Teaching Business Analytics with Virtual Laboratories

Charles Mutigwe, Western New England University, Springfield, Massachusetts, USA
Bruce Mtigwe, National University of Science & Technology, Bulawayo, Zimbabwe
Tendai Chikweche, Western Sydney University, Parramatta, Australia

ABSTRACT

Significant advances in cloud computing technology and the ubiquity of consumer electronics devices, such as, tablets and smart phones, which are connected to the Internet have made it possible to develop flexible and cost effective virtual lab systems to support online learning courses that require a computer laboratory component. In this paper, the authors describe their experience of designing and using a virtual computer lab for teaching an online business analytics class. The virtual lab provides an opportunity for students to do hands-on assignments in navigating business data warehouses, creating data cubes, writing SQL queries, creating dashboards and other visualizations, writing reports, and using descriptive and predictive models for data mining. The authors describe alternative cloud architectures for designing virtual labs and discuss their criteria for selecting the optimal deployment method.

Keywords: virtual lab, remote lab, business analytics education, virtualization, and cloud computing

INTRODUCTION

Virtual computer laboratories (virtual labs) represent an interesting emerging technology for online learning in courses that traditionally required the use of campus-based computer laboratories (Son, Irrechukwu, & Fitzgibbons, 2012; Fulmer & Johnson, 2014). By means of these virtual labs, students can interact with the required software applications anytime of the day, from any location, and using any personal computer or smart mobile device provided that they have Internet access. Furthermore, instructors can remotely assist users in the virtual labs in real-time when required. This emerging development is also part of the blended learning revolution in education which seeks to provide students with convenient options for accessing learning materials (Neacsu & Adascalitei, 2016). All this has been made possible by significant advances in cloud computing technology and the ubiquity of consumer electronics devices, such as, tablets and smart phones. In this paper, the authors describe their experience of designing and using remote labs and then virtual labs over several semesters for teaching an online business analytics class. A virtual lab system provisioned in the public cloud provides a flexible and cost effective solution for students to do hands-on assignments in navigating business data warehouses, creating data cubes, writing SQL queries, creating dashboards and other visualizations, writing reports, and using descriptive and predictive models for data mining. The authors describe alternative cloud architectures for designing virtual labs and discuss their criteria for selecting the optimal deployment method.

LITERATURE REVIEW

Virtual labs, according to Son et al. (2012), are hands-on laboratories that are built using virtualization technology. They describe how virtual labs have been used in technology-based courses, such as information security courses. In other fields, particularly in engineering and the physical sciences, virtual labs have been defined as environments where student experiments can be simulated through the use of software (Chan & Fok, 2009), while remote labs allow students to access the physical lab equipment at a distance via the Internet in order to perform real-time experiments (Auer, 2001). Extension of use of these labs is also now evident in business courses such as Design Thinking and Innovation where visual experiments form a key component of the students' learning experience. We will use the definition by Son et al. (2012) and will describe remote labs as computer labs where all of the software tools for the hands-on exercises are installed directly on a student's personal computer. Burd, Seazzu, and Conway (2009) describe an implementation of a virtual computing lab in one college of a large public university and they compare this type of computer lab to the traditional physical computer labs.

In terms of deployment, there are three types of clouds: public clouds, private clouds, and hybrid clouds (Jadeja & Modi, 2012). With the public clouds all of the computing servers, network infrastructure, and software applications are managed by a service provider. Users select the applications that they need and only pay for the time duration that they use these applications, it is a pay-per-use model. The primary limitation with public cloud deployments is that they are less secure than the other two options. The private clouds on the other hand are provisioned within the organization's internal data center, all of the computing servers, network infrastructure, and software applications are

managed by the organization's IT staff. The hybrid cloud is a combination of the private and public clouds, where primarily the private cloud is used and when additional computing resources are needed then external public clouds maybe engaged on an ad-hoc basis. From an end user's perspective, the deployment model is of no consequence since end users get the same benefits regardless of the deployment model, however from the organization's perspective there are significant differences between the deployment models with regards to costs, security, and time to deploy applications. Despite its security limitations, the public cloud is considered more useful than its private counterpart based on the costs and deployment flexibility (Vikas, Gurudatt, Vishnu, & Prashant, 2013).

