

USING VMWARE VCENTER TO TEACH SYSTEM ADMINISTRATION IN A LAB

Bryan Marshall, Georgia College, bryan.marshall@gcsu.edu
Peter Cardon, University of Southern California, petercardon@gmail.com
Craig Callendar, Georgia College, craig.callendar@gcsu.edu
Philip Robertson, Georgia College philip.robertson1@bobacats.gcsu.edu
Nipul Patel, Purdue University – North Central, npatel@pnc.edu

ABSTRACT

This paper describes our approach to using virtualization in two courses for MIS students. Our approach allows students to learn more about the needed skill of virtualization in today's job market, and to develop programming and networking skills in an environment that does not create any risk to host machines. We contrast our use of bare-metal methods to virtualization with other approaches reported in the literature that employ hosted methods. We also report various approaches, benefits, and challenges of virtualization training based on our four years of experience doing so for our students.

Keywords: virtualization, networking, IS curriculum

INTRODUCTION

Virtualization is now considered a key competency that IS professionals must have—and students who possess expertise in this technology will have a clear advantage when applying for IS positions. In a 2010 Gartner survey of 1,586 CIOs, virtualization was identified as the top priority for their organizations. CIOs state that the benefits of increased security and decreased energy consumption are some of the features that make virtualization so popular. In the public sector, almost 77 percent of federal, state, and local agencies are implementing at least one form of virtualization, and of those, 89 percent are benefiting from the technology, according to CDW's *2010 Government Virtualization Report* [9]. In the public sector, almost 77 percent of federal, state, and local agencies are implementing at least one form of virtualization, and of those, 89 percent are benefiting from the technology, according to CDW's *2010 Government Virtualization Report* [9].

In this paper, we explain our experiences training MIS students in virtualization. Many of our students have reported that they landed their first jobs in system administration based on their familiarity with virtualization. First, in the background and literature review section, we briefly explain the nature of current virtualization technologies, the role of virtualization in the IS 2010 curriculum, and prior research about training MIS students in virtualization. Then, in the next section, we describe our experiences—successes and challenges—using virtualization in courses during the last four years. We contrast our approach, which employs the bare metal method, with that of other MIS courses, which have employed the hosted method.

