considering the spatiotemporal principles of the problems. In order to conduct finer science and better applications, we need computing technologies that can enable us to revisit and include more essential details for models that were simplified for enabling computability.

- **Concurrent intensity**: Recent developments in distributed geographic information processing (Yang *et al.* 2008, Yang and Raskin 2009) and the popularization of web and wireless devices enabled massive numbers of end users to access geospatial systems concurrently (Goodchild 2007). Popular services, such as Google maps and Bing maps, can receive millions of concurrent accesses because of the core geospatial functions and popularity of the geospatial information for making our lives more convenient. Concurrent user accesses and real-time processing require web-based applications to be empowered with fast access and the ability to respond to access spikes – the sudden change in the number of concurrent users (Bodk *et al.* 2010). A study shows that if the response time is longer than three seconds, the users will become frustrated (Nah 2004). With increasing numbers of geospatial systems online, such as real-time traffic (Cao 2007), emergency response (Goodchild 2007), house listings, and the advancement of geospatial cyberinfrastructure (Yang *et al.* 2010b), and other online services based on the framework data, we expect more popular online services and massive concurrent access to become a characteristic of twenty-first century geospatial science and applications. This vision poses great opportunities and grand challenges to relevant scientific and technological domains, such as broadband and cluster computing, privacy, security, reliability issues relevant to the information and systems, and others facing massive numbers of users (Brooks *et al.* 2004).

- **Spatiotemporal intensity**: Most geospatial datasets are recorded in space-time dimensions either with static spatial information at a specific time stamp, or with changing time and spatial coverage (Terrenghi *et al.* 2010). For example, the daily temperature range for a specific place in the past 100 years is constrained by the location (place) and time (daily data for 100 years). The advancement of sensing technologies increased our capability to measure more accurately and to obtain better spatial coverage in a more timely fashion (Goodchild 2007). For example, temperature is measured every minute for most cities and towns on Earth. All datasets recorded for geospatial sciences are spatiotemporal in either explicit (dynamic) or implicit (static) fashion. The study of geospatial phenomena has been described as space-time or geodynamics (Hornsby and Yuan 2008). In relevant geoscience studies such

as atmospheric and oceanic sciences, the space-time and geodynamics have always been at the core of the research domains. And this core is becoming critical in almost all domains of human knowledge pursuant (Su *et al.* 2010). The spatiotemporal intensity is fundamental for geospatial sciences and contributes to other intensities.

Recognizing these geospatial capabilities and problems, the global community realized that it is critical to share Earth observations and relevant resources to better address global challenges. Over 140 countries collaborated to form the intergovernmental Group on Earth Observations (GEO) and proposed a system of systems solution (Figure 1). Within the solution endeavors, GEO organized the process according to information flow stages to better tackle the complex system with various elements including Earth observation and model simulation, parameter extraction, decision support, to social impacts and feedback for improving the system. These steps have been recognized by GEO and other regional and national organizations as practical approaches to solve regional, local, and global issues. Participating organizations in GEO include the geospatial science agencies, such as National Aeronautics and Space Administration (NASA), U.S. Geological Survey (USGS), National Oceanic and Atmospheric Administration (NOAA), Japanese Aerospace Exploration Agency (JAXA), European Space Agency (ESA) of the European Union and the United Nations. Each component within the system is also closely related to the four intensities of geospatial sciences as denoted in Table 1.

The intensiveness issues require us to leverage the distributed and heterogeneous characteristics of both the latest distributed computing and geospatial resources (Yiu *et al.* 2010), and to utilize the spatiotemporal principles to optimize distributed computing to solve relevant problems (Yang *et al.* 2011b) but without increasing much of the carbon footprint (Mobilia *et al.* 2009) and financial budget. This leveraging process has evolved from mainframe computing, desktop computing, network computing, distributed computing, grid computing, and other computing,
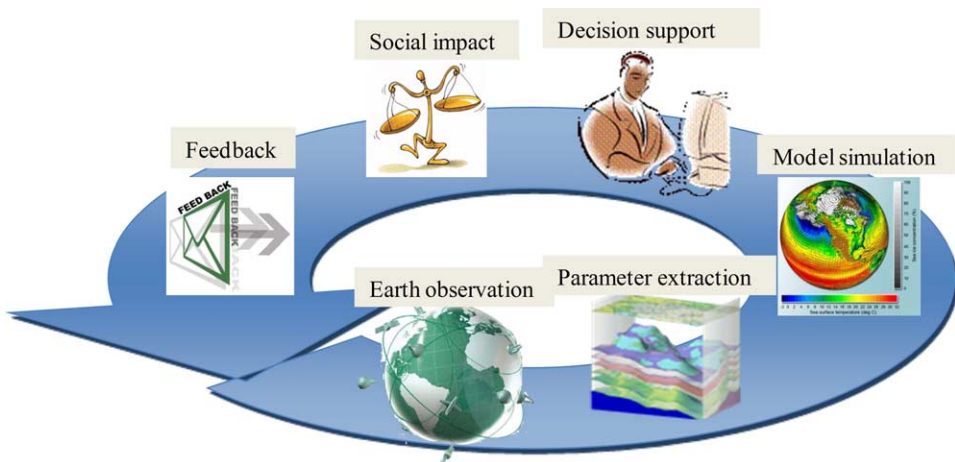

Figure 1. System of systems solution includes Earth observation, parameter extraction, model simulations, decision support, social impact and feedback.

Table 1. The relationship between the elements of geospatial sciences and the issues of data, computing, spatiotemporal, and concurrent intensities.

Intensiveness\elements	Observation	Parameter extraction	Modeling	Information integration/ visualization	Decision-making	Social impact
Data intensive	x	x	x	x		
Computing intensive		x	x	x		
Concurrent access intensive				x	x	x
Spatiotemporal intensive	x	x	x	x	x	x

and recently to cloud computing for geospatial processing (Yang *et al.* 2008, Yang and Raskin 2009). In each of the pioneering stages of computing technologies, geospatial sciences have served as both a driver by providing science-based demands (data volumes, structures, functions and usage) and an enabler by providing spatiotemporal principles and methodologies (Yang *et al.* 2011b) for best utilizing computing resources. Grid computing technology initiated the large-scale deployment of distributed computing within the science community. The emergence of cloud computing brings potential solutions to solve the geospatial intensity problems (Cui *et al.* 2010, Huang *et al.* 2010) with elastic and on-demand access to massively pooled, instantiable and affordable computing resources. The twenty-first century geospatial sciences could also contribute space-time studies (Goodchild *et al.* 2007, Yang *et al.* 2011a) to optimize cloud computing. To capture the intrinsic relationship between cloud computing and geospatial sciences, we introduce spatial cloud computing (SCC) to: (1) enable solving geospatial science problems of the four intensiveness issues; and (2) facilitate the cloud computing implementation and ensure the pooled, elastic, on-demand and other cloud computing characteristics.

2. Cloud computing

Cloud computing refers to the recent advancement of distributed computing by providing 'computing as a service' for end users in a 'pay-as-you-go' mode; such a mechanism had been a long-held dream of distributed computing and has now become a reality (Armbrust *et al.* 2010). National Institute of Standards and Technology (NIST) (Mell and Grance 2009) defines cloud computing as '...a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction'. Because cloud computing is proven to have convenience, budget and energy consumption efficiencies (Lee and Chen 2010), the US government requires that all agencies over the next several years either migrate to cloud computing or explain why they did not use cloud computing. Consequently, it will become the future computing infrastructure for supporting geospatial sciences.

Cloud computing is provided through at least four types of services: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a Service (SaaS), and Data as a Service (DaaS). The first three are defined by NIST and DaaS

is essential to geospatial sciences. These four services are referred to collectively as XaaS.

