The Enhanced Virtual Laboratory: Extending Cyber Security Awareness through a Web-based Laboratory

Michael Black
mblack@southalabama.edu

Debra Chapman
dchapman@ southalabama.edu

Angela Clark
amclark@ southalabama.edu

School of Computing, Information Technology
University of South Alabama
Mobile, AL 33509, USA

Abstract

The Enhanced Virtual Laboratory (EVL) is the product of a Department of Defense grant to enhance outreach, research, and education for cyber security. Through web-based laboratories, EVL allows users to remotely experience interactive content with virtual machines inside a modern web browser providing constructivism-based, student-centered, active-learning instructional activities. Additionally, EVL provides a framework for content centralization that allows designated educators to share their knowledge and techniques of cyber security. By providing users an add-on free environment requiring no additional software to be installed, EVL provides a new and improved method of remotely delivering cyber security content to different users with variable depths of security knowledge.

Keywords: virtual lab, teaching lab, labs, web based

1. INTRODUCTION

It is well known that the United States is experiencing a critical shortage of qualified Information Assurance (IA) professionals. President Barack Obama (2009) stated that the "cyber threat is one of the most serious economic and national security challenges we face as a nation." In February 2014, the Obama Administration announced the launch of the Cybersecurity Framework, through Executive Order 13636, which directed NIST to develop a framework to serve as a guide for organizations hosting critical infrastructure to enhance their cybersecurity (Whitehouse, 2013; Obama 2013). Demand for trained cyber security professionals has outpaced other IT jobs by a significant margin. As noted in the US Department of Labor Occupational Outlook Handbook, the job outlook for 2012-22 for information security analysts is expected to grow about 37%, much faster than average. (Bureau of Labor Statistics, 2014). Formal cybersecurity education programs have been established as part of the National Initiative for Cybersecurity Education (NICE) as well to support increases in STEM literacy and build skills and competencies for the National workforce. (NICE, 2010) One method to attract well-qualified students to the field of IA is through experiential learning opportunities. (Maxim and Elenbogen,

2009; Verma, 2011; National Research Council, 2011).

The Enhanced Virtual Laboratory (EVL) is a content-driven web-based environment that delivers centralized cyber security content alongside live virtual machines for an interactive and seamless educational experience. EVL was developed as part of a Department of Defense capacity building grant - awarded to extend the institution's cybersecurity outreach, research, and education efforts. Through this system, educators can centralize their knowledge and techniques of cyber security. Users can obtain this shared content through a web interface to experience an environment that provides guided tasks alongside live, isolated virtual machines. This experience allows a user to actively learn, in a constructivist-based format, while using the actual environment instead of pre-defined emulators. Improved learning takes place through the use of authentic *real-world* relevant tasks and problems that require students to expand their knowledge through active problem solving (Ertmer & Newby, 2013). The integration of the technology with a structured learning activities provides for technology-based, student-centered educational opportunities that put students in charge of their own learning (Gayton & Slate, 2002-2003). Addtionally, EVL is fully functional inside any modern web browser without requiring extra add-ons or client-side software. By using this virtual laboratory environment, resources are not limited to a single use and can be accessed remotely, thereby extending the institution's Information Assurance outreach capabilities. The virtual lab can at any time be re-tasked to support the desired target audience in a timely manner. EVL provides a new and improved method of educating users about cyber security and computing related subjects.

2. BACKGROUND

Creating environments that facilitate active learning has become a priority for educational institutions, government agencies, and private corporations. Not only does active learning help maintain interest, especially with younger learners, but it also enables reinforcement of the material that is presented and helps provide a meaningful context, facilitating better knowledge retention (Perez-Sabater, et al. 2011; Korwin and Jones, 1990). Specifically, technology-based learning environments have been shown to create student-centered educational environments that provide students the opportunity to develop a sense of control and individual responsibility for their own individual learning (Tsai, 2012). Just as STEM programs in the form of field trips or summer camps provide active learning experiences to expose students to areas of interest such as biology and engineering, laboratory experiences in computing can expose students to the areas of cyber security, promoting inquiry-based learning, allowing exploration of these areas in greater depth. Also, laboratory modules can be used to spread general cyber security awareness, helping to increase knowledge regarding safe computing practices to the public at large.

