Using cloud computing to develop an integrated virtual system for online GIScience programs

Germana Manca  
The Pennsylvania State University, PA, USA  
Nigel W. Waters  
The University of Calgary, Alberta, CA  
Gustavo Sandi  
Washington DC and Costa Rica, USA

Recommended citation:  
Using cloud computing to develop an integrated virtual system for online GIScience programs

Germana Manca*
Geoinformatics and Earth Observations Laboratory
Department of Geography and Institute for CyberScience
The Pennsylvania State University, PA, USA
E-mail: germana.manca@gmail.com

Nigel W. Waters
Professor Emeritus of Geography
The University of Calgary, Alberta, CA
E-mail: nwaters@ucalgary.ca

Gustavo Sandi
Network Consulting Services
Washington DC and Costa Rica, USA
E-mail: Gustavo.sandi@gmail.com

*Corresponding author

Abstract: The variety of offerings of online Geographical Information Science (GIS) programs has been extensively reported in the literature, which describes various types of degrees and certificates offered by institutions all over the world. Most online courses have merely focused on delivering lectures, for which standard presentation tools such as PowerPoint are sufficient to fulfil this task. It is imperative for GIS online courses to deliver instruction as a series of interactive steps. This paper presents how an integrated virtual system based on cloud computing can be developed to enhance GIS online courses, and how such an approach provides an interactive teaching method to improve the quality of communication between students and teachers.

Keywords: Geographical information science (GIS); GIScience; Virtual environment; Online education; SaaS; Cloud computing

Biographical notes: Dr. Germana Manca is a faculty affiliate of the Geoinformatics and Earth Observations Laboratory in the Department of Geography and Institute for CyberScience at Pennsylvania State University. She has taught several GIScience courses, and she has been involved in several educational activities, about GIScience. She published her work in peer review journals, such as Transactions in GIS, Applied GIS, IJGIS, Cartographica among others.

Dr. Nigel W. Waters is a retired Full Professor from the Department of Geography, University of Calgary, where he taught from 1975 to 2007. He was a Full Professor in the Department of Geography and Geoinformation Science and Director of the Center of Excellence for GIS at George Mason University,
Fairfax, Virginia, from 2007 to 2014. He now holds the rank of Professor Emeritus of Geography at the University of Calgary. He has conducted and published peer reviewed research in GIS modelling, spatial analysis, and in medical, environmental and transportation geography applications of GIS. A former Editor of the peer reviewed journal, Cartographica, in 2014 he received the Award of Distinction from the Canadian Cartographic Association and in 2015 the Inaugural Award from the GIScience Study Group of the Canadian Association of Geographers for Lifetime Achievement and GIScience Excellence.

Gustavo Sandi has been working in the IT field since 1997 and holds M.S. and MBA degrees. He is specialized in remote network services and he worked at Tiempos del Mundo as IT Director and at present is employed at the Washington Times as Senior Analyst in Washington D.C. His experience is focused on transforming large scale and expensive projects into more affordable and highly efficient ones.

1. Background

Many institutions in the US are offering online programs, both course-based degrees and certificates. The range of offerings is wide and includes many GIS courses. This explosion is due to the demand shown by the growth in the number of students enrolled in these programs. In some instances, these students are hoping to advance their careers and in other cases they simply desire to broaden their knowledge of GIS. For these reasons, student enrolment in online GIS programs varies depending on the age of the student and on employment commitments that may affect their time schedules.

The explosion of online courses is described by Wikle (2010) who depicts the profound changes that higher education is experiencing. GIS online education has become available at US colleges and universities through programs ranging from traditional (face-to-face) courses to 100 percent online degrees and certificates programs for non-traditional students. In his conclusion, Wikle refers to the GIS&T Body of Knowledge (Waters, 2013), which provides a means for selecting online content, ensuring that students are exposed to the breadth of knowledge needed to develop basic GIS competencies, and assisting institutions in developing strategic plans for implementing new programs on their campuses. A jointly promoted effort among Canadian universities to deliver online courses has been described by Khmelevsky, Burge, Govorov, and Hains (2011). Open education/learning is more than taking an online course (M. A. Peters, 2009). It means fostering a new academic culture that values the core practices of an open science and creating a new cyber infrastructure that facilitates and seamlessly integrates all of the above procedures in open scholarly practices. Millet et al. (2014) describe how a spatial web tool, RacerGISOnline, is an innovative approach to integrating these tools into several courses in the marketing curriculum while avoiding the problems that have constrained adoption.