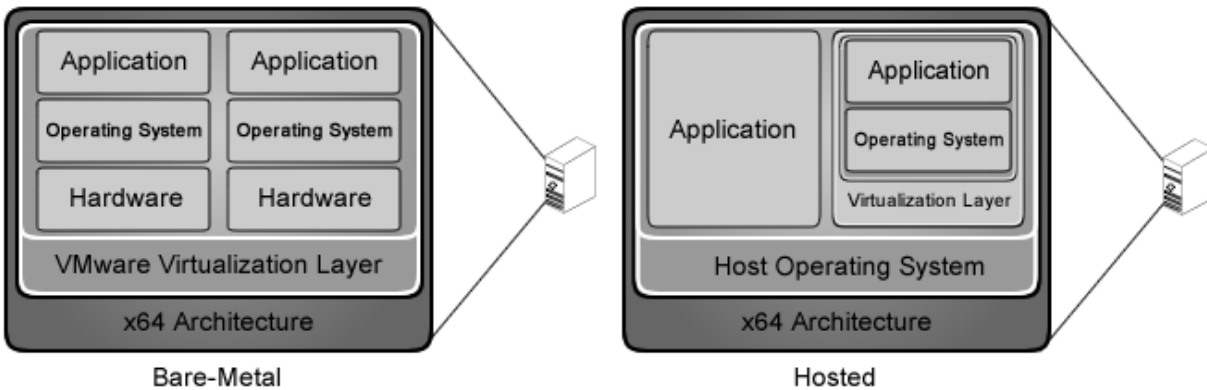
BACKGROUND AND LITERATURE REVIEW

Current Virtualization Technologies

The major virtualization vendors include VMware, Microsoft, and Sun Microsystems. When these vendors promote cloud computing, they are in fact promoting virtualization technologies. There are two virtualization methods: (1) the bare metal method and (2) the hosted method. *Bare-metal* is the most common type for servers, and utilizes a type 1 hypervisor, which runs the virtual machines and interfaces with the hardware on the host system. Typically, the host machine has a very small OS (mostly Linux based) with a tiny footprint on which virtual machines (VMs) are stacked. A *hosted* environment employs a type 2 hypervisor, which does not interface directly with the host system's hardware, but rather with its operating system. The hosted approach generally involves using VMware Player on a workstation and running a virtual machine on top of the original operating systems. This paper focuses on the bare-metal method due to the clear disadvantages shown in the research for using the hosted method [2, 8] (see Figure 1 for depiction of the bare metal and hosted methods).

We introduce several terms related to bare metal virtualization that we will use in the remainder of the paper. The term *ESX* refers to the hypervisor software built by VMware. There are several versions of ESX, including a high-quality, free version called *ESXi*. The interface for this tiny OS is a client called *vSphere*, which is installed on a host machine and

Figure 1. Types of Virtual Systems



accesses the ESX remotely. A *datastore* is essentially a partition on the hard drive where the different VMs can reside. A VM can have files on multiple datastores or even access files (.iso images) on separate datastores. Lastly, the term *vCenter* refers to a software package by VMware that does not actually host any VMs. Instead, it allows for a much tighter control over the ESX (including permissions and user access).

There are three types of virtualization: hardware, operating system, and application. In hardware virtualization, the guest operating system uses virtual hardware created by the hypervisor. In an operating system virtualization environment, the guest OS is not a complete OS, but rather accesses hardware on the host system through the host OS and the hypervisor. Application virtualization provides portability by creating a container with all of the system files and settings belonging to a particular application, and isolating them from those of other applications. Work performed within an application can be transported from one computer to another [8].

Virtualization in the IS 2010 Curriculum

According to the new IS 2010 curriculum, it is important that graduates have a firm understanding of how the IT infrastructure can be used in relation to the strategic missions of the business. Students should be able to manage ongoing information technology operations.

The IS graduates need knowledge and skills related to the management of the ongoing information systems operations within the organization, including the management, operation, and securing of the IT infrastructure. This should include managing the use of enterprise technology resources, managing

application performance and scalability, maintaining existing information systems, managing relationships with technology service providers, securing data and systems infrastructure, and ensuring business continuance (20).

There are two courses in which virtualization could be taught. The first option is in the IT Infrastructure (IS2010.5) course, which “includes computer and systems architecture and communication networks, with an overall focus on the services and capabilities that IT infrastructure solutions provide in an organizational context.” The second option is in the Enterprise Architecture (IS2010.3) course, which “focuses on organizational level issues related to planning, architecting, designing, and implementing IT-based solutions that use the platform technologies covered in the IT Infrastructure (IS 2010.5) course.” The difference between these two courses would be that the “IT Infrastructure covers the core technical foundations whereas Enterprise Architecture provides an integrated organizational perspective to planning and designing institutional solutions.” [11]

Training MIS Students in Virtual Systems

Benefits of a Virtualized Training Environment for MIS Students

Allowing students to work in a virtualized environment gives them administrative privileges without the risk. It allows students to practice working in hands-on situations, without risking host machines. Because computers in labs must remain available for all students to use, students in IT security, system administration, operating systems, and similar courses typically do not have administrative privileges that would allow them to manipulate system settings or engage in behavior that could make lab computers unavailable to other students. This limits the types of activities that can be practiced in these courses. Virtualization provides a safe environment for experimentation and system-level manipulation, without danger to physical machines or their operating systems [4, 6, 7, 8]. In fact, one set of professors found that their students were able to update a Linux kernel, add a system call to it, and in doing so, learn not only about the structures of an operating system, but also improve their programming skills in C, all without risk. In contrast to a host OS, a virtual one is easy to reinstall if it is damaged or corrupted [4]. It is even possible for students to learn about how viruses and malware work without risk [1].