Constructivist learning theory contends that students are responsible for creating their own knowledge and learning based on their individual experiences. Each student creates their own meaning through individual interpretation of the instructional materials (Ertmer & Newby, 1993). Meaning and knowledge is situationally determined as it is created by each learner as new information is integrated into their existing knowledge base. The student determines what they learn based on the information received from the outside world (Koohand, Riley, & Smith, 2009). Effective learning takes place when students apply existing knowledge to new or different situations. This requires learning tasks to be considered relevant and realistic to the student (Ertmer & Newby, 1993). Learning is an adaptive process that is based on prior learning, new experiences, and social interactions. The primary focus of constructivism is problem solving and higher order processing, not simply data representation (Koohang et al., 2009).

Hands-on learning in a lab environment provides constructivist-based activities that allows learners to experiment with different relevant *real-world* scenarios and to test configurations without the detrimental consequences of practicing security techniques in a live domain. Realistic experimentation of security concepts can be costly regarding hardware and software. While there are a myriad of third-party training solutions that are available, many of these products are not customizable for institutions to tailor as they wish. Also, many of these products require the user to install additional software to their machine and may be costly to obtain. Current products are built either to guide users through a scripted scenario or to simulate security techniques, thus limiting the ability to explore and experiment. These solutions are not practical when being used to provide an introduction to cyber security due to their heavy initial investment.

Physical laboratories require substantial resources regarding the financial investment in hardware and software as well as setup and configuration. Once a physical lab has been configured, it is usually dedicated to a small group of individuals and a specific learning objective. This does not allow for repurposing a lab promptly for other learning opportunities. Physical space and energy requirements paired with computing equipment cost makes it difficult for many institutions to justify the capital expenditure, especially when considering the respectively small number of students which would receive benefit from the physical computer lab. Also, individuals must be present to use the resources in the physical lab.

Another important caveat in teaching security-related concepts is the need for isolation. Isolation is extremely important to prevent any malicious use of the machines to the public Internet. Also, misconfiguration of network services by users learning security techniques also poses vulnerabilities. However, physical isolation means that internal machines must be accessed onsite, which creates a geographical limitation in delivering training and education.

Related work

The use of virtual laboratories is not a new concept (Hay, Dodge, & Nance, 2008). Organizations have leveraged the use of virtual laboratories for quite some time, employing a wide variety of approaches. As educational institutions with cyber security programs have also had to address increased growth in enrollment and constrained budgets, virtualization has enabled many to extend their learning experiences to support a larger number of individuals. These solutions not only allow these institutions to extend their resources to serve more students, but they also provide a method of extending the classroom and physical laboratory curriculum with a distance learning approach.

Training well-qualified cyber security professionals requires significant, meaningful hands-on active-learning laboratory experiences to build cyber skills to defend networks and to protect the nation's infrastructure. As noted by Zlateva et al., (2008) "…in few fields is the contrast between theory and practice as stark and as important to reconcile as it is in information security: cryptographic algorithms draw on the most abstract branches of mathematics while their correct (or incorrect) application decides vital problems ranging from the confidentiality of the nation's critical infrastructure to the privacy of personal information."

Padman et al. (2002) stated that "One of the main impediments to establishing an IA program is the requirement of a laboratory facility that will reinforce concepts taught in class with hands-on experiences." One early implementation illustrating the need to virtualize the cyber security curriculum started in 2001 by the United States Military Academy at West Point (USMA). The Information Warfare Analysis and Research Lab (IWAR) allowed students to practice theoretical topics on virtual machines inside a private network. However, IWAR was only accessible through local workstations inside a closed environment.

Other implementations, such as the Rochester Institute of Technology's solution, have made use of remote desktop protocol (RDP). This solution required the use of Microsoft Terminal Services and Remote Assistance to be employed for students to connect to running virtual machines. This configuration necessitated extra client-side software and for users to know the IP address of a specific virtual machine to connect to the system (Border, 2007).