A new online system for teaching GIS was implemented and evaluated at the University of Georgia, US, (Rivero & Buchanan, 2014) for its potential for implementation in other university marketing departments. These authors describe experiences and observations from transitioning such a lab-intensive, face-to-face course in “Advanced Geographic Information Systems” to a fully online course, using technologies already available, such as ArcGIS, and the Learning Management System.
(LMS) known as eleNew. Implementation involved putting together these software packages to provide a powerful learning environment.

A centralized ArcGIS desktop server has been described by D. Peters (2009). He lists several choices, including the client’s use of Citrix XenApp terminal clients for optimum computing and display performance. This framework for the learning environment enables a more efficient and independent computing architecture communication protocol to support communication between the server and client’s platform. The system functions as a framework for delivering the technology. Nevertheless, any effective way to provide “education” should exploit a diverse set of technologies.

Course delivery technologies are analysed by Johnson, Corazzini, and Shaw (2011). Three different online modalities of learning, including the Learning Management System, Webinar, and Virtual Environment approaches were compared in order to understand the students’ perception of learning. Two concurrent themes arose from the three platforms: the technical challenges inherent in the technology and the students’ various preferences for synchronous web-based learning. The Virtual Environment approach emerged as the preferred distance based education methodology.

A conceptual framework in the GIS environment has been described by Schultz (2012) for an adult, GIS online course. Essentially he described the advantages of online courses for GIS, especially the use of a virtual GIS Server and having a professor as a facilitator, delivering the GIS courses.

MaKinster and Trautmann (2014) refer to the concept of “evolution”, when describing the ways in which teachers contribute to and influence the design and direction of their professional development experiences and project outcomes. An evolutionary approach is critical in enabling teachers and educational leaders to have significant input into shaping the nature and direction of the project. It occurs also when teachers work with the project team to adapt resources, develop complementary ones, and share lessons and teaching experiences with one another using Web-based, courseware tools. Moreover, using an integrated approach with those tools in order to deliver courses/education and technologies, is the key to lecture development, requiring the authors to develop a complex technological framework.

2. Methods and computational environments for GIS online courses

Any method that is applied to create a virtual educational framework for GIS online courses requires On-demand Application Delivery Software or SaaS (Software as a Service, the application layer of cloud computing), as well as Webinar applications and a Virtual GIS Environment. The Learning management system known as Blackboard, works as a collector of class materials, as a grading book and supplier of datasets, but it is not directly involved in the learning process. SaaS is the first building block in an integrated system. SaaS is a model where the client software is hosted by a remote server which can be setup as a stand-alone server or a cluster of servers to support from two to virtually an unlimited number of users. This structure uses a client-server architecture, enabling the delivery of a very powerful learning experience that can be accessed remotely or locally. On the Server Side, the need to have On-demand Application Delivery software ensures that multiple users can connect to the server and share the resource. Once this has been set up the GIS application can then be installed and run locally in order to be shared by multiple users. Security can be guaranteed through the SaaS environment or using a Virtual Private Network (VPN) model. The VPN is always
the most secure and easy method to implement, avoiding exposure on the Internet. On the client side, the variety of plugins available for the different operating systems on the market, such as Android, iOS, Windows, Linux and MacOS, ensures that all commonly available hardware devices will have access to the required functionality. However, if the server is behind a firewall then a VPN client needs to be installed.

