ISSN: 2167-1435
v30 n2522

Creating a Virtualized Environment for Large-Scale Hands-On IA Education

Kyle Cronin
kyle.cronin@dsu.edu

Wayne Pauli
wayne.pauli@dsu.edu

Michael Ham
Michael.ham@dsu.edu

Dakota State University
Madison, SD 57042

Abstract

Hands-on security labs have become an essential component in any technical information system or cybersecurity program. Critical thinking skills are gained through reinforcing discussion and lecture, reinforcing and improving the skillset of the student. Improving our cybersecurity workforce starts at our college graduates, properly preparing students for today's work environment requires hands-on lab exercises. Many technical hurdles exist when creating a hands-on lab environment to be used for educational purposes. We discuss several solutions that are currently available on the market, both free and commercial. Finally, we discuss the design of our virtual lab, both hardware and software, and how we use it to support the academic needs of 400+ students while having little or no overhead in terms of faculty development time.

Keywords: cybersecurity, hands-on learning, remote lab, virtualization, curriculum.

1. INTRODUCTION

The realm of cybersecurity and information assurance (IA) has brought challenges upon educators that were previously unimaginable. This challenge has stemmed from the notion that providing technology skills requires a hands-on component. Simply put, students need to learn by doing in these situations.

In this paper, we will describe various solutions that currently exist for exposing students to hands-on lab environments, describe the solutions that we have implemented, and describe the challenges we faced with our requirements at Dakota State University. To support our findings, we will establish a basis for the needs of hands-on educational settings. The end goal of building such environments is to improve the success of our students, and we feel that others can take away from our lessons learned to improve the cybersecurity posture of students across the nation.

The federal government has attempted to answer the call for an increase in cybersecurity professionals by implementing the National Initiative for Cybersecurity Education (NICE) framework. As described by Paulsen, McDuffie, Newhouse, and Toth (2012), NICE is established in order to create a sustainable cybersecurity program aimed at improving skills and

awareness of the workforce. The long-term goal of the NICE framework seeks to improve the overall cybersecurity posture of the United States.

Developing and maintaining cybersecurity skills starts at the education level. The NICE framework establishes that those involved in cybersecurity need to have a practical experience, a hands-on experience. This allows students to have a more in-depth knowledge of their area of study when they enter the workforce.

2. THE NEED FOR HANDS-ON EDUCATION

The growth of the cybersecurity field is not unknown. The demand for well trained professionals in this field is still exponentially growing with little or no end in sight (Ciampa, 2012). The challenges faced with training these professionals are clear. No longer is a policy or a contract stopping attackers. Hackers can easily hide their traces, or in the case of "hacktivist" groups such as Anonymous and Lulzsec, publish their attacks to the world (Olson, 2012).

A shift in mentality occurred resulting from the publicized success of hackers from being simply malicious to trying to steal data or damage an entity's reputation. In order to protect networks, individuals must be trained in highly technical fields of study (Lukasik, 2011). This training must coincide with practical experiences. Simply put, if a learner does not know *how* to apply security methodologies and best practices, their security implementations will fail. This brings forth the need for hands-on education.

Typically, hands-on learning has meant that a student would verify or test a concept derived from a textbook in a real world setting (Bork, 2000). Bork (2000) does establish that this paradigm should be challenged; repetition of a textbook's content does not necessarily equate to a student understanding a concept. A shift in learning methods by which students are engrossed in the theories, concepts, and practical applications is essential in order to properly prepare students for their future careers.

Due to the overwhelming number of variables and lack of standardized evaluation of success, the overall impacts of hands-on learning can be nearly immeasurable. Martin and Woodward (2012) establish that hands-on learning benefits students in the cybersecurity field by not only verifying concepts, but by engrossing students in problem solving situations. Cybersecurity education lends itself to real world exercises. Academic or simulated labs, while beneficial, still lack a critical problem-solving component. In essence, students will learn technology better by becoming fully immersed in the good and the bad.