Other scholars have noted the advantages of virtualization for teaching networking. One professor found that virtual networks can be created using TUN/TAP, which simulates an Ethernet cable, and can be used to connect virtual machines together into a network on the host system [2]. Another set of MIS professors argue that, while physical hardware is preferable in courses designed to train students to work as entry-level installers or maintainers of networks, virtual systems allow them to work with larger networks, and therefore to focus on higher-level concepts and theory [5]. Virtual systems also allow students to build their networks and have them graded much faster than would be possible in physical environments. They can then quickly proceed to the next iteration, which facilitates faster development. Another professor also reported on the advantages of virtualization in a networking course, which used a system that he and his colleagues had developed based on four existing technologies [1].

A final benefit for students is the ability to work and study remotely with virtual machines, using tools such as Putty [2]. This allows students to work on projects on their own time, unhindered by lab schedules [6].

Challenges Reported when Using Virtualized Training Environments

A number of MIS faculty members have reported their experiences training MIS students to use virtualization. Notably, Lunsford and Bower explained their experiences [2, 8]. Lunsford used VMware and Bower used VMware and User Mode Linux (UML). They reported the following challenges:

- Getting students accustomed to using the virtual environment
- Choosing a platform
- Programs not working correctly in the virtual environment
- Issues booting from CD in the virtual environment due to the fast BIOS, which he found out later could be addressed with a delay in the boot process
- Disk space limitations
- Access to lab computers if the VM is stored on a particular machine
- Backing up VMs
- With the VMware Player, students had to store their work on USB drives because the lab computers were configured with Deep Freeze, which restored the systems to their previous states each time students logged out. This also meant that students had to repeat changes in configuration when they logged back in to work on a previous assignment.
- VMware did not support direct printing, so students had to save their work to the host system or USB drive and print it later.

Alternatives for Virtualization Training

A number of alternatives to virtualization in education have been considered. Bower experimented with using dual boot configurations, but they suffer from several disadvantages over virtual machines. For one, disk partitions on unbooted OSs are unprotected in dual boot mode. Bower also reported that students found switching OSs to be a slow and confusing process [2]. Another possibility is the implementation of dedicated labs intended for specific courses, but such labs are not utilized to capacity when the relevant courses are not in session [10]. Cranitch and Rees considered two further alternatives: one was a time sharing option, but this had the disadvantage that students could only work on projects for a specific period of time before turning over the machines to the next group. Another was to have specific computers dedicated to database work, but this meant that students were forced to use the same computer for their work every session [3].

Cranitch and Rees considered four virtual solutions. The first was to have students store virtual machines on removable devices, which would work well for small virtual machines. Another option was to download the virtual machine from the main server at the beginning of every class session. The major disadvantage to this is that at the beginning of each session, the network would be flooded with traffic as students downloaded their machines, leading to degraded performance. A third possibility was to store the virtual machines on a central NAS. In the end, the authors' university opted to create a central server to run the students' virtual machines. The authors justified the use of a central server rather than VMWare on each workstation because students using VMware "could create their own virtual machines from scratch with whatever network options and operating system they choose and then run these from within the university network." They thought that the security risk was not worth the savings gained from VMware [3].

Stackpole and his colleagues were able to quantify the time loss associated with their use of virtualization in their courses. They found it problematic that students had to download their OS disk images from a central server and then store them again at the end of the lab session. Since retrieval and storage both take 20 minutes, students spend a total of 40 minutes retrieving and storing their work each session, which cuts the total available work time from 110 minutes to 70 minutes. As the authors put it: "Students who spend less time preparing and saving their infrastructures spend more time learning." Furthermore, because students are imaging much of the same data, there is a good bit of redundant traffic on the network. Storing VMware on each machine would avoid the problems of redundant network traffic and downloading time associated with downloading OS disk images from the central server each class period. Rather than downloading and storing an entire instance of the OS each class period, only the modifications, in the form of linked clones, would be downloaded from and stored to the central server [10].