The next building block required in the learning framework is the Webinar tool. These tools are now attracting increased attention due to their ability to facilitate synchronous communication in online learning environments (Wang & Hsu, 2008). Several software programs are freely or commercially available. Among them are Joinme (LogMeIn, Boston, MS), Anymeeting (Anymeeting, Huntington Beach, CA), GoToWebinar (Citrix Systems, Santa Barbara, CA), Elluminate (Elluminate, Inc., Calgary, Canada), or Adobe Connect (Adobe Systems Inc., San Jose, CA). These applications enable many-to-many interaction between users, have the ability to transmit and record audio and video, offer access to the Internet, and provide opportunity for information exchange via whiteboards and application sharing (Wang & Hsu, 2008). The advantages of this technology include affordability, multi-level interaction (Wang & Hsu, 2008) and real time interaction between faculty and students, providing opportunities to learn new technologies or concepts over a semester. Online discussion both synchronous and asynchronous, typically creates an environment in which participants engage with one another in more equitable ways by giving equal voice to those who tend to be more reserved in face-to-face settings (Bonk, Hansen, Grabmer, Lazar & Mirabelli, 1998). Finally, the last building block in the framework is the Virtual GIS Environment, which is based on the technology described in Fig. 1. The client side is represented on the left of the drawing where multiple devices can either access the servers via VPN or internally in a Local Area Network or LAN (this might comprise a department, faculty or campus wide environment).

![Fig. 1. The Citrix Xenapp environment](image-url)
If students are connected remotely, they use a VPN and if they are connected locally, they connect directly to the server. On the right side of Fig. 1, there are several servers that perform different functions, allowing many users to be served concurrently. The number of application servers that can be configured can range from 1 to a great many. For a GIS application it is recommended that there should not be more than 10 users per server. A Citrix “farm” allows the control and management of interaction issues arising from the user or student. Moreover, the shadow desktop functionality provides a useful application to control the user session.

One of the most popular VPNs is Hamachi because of its reasonable price and full internet support. It is very easy to manage and deploy. Citrix, on the other hand, is one of the most expensive platforms on the market for application virtualization. Nevertheless, it provides the best performance; it runs on top of remote services from Microsoft and provides a better and more efficient protocol to deal with remote applications, including the way the video is handled, which is better than its competitors.

The licensing and installation of the applications is a subject that changes from vendor to vendor. An instructor implementing such a learning environment must check with the vendor concerning the licensing rights and limitations, before any installation is adopted.

3. Learning procedures in GIS online courses

This section describes the philosophy behind the GIS Virtual Environment, how it was implemented and how it operates in online learning practice.

In order for a GIS online course to work successfully, the professor must use a web conferencing program to invite the students to attend the class online. Ideally the professor must have presentations pre-recorded with both audio and video. Each student must login to a password protected, presentation website. They will then be entitled to participate as both listener and presenter in the course conferences. In an ideal online environment, the professor will be able to request that a student present material to the class. The student should be able take over the instructor’s role and will only relinquish control once his or her presentation is complete. The instructor should be able to make PowerPoint presentations, explain exercise material and run videos in a manner that all students will be able to participate in because all students will share a view of the instructor’s screen.

When the professor ends the class, the session is then closed but all of the students should have access to the session at a later time. If for some reason one of them could not attend class that day, they can watch the recorded session at another time. Video conferencing can be provided at any time in the session; however, it has been discovered that this is not very practical since it will require one to one video and the screen could be filled up with the student faces, reducing the screen size that is dedicated to displaying the presentation.

This process improves the overall learning experience, offering the student the ability to attend class without being physically present in a classroom, avoiding the difficulty of transportation to the campus premises. Also it may improve the learning experience since the student can access past material online and review the material at any subsequent time. Once the professor dismisses the class and distributes homework assignments, the students will have full access to the GVEnvironment, which stands for
GIS Virtual Environment and is a combination of technologies specifically tailored for teaching GIS.

The following description indicates how the GVEnvironment works. First of all, students will require a fast internet connection for the most rewarding experience. The students also have to install two client programs that are free and can be uninstalled at any time. One of these programs is the VPN software and the other is client software that permits access to the GVEnvironment. The VPN software is needed to create a secure connection between the student’s computer and the server computer while the client software is needed to run the GIS software remotely. The installation of the client software is completed by the student. In less than 15 minutes they can have access to the GVEnvironment without any additional technical knowledge. Thus the complexities of installing the GIS software on their computers is circumvented and they experience none of the technical installation problems such as their computers being too slow, or having insufficient memory, or that they are using a Macintosh that is not configured for ArcGIS among other problems. All of these difficulties are avoided when using a GVEnvironment because all of the software is pre-installed on the server.