The Advanced System Security Education Research and Training (ASSERT) lab created by the University of Alaska at Fairbanks provided cyber security education through distance learning by accessing virtual machines using a web-based portal (Nance, et al. 2009). ASSERT gave users the ability to practice cyber security techniques but did not provide detailed instructions alongside the virtual machine to assist users. A system that could provide users with detailed lesson plans would allow that user to further their understanding of topics, which are being practiced inside the virtual machine. The Open University of Catalonia employed a system to teach networking to remote students through the development of the Virtual Networking Laboratory (VNLab). This system employs a combination of several different systems to provide a user with a laboratory experience, which includes Cisco NETLAB+ (Prieto et al., 2008). Although many individual systems can be combined to create an e-learning system, a remote laboratory can be built to deliver all the necessary functionality without depending on other systems and vendors.

3. OBJECTIVES

A fundamental goal of the EVL project was to provide the institution with a means to support

active-based learning activities as outreach to prospective and current students, as well as the DoD and other federal, state, and local agencies through our cyber security center. Once fully completed, training modules will be available for DoD training centers, government agencies, and other CAE institutions that do not have their own virtualization infrastructure. By using a virtual environment, the lab is not limited to a single use and also allows the institution to avoid many of the hardware and geographic limitations of physical computer lab environments.

To accomplish these goals, the EVL project had three main objectives: 1) provide a fully web-based interface, 2) deliver authentic and isolated virtual machine environments, and 3) centralize and deliver content inside an environment with minimal complexity. Each of these is discussed further in the following subsections.

Web-Based Interface

EVL's web-based interface was designed to be accessible by any user, provides a control mechanism for virtual machines, and delivers a centralized content driven environment. EVL can operate in all modern web browsers that support HTML5 elements. The web-based interface allows all users to access the interface without needing specialized software or hardware remotely. Through the web interface, EVL provides server-side controls and client-side initiation for the virtual machine sessions by utilizing a customized engine, web-sockets proxy, and the noVNC project. Through this implementation, users experience a seamless, platform-independent session using isolated virtual machines without needing extra client-side software or add-ons. This environment prevents EVL from needing client-side administrative permission to operate. Lastly, the EVL interface allows educators to contribute and collaborate on content that can be shared to authenticated users. Additionally, these contributors are provided a specialized interface for customizing virtual machines to be connected with their content. The EVL interface is a fully featured product that allows users to receive dynamically rendered content that connects with live isolated virtual machines. Figure 1 (appendix) provides a graphical depiction of a content example and Figure 2 (appendix) illustrates the content creation interface. As shown in Figure 1 (appendix), the user is provided with detailed lesson plans explaining the topics which are being practiced inside the virtual machine. This allows the user to have the machine interface and the instructions side-by-side, eliminating the need for separate online or physical documentation.

True Virtual Environment

EVL uses a web-based interface to deliver embedded, isolated virtual machines to any authenticated user. Emulation can be used imitate aspects of a computer operating system to support specific learning objectives providing relevancy. Emulation can be a functionally acceptable technique but carries many far-reaching consequences. The administrative process associated with creating lab exercises when using emulation mandates that each exercise be manually developed to fulfill the goals of the exercise. Each lab exercise would be limited to a finite set of predetermined commands and responses. Virtualized operating systems give the user a *real-world* experience as they respond exactly as if the operating system was running on dedicated, physical hardware resources. Also, the user is presented with the same errors and responses, just as they would by using a physical machine. The web interface achieves this feature without needing additional client-side software or add-ons due to its customized engine. The customized engine masks the complexities of virtualization from the client-side to allow the user to focus on the learning objective. The engine is responsible for automating tasks associated with the virtual infrastructure. Traditionally, the overhead associated with the configuration of the educational environment has been a burden on faculty and students alike. In standalone virtual environments, educators usually provision virtual machines on a per student basis. The engine simplifies this process by dynamically provisioning virtual machines based on learning objectives rather than a per student basis.

Minimal Complexity

A top priority of the Enhanced Virtual Laboratory engine is removing the complexities found when interacting with virtual infrastructure. The ability to remotely access a standalone virtual environment can sometimes require the use of virtual private networks or utilize public IP addresses for connectivity. The engine simplifies this issue by aggregating connectivity to virtual machines through a dedicated proxy. The proxy manages the connections between the client and virtual machine without the overhead of specialized networking or use of proprietary applications distributed with the virtualization platform. EVL abstracts the complexities of delivering virtualization providing a seamless experience for any depth of user knowledge and centralizes contributed content from educators into one location. EVL's web interface uses its customized engine to auto generate and remove virtual machines as a lesson is started and

completed. This abstraction prevents the user from needing to have any pre-existing knowledge of virtualization. Instead, EVL handles all functionality on the server side to give the user a straightforward experience. In standalone virtual environments, educators must provision virtual machines for each user. The engine simplifies this process by dynamically provisioning virtual machines based on a selected learning object rather than a per student basis. An educator is only required to associate a virtual machine with a learning exercise at the time the exercise is created. EVL is content driven by educators who collaborate and create learning modules in one centralized location. This allows educators to create the content once and deliver to many students without being constrained to certain types of client operating systems or configurations.