OUR EXPERIENCE IMPLEMENTING VIRTUALIZATION TRAINING

For the last four years we have been using virtualization in both the *Introduction to Networking* (IS2010.5) course and the *System Administration* (IS2010.3) course. In the networking course, our main use was in a hosted environment. In the traditional lab we were very limited to the tools we could use to teach the fundamentals of networking. Much like the previous authors, we found that by using the VMs in a hosted environment we had much more freedom [2,3,10]. In the system administration course we started by using a hosted environment, but found the limitations were too much to really reach the outcomes of the course. By moving to the bare-metal environment we found that many of the shortcomings previously listed by other authors were easily overcome.

Throughout the rest of this paper we hope to make some recommendations based on our experience using this new vCenter bare-metal technology. We start by recommending some hardware and software. Then we make some recommendations and observations based on our recent experience. Most of the previous research has focused mainly on a hosted environment. Here we hope to show some clear advantages to using a bare-metal environment.

Hardware and Software Requirements

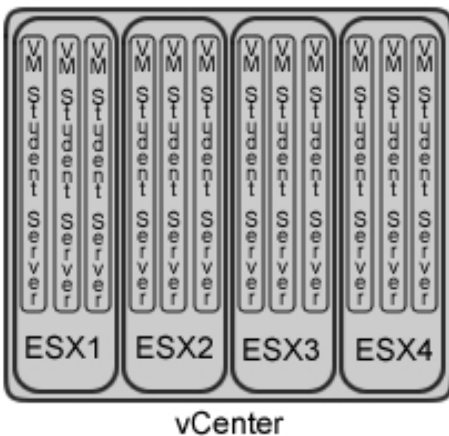
Based on our experience, to take full advantage of a virtual environment, two servers are necessary. The first needs to be a high performance machine that will run the ESX hypervisor and function as the engine that drives the course. The second server could be a less expensive, lower performance machine, which should run Windows Server 2008

and house the vCenter software. Our first server happens to be a Dell PowerEdge 2970. It has 32 GB of RAM and two 64 bit processors. We initially had a RAID 5 configuration on the machine and since have gone to a RAID 0 (due to how the datastores need to be set up for permissions). Because we back up critical data regularly, it was not mission critical to have a RAID 1 or 5 configuration. Our second server is only used to run the vCenter software. The most basic need was to have a server that supported 64 bit processing. We found that Windows Server 2008 R2 was probably the easiest server on which to administer the user logins and the vCenter software. This server is comprised of 4 GB of RAM and a simple 64 bit processor.

An ESX hypervisor is relative easy to obtain and install. The real value of this paper is showing the power of vCenter. Some of the drawbacks of just using ESX and vSphere are the permissions that are given to every user created on the ESX box. When vCenter is used, it is the systems manager of the ESX box. Users and permissions are added to the vCenter software on server 2. Using the vSphere to login to server 2, server 1 (ESX) is simply added as a host to vCenter. With vCenter it is possible to manage multiple ESX servers, but each person who will have a user login on the Windows 2008 server needs to be added. The password assigned to them in Windows 2008 will be the one they use when logging in via vSphere.

Figure 2 shows the relationship between vCenter and the ESX servers. Each ESX host can house as many VMs as you see fit. The power of the vCenter is the user permissions that can be implemented on each ESX host.

Figure 2. vCenter Relationship to ESX



Key Learning Outcomes and Advantages of a Bare-Metal Virtualization Environment

Performance

As the other authors have mentioned, the largest drain on the students was the time to move and save their completed work. One professor reported that it would take students 20 minutes to move the files over at the beginning and the end of the class period. This means students are wasting an inordinate amount of time just waiting for files to move. When we use the hosted method and store the VMs on USB devices, we find that the performance is extremely degraded. Even with a USB 2.0 controller, the files are so large that the hosted system has a hard time reading them. By using the bare-metal environment, students have instant access to their production servers by simply logging in using the vSphere client. Also, when students use the bare-metal system, the processing is taking place on the server, which has been optimized to work within the virtualized environment. Even with 30 VMs running, the advances in virtualization allow for a high performance experience. Students were able to install a new version of Ubuntu Server 10.04 LTS in less than 4 minutes.