When the user uses the server, many additional benefits are available such as interaction in real-time between the professor and student. If the professor is online and the student has questions or problems, the professor can oversee and control the user session in the server and help him/her understand the problem. The professor never takes control of the user’s computer, only the user’s session that lives in the server. This is very important because the privacy of the student’s equipment remains intact. The professor will have access to all the student sessions in real-time (i.e. there are no email or blackboard downloadings) and can place homework files, assignments and even review students’ progress on different assignments in true real time. It is the same as having the professor physically present in the lab where a student would have raised his or her hand to ask the professor to visit his or her station. This is something that rarely happens in the real world, but the virtual concept allows it.

4. An experiment

An experiment was conducted to test the effectiveness of the GIS Virtual Environment. Fig. 2 shows a screenshot of the Citrix farm with 11 users connected. In addition, it portrays how the sessions are displayed for the professor in an environment where every user is running the GIS software individually, and the hardware resources are assigned based on demand using the Citrix software. In our test environment the applications ArcGIS and Citrix Xenapp were run on a computer using an Intel i5 chip running at 2.7 Ghz with 6Gb of RAM, preferably on Solid-State Drives (SSD) but the system performs well using regular Hard Drives with a minimum of 80Gb of memory available.

Fig. 3 shows the server performance. In our test environment it never reached its peak even with 11 users connected concurrently; memory usage was around 50% and processor around 6%. However, this was with the majority of users idle. When the server was rendering or processing a map, processor use can easily reach 100%, which will be distributed among the users who require it. Statistically it is highly improbable that all the users at the same time will require 100% of the processor’s operating power.

Moreover, this technology allows multiple users to be connected concurrently to different GIS programs. GVEnvironment has been tested thoroughly using the ArcGIS software (Fig. 4) since ArcGIS is the most widely accepted software for teaching GIS.
Nevertheless, the GVEnvironment can be used with other GIS software packages. Therefore, QGIS and GRASS, which are free GIS packages, have been tested and shown to run with no issues.

Fig. 2. Screenshot of the Citrix Farm

Fig. 3. Screenshot showing server performance
In regard to the technical performance of the system the bandwith, suggested by Citrix literature (Ben-Chanoch, 2013), ranges between 100k and 200k (with video applications). In our case study, we did not stream video therefore anything below the 200kbps was sufficient. The system ran with 14 users with a 5Mbps bandwidth, and according to the rule of three, this should be more than enough to sustain all of the users connected simultaneously. The conferencing system was hosted separately from the Citrix server so that the two systems would not conflict.

The server provision for teaching is very different from a server used for production. In our case, we focused on a server provision that would be able to run GIS applications in real time without too many delays. ArcGIS for instance requires, at a minimum, an Intel Pentium 4 at 2.2Ghz Processor with 2GB Ram for a single user, while the setup of the Learning Server was an Intel i5 at 3.4Ghz with 6Gb Ram. This meant 3 to 4 concurrent users could run the system at close to the 100% of computing resources. This, however, happened only when an image was rendered. Most of the time, the processor was idle which was good for maintaining a learning environment where users might come and go at different times during the day. Although the technology appears to be easily supported by a standard server, the primary source of workload uncertainty was the student’s response to the application and his or her fluency in the use of the technology.

The GVEnvironment was tested in a class that in past years had a high level of student appreciation and excellent evaluations. The comparison group was an introductory GIS class, which was ideal as a test to determine if the GVEnvironment could help the students reach the same level of knowledge as the control group. The material was already tested by the control group students, who gave an overall positive evaluation. In the test of the GVEnvironment some of the course content was slightly modified to fit the display and the online sequences. Consequently, the class could be improved only by the use of this technology. Therefore, considering that the students have no initial knowledge of GIS, which could be misleading in regard to students...
already trained in GIS science, the educational benefit of an improved technology can be highlighted.