4. SOLUTION

EVL uses the latest web technologies to deliver a fully featured web interface and content management system (CMS). These technologies include the latest web programming languages including HTML5, CSS3, JavaScript, and JQuery. Through these updated web languages, EVL can deliver a responsive learning system.

Additionally, HTML5 utilizes new elements that are essential to EVL. Canvasses and web-sockets are the keys to the open source project noVNC by Kanaka. The EVL system makes use of another open source project as its CMS framework, Django. This Python-based framework allows EVL to render client-side HTML5 web pages through server-side Python functions dynamically. Django also provides database abstraction and standardization for EVL. Through Django, EVL users cannot directly query the database for security purposes, and EVL can continue to be updated without needing content developers to understand proprietary code.

Using Python as its server-side scripting language, EVL natively imports its engine's python-based classes. Through the engine's python classes, EVL generates, edits, and removes virtual machines on a XenServer hypervisor. Without needing any client-side action, the engine allows EVL to deliver any virtual machine to the user through a web sockets proxy. The proxy accepts and directs requests based on a unique identifier that is generated by the engine.

The entire process from request to virtual machine connection is completed in four steps. First, an authenticated user request specific content from EVL. Then EVL activates its engine by requesting it to generate virtual machines that have been assigned to the requested content. Third, the engine responds with the IP address, port number, and unique token identifier for the requested virtual machines. Then EVL responds to the client with the retrieved information in the form of an HTML5 web page. Lastly, the noVNC JavaScript code is executed in the background on the client side to start all VNC sessions into the virtual machines. When all the steps are completed, the client side web page has both the content and virtual machines embedded alongside each other.

As figure 3 demonstrates, a user activates the EVL CMS by navigating to the URL of the web page. Following the user properly authenticating with the system, data is retrieved for the user to make a task selection. This selection determines what specific data is retrieved and which virtual machines the EVL CMS will ask the EVL Engine to create. Once all of the data is retrieved a page is rendered and allows the user to interactively use the content alongside the virtual machines until that user is ready to start over. Together the CMS and engine work in unison to deliver a fully featured virtualization experience without being dependent on any extra client-side software or add-ons.

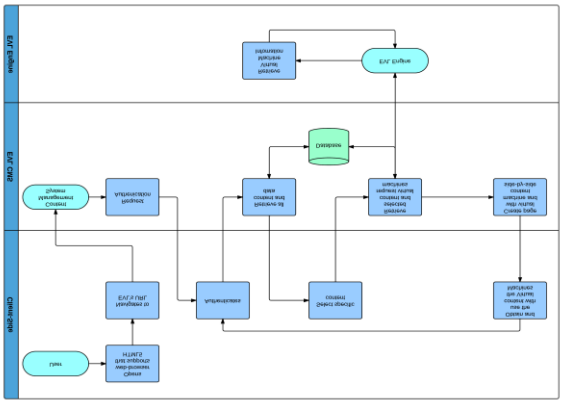

Figure 3: Conceptual Model

5. DISCUSSION

The EVL development group utilized a team-based approached that divided the project into two equal parts. One team designed the interfaces or the content management system (CMS), and the other team designed the engine. Working in unison, the teams would build features that were tested against the other team's components until no errors or design flaws were

found. This partitioned but collective - design process allowed each team to focus on separate objectives but still gather feedback and requests from each other. As shown in Figure 4, the content management system and the engine are two separate project components that combine to form the Enhanced Virtual Laboratory.