Snapshots

Although snapshots are also available in the hosted environment, the biggest advantage is the centrally stored location of the snapshot. Students can use the vSphere client from any computer in the world as long as they have Internet access. When students take snapshots of their servers, they are stored in the datastore on the ESX machine. Snapshots could be used by both faculty and students to create interesting lab scenarios. Students could take snapshots of their servers after completing tasks, so that mistakes on future tasks do not result in them losing all of their work. Faculty can set up an exam machine and push the appliance (VMs) out to the student folders and have them troubleshoot the VM as part of a simulated test or lab.

Create / Delete Servers

By using vCenter we could give the students permissions to create and delete their own servers. They could potentially have as many servers as they want. Indeed, some of our students created multiple servers, allowing them the opportunity to test real email, web, and other types of servers. Furthermore, virtual machines are not just limited to servers. Students can install and configure workstation machines. This is an easy way for them to practice installing operating systems, which many of them have never done before. Unfortunately, if only ESX is used, students cannot be given permissions to create or delete, or they will be able to delete other student VMs, along with any exam / lab VMs that the instructor may have on the system.

Add Hardware

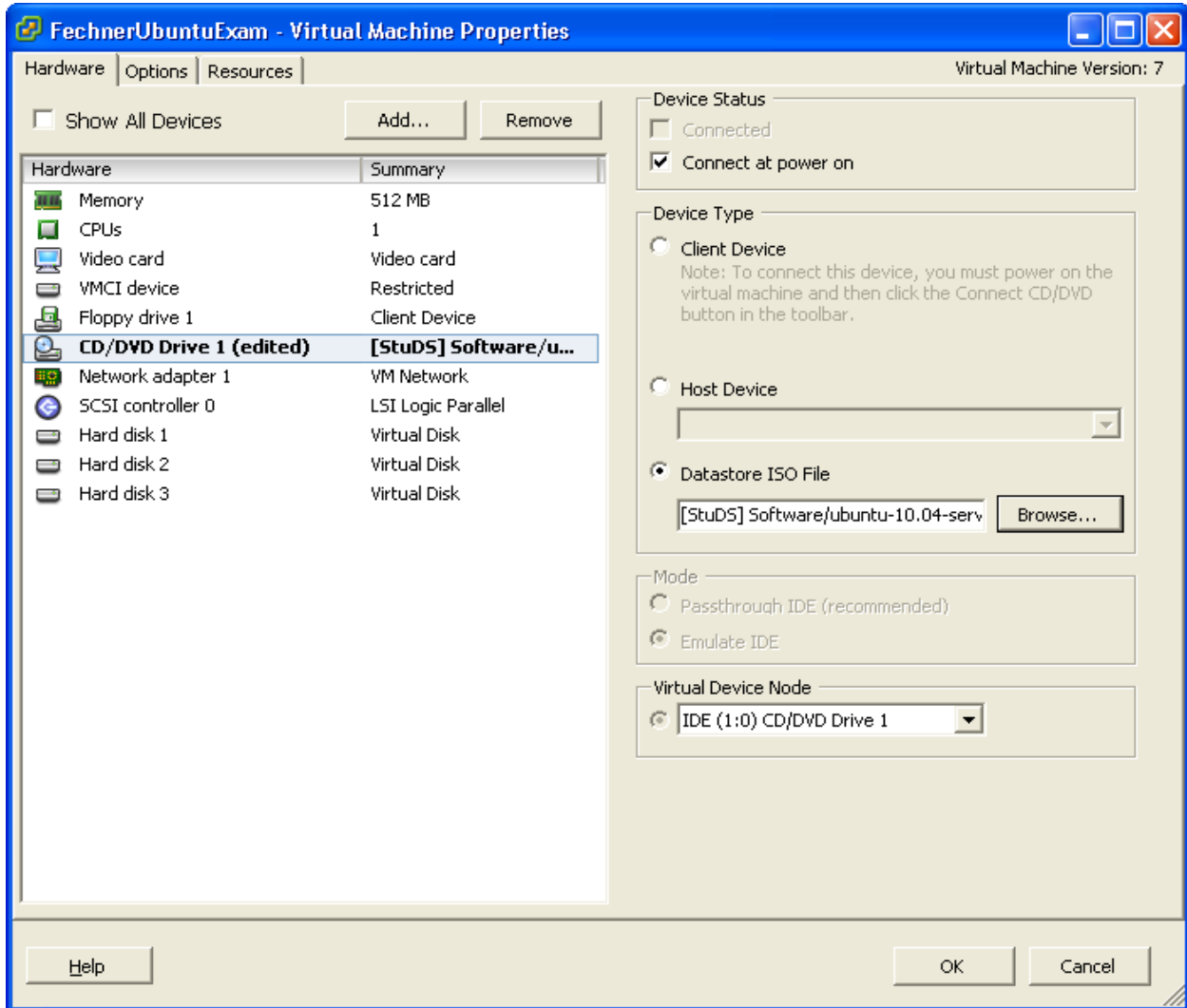
Adding hardware in the bare-metal environment is similar to doing so in a hosted environment. However, because students are working off one server there are very few issues with compatibility when changing hosts. When we used the hosted environment in the past, we found that when students changed from one host machine to another, there were many issues with hardware that had been added. By using the bare-metal environment this problem has been eliminated.