Martin and Woodward (2012) further support these notions, students should learn in situations where they are required to engage several critical thinking skills. Chatmon, Chi and Davis (2010) corroborate this concept by defining the term "active learning". Active learning is essentially the series of steps that students take when increasing their levels of understanding from learning about a concept, actively repeating the concept, modifying the concept, until they finally have a creative understanding of the concept. By immersing students in these real world, professional, situations, the future workforce in the cybersecurity field will be greatly enhanced.

Challenges to Learning-by-doing

Hands-on labs are clearly established as the best method for teaching technology to students. However, as great as the methods may be, they are not without cost. Establishing the lab, maintaining the operational status, discovering budgetary or other tangible assets, and scheduling resources are often tasks that fall upon faculty in an academic setting. These challenges can be observed through a review of lab implementations of other universities (Ayers, 2010; Bhagyavati, 2006; Chatmon et al., 2010; Sharma, Murphy, Rosso, & Grant, 2012).

Practical Experiences

Many universities, ourselves at Dakota State University included, start their information assurance labs with a physical lab of extra or out-of-date PCs. Typically phased out by the IT department, these machines have been written off and are ready to be poetically "thrown to the hackers". The decommissioned machines, paired with some old networking gear compose a security lab. Cheap, reasonably reliable, and simple, all things that educators want to hear when considering adoption of a security lab.

These labs are not met without their own unique challenges. Provisioning of the lab typically requires drive imaging (the process of cloning the data from one hard disk to another), which

is easy but time consuming. Considering the time it takes to image a drive verses compromising an operating system with an offensive penetration utility, classrooms spend more time rebuilding the lab than they actually do using the lab.

Aside from the simple issues of time spent in the classroom, several other issues exist with this model including scalability and delivery. A physical lab requires more machines in order to grow. In our situation, extra classroom space for an expanded lab of old computer hardware was not easy to find. Classroom enrollment continually grows, which contends with locations purposed for cyber security labs. Breaking individual classes into multiple sections was an ineffective means to ease this issue because the faulty member would now be doubling the amount of preparation time in order to prepare a lab environment.

The advent of inexpensive and easy to use desktop virtualization was a simple solution to this problem. Instead of installing and attacking the host operating system, running a type 2 hypervisor such as VMware Workstation or VirtualBox seemed like a viable solution. This eased our growing pains for some time; lab preparation was as simple as copying a virtual machine down to the desktops in a lab.

Unfortunately, with all solutions there are still setbacks. Having students use dedicated lab machines did help to reduce the amount of licensing issues encountered, but students were still able to walk out the door with a virtual machine containing both host and software licenses. Solutions to this problem exist but they ultimately did not alleviate the ultimate problem with using type 2 hypervisors: distance students.

Running virtual machines on a desktop is a great solution, until you try to distribute multiple copies of operating systems across the internet. Couple the distance issue with the notion that distance students are introducing compromised virtual machines onto their home network and issues begin to arise. Additionally, all control is lost over the actual licenses contained within the virtual machine.

The Requirements

With years invested into finding an effective means of hands-on education, a better solution was required. Key sets of requirements were developed. At first, these requirements were more of a wish list, a list of features that an ideal hands-on lab would contain. As further research was conducted however, this wish list started to become more and more attainable:

- The lab must be internet accessible.
- The lab must be the same for on and off campus students.
- The lab must be self-contained.
- The lab must allow self-provisioning.
- The lab must perform well.
- The lab needs to be easy to use.

These requirements, while daunting, were very practical for the needs of our on-campus and online education needs. The notion of delivering the lab over the internet, yet keeping all network traffic for the lab self contained posed our biggest challenge to overcome.

3. AVAILABLE OPTIONS

Virtualization was a key winner in this debate of using a physical lab or a virtualized lab environment because of the accessibility over the internet and time savings in lab preparation. The critical issue came over deciding which virtualization platform to go with. In an effort to broaden the knowledge in this area, we will overview several popular virtualization platforms that we evaluated and give our evaluations of each platform. Ultimately, we will describe our choice, the components and the hardware used, and why.