Two classes were chosen to compare and evaluate the teaching class methods. An online class, taught in 2014, and the control group, taught in 2013. In 2013 there were 16 students enrolled in the class, while in 2014, there were 12. The choice of these classes was based upon the characteristics that they had in common: namely that they had used the same version of ArcGIS and were given the same assignments. Previous classes could not be compared, because of the different versions of the software, and consequently their different homework assignments which were designed to be software specific. Both classes were split into two parts: the theoretical and practical one. The practical one proposed the application of modules and theory that had been explained during the theoretical part. Generally, it lasted one and a half hours and the exercises were calibrated over the acquired knowledge of the theory class. This framework worked well in both classes, and student appreciation was shown by the unsolicited comments in the students’ emails. Moreover, the virtual interaction in an online session was both more effective and explicit, because of the instructor’s ability to share and control the student’s desktop, software and mouse. The instructor was able to direct the students through a sequence of steps, retaining the student’s attention on the computing and GIS processes. Furthermore, thanks to the Citrix farm, the instructor was able to detect any anomalies in the software or processing being performed by each student, and was then able to restore the initial state of the software or overcome the difficulties that the student had encountered.

Fig. 5 and Fig. 6 show a comparison between the two classes described above based on the students’ homework performance.

Assessment of student results was based on the completion of nine homework assignments during each of the courses. Each assignment was worth 50 points. The nine learning tasks were designed to deliver GIS knowledge from a beginning up to an intermediate level. The x axis shows the sequence of the homework, while the y axis shows the average percentage for each of the homework assignments for the 2013 and 2014 classes. These nine assignments were designed to measure a series of increasingly sophisticated learning tasks.
The online class in 2014 shows a slightly better performance as measured by the percentage of students who reached the class goal, for each assignment during the semester. In 2014 students reached a peak in the final assignment. In 2013 the control group also achieved good performances too, but they did not match the online class. Essentially this comparison is based purely on the percentage of the students that achieved the assignment’s goal. Statistical tests of the means (t tests) and ANOVA comparisons of the two data sets showed that in most cases there was no significant difference between the control group and the class taught using the online technology. So it may be concluded that the average results for the students over the nine assignments showed that the online course had not impaired the students’ performance and the variance in student’s performance was in some exercises reduced. More importantly the online course provided a number of additional benefits including removal of the need for scheduled office hours for which both instructor and students needed to be physically present (including the provision of weekend office hours) and a reduction in unsolicited emails from the students.

More specific improvements in the implementation of the online course can be seen for the HM IV assignment. In this assignment the students were required to provide a report of an industrial model that had already been developed. Thus the students were required to investigate the materials available, and then download and execute the model. The 2013 (control group) and the 2014 online courses differed in their execution of the GIS code. In the online class the instructor tested the running code of each student in his/her account, and checked to see if it worked properly. In the control group (the 2013 class), the operating systems used by the students were diverse, and the students reported difficulties in compiling the code which could only be resolved in class while they were physically present.

HM VIII resulted in similar performances in both classes. Again the assignment was based on the identification of land use characteristics through GIS visualization, and thus it required the use of only the display command. By contrast, HM VII showed the worst performance for the online class, despite the fact that the homework was identical. In this case, the reason for the difference in performance was that few students in the online class received the highest grade, while some students did not complete the assignment. Why the online students fared worse in this particular instance bears further investigation.
The control group in 2013 had scheduled office hours once a week and could, in addition, schedule meetings at mutually convenient times. This meant that the time used to satisfy the students’ requests and questions was considerably higher than for the online class. The reason for the differential in the amount of student interaction is the immediacy of help that could be provided for the online class. Quite simply put, this immediacy was far more efficient. Students in the control group that had to wait for scheduled office hours might not receive assistance for several days. In a busy semester, a prompt answer solves quickly the issue and clears the way for the student to proceed with the next steps to be undertaken in the assignment. As a result, the number of hours used to explain to students in the control group on how to proceed in a given assignment was increased. It may be summarized that for the control group the support was only available during class or during office hours, while for online class, support was both personalized and provided immediately during class hours, office hours and outside of these times as well.

5. Discussion

Geospatial technology plays a fundamental role in training geospatial scientists for private industry and for preparing students for academic careers. As such the GVEnvironment framework benefits both the business and scientific communities. These benefits can be divided into economic and educational advantages.