The bare-metal environment allows students to add pretty much any type of hardware a physical machine could have, including a serial port, parallel port, floppy drive, CD/DVD drive, USB controller, Ethernet adapter, hard drive, SCSI device, and PCI device. As seen in Figure 3, adding hardware is very simple. After an instructor adds hardware to a VM, the students can then configure the device as if they were sitting in front of the physical machine. For example, we have the students add multiple hard drives to their Linux servers. They then learn how to format, partition and mount the drives.

Mount ISO (including software in one location)

Many instructors spend countless hours trying to distribute or have students download operating systems. By going to the bare-metal system, the instructor can simply place one copy of the software (as an .iso file) onto the public datastore, and all the students simply mount the same .iso file onto their individual machines. Since the processing power is done on the server, each student takes full advantage of the server performance. As seen in Figure 3, the option for the .iso files is available in the configuration of the CD/DVD drive. As an added benefit, students learn the power of .iso files and how these large images are used in industry.

Figure 3. Virtual Machine Properties



Pausing VMs

Many times in lab environments students will run out of time. This causes the dilemma of what to do with a computer that is only halfway installed. Using virtualization, students can pause the VMs and come back to them at their convenience. Many times in a hosted environment, students complain that their VMs have been removed or changed by students from other classes. And like most lab environments, setting up user permissions for an individual class can become very difficult semester after semester. When the students pause a VM in a bare-metal environment they can pick up where they left off from any computer that has the vSphere client installed.

Isolated Environment (Permissions)

Using the vCenter software, student permissions can be locked down so that they cannot see anyone else on the system. This allows users to work freely without fear of files being removed or manipulated. In the past, students have had issues with other students hacking their project servers due to security issues (mostly leaving them logged in). With virtualization the servers still run even when the client (vSphere) has been shut down.

Challenges to Training in a Virtualized Environment

Learning Curve

Like the other authors, we found that some students had difficulty understanding the concept of virtualization. It helped when we used the term “appliance” and showed how the VM is similar to a Word document, which can be opened and closed. After a couple of weeks even our weakest students grasped the power of virtualization and enjoyed the class. We have found that students have a tendency to understand the bare-metal virtualization environment easier than the hosted environment. We think this is because the students are logging into the server to work on their projects, rather than keeping the VMs on a USB drive or local computer.

On Campus vCenter Logins Only (until SSH tunnels were covered)

Because of how our campus firewall is configured, and out of security concerns, we only allow access to our vCenter server from on campus. In the beginning of the course this was more of an issue as our students set up the vSphere client. However, as the students had their servers running, they simply putted into a public facing box and from there had access to their servers. Later in the course, as we covered ssh tunnels, students were able to configure the vSphere client to tunnel into campus as well.