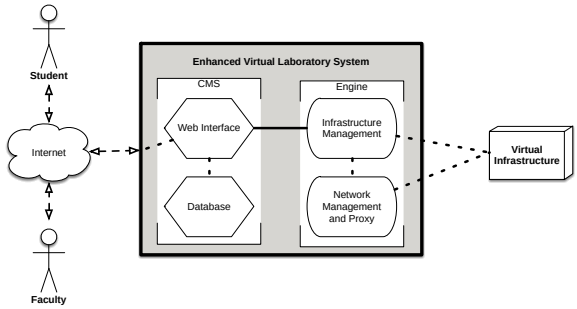

Figure 4: Project Architecture

The teams have created a product that is usable from any location, system, or browser. However, EVL still has a few requirements for operation. First, the user must have an Internet connection. The system is fully web-based, meaning that to access the system a connection must be established via the Internet. Second, the user must be using a recent web browser. Updated browsers utilize HTML5 which is essential to delivering the embedded virtual machines. Third, EVL is content driven and must have educators adding content for learners to use. Fourth, the client-side firewall must allow requests and responses on port 8080. As all requests to the virtual machine VNC connections are routed through a proxy, a port other than the default web traffic port is used to make those connections.

6. USABILITY AND PERFORMANCE TESTING AND EVALUATION

EVL was pilot tested in a classroom of twenty students. During this test session, the students were instructed to complete a series of *real-world* relevant exercises that would guide them on how to set up a network with different operating systems. Each constructivist-based lesson progressively built on the previous one and increased the virtual machine count by one. Therefore, on the first lesson, the students were tasked with configuring a Linux network card using one virtual machine. Next, they configured Windows network cards using two virtual machines, a Windows 7 and Windows XP. Lastly, the students used a lesson that instructed them to configure both Linux and Windows network cards using three virtual machines: Ubuntu Linux, Windows 7, and Windows 8. Almost all students were able to complete the first and second lessons successfully. As students began to move into the third and final lesson, a bottleneck occurred where the engine did not connect the users and virtual machines. After reviewing the data collected during testing, within 31 minutes, 42 virtual machines were created and successfully connected to by students. However, it was found that 122 virtual machines were created that were never successfully connected to by students. The testing session revealed an issue in the engine's proxy service that has now been resolved. Future testing has been scheduled to confirm that the problem is resolved and that no other performance issues exist with the current configuration.

7. LESSONS LEARNED

EVL went through many design changes throughout its creation; however, the main goal to extend cybersecurity awareness stayed consistent. Each feature of EVL has been made to enhance accessibility, improve reliability, and simplify future contributions. For example, the EVL original design was completely propriety. This design provided all the needed features but made contributing to the system a complex and time cumbersome task. Therefore, the EVL scope switched to using an open source framework that would simplify the process of allowing multiple parties to add to the system in the future.

The EVL design still has certain limitations. With a finite amount of hardware resources, EVL can only deliver a certain number of virtual machines before exhausting the available resources. Also since the system is primarily built for outreach, there are no existing mechanics for grading or assuring that a user completed the content. Additionally, EVL does not allow virtual machines to persist among user sessions. This means when the user switches between content the virtual machines are refreshed. Although these limitations are inside the current version, with EVL's modular design updates can be implemented to lengthen the scope of the project.

8. CONCLUSIONS AND FUTURE WORK

EVL is a new and improved method of delivering specialized content alongside virtual machines for effective education, research, and outreach. EVL provides a fully OS independent and modern web-browser independent interface that educators can use to contribute or collaborate on content to build student-centered, active-learning, constructivist-based educational activities. EVL is

a system that delivers a seamless experience since it uses the latest web technologies and server-side scripting to require very minimal resources from the audience's machine. The active-learner, constructivist-based, relevant activities provide effective hands-on learning. Additionally, EVL has a standardized framework which allows updates and feature additions to be contributed. These contributions can come in different forms. With more hardware resources, EVL can deliver more virtual machines to larger audiences. Since the EVL system works as a tool for centralization, contributors could add resources to the environment. Then they could use the system for distributing to their audiences without needing individual infrastructures. With more content contributors adding to EVL, the system can provide different types of education to a wider range of audiences. The EVL system was built for cybersecurity awareness outreach, but due to its nature of using live virtual machines, many different disciplines could be taught within the system, such as secure software engineering or applications development. EVL can support a community of educators from different disciplines working together to expand outreach, research, and education.