While several virtualization platforms are available several key features set each apart. Most vendors (Microsoft, VMware, and Citrix) offer their base hypervisor at little or no cost. This discussion is not meant to compare one hypervisor's specific features to another, but the aspects of how they can be applied to our particular educational environment. To this end, in order to meet our needs a management solution was as important, if not more so than the actual hypervisor.

Microsoft

Microsoft's hypervisor, Hyper-V, was initially released over windows update in the summer of 2008 (Howard, 2008). Unfortunately, our project was already getting off the ground at this time.

ISSN: 2167-1435
v30 n2522

Selecting a newly-released hypervisor that did not have large-scale management capabilities was a choice we decided against.

Hyper-V is managed through Microsoft's System Center Virtual Machine Manager (SCVMM) suite. This application is designed to manage multiple Hyper-V systems from a single interface. In order to allow users to self-provision virtual machine configurations, an additional component called Self-Service Portal is needed.

As previously stated, Microsoft's solution was passed over simply due to timing. If an institution that is currently a Microsoft-only institution was considering a large scale implementation, Microsoft's solution would be highly recommended due to its feature set and level of support.

Netlab

NDG's NETLAB+ is a hardware-based appliance that provides management of both virtualized systems in addition to outside lab devices. These devices include switches, routers, switched outlets, and firewalls ("NDG NETLAB+ System Requirements," n.d.). For institutions looking to teach specific curriculum, such as: Cisco Networking Academy, VMware IT Academy, Linux+, and many others, the NDG NETLAB+ platform is a great solution.

NETLAB+ requires VMware's ESXi and vCenter management software in order to virtualize the operating systems. It should be noted that NETLAB+ is specifically a management and access interface for other hardware devices. Faculty and students can leverage this centralized appliance to complete a lab using physical firewalls integrated with virtualized operating systems and software applications.

NETLAB+ does have limitations, which were the primary reasoning behind our choice against the platform. While the features offered would be a great benefit to our program, a single NETLAB+ device is limited to 32 active "pods", which equates to a limitation of 32 concurrent labs being completed at one time("NDG NETLAB+ Configuration Maximums,"). In addition, the device is limited to a maximum of 160 active virtual machines.

Overall NETLAB+ is a great product, especially given its integration with hardware devices. However, because of the size and scaling limitations, NETLAB+ was not a good fit for our environment. NETLAB+ would be a great appliance to review for any institution that is not planning on large-scale deployment of lab environments.

VMware

VMware's basic offerings in the virtualization environment provide the basic features that would be expected of a lab environment. VMware's Hypervisor, called ESXi, is often teamed with its management application, vCenter. Together these components are marketed as a product unit called vSphere. Upon initial evaluation, the vSphere product line did not exactly meet the needs we were seeking; however, an additional product, Lab Manager, exactly met our needs.

The Lab Manager product is not longer available from VMware, but it has been replaced by VMware's vCloud Director. vCloud Director aims to abstract the management of the virtual infrastructure from the end user experience (Krieger, McGachey, & Kanevsky, 2010). This level of abstraction is necessary in order to fully automate and manage back end hardware while providing users a simple and easy to use interface.

The product lines of vSphere and vCloud Director have near limitless potential in terms of growth. Limited by 10,000 active virtual machines and 2,000 physical hypervisors, the vCloud platform is designed for scalability ("vCloud Director 5.1 Configuration Maximums," 2012). vCloud does not natively interface with outside hardware devices, such as switches and firewalls, but vCloud is able to natively isolate network traffic by creating virtual switches for all virtual machine traffic. This means that the network traffic of one student's lab will not interfere nor be detectable with another.

4. OUR APPROACH

Software

The approach we currently operate runs VMware's vSphere platform with vCloud Director. vSphere itself is an excellent management application for hypervisors, but does not easily support user self-provisioning or remote access solutions. vCloud Director supplements the missing functionality from vSphere.