- IaaS is the most popular cloud service, which delivers computer infrastructure, including physical machines, networks, storage and system software, as virtualized computing resources over computer networks (Buyya *et al.* 2009). IaaS enables users to configure, deploy, and run Operating Systems (OS) and applications based on the OS. IaaS users should have system administrative knowledge about OS and wish to have full control over the virtualized machine. The most notable commercial product is the Amazon Elastic Compute Cloud (EC2, http://aws.amazon.com/ec2/).
- PaaS is a higher level service than IaaS and provides a platform service for software developers to develop applications. In addition to computing platforms, PaaS provides a layer of cloud-based software and Application Programming Interface (API) that can be used to build higher-level services. Microsoft Azure (http://www.microsoft.com/windowsazure) and Google App Engine are the most notable examples of PaaS. Users can develop or run existing applications on such a platform and do not need to consider maintaining the OS, server hardware, load balancing or computing capacity. PaaS provides all the facilities required to support the complete lifecycle of building and deploying web applications and services entirely from the Internet (Bernstein *et al.* 2010).
- SaaS is the most used type of cloud computing service and provides various capabilities of sophisticated applications that are traditionally provided through the Web browser to end users (Armbrust *et al.* 2010). Notable examples are Salesforce.com and Google's gmail and apps (http://www.google.com/apps/). The ArcGIS implementation on the cloud is another example of a spatial SaaS.
- Of the four types of cloud services, DaaS is the least well defined. DaaS supports data discovery, access, and utilization and delivers data and data processing on demand to end users regardless of geographic or organizational location of provider and consumer (Olson 2010). Supported by an integrating layer of middleware that collocates with data and processing and optimizes cloud operations (Jiang 2011), DaaS is able to facilitate data discoverability, accessibility and utilizability on the fly to support science on demand. We are currently developing a DaaS based on several cloud platforms.

Besides the cloud platforms mentioned, Hadoop and MapReduce can also be leveraged as open source resources for expansion to provide elastic and on demand support for the cloud services. The cloud services could be used to support the elements in geospatial sciences according to their respective characteristics:

- Earth observation (EO) data access: DaaS is capable of providing fast, convenient, secure access and utilization of EO data with storage and processing needs.
- Parameter extraction: Extracting parameters, such as vegetation index (VI) or sea surface temperature (SST), from EO data involves a complex series of geospatial processes, such as reformatting and reprojecting, which can be best developed and deployed based on PaaS.

- Model: IaaS provides users with full control of computing instances to configure and run a model, however, network bottlenecks would be a great challenge for IaaS to utilize multiple computing instances to support the model running when massive communication and synchronization is required (Xie *et al.* 2010, Yang *et al.* 2011a). This is where cloud computing can be complemented by high-end computing to solve computing intensive problems.
- Knowledge and decision support: Knowledge and decision support are normally provided and used by domain experts, managers, or the public. Therefore, SaaS would provide good support.
- Social impact and feedback: Social impacts are normally assessed by providing effective and simple visual presentation to massive numbers of users, and feedback can be collected by intuitive and simple applications. Therefore, SaaS, such as Facebook and email, can be best utilized to implement and support social impact and feedback.

NIST denotes five characteristics of cloud computing (Mell and Grance 2009, Yang *et al.* 2011a, b): (1) on-demand self-service (for customers as needed automatically); (2) broad network access (for different types of network terminals, e.g. mobile phones, laptops and personal digital assistants [PDAs]); (3) resource pooling (for consolidating different types of computing resources); (4) rapid elasticity (for rapidly and elastically provisioning, allocating, and releasing computing resources); and (5) measured service (to support pay-as-you-go service). *These five characteristics differentiate cloud computing from other distributed computing paradigms, such as grid computing*. Normally, an end user will use cloud computing by: (1) applying for an account and logging in; (2) testing the scientific or application logic on a local server; (3) migrating to the cloud computing by either customizing a virtual server in a cloud (IaaS), redeveloping on a cloud supported developing environment, such as Microsoft visual studio, and deploying to the cloud (PaaS), or accessing software level functions, such as email process (SaaS). Traditional procedures can take up to months to: (1) identify requirements; (2) procure hardware; and (3) install OS and set up network and firewall. By comparison, cloud users can finish the procedure from a few minutes to a few hours depending on the cloud platform. The deployment modes include private, public, hybrid and community clouds. The integration or interoperation of cross cloud platforms is an active research and development area.

These different concepts are applicable to different roles of users in cloud computing. If we differentiate the user role as: end user, system administrator, developer, designer, manager, operator and developer, we can map each role to the four modes of services, and the elements of geospatial sciences can also be matched to the service modes. Most end users will be using SaaS to relieve them of IT tasks: Earth observation end users are normally engineers who collect, archive and serve EO products, such as MODerate-resolution Imaging Spectroradiometer (MODIS) sensor images, with SaaS and DaaS. Scientists may use the products to extract parameters and conduct modeling hypothesis testing in a SaaS fashion and will require configuration or may develop systems in collaboration with system administrators, designers, and developers using PaaS, DaaS or IaaS. Decision-makers would normally use popular interfaces and need well mined and prepared

information or knowledge for decision support; therefore, they would only use SaaS. To produce social impact, information and knowledge should also be disseminated in web services so that the largest number of users can access them (Durbha and King 2005). The end user's access to SaaS in a convenient fashion is ensured by support from and collaborations among system administrators, developers, designers, managers, and cloud operators and developers.

Typically, only system administrators are granted access to manage underlying virtual computing resources and other roles that are restricted to direct control over the computing resources. The system administrators are usually in charge of hardening virtual machine images, setting up the development environments for developers, and maintaining the virtual computing resources. PaaS provides a platform for a software developer to develop and deliver algorithms and applications. The designer should have an overview of all types of cloud computing models (XaaS) and determine which model is the best solution for any particular application or algorithm; therefore, a good designer is an expert across different types of services. The manager for the whole project can use SaaS, such as an online project management portal, to control and manage the entire procedure from design and development to maintenance. The cloud operator grants permissions to operations for all other roles in all projects. Within the geospatial science element loop from Earth observation to social impact, the cloud developer does not have to be involved if the cloud is well designed and no special requirements are added. However, when organizations want to develop individual cloud platforms with specific requirements that cannot be satisfied by commercial or open cloud platforms, e.g. the USGS Earth Resources Observation Systems (EROS) project, the cloud designers and developers are required to be familiar with XaaS to provide a good solution.

Although cloud computing has been publicized for three years and we have notable successes with Web services best migrated to cloud computing, its potential has been only partially achieved. Therefore, research is still needed to achieve the five characteristics of cloud computing to enable the geospatial sciences in a SCC fashion. This capability can be as simple as running a GIS on a cloud platform (Williams 2009) and using cloud computing for GIServices (Yang and Deng 2010) or as complex as building a well optimized cloud computing environment based upon sophisticated spatiotemporal principles (Bunze *et al.* 2010, Yang *et al.* 2011b).

3. Spatial cloud computing (SCC)

Cloud computing is becoming the next generation computing platform and the government is promoting its adoption to reduce startup, maintenance and energy consumption costs (Buyya *et al.* 2009, Marston *et al.* 2011). For geospatial sciences, several pilot projects are being conducted within Federal agencies, such as FGDC, NOAA and NASA. Commercial entities such as Microsoft, Amazon and Environmental System Research Institution Inc. (ESRI) are investigating how to operate geospatial applications on cloud computing environments and learning how to best adapt to this new computing paradigm. Earlier investigations found that cloud computing not only could help geospatial sciences, but can be optimized with spatiotemporal principles to best utilize available distributed computing resources (Yang *et al.* 2011b). Geospatial science problems have intensive spatiotemporal constraints and principles and are best enabled by systematically considering the

general spatiotemporal rules for geospatial domains (Goodchild 1992, De Smith *et al.* 2007, Goodchild *et al.* 2007, Yang *et al.* 2011b): (1) physical phenomena are continuous and digital representations are discrete for both space and time; (2) physical phenomena are heterogeneous in space, time, and space-time scales; (3) physical phenomena are semi-independent across localized geographic domains and can, therefore, be divided and conquered; (4) geospatial science and application problems include the spatiotemporal locations of the data storage, computing/ processing resources, the physical phenomena, and the users; all four locations interact to complicate the spatial distributions of intensities; and (5) spatiotemporal phenomena that are closer are more related (Tobler' first law of geography). Instead of constraining and reengineering the application architecture (Calstroka and Watson 2010), a cloud computing platform supporting geospatial sciences should leverage those spatiotemporal principles and constraints to better optimize and utilize cloud computing in a spatiotemporal fashion.

> Spatial cloud computing refers to the cloud computing paradigm that is driven by geospatial sciences, and optimized by spatiotemporal principles for enabling geospatial science discoveries and cloud computing within distributed computing environment.