BACKGROUND

This paper discusses the use of virtual labs for an online class in the MBA program at a college of a large public university in the spring semester of 2016. The college's MBA program offered both a full-time residential MBA option and a completely online MBA option. Since the fall semester of 2013 the residential option had been offering a business analytics course as a core course. This course was hands-on based and as such it required the use of a computer lab for all class meetings. A couple of factors drove the need to explore ways to offer this course in the online option. Firstly, in order to remain competitive in the online MBA space there was a growing need to quickly offer one or more analytics courses as part of the core. Secondly, a strategic decision was taken by the college to harmonize the core course offerings in the two MBA options.

Three technology options on how to take the residential business analytics course and make it available to students taking classes via the online option were studied. Below we discuss these options.

The first option was to try and use university-hosted virtual labs (UVlabs). This would involve the provisioning of virtual labs using desktop virtualization (VDI) technology that would be hosted in the university's data center (University of Southern California, 2017; University of Arkansas, 2017). In other words, with this option a virtual lab would be created in a private cloud. The advantages with this option were that it would facilitate compliance with the university's data policies and if successful could be scaled to the online offerings campus-wide. However, the campus IT department would need to acquire additional IT personnel to support this options, as well as, additional computer hardware and software resources. As such, the challenges with this option were the high direct costs, together with the long lead time required to get this option approved via the university inter-departmental processes.

The second option was to try and use university-supplied remote labs (URlabs). This would involve the campus IT department creating images of the campus computer labs and placing those images as virtual machine files on portable USB (flash) drives or DVDs (Casini, Prattichizzo, & Vicino, 2007). Each student taking the business analytics course via the online option would then be mailed a flash drive together with instructions on how to access their personal computer lab using the home machine. This option had the advantage that students can use their personal computer labs in situations where they have limited Internet connectivity and all of the students' data would be stored locally and so minimizing the risk of violating the university's data policies. The challenges with this option centered around the cost of creating and maintaining the virtual machine images over time, together with the licensing for the operating system and productivity applications, such as, Microsoft Office® that were installed on the image. There was also a concern about the potentially long lead time required to get this option approved via the university inter-departmental processes.

The third option that was reviewed involved the use student-built remote labs (SRlabs). This required the development of detailed install instructions for all of the software applications to be used in the course. The 10 software applications that were used for the exercises in this course are described in the Implementation and Course Description section below. Using this option students would then be required, with some support from the instructor, to install all of the software applications onto their personal computers before the start of the course. This option meant that students were responsible for acquiring the software licenses for the operating system and productivity applications, such as, Microsoft Office® on their personal machine. The install software and software licensing information for the analytics applications application for the course would be provided as part of the install instructions. This option could be implemented quickly as it did not require any University inter-departmental approval processes. This option required that the course instructor have a background or some experience in IT support. Furthermore, both students and the instructor would spend a significant amount of time working on issues that were not related to the course content. A major limitation of this option was that the business analytics course could not be offered as a core course since there was the possibility that some students due to technical or personal limitations may not be able to build their own remote labs.

The SRlabs option was selected primarily due to its lower direct costs and the shorter timeframe for implementation, as shown in Table 1. The direct costs for the UVlabs and URlabs options could have been offset by adopting a technology fee for students enrolling in the business analytics course. However, such a decision would have required approval at the university-level. The UVlabs and URlabs options were also dependent on the IT department's resource availability. Phase 1 section below discusses the implementation of the SRlabs option and the issues that were encountered.

Table 1: Remote Labs versus Virtual Labs in the Private Cloud

	UVlabs	**URlabs**	**SRlabs**
Pros	- Students can use any device - No software installs by students - University's data policies can be strictly enforced - The instructor does not provide tech support.	- Students only have to install a VMM - The instructor does not provide tech support.	- IT department is not required to setup the remote labs - No 'technology fee' is required for the course
Cons	- Costs of setting up private cloud - IT department required to setup each virtual lab - Costs have to be offset by a 'technology fee'	- Costs of distributing the drives - IT department required to setup & duplicate the imaged drives - Costs have to be offset by a 'technology fee' - Licensing issues since software in VMs not installed by the user.	- Students have to install 10 applications - The instructor has to provide tech support.