ISSN: 2167-1435
v30 n2522

VMware vCloud Director is used to manage and provision resources. Multiple "Organizations" exist, granting different permissions and resources to users. The "Learn" organization grants faculty access to provision and configure groups of virtual machines for in-class labs or assignments. As a part of this configuration, faculty can configure a lab, referred to as a vApp, to use a class-wide network (essentially one large broadcast domain) or individual isolated networks. The decision is up to the faculty member at the time of the lab creation.

Once a faculty member has created a base configuration it is copied into a catalog. Multiple catalogs exist, one for each class, with read permissions set. This prevents a student from accessing a lab from another class. From this point, the faculty member's work is essentially finished. The management and provisioning of the individual student labs is a student-driven process.

When students access the Learn organization, they can only provision vApps from the catalog. This prevents students from using essential hardware resources needed for classroom virtual machines for personal use. Students can check out and deploy any vApp that they have permission to access at any time. Once the vApp has been checked out to the student, timers prevent the vApp from using resources for an extended period of time. In our situation, a student's virtual machines are only allowed to be online for a period of up to 12 contiguous hours. If the student chooses, they can renew this lease. Otherwise, their virtual machines are powered off to release the CPU and memory resources. After the virtual machines are unused for a 2-week period, they are automatically purged from the system releasing the storage resources consumed.

The key takeaway from this design is the amount of faculty effort required is minimized. The faculty member's only requirement is to create the lab, install any software needed in the virtual machine and copy it to the catalog. The process of actually provisioning the lab to students is student-driven.

The time taken for a student to deploy a lab is relatively minimal. A lab configuration consisting of 5 virtual machines with three separate virtual switches will take 2 to 3 minutes to deploy and fully power on. This gives the instructor ample time to give any directions and precursors to the hands-on lab.

Finally, the issue of student access: vCloud Director is easily accessed over the internet. In order to prevent direct interactions between internet users and the hypervisor, vCloud implements a proxy service. Users connect over an SSL encrypted session to the vCloud server which authenticates and proxies the connection to the hypervisor. Once connected to the hypervisor, the user is sent the rendering of the virtual machine's current screen. This eliminates the need for remote connection software to pass packets into the virtual machine, such as with Microsoft's RDP or the VNC protocol.

Virtual machines are fully isolated preventing outside entities from being passed in and also preventing anything inside the virtual environment. In addition, the virtual machines are running and are stored on the servers at the university, not the users machine. This prevents software or OS licenses from being misappropriated.

Hardware

Cloud computing has brought us to the realization that software can exist independently of hardware. With this notion, choice of server vendor is up to the end-user.

Our hardware environment consists of:

- 16X HP DL380 Servers
- 3X Apple Mac Pros
- 144TB RAW disk storage
- 2X Cisco 9148 Fibre Channel Switches

Each DL380 is configured with between 72 and 144 GB of RAM and dual quad or hex-core processors. In addition to the DL380s, 3 Mac Pros are available with 64GB of RAM and dual hex-core CPUs. The Mac servers are used to legally virtualize MacOS for our students.

All servers, including the Mac Pro, connect to our fibre channels switches using dual 8GB fibre channel. Our storage environment was initially iSCSI but because of our massive expansion we migrated to fibre channel.

5. CONCLUSION

End User Experience

All students access the same interface via web browser, no matter if they are on or off campus. Deployment times for in-class labs still remains less than 10 minutes for a class of 50 students.

One major consideration for the success of virtual environments is the performance. The hardware we utilize is able to run approximately 400-500 virtual machines simultaneously with limited performance impact. It should be noted that this value varies significantly from published specifications for the business realm. This is due to the fact that our student virtual machines are not running necessary business software, such as anti-virus. Typically AV software causes a huge demand on back-end storage. Since each lab environment is fully isolated (and often deliberately infected with malware) antivirus software is not necessary.