SCC can be represented with a framework including physical computing infrastructure, computing resources distributed at multiple locations and a SCC virtual
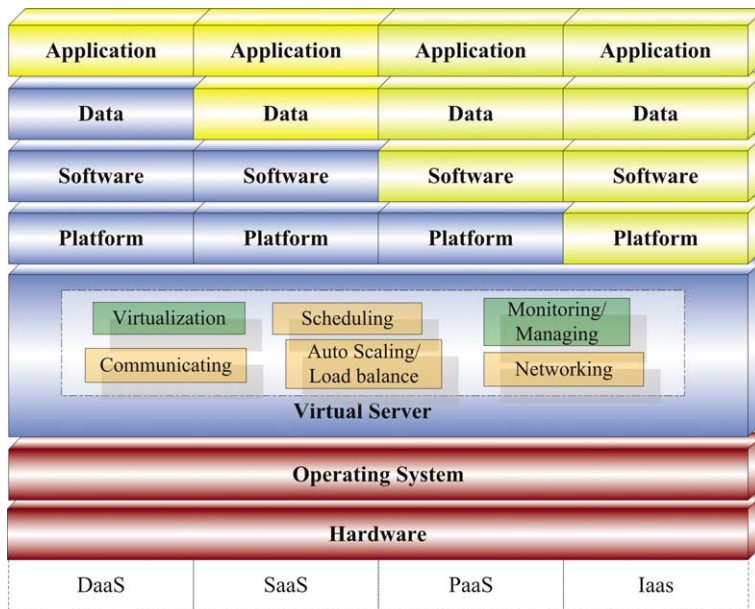

Figure 2. Framework of SCC: red colored components are fundamental computer system components. Virtual server virtualizes the fundamental components and support platform, software, data, and application. IaaS, PaaS, SaaS and DaaS are defined depending on end users' involvements in the components. For example, end user of IaaS will have control on the virtualized OS platform, software, data, and application as illustrated in yellow colour in the right column. All blue colored components will require spatiotemporal principles to optimize the arrangement and selection of relevant computing resources for best ensuring cloud benefits.

server that manages the resources to support cloud services for end users. In Figure 2, the components highlighted in blue are amenable to optimization with spatiotemporal principles to ensure the five characteristics of cloud computing. A virtual server should: (1) provide the functionality of virtualization and support virtual machines above the physical machine with the most important enabling technologies of cloud computing; (2) optimize networking capabilities to best provide and automate public and private IPs and domain names based on the dynamic usage and spatiotemporal availability distribution of the computing resources; (3) determine which physical machine to use when a cloud service is requested, based on scheduling policies optimized by spatiotemporal principles; (4) maintain the spatiotemporal availability, locality, and characteristics of memory and computing resources by communicating, monitoring and managing the physical computing resources efficiently; (5) automate the scalability and load balance of computing instances based on optimized user satisfaction criteria and spatiotemporal patterns of computing resources (Chappell, 2008); and (6) connect to public cloud resources such as Amazon EC2, to construct hybrid cloud computing to serve multiple cloud needs to ensure the five cloud computing characteristics.

The core component of a SCC environment seeks to optimize the computing resources through the middleware with the spatiotemporal principles to support geospatial sciences. Based on the capabilities of the generic cloud computing platform, core GIS functions, such as on-the-fly reprojection and spatial analysis, can be implemented. Local users and system administrators can directly access the private cloud servers through the middleware management interface and cloud users can access the cloud services through spatial cloud portals. Further research is needed in alignment with the IaaS, PaaS, SaaS and DaaS to implement the bidirectional enablement between cloud computing and geospatial sciences (Yang *et al.* 2011b).

4. SCC scenarios

To illustrate how cloud computing could potentially solve the four intensity problems, we select four scientific and application scenarios to analyze the intrinsic links between the problems, spatiotemporal principles and potential SCC solutions.

4.1 Data intensity scenario

Data intensity issues in geospatial sciences are characterized by at least three aspects: (1) multi-dimensional – most geospatial data reside in more than two dimensions with specific projections and geographic coordinate systems. For example, air quality data are collected in four dimensions with 3D space and time series on a daily, weekly, monthly or yearly basis. (2) Massiveness – large volumes of multi-dimensional data are collected or produced from multiple sources, such as satellite observations, camera photo taking, or model simulations, with volumes exceeding terabytes or petabytes. Geospatial science data volume has increased six orders of magnitude in the past 20 years, and continues to grow with finer-resolution data accumulation (Kumar 2007). (3) Globally distributed-organizations with data holdings are distributed over the entire Earth (Li *et al.* 2010b). Many data-intensive applications access and integrate data across multiple locations. Therefore, large
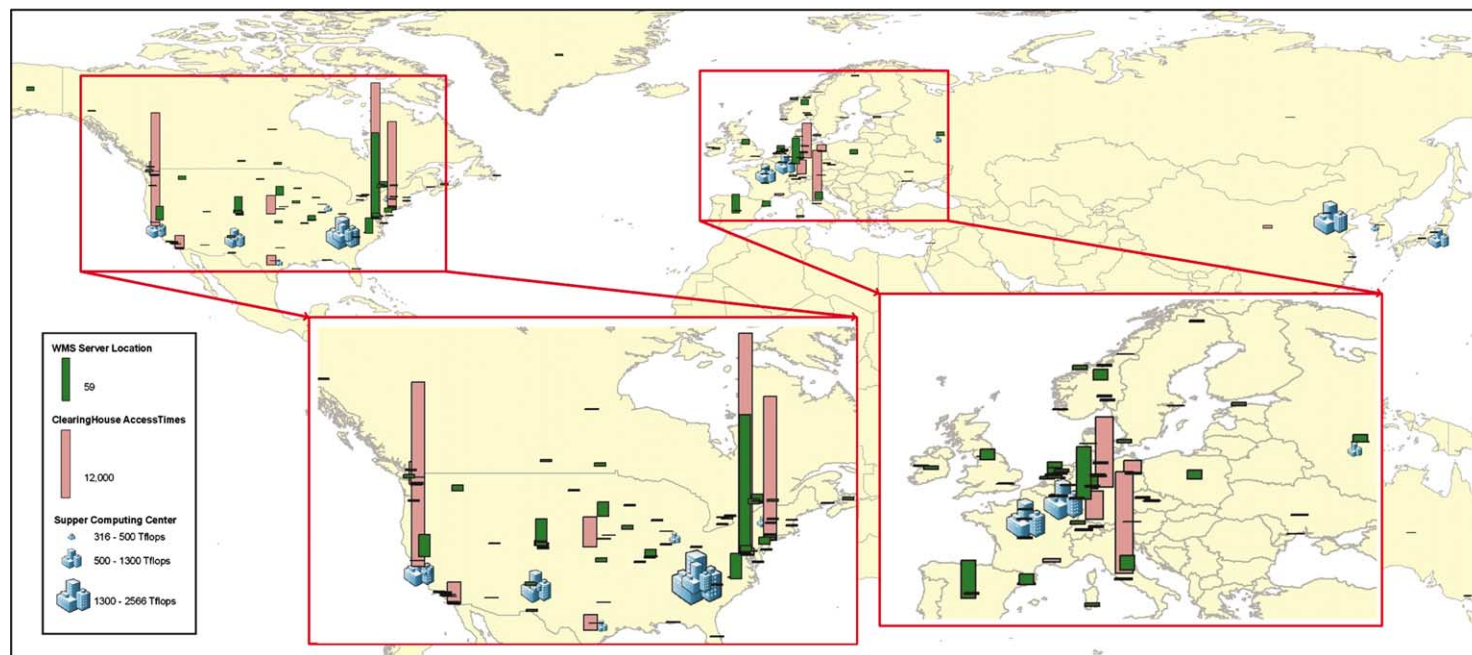
Figure 3. The data services, computing resources, and end users are globally distributed and dynamic. SCC should consider maintaining and utilizing the information of the locality, capacity, volume, and quality of data, services, computing, and end users to optimize cloud computing and geospatial science and applications using spatiotemporal principles.

volumes of data may be transferred over fast computer networks, or be collocated with processing to minimize transmitting (Figure 3).