IMPLEMENTATION AND COURSE DESCRIPTION

The course offered an introduction to business intelligence (BI) with emphasis on the BI infrastructure, as well as, the processes used to get data into data warehouses and the data mining tools used to search for patterns in the data. It was a single semester 14-week course. Concepts were introduced through lectures and assigned readings, and they were assessed using quizzes. Application of the concepts was taught via twelve in-class hands-on exercises using different commercial and open-source BI tools. Students were first introduced to Structured Query Language (SQL) to extract and manipulate data in production relational database systems and data warehouses. Queries were written and executed using the Oracle SQL Developer tool (Oracle, 2017). Next, students studied the need for data warehouses and they learned about the different architectures that were available for data warehouses. In the hands-on exercises students used a SAP Business Warehouse (https://www.sap.com/products/business-warehouse.html) to create a data cube and to load data into the cube from external data sources. Students were then able to slice-and-dice the data in their cubes using the SAP Business Explorer Query Designer (BEx) and SAP BusinessObjects Analysis for Microsoft Office (https://www.sap.com/community/topic/businessobjects-analysis-ms-office.html). Students then developed two dashboards using SAP BusinessObjects Dashboards (https://www.sap.com/products/dashboards.html). Using data from a database students wrote a report using SAP Crystal Reports (https://www.sap.com/products/crystal-reports.html). Several data mining algorithms were introduced: apriori algorithm for association mining problems, decision tree algorithm for classification problems, linear regression and exponential smoothing algorithms for estimation problems, and k-means algorithm for clustering problems. Students completed a market basket analysis exercise using a Teradata data warehouse with historical retail data from Sam's Club and the SAP Predictive Analytics application (https://www.sap.com/products/analytics/predictive-analytics.html). Several scenarios were presented with different data sets for the data mining hands-on exercises that were completed using SAP Predictive Analytics and R software. Text mining was introduced using the concept of regular expressions (Friedl, 2006) and students used the open-source GNU Grep tool to mine text files and emails for patterns of interest (https://www.gnu.org/software/grep).

Phase 1: Using Student-Built Remote Labs

The online version of the Business Analytics course was rolled out in the fall semester of 2014. The course was taught by an instructor whose experience included IT technology support and database development. The course was also taught by the same instructor in the spring and fall semesters of 2015.

In all three semesters the course was taught using the student-built remote labs. Students could email questions to the instructor or post them on a discussion board in the online classroom. We categorized these inquiries into two categories; lab-setup related and course-content related. Table 2 shows the student inquiries data, together with the course enrollments for each semester.

In the fall of 2014 students were encouraged to install all of the applications and thus build their complete remote lab at the start of the course, however, some students chose to install the applications separately. They installed an application as and when it would be needed for each upcoming hands-on exercise. In subsequent semesters the setting up of the complete remote lab by the students was made into a graded assignment which was due in the first week of the course.

Students used personal computers (Windows and Mac), as well as, employer-provided laptops (Windows). Since most of the commercial analytics applications used in the course only run on the Windows platform, students using Macs needed to have a virtual machine running Windows 7 or later hosted in a virtual machine monitor (VMM), such as VirtualBox (https://www.virtualbox.org/wiki/VirtualBox) or Parallels Desktop for Mac (http://www.parallels.com/products/desktop). Common problems encountered with the setups of the remote labs were:
Firewalls and anti-virus software restricting the installs and access to the hosted data warehouse servers.
Lack of administrative rights to perform the installs on employer-provided laptops.
Misconfigured environment variables
Missing operating system device drivers or incorrect versions of device drivers.
Missing pre-requisite applications, such as, Java or Microsoft .NET frameworks
Most problems were diagnosed via screen shots in email exchanges between the instructor and students. A few more difficult cases required that the instructor remotely access the student's computer using a WebEx session. On a few occasions a student's computer with a functional remote lab broke down during the semester which resulted in that student needing to rebuild their remote lab.