Growth to Today

The key maintenance issue we encountered with growth was the performance of our storage array. At the time we were operating 24X 600GB SAS drive in a RAID 5 volume. Unfortunately this configuration caused an enormous impact on virtual machine performance. Since then, we have migrated our storage to connect over 4X 8GB fibre channel connections. Additionally we now operate 72 hard disks, configured in 5 RAID10 pools. This gives us the fastest drive performance possible.

Forward looking, we will be upgrading our storage to enable the use of storage tiering with solid-state disks. This transition will be slow due to extremely high costs of solid-state drives at the time of this writing.

As of spring 2013, our lab environment supports approximately 300 students in 8 separate courses. The lab environment's future is expected to continue to grow to approximately 400 students in 14 courses in the fall of 2013. This current workload is roughly 60% online/distance education students and 40% on-campus students. Today's environment gives both on-campus and online students the same end-user experience.

Future growth is expected as our classes continue to grow. Due to the modularity of the lab's design, additional hardware can easily be added. Additional storage can be configured in the SAN, additional memory and CPU power can be added with additional servers.

6. REFERENCES

Ayers, Duke. (2010). Instruction, Exercise, Competition and Certification: The Cyber Defense Training Continuum. Paper presented at the Colloquium for Information Systems Security Education, Baltimore, Maryland. http://www.cisse.info/archives/category/14-papers?download=163:1503-2010

Bhagyavati. (2006). Laboratory exercises in online information assurance courses. Journal on Educational Resources in Computing, 6(4), 4. doi: 10.1145/1248453.1248457

Bork, Alfred. (2000). Learning. Educase Review, 35(1), 74-81.

Chatmon, Christy, Chi, Hongmei, & Davis, Will. (2010). Active learning approaches to teaching information assurance. Paper presented at the 2010 Information Security Curriculum Development Conference, Kennesaw, Georgia.

Ciampa, Mark. (2012). Security+ Guide to Network Security Fundamentals (S. Helba Ed. 4th ed.): Course Technology.

Howard, Jeff. (2008). Hyper-V RTM announcement. Available today from the Microsoft Download Centre. Retrieved 15 May, 2013, from http://blogs.technet.com/b/jhoward/archive/2008/06/26/hyper-v-rtm-announcement-available-today-from-the-microsoft-download-centre.aspx

Krieger, Orran, McGachey, Phil, & Kanevsky, Arkady. (2010). Enabling a marketplace of clouds: VMware's vCloud director. SIGOPS Oper. Syst. Rev., 44(4), 103-114. doi: 10.1145/1899928.1899942

Lukasik, Stephen J. (2011). Protecting users of the cyber commons. Commun. ACM, 54(9), 54-61. doi: 10.1145/1995376.1995393

Martin, Nancy, & Woodward, Belle. (2012). Building a Cybersecurity Workforce with

ISSN: 2167-1435
v30 n2522

Remote Labs. Paper presented at the Information Systems Educators Conference, New Orleans, LA.

NDG NETLAB+ Configuration Maximums.

NDG NETLAB+ System Requirements. (n.d.). from http://www.netdevgroup.com/products/requirements/

Olson, Parmy. (2012). We Are Anonymous: Inside the Hacker World of LulzSec, Anonymous, and the Global Cyber Insurgency. New York, NY: Back Bay Books.

Paulsen, C., McDuffie, E., Newhouse, W., & Toth, P. (2012). NICE: Creating a Cybersecurity Workforce and Aware Public. Security & Privacy, IEEE, 10(3), 76-79. doi: 10.1109/MSP.2012.73

Sharma, Aditya, Murphy, Marianne C., Rosso, Mark A., & Grant, Donna. (2012). Developing an Undergraduate Information Systems Security Track. Paper presented at the Information Systems Educators Conference, New Orleans, LA.

vCloud Director 5.1 Configuration Maximums. (2012). Retrieved 15 May 2013, 2013, from http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=2036392

ISSN: 2167-1435
v30 n2522