To address these data intensity problems, we are developing a DaaS, a distributed inventory and portal based on SCC to enable discoverability, accessibility and utilizability of geospatial data and processing to enable geospatial sciences and application. The DaaS is designed to maintain millions to billions of metadata entries (Cary *et al.* 2010) with data locations and performance awareness to better support data-intensive applications (Li *et al.* 2010a). Spatiotemporal principles of the applications that need the data will play a large role in optimizing the data and processing to support geospatial sciences while minimizing the computing resource consumption (e.g. central processing unit [CPU], network and storage) to address how to (Jiang 2011, Nicolae *et al.* 2011): (1) best collocate data and processing units; (2) minimize data transmitting across sites; (3) schedule best sites for data processing and computing optimized by mapping computing resource capacity to demands of geospatial sciences; and (4) determine optimized approaches to disseminate results. The DaaS is being developed and tested based on Microsoft Azure, Amazon EC2 and NASA Cloud Services for the geospatial community.

4.2 Computing intensity scenario

Computing intensity is another issue that needs to be addressed in geospatial sciences. In the elements of geospatial science, computing-intensive issues are normally raised by data mining for information/knowledge, parameter extraction, and phenomena simulation. These issues include: (1) geospatial science phenomena are intrinsically computing-expensive to model and analyze because our planet is a large complex dynamical system composed of many individual subsystems, including the biosphere, atmosphere, lithosphere and social and economic systems. Interactions among each other within spatiotemporal dimensions are intrinsically complex (Donner *et al.* 2009) and are needed for designing data mining, parameter extraction and phenomena simulation. Many data-mining technologies (Wang and Liu 2008) have been investigated to better understand whether observed time series and spatial patterns within the subsystems are interrelated such as to understand the global carbon cycle and climate system (Cox *et al.* 2000), El Niño and climate system (Zhang *et al.* 2003), and land use and land cover changes (DeFries and Townshend 1994); (2) parameter extraction is required to execute complex geophysical algorithms to obtain phenomena values from massive observational data, the complex algorithmic processes make the parameter extraction extremely computational intensive. For example, the computational and storage requirements for deriving regional and global water, energy and carbon conditions from multi-sensor and multi-temporal datasets far exceed what is currently possible with a single workstation (Kumar *et al.* 2006); (3) simulating geospatial phenomena is especially complex when considering the full dynamics of Earth system phenomena, for example, modeling and predicting cyclic processes (Donner *et al.* 2009), including ocean tides (Cartwright 2000), earthquakes (Scholz *et al.* 1973), and dust storms (Xie *et al.* 2010). Such periodic phenomena simulation requires the iteration of the same set of intensive computations for a long time and high-performance computing is usually adopted to speed up the computing process. More importantly, spatiotemporal principles of the phenomena progressions should be utilized to optimize the

organization of distributed computing units to enable the geospatial scientific simulation and prediction (Govett *et al.* 2010, Yang *et al.* 2011b). These principles are also of significance to cloud computing for optimizing computing resources to enable the data mining, parameter extracting and phenomena simulations (Ramak-rishnan *et al.* 2011, Zhang *et al.* 2011) by: (1) selecting best matched computing units for computing jobs with dynamic requirements and capacity; (2) parallelizing processing units to reduce the entire processing time or improve overall system performance; and (3) optimizing overall cloud performance with better matched jobs, computing usage, storage and network status. Because of the diversity and dynamics of scientific algorithms, the best implementing platforms are PaaS and IaaS.

Figure 4 illustrates an example of dust storm simulations, which utilize massive data inputs from both static and dynamic data sources in real-time; the simulation itself is decomposed to leverage multiple CPU cores connected with a computer network and supported by large memory capacity (Chu *et al.* 2009, Xie *et al.* 2010). In this process, the network bandwidth, the CPU speed, and the storage (especially random-access memory [RAM]) play significant roles. The test uses the Nonhydro-static Mesoscale Model (NMM) dust model (Xie *et al.* 2010) for the southeast United States (US) to find how cloud computing infrastructure parameters, such as network speed, CPU speed and numbers and storage impact the predictability of a dust storm. The experiments are conducted with 14 nodes with 24 CPU cores, 2.8 GHz CPU speed and 96 Gbytes memory per node from one data center, and one node with eight CPU cores, 2.3 GHz CPU speed and 24 Gbytes memory from another data center located at a different place. A better connection, faster CPU speed, more memory, and local storage will speed up the simulation and enable prediction. However, compared to CPU and memory factors, network connection is more important as the performance of two nodes each located at a different data center has much worse performance than that of two nodes located at the same data center. During the simulation, every process will produce temporary files for its subdomain to integrate after simulation. The experiment results show that much better performance can be obtained by using a local file system to store the temporary files than by using a Network File System (NFS) share-file system, where all processes will access the same remote storage and transfer data to the storage in real-time. The relationship between these parameters and the predictability across geographic scope, time coverage, and spatiotemporal resolutions (Yang *et al.* 2011b) is critical in providing elastic computing resources for on-demand dust storm forecasting using IaaS or PaaS. It is also apparent that generic cloud computing itself is not enough to solve the problem, but could be complemented by well-scheduled high-performance computing to solve the computing-intensive problem. Also, different job sizes will demand different types of computing environment (Kecske-meti *et al.* 2011).

4.3 Concurrent-access-intensity scenario

The growth of the Internet and the notion to 'provide the right information to any people, anytime and anywhere' makes geospatial services popular to provide location-based services (Jensen 2009) and enable thousands to millions of users to access the system concurrently (Blower 2010). For example, Google Earth supports
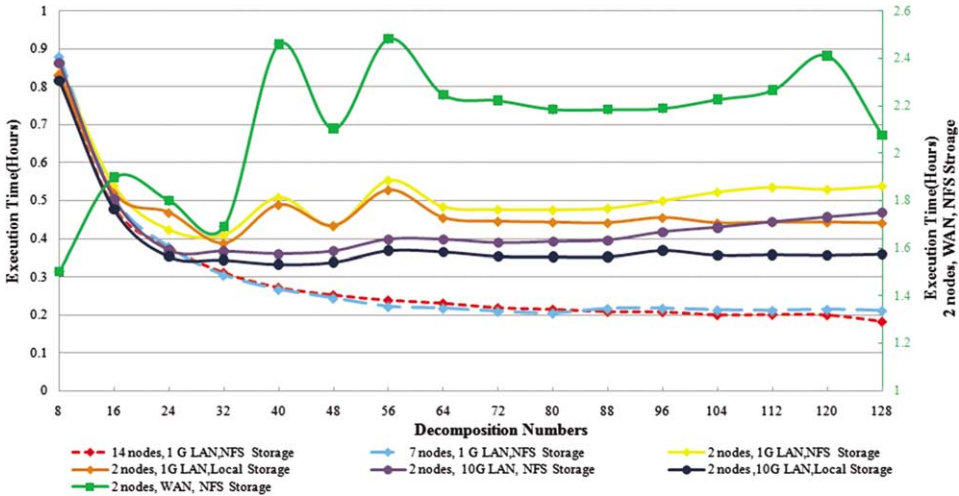
Figure 4. Scalability experiment as a function of CPUs employed, network bandwidth, and storage models to run the NMM dust storm model over a domain of 5.5×9.1 degree in the southwest US at 3 km resolution – a resolution that is acceptable to public health applications for three-hour simulations.

millions of concurrent accesses internationally through its SaaS. These concurrent-intensive accesses may be very intensive at one time (such as the earthquake and tsunami of Japan in March 2011) and very light at other times. To better serve these concurrent use cases, SCC needs to elastically invoke more service instances from multiple locations to respond to the spikes.

In contrast to a constant number of instances, Figure 5 illustrates how the cloud responds to massive concurrent user requests by spinning off new IaaS instances and by balancing server instances using the load balancer (http://aws.amazon.com/elasticloadbalancing/) and auto scalar (http://aws.amazon.com/autoscaling/) of Amazon EC2 to handle intensive concurrent user requests. The example illustrates varying numbers of requests to the Global Observation System of Systems (GEOSS) clearinghouse. The Amazon EC2 load balancer automatically distributes incoming
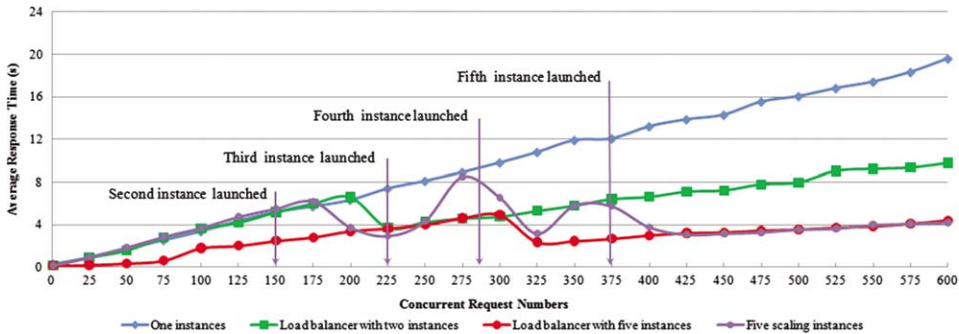

Figure 5. GEOSS clearinghouse GetRecords performance comparison by single, two and five load balancing instances and five autoscaling instances.

application traffic across multiple Amazon EC2 instances. Every instance includes two virtual CPU cores and 7.5 G memory. The load balancer is set up to integrate the computing instances to respond to incoming application traffic and then to perform the same series of tests. Figure 4 shows the response time in seconds as a function of concurrent request numbers when there are one instance, two service instances, five service instances and autoscaling five instances. All instances are run from the beginning except the autoscaling case, which has one instance running at the beginning and elastically adds instances when needed from concurrent requests. It is observed that when more computing instances are utilized, higher gains in performance can be obtained. The elastic automated provision and releasing of computing resources allowed us to respond to concurrent access spikes while sharing computing resources for other applications when there were no concurrent access spikes.