Phase 2: Using a Public Virtual Lab

The student-built remote labs solution that was implemented in Phase 1, while successful still had a significant downside; the solution was not scalable. It required that the instructor have the dual roles of teaching and providing technical support, and that the students spend time doing software installs which were not part of the course content. Furthermore, students who were not using the latest versions of Microsoft Windows on their personal computers faced significant challenges in taking this course online. All of these limitations did not affect students taking the course in the residential MBA option since the course was taught in a computer lab that was supported by the IT Department.

Table 2: Student Inquiries by Type and Semester

	Fall 2014	**Spring 2015**	**Fall 2015**	**Spring 2016**
Enrollment	19	31	18	16
Lab Setup Inquiries	140	69	64	29
Course Content Inquiries	50	109	75	77

In the spring 2016 semester a public cloud-based virtual lab solution was trialed for this course in the online MBA option. This solution did not require the IT Department's resources and it addressed the limitations discussed above. The public cloud-based virtual lab for the course was provisioned and managed by a vendor, vLab Systems (http://www.vlabsystems.net). The service was provided at no cost as a trial in the spring 2016 semester, however for future semesters students would be required to pay for access to the virtual lab at a cost of $48.86 per 14-week semester. These payments would be made directly to the vendor by the students. The virtual lab access course requirement and costs would be posted in the syllabus alongside information about the required course textbook. All of the required software was installed in the virtual lab by the vendor and each student only had to configure a single app on their personal computer or mobile device in order to access virtual lab, as shown in Figure 1. The virtual lab could be accessed from any computer or mobile device that was running any of the following operating systems: Microsoft Windows 7® or later, Mac OS X®, Android® or iOS®. Using the Microsoft Remote Assistance® feature in the virtual lab the instructor could access a student session in the virtual lab in real-time via a student-initiated invitation. This allowed students to receive personalized assistance with the hands-on exercise when needed.

Security

Potential security threats were the main issue noted against the public cloud deployment model (Jadeja, & Modi, 2012). Below we discuss some of the actions were taken in the implementation of the virtual lab in Phase 2 in order to mitigate against some of these potential security threats:

Firstly, the login credentials (usernames and passwords) that the students used to access the virtual lab were not associated with any of the other university-supplied computer accounts. The virtual lab vendors generated the login credentials for the class and provided them to the instructor via a secure download. The instructor then assigned the individual accounts to the students in the class. This provided a fail-safe mechanism in that, in the event of any security compromises to the virtual lab, such a security compromise would not lead back to other university systems used by the students. Furthermore, the assigned passwords were 'strong' passwords created using a password generator and students were not permitted to change these passwords.

Secondly, students were required to store all of their data files on the university cloud drive or on other cloud storage systems, such as Google Drive® or Microsoft OneDrive®. This ensured that students would have access to their data files even if they did not have access to the virtual lab. Storing the data files on the university cloud drive ensured that the data storage for the virtual lab solution met the university's data policies.

Finally, a sandbox environment was created where all of the analytics applications and utilities which the students needed to complete the hands-on exercises were only available via icons on the desktop. Students could not start any applications using other methods, such as, the command line or any other menus. This measure provided for easy access to the required applications and protected the virtual lab system against any potential misuses.

Figure 1: Accessing the Virtual Lab from an Android Phone.

RESULTS AND DISCUSSION

The change of making the setting up of the student-built remote labs into a graded assignment starting in the spring 2014 semester resulted in a 50.7% drop in number of lab setup inquiries. The number of lab setup inquiries remained relatively steady between the spring 2015 and fall 2015 semester, despite the differences in the enrollments in those semesters. This might suggest common lab setup issues with each cohort and that the use of the discussion boards to disseminate solutions to the issues may have helped lower lab setup inquiries. The use of the virtual labs starting in the spring of 2016 resulted in a further 54.7% drop in the number of lab setup inquiries. Many of these lab setup inquiries were resolved by the virtual lab solution vendor. From the student survey in Appendix A, all of the students found the instructions on how to setup access to the virtual labs easy to follow. From a scalability perspective, the virtual lab solution doesn't require the use of the IT Department and the students only have to setup access to the virtual lab from any of their personal computing devices. Furthermore, students only have to setup access to the virtual lab from any of their personal computing devices just once. Students in the spring 2016 did not have to spend anytime installing the analytics tools onto personal computing devices. With the virtual lab solution,

the course instructor is not required to have any experience with technical support. The course content inquiries, however increase with the class size.