An additional area of interest for future study could be to examine how students feel use of this type of system impacts their learning and helps maintain or further their interests in the study of cyber security. Once EVL is expanded with additional training modules, this is the goal of the research team in understanding its effectiveness and impact on student learning outcomes.

9. REFERENCES

Border, C. (2007). The development and deployment of a multi-user, remote access virtualization system for networking, security, and system administration classes. SIGCSE Bull., 39(1), 576–580. doi:10.1145/1227504.1227501

Bruce R. Maxim and Bruce S. Elenbogen. (2009). Attracting K-12 students to study computing. In *Proceedings of the 39th IEEE international conference on Frontiers in education conference* (FIE'09). IEEE Press, Piscataway, NJ, USA, 119-123.

Bureau of Labor Statistics, U.S. Department of Labor, *Occupational Outlook Handbook, 2014-15 Edition*, Information Security Analysts, on the Internet at http://www.bls.gov/ooh/computer-and-information-technology/information-security-analysts.htm (last visited *March 03, 2014*).

C Pérez-Sabater, B Montero-Fleta, M Pérez-Sabater, B Rising (2011). "Active learning to improve long-term knowledge retention" Actas del XII Simposio Internacional de Comunicación Social, 75-79.

Ertmer, P., & Newby, T. (1993). Behaviorism, cognitivism, constructivism: comparing critical features from an instructional design perspective . *Performance Improvement Quarterly*, 50-72.

Gaytan, J. A., & Slate, J. R. (2002-2003, Winter). Multimedia and the college of business: A literature review. Journal of Research on Technology in Education, 35(2), 186-205.

Hay, B., Dodge, R., & Nance, K. (2008). Using Virtualization to Create and Deploy Computer Security Lab Exercises. In S. Jajodia, P. Samarati, & S. Cimato (Eds.), Proceedings of The Ifip Tc 11 23rd International Information Security Conference (Vol. 278, pp. 621–635). Springer US. Retrieved from http://dx.doi.org/10.1007/978-0-387-09699-5_40

Koohang, A., Riley, L., & Smith, T. (2009). E-learning and constructivism: From theory to application. *Interdisciplinary Journal of E-Learning and Learning Objects*, 91-109.

Korwin, Anthony R.; Jones, Ronald E. (1990). "Do Hands-On, Technology-Based Activities Enhance Learning by Reinforcing Cognitive Knowledge and Retention?" *Journal of Technology Education*, v1 n2 p26-33 Spr 1990.

Nance, K., Hay, B., Dodge, R., Seazzu, A., & Burd, S. (2009). Virtual laboratory environments: methodologies for educating cybersecurity researchers. Methodological Innovations Online, 4(3), 3–14.

National Initiative for Cybersecurity Education (NICE) (2010\) Relationship to President's Education Agenda, 19 April 2010, (last visited *March 18, 2014*) http://www.whitehouse.gov/sites/default/files/rss_viewer/cybersecurity_niceeducation.pdf

National Research Council. (2011) *Learning Science Through Computer Games and*

Simulations. Washington, DC: The National Academies Press, 2011.

Obama, Barack (2013). "Executive Order – Improving Critical Infrastructure Cybersecurity" Retrieved from http://www.whitehouse.gov/the-press-office/2013/02/12/executive-order-improving-critical-infrastructure-cybersecurity, (last visited *March 18, 2014*).

Office of the Press Secretary (2014). "Launch of the Cybersecurity Framework" Retrieved from http://www.whitehouse.gov/the-press-office/2014/02/12/launch-cybersecurity-framework, (last visited *March 18, 2014*).

Padman, Vikram, and Nasir Memon (2002). "Design of a Virtual Laboratory for Information Assurance Education and Research." Workshop on Information Assurance and Security. Vol. 1. 2002. 1555.

Prieto-Blázquez, J., Arnedo-Moreno, J., & Herrera-Joancomartí, J. (2008). An Integrated Structure for a Virtual Networking Laboratory. IEEE Transactions on Industrial Electronics, 55(6), 2334–2342. doi:10.1109/TIE.2008.921231

Schafer, J., Ragsdale, D. J., Surdu, J. R., & Carver, C. A. (2001). The IWAR range: a laboratory for undergraduate information assurance education. Journal of Computing Sciences in Colleges, 16(4), 223–232.