4.4 Spatiotemporal intensive scenario

To better understand the past and predict the future, lots of geospatial data collected are time series and efforts have been made to rebuild time series data from existing observations, such as climate change records (NRC 2010). The importance of spatiotemporal intensity is reflected by and poses challenges to spatiotemporal indexing (Theodoridis and Nascimento 2000, Wang *et al.* 2009), spatiotemporal data modeling methods (Monmonier 1990, Stroud *et al.* 2001), Earth science phenomena correlation analyses (Kumar 2007), hurricane simulation (Theodoridis *et al.* 1999), and the computer network itself that is fast changing in transmitting loads and topological complexities (Donner *et al.* 2009). One popular relevant application is real-time traffic routing (Cao 2007), where massive amounts of geospatial data are collected and preprocessed, route status is predicted, and routing is executed in real-time. The real-time processing requires an infrastructure that can ingest real-time data flow and simulate potential link travel times, as well as conduct real-time traffic routing according to predicted link travel time.

For data collection, different route sensors, cameras and citizen sensing technologies are used to obtain real-time traffic conditions (Goodchild 2007). Existing route links and route nodes are also added as base data. Model simulations are conducted with high-performance computing. Unlike static routing that can be solved by the Dijkstra algorithm, near real-time routing (Cao 2007) have to be conducted routing for every routing request in near real-time. This complexity poses grand challenges to computing and geospatial sciences. Because of the dynamics of routing requests, we cannot maintain the largest capacity needed for responding to the largest number of users because we typically will not need the full computing capacity. The elasticity and on-demand characteristics provided by cloud computing can be used to address this problem and PaaS would be most proper to support this application. The computing power can be shared across metropolitan regions to best optimize the computing process because: (1) traffic peak periods will vary with time zones; (2) collecting, simulating, and routing are data and computing intensive, but the results include only limited information, producing volumes that can be easily transferred across regions; (3) routing tasks are related to dynamic traffic network topology and can be data intensive; and (4) routing requests have significant spikes with dynamic, changing number of requests.

A real-time traffic network with rapid flow, large volume, and multidimensional data for each edge, is generated by location-aware devices and traffic simulation models (Cao 2007). For a metropolitan region such as D.C., when considering static routing only, there will be about 90k nodes, 200k links, 90k*90k potential origin and destination (OD) requests and several optimized routes for each OD request pair, and all of the solutions can be stored with less than 1 Gbyte of storage. But when considering dynamic real-time routing, a routing condition will change for every minute and for each link and node. The volume increases by about (24×60) 1TB for a daily basis, about $(24 \times 60 \times 7)$ 10TB for a weekly basis, or about $(24 \times 60 \times 365)$ 1PB for a yearly basis to retain historical records.

5. Opportunities and challenges

This paper laid out the grand challenges that geospatial science faces in the twenty-first century: the intensiveness of data, computing, concurrent access and spatio-temporal. We argue that the latest advancements of cloud computing provide a potential solution to address these grand challenges in a SCC fashion. Further, the spatiotemporal principles that we encounter in geospatial sciences could be used both to enable the computability of geospatial science problems and to optimize distributed computing to enable the five characteristics of cloud computing. Through four examples, we illustrate that spatiotemporal principles are critical in their abilities to: (1) enable the discoverability, accessibility and utilizability of the distributed, heterogeneous and massive data; (2) optimize cloud computing infrastructure by helping arrange, select and utilize high end computing for computing intensive problems; (3) enable the timely response to world-wide distributed and locally clustered users through geospatial optimization; and (4) assist the design of spatiotemporal data structure, algorithms, to optimize the information workflow to solve complex problems (Herath and Plale 2010). Although these examples are geospatial-centric, spatiotemporal principles can also be utilized to enable the characteristics of cloud computing to support other science discoveries, such as biological and physical sciences where spatiotemporal principles provide driving forces at scales ranging from molecular to the universe.

The success of SCC depends on many factors, such as the outreach of SCC to geospatial scientists who can employ the cloud solutions and to computing scientists and engineers to adapt spatiotemporal principles in designing, constructing, and deploying cloud platforms. We enumerate several aspects including: (1) spatiotemporal principle mining and their mathematical representations for utilization in computing processes with both application-specific forms and generalized forms that can be easily specified and implemented for specific problems; (2) bigger context investigations for considering global challenges, such as the construction of digital earth and responding to tsunami; (3) applications in important complex environments, such as real-time and predicted traffic routing; (4) monitoring of the internal structure and operational status of cloud computing (Yang and Wu 2010) for the utilization of the spatiotemporal principles to optimize the scheduling of cloud computing resources for geospatial and other science demands. Mapping mechanisms and algorithms needs to be researched to help link spatiotemporal characteristics of computing resources in computing capacity and domain problems in computing demands; (5) security and trustworthy issues that emerge in the

virtualized world and are magnified in the cloud computing arena; and (6) ethical and social issues includes privacy and other aspects (Song and Wang 2010).

5.1 Spatiotemporal principle mining and extracting

Geospatial phenomena are ever-changing in time and space and it is possible to use four or more dimensions to represent or describe their evolution. We have established Euclidean and other spaces to describe the phenomena. Due to the complexity of the phenomena and the massiveness of the four-plus dimensions, we have tried to simplify the dimensions and introduce the characteristics or patterns of the phenomena to help better represent the phenomena in both theory and a computing environment to make them computable. For example, we use solid physics and mechanics to describe the Earth's internal structure, fluid dynamics to describe the atmospheric environment, and road networks and topology to describe traffic conditions. These science domains are defined by the principles that govern the evolution of the phenomena.

In geospatial sciences, some of the representation needs revisiting because of the globalization and expansion of human activities. For example, we need to integrate the domains of land, ocean, and atmosphere processes to better understand how the climate is changing. On the other hand, we need to better describe how the geospatial phenomena are impacting our lives, for example, how snow and rainfall impact driving habits and traffic, how earthquakes trigger tsunamis, and how Earth phenomena anomalies indicate a potential earthquake. These spatiotemporal relationships will help us to form better spatiotemporal principles and develop better spatiotemporal examples within multiple dimensions. The crosscutting applications will require scientists from multiple domains with diverse backgrounds to collaborate. Socially, the blending of scientists across domains and geographically dispersed teams is a grand challenge, as has been observed by various geospatial cyberinfrastructure projects, such as Linked Environments for Atmospheric Discovery (LEAD) (https://portal.leadproject.org). Theoretical, experimental, developmental and applied research is needed to: (1) understand the body of knowledge of spatiotemporal principles; (2) formalize the knowledge accordingly to computing capability and domain principles; (3) integrate and interoperate scientific domains with spatiotemporal principles; and (4) evolve cross-cutting computing solutions for integrated domain discoveries.

5.2 Important digital earth and complex geospatial science and applications

Digital earth calls for the integration of digital information about our home planet and the development of solutions for geospatial problems. Some of these problems are of significance to massive numbers of people spanning local, regional, to global geographic scopes, for example, tsunami and earthquake response and real-time traffic routing. Many users will access the system at different times with access spikes, which are mostly predictable, but with frequent anomalies. It is of essential importance to understand the predictable patterns and provide best solutions under specific circumstances. Timely information should also be available to respond to real-time or emergency events (Cui *et al.* 2010). Solving these problems not only

provides convenience to people in need but contributes to the process of improving the quality of life in the long-term.