The economic benefits refer to the gap between the traditional classroom and the online SaaS virtual environment models. In the traditional model, hardware and software needs to be refreshed every 3 or 4 years. During this period of time the instructor needs to upgrade their hardware and software once they become obsolete and no longer receive technical support. In order to stay up to date, it is recommended by the software providers that users purchase annual maintenance programs, assuring that the software and hardware is current. This model requires licenses for every individual computer which poses a 1:1 ratio for users and licenses. The SaaS model saves resources in two ways: first, the client software can be run on aging hardware, extending the lifespan of the equipment for possibly 3 of 4 years more. The GVEnvironment software is supported by many systems and configurations, and does not require high capacity computers. Second, it reduces the need for large numbers of software licenses. Since it is a client/server model, only the cost of the server’s concurrent users will be paid. In such situations the user:license ratio is about 3 to 1. Servers can host a number of licenses thus reducing the number of licenses to about a third of the normal requirements.

In summary, a SaaS virtual environment for online GIS teaching has proven to be superior in terms of the use of resources, more efficient management than a traditional model, and less expensive in hardware and software requirements. The human resource investment in term of technical knowledge to implement SaaS has to be carefully taken into account. Nevertheless, when it works at full performance, the Return-On-Investment (or ROI), can be obtained in less than one year.

Educational advantages include the increased interaction between students and teachers in a manner which is beneficial in terms of the reduced time for responses to student questions and more extensive visualization options (including screen and mouse sharing). These benefits were confirmed by the unsolicited emails sent by the students to the instructor which emphasized their appreciation of the rapid, positive feedback and their support for receiving their instruction within the GVEnvironment.
Other advantages of the system include the fact that the professor fully monitors the students and is able to observe when they are having difficulties with the assigned exercises. The ability to observe the student’s screen is an effective method for interacting directly with the student. It is more productive, because the professor focuses on a specific student, until he/she understands. In a traditional classroom setting, the professor cannot provide this as individualized attention. The GVEnvironment application permits the professor to focus on the needs of a single student. Consequently, each student interacts directly with the professor and the professor can then focus on helping the specific student’s needs. Moreover, the student’s ability to use the server over the time requested to complete their work and assignments, is a sign of their familiarity and fluency with the online system.

6. Conclusion

SaaS technologies for the integration of software and hardware have been used and developed in other disciplines and business environments. For instance, online editorial systems throughout the academic world have used these technologies for a number of years. Moreover, several commercial enterprises and public institutions have applied the SaaS model for geospatial data, achieving impressive ROIs over short periods of time (Smith & Turner, 2010; Maguire, Kouyoumjian, & Smith, 2008).

The same concepts and computing technologies can be applied to education. It has been estimated that this architecture saves around 20% of yearly expenditures for technical infrastructure. This saving has wide application since this technology can work on almost any internet connection available. Tests have been carried out in central and South America and the results were satisfactory. Those times when the technology described here did not work were due to external factors that in a dedicated environment should not occur.

While the economic and technical benefits of these technologies is now clearly apparent, the educational advantages for geospatial education has been demonstrated by the gains described above, that were made by students involved in the experimental use of this infrastructure in an online course taught during 2014. Using the GVEnvironment, students sent in unsolicited emails that described their positive feedback, regarding this specific course and the curriculum in general.

Nevertheless it is worth re-emphasizing the following educational aspects that have emerged from our online teaching within the GVEnvironment: first, the students’ ability to use the server over the time requested to complete their work and assignments is a sign of their familiarity and fluency with the online system; second, the classes that were taught using this methodology met the students’ expectations, which were: a) to be able to work from any location; b) to be able to use a variety of platforms; c) to be able to interface easily and constantly with their professor; and d) to be able to access online support for technical issues; third, the instructor was able to directly monitor the workflow of the students, to check their progress, to assess their results, to balance their educational activities, and to measure their achievements; fourth, it is not possible to make a comparison with existing literature on online education, because this particular GVEnvironment approach is new and has not yet been used for the teaching of geospatial concepts. Basically the GVEnvironment overcomes the drawback of traditional in-class teaching, thereby improving the learning process.
These observations reveal how this SaaS architecture facilitates the interaction of the instructor-student relationship in a manner which was not possible before, due to the limitations and constraints of the former technologies. The educational outcome, obtained using these technologies, has produced results that provided significant benefits for both teachers and students.

Considering the outcomes, both economic and educational, this system represents a powerful tool, not only for online teaching, which is actually well established, but also for software interaction at the student-instructor interface.

References