Based on the survey results 80% of the students had previous experience with the use of remote desktops, which are the foundational technology for VDI-based virtual labs. This suggests that wider adoption of virtual labs in higher education may not be accompanied with a high learning curve on the use of such labs by the students. 20% of the students used multiple devices to complete a hands-on assignment in the virtual lab, this is a convenience not available with traditional computer labs. 80% of the students found that the virtual lab provided an environment that was: easy to navigate, easy to locate their applications and data files, and the applications responded quickly. All of the students surveyed found it easy to exchange their data files between the virtual lab and external systems, such as, Google Drive® or Microsoft's OneDrive®.

CONCLUSIONS

In this paper, the authors have described their experience with teaching a hands-on based online business analytics class by first using student-built remote labs, then followed by the use of a virtual computer lab. The challenges for both the instructor and the students with the remote labs were highlighted. This was then followed by a discussion of how those challenges were mitigated through the use of virtual labs based on the public cloud. Based on their experience, the authors concluded that public cloud-based virtual labs provide a cost effective and flexible avenue for teaching computer labs based courses in online programs. For student-based use of virtual laboratories there were no data security concerns identified by student relating to their own files or to the university's own systems, since the public cloud-based virtual labs operate in much the same way that the email system provided on public platforms such as Google Mail®, Yahoo Mail® or Hotmail®, for example. Public cloud-based virtual labs apply the same public email concept to traditional campus labs and therefore making the traditional campus lab portable albeit at a nominal fee. The efficacy of virtual labs for teaching, experimentation requires more trials beyond what this study was able to do in order to make more data available, before definitive results can be determined. However, preliminary evidence of their efficacy for students and their cost-sharing benefit between students and institutions is convincing.

Public cloud-based virtual labs have been in existence for more than a decade and they offer some attractive advantages over traditional campus labs and yet universities have been slow to adopt them. The question is what accounts for their lack of traction within the academic community when this case study has shown their attractiveness and at a time when more and more higher education institutions are offering online programs and competing for students worldwide? Intuitive logic would suggest that public cloud-based virtual laboratories should be a vital tool that enables institutions to offer previously restricted on-campus only lab-based courses, to online students worldwide. As such, they should be more attractive than they have been to date. Future research on the subject should perhaps examine the nature and causes of barriers to the adoption of public cloud-based virtual laboratories compared to traditional campus-based laboratories which at face value are more costly to acquire and maintain. Given their benefit of portability for all students regardless of whether they are online or on-campus students and their ability to widen access to lab-based courses for online students worldwide, the question that remains to be answered is; are public cloud-based virtual laboratories the future of lab-based teaching and learning or are they a passing fad?

REFERENCES

Auer, M. E. (2001). Virtual Lab versus Remote Lab. *20th World Conference on Open Learning and Distance Education*. Dusseldorf, Germany: ICDE.

Burd, S. D., Seazzu, A. F., and Conway, C. (2009). Virtual Computing Laboratories: A Case Study with Comparisons to Physical Computing Laboratories. *Journal of Information Technology Education: Innovations in Practice*. V. 8. Retrieved on September 6, 2017, from http://jite.org/documents/Vol8/JITEv8IIP055-078Burd693.pdf

Casini, M., Prattichizzo, D., and A. Vicino, A. (2007). Operating Remote Laboratories through a Bootable Device. *IEEE Transactions on Industrial Electronics*. V. 54, No. 6.

Chan, C., and W. Fok, W. (2009). Evaluating Learning Experiences in Virtual Laboratory Training through Student Perceptions: A Case Study in Electrical and Electronic Engineering at the University of Hong Kong. *Engineering Education*. V. 4, No. 2, pp 70-75. http://dx.doi.org/10.11120/ened.2009.04020070

Friedl, J. E. F. (2006). *Mastering Regular Expressions: Understand Your Data and Be More Productive*. Sebastopol, CA: O'Reilly Media.