To address these issues, research is needed to: (1) identify applications of massive impact, of fundamental importance and needed computing support; (2) analyze the four intensiveness problems of the application by mapping to the computing capacity that can be provided by distributed computing; (3) expand or specify the mathematical and conceptual models to computer models to enable the computability of applications by considering both cloud computing capacity and spatio-temporal principles; (4) implement or address the problem with decision-makers and other end users; (5) improve the applications by improving sensor technologies, data processing algorithms, data structures, and model simulations; and (6) summarize the lessons learned and experience that can be leveraged to optimize generic cloud computing that enable generic geospatial sciences or other science domains.

5.3 Supporting the SCC characteristics

The Amazon EC2 Service Level Agreement (SLA) guarantees 99.95% availability for all Amazon EC2 regions, including US Standard, EU (Ireland), US West (Northern California) and Asia Pacific (Singapore). However, Amazon Simple Storage (S3) suffered an outage lasting about two hours in 2008 (http://www.informationweek. com/news/services/storage/showArticle.jhtml?articleID = 209400122) and a major outage in April 2011. The breakdowns caused outages of web services and applications and Amazon EC2 instances relying on S3 for file storage. There is trust that the cloud provider will provide their services for perpetuity. However, Coghead, a cloud vendor closed its business in February 2009 and customers needed to rewrite their applications with other vendor services. The online storage service 'The Linkup' closed July 2008, causing 20,000 paying subscribers to lose their data.

SCC relies heavily on the dynamics of a computing infrastructure, including the network bandwidth, storage volume and reliability, CPU speed and other computing resources. It is hard to ensure all of these characteristics within a reasonable budget. Besides engineering research and assurance of the characteristics of the computing infrastructure, dynamic information is important on the usage/status of network, CPU, RAM, hard drive, software license and other resources to provide a basis for optimizing cloud computing using spatiotemporal principles.

In investigating the characteristics of cloud computing for the four intensive geospatial issues, extensive research is needed to better understand the spatiotemporal behavior of the computing infrastructure and applications, and the optimized scheduling of applications and computing resources will be critical (Rafique *et al.* 2011). Cloud computing platforms can facilitate the sharing, reusing and communicating of knowledge of the scientist and framework of applications from multiple domains (Huang *et al.* 2010). Across-cloud tools and middleware will be available in the future to enable interoperability and portability across clouds, organizations, data, and models.

5.4 Security

Security has always been the biggest concern in migrating to cloud computing in that the entire computing infrastructure is maintained and controlled by third

parties (Subashini and Kavitha 2011, Zissis and Lekkas 2011) rather than by providers and users. Not knowing where our data, applications and users are located, can scare away some potential cloud computing adopters. While cloud computing companies usually utilize authentication and authorization techniques to protect client privacy, it is essential for cloud service providers to ensure that their infrastructure is secure and has proper solutions to protect client data and applications.

Usually, the security requirement baseline can be summarized as (Brodkin 2008):

- Privileged users at cloud computing companies should have separating duties to prevent data leaks or access by other third parties. For instance, computing resource maintainers that have control over computing infrastructure cannot access user accounts, while user account staff should not be able to access the physical machine.
- Cloud computing providers should ensure the functionality and availability of the cloud services.
- Cloud computing providers should provide possible solutions to protect data loss because of failure of cloud services, and have back-up strategies when the cloud service fails to enable data transfers securely from one location to another.
- Each end user should have its own level-based identity management system to control access to cloud data and resources. Users can only access and control their own jobs.

The US Federal Chief Information Officers Council (CIO) is trying to consolidate security assessment and authorization into one function with three steps (CIO 2010): (1) security requirement baseline; (2) continuous monitoring; and (3) potential assessment and authorization. Further research is needed to compare, analyze, test and form security solutions for cloud computing in comparison with other computing platforms (Subashini and Kavitha 2011, Zissis and Lekkas 2011).

5.5 Citizen and social science

SCC is targeting the geospatial sciences and applications with the four intensity problems. When massive users access the data and applications through location-based services, and also contribute to the data and applications, it becomes a paradigm shift in providing convenient electronic media for citizens to both provide and receive information, opinions, data, and knowledge, and therefore democratize the information channels. This shift brings in significant social and ethical concerns in several dimensions:

- Trustworthy: if the data and information are provided officially, it would be easy for users to track the data quality and information accuracy. If any citizen can collect and contribute information, it is hard to guarantee its authority. Sometimes, it becomes a balance of trusting the information or waiting for official information but losing valuable time, e.g. in emergency response where any information may be taken to save human lives.

- Privacy: with data and services deployed over the Internet and on cloud services, protecting provider infrastructure, user privacy and security would be a great challenge (Hayes 2008). One excellent feature of cloud computing is location and device independent access to cloud data and services, which in turn results in a privacy issue when everyone is in an open environment to provide or receive services. And anyone can have access or track the behavior of other individuals.
- Ethical: The advancement of location technologies, such as global positioning system (GPS) and location-based services (Blunck *et al.* 2010) will bring up numerous privacy and ethical issues when sharing information across religious groups, jurisdiction boundaries, and age groups. These and other differences may cause confusion, interference, and side effects for the data & information providers and end-users (e.g. for decision support).

Citizen and social sciences should be investigated in a virtualized cloud computing fashion to analyze the problems, form solutions, and produce best social impacts for human kind.

Acknowledgements

Notes on contributors

Chaowei Yang is an associate professor and he co-directs the Center of Intelligent Spatial Computing for Water/Energy Sciences, which he founded at George Mason University, Fairfax, VA in 2006. His research interest is utilizing spatiotemporal principles to optimize computing for enabling science discoveries. He has published over 60 peer reviewed articles and served as a guest editor for special issues of five international journals. He co-founded the AAG Cyberinfrastructure Specialty Group (CISG) and acts as the chief architect of NASA Cloud Computing and Climate @ Home initiatives at Goddard Space Flight Center.

Michael Goodchild is Director of the Center for Spatial Studies and the Jack and Laura Dangermond Professor of Geography at the University of California, Santa Barbara. He coined the term 'geographic information sciences' and helped to solidify it as a field through his more than 500 publications and tens of millions dollar of research funding.

Qunying Huang is a Ph.D candidate at George Mason University with research focused on computing issues of geospatial sciences. She has published over 10 peer reviewed articles in various journals and conferences.

Doug Nebert is the FGDC secretariat and the lead of GeoCloud Initiative among FGDC and other relevant agencies. He has led the design of architectures of most FGDC initiatives.

Robert Raskin is Supervisor of the Science Data Engineering and Archiving Group at NASA Jet Propulsion Laboratory. He is co-founder of the AAG Cyberinfrastructure Speciality Group and a strong advocate of data interoperability.

Yan Xu is a Senior Research Program Manager of Earth, Energy, and Environment at Microsoft Research Connections, Microsoft Corporation. She is responsible for the Environmental Informatics Framework (EIF), a Microsoft eScience initiative aiming at interdisciplinary computational research that engages Microsoft technologies with environmental sciences.

Myra Bambacus is a program manager for NASA Cloud Services and Climate@Home project. She has served as the manager for many geospatial interoperability and innovation initiatives, such as FDGC Geospatial One Stop and NASA Interagency Digital Earth Office.

Daniel Fay is the Director of Earth, Energy, and Environment for Microsoft Research Connections, Microsoft Corporation, where he works with academic research projects focused on utilizing computing technologies to aid in scientific and engineering research. Dan has project experience working with High Performance Computing, Grid Computing, collaboration and visualization tools in scientific research. Dan was previously the manager of eScience Program in Microsoft Research where he started Microsoft's engagements in eScience including the MSR eScience workshop.

References

Armbrust, M., *et al.*, 2010. Above the clouds: a Berkeley view of cloud computing. *Communications of the ACM*, 53 (4), 50–58.

Armstrong, M.P., Cowles, M., and Wang, S., 2005. Using a computational grid for geographic information analysis. *Professional Geographer*, 57 (3), 365–375.

Bernstein, Q.T., Vidovic, N., and Modi, S., 2010. A cloud PAAS for high scale, function, and velocity mobile applications – with reference application as the fully connected car. *In: Proceedings of the 2010 Fifth International Conference on Systems and Networks Communications (ICSNC'10)*. Washington, DC, USA: IEEE Computer Society, 117–123. doi: 10.1109/ICSNC.2010.24. Available from: http://dx.doi.org/10.1109/ICSNC.2010.24 [Accessed 30 March 2011].