The White House. President Barack Obama (2009). Remarks by the President on Securing Our Nation's Cyber Infrastructure." Retrieved from http://www.whitehouse.gov/issues/foreign-policy/cybersecurity (last visited 03 Mar. 2014).

Tsai, S. C. (2012). Integration of multimeida courseware into ESP instruction for technological purposes in higher-education technology. *Educational Technology and Society, 15*(2), 50-61.

Verma, A; M. Talaiver; S. McKinney; D. Dickerson; S. Dwivedi; D. Chen (2011) "Attracting K-12 Students towards Engineering Disciplines with Project Based Learning Modules." *Proceedings of the ASEE Conference*, Vancouver, Canada, June 26-29, 2011.

Zlateva, T, L. Burstein, A. Temkin, A. MacNeil, and L. Chitkushev (2008). Virtual Laboratories for Learning Real World Security, Proceedings of the Colloquium for Information Systems Security Education, University of Texas, Dallas, TX, June 2008.

Appendices and Annexures

Figure 1: Content Example

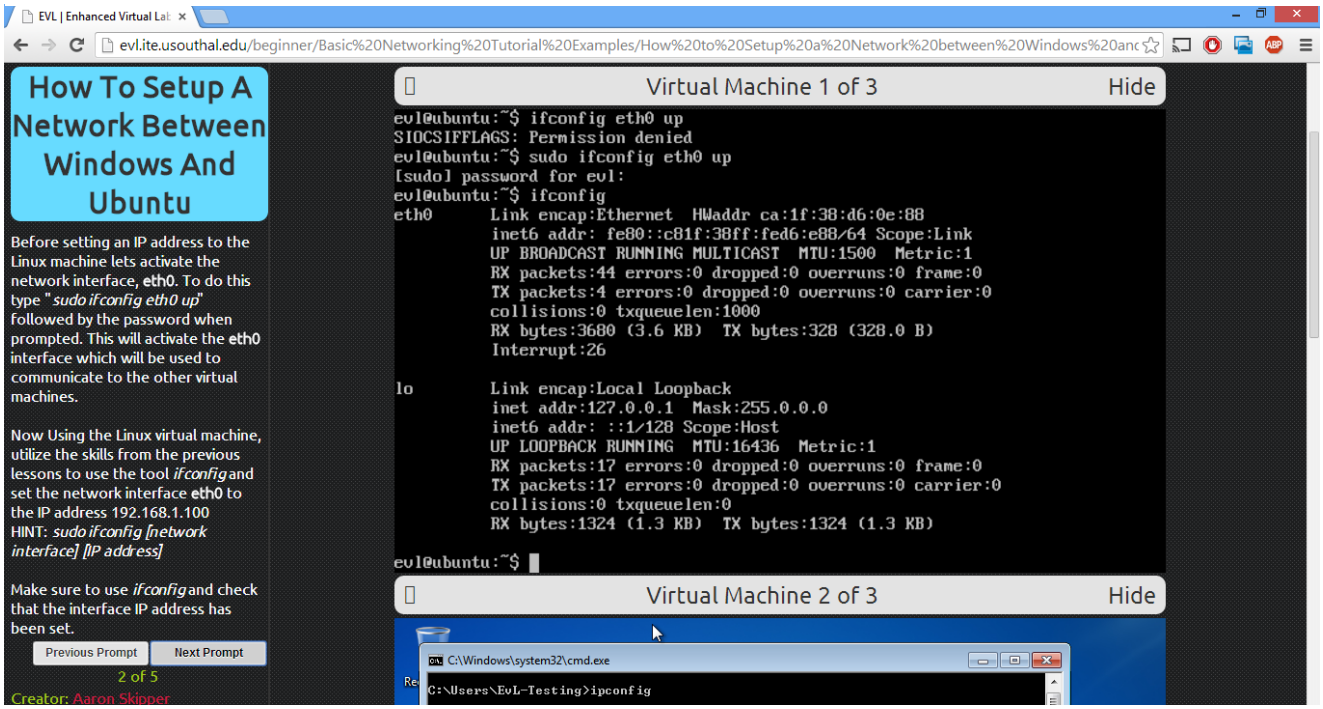

Figure 2: Content Creation Example