Jadeja, Y., and Modi, K. (2012). Cloud Computing-Concepts, Architecture and Challenges. *International Conference on Computing, Electronics and Electrical Technologies*. Kumaracoil, India. https://doi.org/10.1109/ICCEET.2012.6203873

Neacsu, M. G., and Adascalitei, A. (2016). New Trends for Developing Blended Learning Courses. *The International Scientific Conference eLearning and Software for Education*. V. 3.

Oracle (2017). *Oracle SQL Developer*. Retrieved from Oracle Corporation website http://www.oracle.com/technetwork/developer-tools/sql-developer/overview/index.html

Son, J., Irrechukwu, C., and Fitzgibbons, P. (2012). Virtual Lab for Online Cyber Security Education. *Communications of the International Information Management Association*. V. 12, No. 4. Retrieved on September 6, 2017, from http://scholarworks.lib.csusb.edu/ciima/vol12/iss4/5

University of Arkansas. (2017, July 31). *Virtual Desktops*. Retrieved from https://its.uark.edu/services/network-devices/virtual-desktops/index.php

University of Southern California. (2017, February 28). *Viterbi MyDesktop*. Retrieved from https://viterbi.usc.edu/resources/vit/services/vdi.htm

Vikas, S., Gurudatt, K., Vishnu, M., and Prashant, K. (2013). Private vs Public Cloud. *International Journal of Computer Science & Communication Networks*. V. 3, No. 2, pp 79-83, 2013. Retrieved on September 6, 2017, from http://www.ijcscn.com/Documents/Volumes/vol3issue2/ijcscn2013030203.pdf

Wu, D., Fulmer, J., and Johnson, S. (2014). Teaching Information Security with Virtual Laboratories. In J. Carroll. (Eds.), *Innovative Practices in Teaching Information Sciences and Technology*. Cham: Springer. https://doi.org/10.1007/978-3-319-03656-4_16

Charles Mutigwe is an assistant professor of business information systems at Western New England University in Springfield, Massachusetts, USA. His research interests are cloud computing, RFID systems, and blockchains. In 2016 Dr. Mutigwe was an Innovation Fellow at the University of Massachusetts Amherst. The authors recently presented an abridged version of this paper at the Northeastern Association of Business, Economics and Technology (NABET 2017) Conference under the title "Using a Virtual Lab to Teach an Online Business Analytics Program".

Bruce Mtigwe is a senior lecturer in the Graduate School of Business at the National University of Science & Technology in Bulawayo, Zimbabwe. His research interests are technology entrepreneurship and international entrepreneurship.

Tendai Chikweche is a lecturer in the School of Business at Western Sydney University in Parramatta, New South Wales, Australia. His research interests are primarily on enquiry on innovation and entrepreneurship, small to medium size enterprise and emerging markets. He has extensively published refereed conference and journal papers in these subject areas.

APPENDIX A: Student Survey Questions and Responses

Response Rate: 31.25%

Did you find the instructions to setup the connection to the virtual lab from your device(s) easy to follow?

Answer	Response
Yes	100%
No	0%

What devices(s) did you use to access the virtual lab?

Answer	Response
Android Phone	0%
Android Tablet	0%
iPad	0%
iPhone	0%
Mac	20%
Windows Desktop Computer	40%
Windows Laptop	40%
Windows Phone	0%

Did you work on an assignment using two different devices?

Answer	Response
Yes	20%
No	80%

Please tell us if you agree or disagree with the following statements about your experience once you successfully logged on to the virtual lab.

Question	Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree
The environment was easy to navigate	40%	60%	0%	0%	0%
It was easy to locate what you needed	40%	40%	10%	0%	0%
It was easy to exchange my data files with other systems	20%	80%	0%	0%	0%
The applications responded quickly	20%	60%	0%	20%	0%

Before your experience with this virtual lab, had you used any other remote desktop solution to complete an assignment for school or work?

Answer	Response
Yes	80%
No	20%