Blower, J.D., 2010. GIS in the cloud: Implementing a web map service on Google app engine. *In: Proceedings of the 1st International Conference and Exhibition on Computing for Geospatial Research & Application, COM.Geo'10*, 21–23 June 2010, New York, NY, USA, ACM, Article 34. Available from: http://doi.acm.org/10.1145/1823854 [Accessed 30 March 2011].

Blunck, H., *et al.*, 2010. PerPos: a platform providing cloud services for pervasive positioning. *In: Proceedings of the 1st International Conference and Exhibition on Computing for Geospatial Research & Application, COM.Geo'10*, 21–23 June 2010, Washington DC, USA. New York: ACM, 11, 1–8.

Bodk, P., *et al.*, 2010. Characterizing, modeling, and generating workload spikes for stateful services. *In: Proceedings of the 1st ACM Symposium on Cloud Computing, SoCC'10*, 6–11 June 2010, UC Berkeley, Berkeley, CA, USA, 241–252.

Brenner, N., 1999. Beyond state-centrism? Space, territoriality, and geographical scale in globalization studies. *Theory and Society*, 28 (1), 39–78.

Brodkin, J., 2008. *Gartner: seven cloud-computing security risks* [online]. Available from: http://www.networkworld.com/news/2008/070208-cloud.html [Accessed 31 March 2011].

Brooks, C., *et al.*, 2004. The Massive User Modelling System (MUMS). *Intelligent Tutoring Systems: Lecture Notes in Computer Science*, 3220 (2004), 73–112.

Bunze, K., Ager, A., and Schrader, P.C., 2010. Up in the air: adventures in serving geospatial data using open source software and the cloud. *In: Proceedings of the 1st International Conference and Exhibition on Computing for Geospatial Research & Application, COM.Geo'10*, 21–23 June 2010, Washington, DC, USA. New York: ACM, 35, 1–4.

Buyya, R., Pandey, S., and Vecchiola, S., 2009. Cloudbus toolkit for market-oriented cloud computing. *Cloud Computing, Lecture Notes in Computer Science*, 5931 (2009), 24–44. doi: 10.1007/978-3-642-10665-1_4.

Calstroka, J. and Watson, P., 2010. Automatic software deployment in the Azure cloud, distributed applications and interoperable systems. *In: Proceedings of 10th IFIP WG 6.1 International Conference, DAIS 2010*, 7–9 June 2010, Amsterdam, Netherlands, 155–168.

Cao, Y., 2007. *Transportation routing with real-time events supported by grid computing*. Thesis (PhD). George Mason University, Fairfax, Virginia, USA.

Cartwright, D.E., 2000. *Tides: A Scientific History*. Cambridge: Cambridge University Press.

Cary, A., et al., 2010. Leveraging cloud computing in Geodatabase Management. *In: Proceedings of the 2010 IEEE International Conference on Granular Computing*, 14–16 August 2010, San Jose, CA, USA, IEEE, 73–78.

Chappell, D., 2008. *Introducing the azure services platform*. Microsoft Corporation document. Available from: http://www.davidchappell.com/writing/white_papers/Introducing_Windows _Azure_v1-Chappell.pdf [Accessed 20 February 2011].

Chu, S., Yeh, C., and Huang, C., 2009. A cloud-based trajectory index scheme. *In: Proceedings of IEEE International Conference on e-Business Engineering*, 21–23 October 2009, Macau, China, 602–607.

CIO Council, 2010. *Proposed Security Assessment & Authorization for U.S. government cloud computing* [online]. Federal CIO Council White paper, ver.0.6. 90 pp. Available from: https:// info.apps.gov/sites/default/files/Proposed-Security-Assessment-and-Authorization-for-Cloud-Computing.pdf [Accessed 20 February 2011].

Cox, P. M., et al., 2000. Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature*, 408, 184–187.

Cui, D., Wu, Y., and Zhang, Q., 2010. Massive spatial data processing model based on cloud computing model. *In: Proceedings of the Third International Joint Conference on Computational Sciences and Optimization*, IEEE Computer Society, Los Alamitos, CA, 28–31 May 2010, Huangshan, Anhui, China, IEEE, 347–350.

Defries, R.S. and Townshend, J.R.G., 1994. NDVI-derived land cover classifications at a global scale. *International Journal of Remote Sensing*, 15 (17), 3567–3586.

De Smith, M.J., Goodchild, M.F., and Longley, P., 2007. *Geospatial analysis: a comprehensive guide to principles*. Leicester, UK: Troubador Publishing Ltd.

Donner, R., et al., 2009. Understanding the Earth as a Complex System – recent advances in data analysis and modelling in Earth sciences. *European Physical Journal Special Topics*, 174, 1–9.

Durbha, S.S. and King, R.L., 2005. Semantics-enabled framework for knowledge discovery from Earth observation data archives. Geoscience and Remote Sensing. *IEEE Transactions on*, 43 (11), 2563–2572.

Gonzalez, H., et al., 2010. Google fusion tables: data management, integration and collaboration in the cloud. *In: Proceedings of the 1st ACM Symposium on Cloud Computing, SoCC'10*, 6–11 June 2010, UC Berkeley, Berkeley, CA, USA. New York: ACM, 175–180.

Goodchild, M., 1992. Geographical data modeling. *Computers & Geosciences*, 18 (4), 401–408.

Goodchild, M., 2007. Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69 (4), 211–221.

Goodchild, M., Yuan, M., and Cova, T., 2007. Towards a general theory of geographic representation in GIS. *International Journal of Geographic Information Science*, 21, 239–260.

Govett, M., Middlecoff, J., and Henderson, T., 2010. Running the NIM next-generation weather model on GPUs. *In: Proceedings of 10th IEEE/ACM International Conference on Cluster, Cloud, and Grid Computing, CCGrid 2010*, 17–20 May 2010, Melbourne, VIC, Australia. New York: ACM, 792–796.

Hayes, B., 2008. Cloud computing. *Communications of the ACM – Web science*, 51 (7), 9–11.

Herath, C. and Plale, B., 2010. Streamflow-Programming model for data streaming in scientific workflows. *In: Proceedings of 10th IEEE/ACM International Conference on Cluster, Cloud,*

and Grid Computing, CCGrid 2010, 17–20 May 2010, Melbourne, VIC, Australia. New York: ACM, 302–311.

Hey, T., Tansley, S., and Tolle, K., 2009. *The fourth paradigm: data-intensive scientific discovery* [online]. Microsoft Research, Redmond, WA. Available from: http://research.microsoft.com/en-us/collaboration/fourthparadigm/ [Accessed 1 January 2011].

Hornsby, K. and Yuan, M., 2008. *Understanding dynamics of geographic domains*. Boca Raton: Taylor & Francis: CRC Press.

Huang, Q., *et al.*, 2010. Cloud computing for geosciences: deployment of GEOSS clearing-house on Amazon's EC2. *In: Proceedings of the ACM SIGSPATIAL International Workshop on High Performance and Distributed Geographic Information Systems, HPDGIS'10*, 3–5 November 2010, San Jose, CA, USA. New York: ACM, 35–38.

Jensen, C.S., 2009. Location, location, location. *In: Proceedings of 2009 tenth international conference on mobile data management: systems, services and middleware, MDM 2009*, 18–20 May 2009, Taipei, Taiwan, IEEE.

Jiang, J.R., 2011. Nondominated local coteries for resource allocation in grids and clouds. *Information Processing Letters*, 111, 379–384.

Kecskemeti, G., *et al.*, 2011. An approach for virtual appliance distribution for service deployment. *Future Generation Computer Systems*, 27, 280–289.

Kumar, S.V., *et al.*, 2006. Land information system: an interoperable framework for high resolution land surface modeling. *Environmental Modelling & Software*, 21 (10), 1402–1415.

Kumar, V., 2007. High performance data mining – application for discovery of patterns in the global climate system. *In*: S. Aluru, M. Parashar, R. Badrinath, and V.K.. Prasanna, eds. *Proceedings of the 14th international conference on High performance computing (HiPC'07)*. Berlin, Heidelberg: Springer-Verlag, 4.

Lee, Y. and Chen, K., 2010. Is server consolidation beneficial to MMORPG? A case study of world of warcraft. *In: Proceedings of 2010 IEEE 3rd international conference on cloud computing, CLOUD 2010*, 5–10 July 2010, 435–442.

Li, Z., *et al.*, 2010a. An optimized framework for OGC web service seamlessly integration to support Geospatial Sciences. *International Journal of Geographic Information Science*. doi: 10.1080/13658816.2010.484811.

Li, W., Yang, C., and Yang, C., 2010b. An active crawler for discovering geospatial Web services and their distribution pattern – a case study of OGC Web Map Service. *International Journal of Geographical Information Science*, 24 (8), 1127–1147.

Liu, Y., *et al.*, 2009. Research of remote sensing service based on cloud computing mode. *Application Research of Computers*, 26 (9), 3428–3431.

Marston, S., *et al.*, 2011. Cloud computing – the business perspective. *Decision Support Systems*, 51, 176–189.

Mell, P. and Grance, T., 2009. *The NIST definition of cloud computing* Ver. 15. [online]. NIST.gov. Available from: http://csrc.nist.gov/groups/SNS/cloud-computing/ [Accessed 22 November 2010].

Mobilia, S., *et al.*, 2009. Sustainability on-orbit: space solar power and cloud computing constellation two examples of international offset projects. *In: Proceedings of 60th International Astronautical Congress 2009, IAC 2009*, 8:6141–6147, 12–16 October 2009, Daejeon, Republic of Korea, Curran Associates Inc., 6141–6147.

Monmonier, M., 1990. Strategies for the visualization of geographic time-series data. *Cartographica: The International Journal for Geographic Information and, Geovisualization*, 27 (1), 30–45.

Montgomery, K. and Mundt, C., 2010. A new paradigm for integrated environmental monitoring. *In: Proceedings of the 1st International Conference and Exhibition on Computing for Geospatial Research & Application (COM.Geo'10)*, 21–23 June 2010, New York, NY, ACM, USA.

Nah, F., 2004. A study on tolerable waiting time: how long are Web users willing to wait? *Behavior & Information Technology*, 23 (3), 153–163.

Nicolae, B., *et al.*, 2011. BlobSeer: Next-generation data management for large scale infrastructures. *Journal of Parallel Distributed Computing*, 71, 169–184.

NRC, 2003. *Living on an active earth: Perspectives on earthquake science*. Washington DC: The National Academies Press.

NRC, 2009a. *Global environmental health: research gaps and barriers for providing sustainable water, sanitation, and hygiene services: workshop summary.* Washington DC: The National Academies Press.

NRC, 2009b. *Informing decisions in a changing climate.* Washington DC: The National Academies Press.

NRC, 2010. *The rise of games and high performance computing for modeling and simulation.* Washington DC: The National Academies Press.

NRC, 2011. *Tsunami warning and preparedness: an assessment of the U.S. Tsunami Program and the Nation's Preparedness Efforts.* Washington DC: The National Academies Press.

Olson, A.J., 2010. Data as a service: Are we in the clouds? *Journal of Map & Geography Libraries*, 6 (1), 76–78.

Plaza, A.J. and Chang, C.I., 2008. High performance computing in remote sensing. Gainesville, USA: CRC Press.

Rafique, M., Butt, A.R., and Nikolopoulos, D.S., 2011. A capabilities-aware framework for using computational accelerators in data-intensive computing. *Journal of Parallel and Distributed Computing*, 71, 185–197.

Ramakrishnan, L., *et al.*, 2011. Deadline-sensitive workflow orchestration without explicit resource control. *Journal of Parallel and Distributed Computing*, 71, 343–353.

Scholz, C.H., Sykes, L.R., and Aggarwal, Y.P., 1973. Earthquake Prediction: A Physical Basis. *Science*, 181 (4102), 803–810.

Song, W. and Wang, X.S., 2010. In-device spatial cloaking for mobile user privacy assisted by the cloud. *In: Proceedings of 11th International Conference on Mobile Data Management (MDM 2010)*, 23–26 May 2010, Kansas City, MO, IEEE Computer Society, 381–386.

Stroud, J.R., Müller, P., and Sansó, B., 2001. Dynamic models for spatiotemporal data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63, 673–689.

Su, D., *et al.*, 2010. Research of spatial decision support system construction based on cloud model. *In: Proceedings of 2010 International Conference on E-Business and E-Government (ICEE 2010)*, 7–9 May 2010, Guangzhou, China, IEEE, 691–694.

Subashini, S. and Kavitha, V., 2011. A survey on security issues in service delivery models of cloud computing. *Journal of Network and Computer Applications*, 34, 1–11.

Terrenghi, L., *et al.*, 2010. CloudRoom: a conceptual model for managing data in space and time. *In: Proceedings of The 28th Annual CHI Conference on Human Factors in Computing Systems (CHI 2010)*, 10–15 April 2010, Atlanta, GA. New York: ACM, 3277–3282.

Theodoridis, Y., Jefferson, S., and Nascimento, M.A., 1999. On the generation of spatiotemporal datasets. *Lecture Notes in Computer Science* 1651 (1999), 147–164.

Theodoridis, Y. and Nascimento, M.A., 2000. Generating spatiotemporal datasets on the WWW. *SIGMOD Record*, 29 (3), 39–43.

Wang, J. and Liu, Z., 2008. Parallel data mining optimal algorithm of virtual cluster. *In: Proceedings of 2008 fifth international conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, 18–20 October 2008, Jinan Shandong, China, IEEE Computer Society, 358–362.

Wang, Y., Wang, S., and Zhou, D., 2009. Retrieving and indexing spatial data in the cloud computing environment. *In: Proceedings of First International Conference on Cloud Computing, CloudCom 2009*, 1–4 December 2009, Beijing, China. Berlin, Heidelberg: Springer-Verlag, 322–331.

Williams, H., 2009. *Spatial cloud computing (SC2) white paper: a new paradigm for geographic information services* [online]. Available from: http://www.skeinc.com/pages/SC2/SKE_SC2_White_Paper.pdf [Accessed 21 November 2010].

Xie, J., *et al.*, 2010. High performance computing for the simulation of dust storms. *Computers, Environment, and Urban Systems*, 34 (4), 278–290.

Yang, C., *et al.*, 2008. Distributed geospatial information processing: sharing earth science information to support digital earth. *International Journal of Digital Earth*, 1 (3), 259–278.

Yang, C. and Raskin, R., 2009. Introduction to distributed geographic information processing research. *International Journal of Geographic Information Science*, 23 (5), 553–560.

Yang, X. and Deng, Y., 2010. Exploration of cloud computing technologies for geographic information services. *In: Proceedings of 18th International Conference on Geoinformatics*, 18–20 June 2010, Wuhan, China, IEEE Computer Society, 1–5.

Yang, J. and Wu, S., 2010. Studies on application of cloud computing techniques in GIS. *In: Proceedings of 2010 second IITA International Conference on Geoscience and Remote Sensing*, 28–31 August 2010, Fuzhou University, Fuzhou, China, IEEE Computer Society, 492–495.

Yang, C., *et al*., 2010. Geospatial cyberinfrastructure: past, present and future. *Computers, Environment and Urban Systems*, 34 (4), 264–277.

Yang, C., *et al*., 2011a. WebGIS performance issues and solutions. In: S. Li, S. Dragicevic, and B. Veenendaal, eds. *Advances in web-based GIS, mapping services and applications*. London: Taylor & Francis Group, ISBN 978-0-415-80483-7.

Yang C., *et al*., 2011b. Using spatial principles to optimize distributed computing for enabling physical science discoveries. *Proceedings of National Academy of Sciences*, 106 (14), 5498–5503. doi: 10.1073/pnas.0909315108.

Yiu, M.L., *et al*., 2010. Enabling search services on outsourced private spatial data. *VLDB Journal*, 19 (3), 363–384.

Zhang, P., *et al*., 2003. Correlation Analysis of Spatial Time Series Datasets: A Filter-and-Refine Approach. *In*: *Proceedings of the 7th Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Berlin, Heidelberg: Springer-Verlag, 532–544.

Zhang, T., *et al*., 2011. Typical virtual appliances: an optimized mechanism for virtual appliances provisioning and management. *Journal of Systems and Software*, 84, 377–387.

Zissis, D. and Lekkas, D., 2011. Addressing cloud computing security issues. *Future Generation Computer Systems*. doi: 10.1016/j.future.2010.12.006