When using vCenter, you have on point of access through that server. It doesn't matter how many ESX servers you have, the only security focus really needs to be on the Windows Server hosting the vCenter software. The other ESX machines are essentially pooled together under the vCenter software.

Static IP Addresses (do not turn off computer from off campus)

Adding 20-30 new servers with non-static public IP addresses naturally has some potential for problems, especially if the students leave their VMs turned off over an extended period of time. Students quickly became aware of the need for static IP addresses and only restarted their servers when they were able to access them using the vSphere client to obtain a new IP address if the old one changed. A side benefit of this was that students were able to use their networking knowledge of IP addresses in a second course, thereby reinforcing networking principles in the system administration course.

Limited Hard Drive Space

Obviously one challenge would be the sheer size of the VMs on the server. We only had 30 VMs installed at any one time. An instructor with more students, who needs a substantial amount of VMs, may want to limit the number of snapshots students are allowed to take, since a large file is created with each one. This being said, space is so cheap these days that it is almost a null issue. In our experience, even with 30 VMs working, the performance was great. Most servers come pre-configured and ready to be used in a virtualized environment.

Combined Datastore

We only set up three datastores (locations where physical files reside). In an ideal environment it would have been nice to fragment this out more. The three datastores we created were a production DS (where we stored several production VMs), a student DS where all of the student files were stored, and a backup DS where we put some temporary backup files before we moved them off the physical server. Unfortunately it would have been possible for students to delete each other's physical files. Ideally all students should have had their own DS, but due to partition constraints we did not feel this was vital. The vCenter software does not allow for folder permissions from within the datastore.

CONCLUSION

In conclusion, we are excited about this new technology. We have found that our students are accepting the challenge to learn virtualization and are benefiting greatly from their experiences. Although other researchers have

reported using this type of software, we feel we have added to the research by introducing the advantages of using vCenter and some of the strengths of using the bare-metal method.

Note of Gratitude: We would like to thank Jamie Defoor and the IT staff at Georgia College for their patience and support while we labored through the implementation of this system.

REFERENCES

1. Border, C. (2007). The development and deployment of a multi-user, remote access virtualization system for networking, security, and system administration classes. *Proceedings of the SIGCSE Technical Symposium on Computer Science Education*.
2. Bower, T. (2010). Experiences with virtualization technology in education. *Journal of Computing Sciences in Colleges*, 25(5), 311-18.
3. Cranitch, G., & Rees, M. (2009). Virtualisation: A case study in database administration laboratory work. In *Same places, different spaces. Proceedings of Ascilite Auckland 2009*, 168-74.
4. Gaspar, A., Langevin, S., & Armitage, W. D. (2007). Virtualization technologies in the undergraduate IT curriculum, *IT Professional*, 9(4), 10-17.
5. Gaspar, A., Langevin, S., Armitage, W., & Rideout, M. (2008a). Enabling new pedagogies in operating systems and networking courses with state of the art open source kernel and virtualization technologies, *Journal of Computing Sciences in Colleges*, 23(5) 189-98.
6. Gaspar, A., Langevin, S., Armitage, W., & Rideout, M. (2008b). March of the (virtual) machines: Past, present, and future milestones in the adoption of virtualization in computing education. *Journal of Computing Sciences in Colleges*, 23(5), 123-32.
7. Grammer, K., Stoleran, J., & Yi, B. (2011). Introduction of virtualization in the teaching of operating systems for cs undergraduate program. *Journal of Computing Sciences in Colleges*, 26(6), 44-50.
8. Lunsford, D. (2009). Virtualization Technologies in Information Systems Education.' *Journal of Information Systems Education*, 20(3), 339-48.
9. Montalbano, E. (2010, July 16). Government sees virtualization benefits. *InformationWeek* [online]. Available at www.informationweek.com/news/government/enterprise-apps/225900083.
10. Stackpole, B., Koppe, J., Haskell, T., Guay, L., & Pan, Y. (2008). Decentralized virtualization in systems administration education. *Proceedings of the 9th ACM Conference on Information Technology Education*. 249-53.
11. Topi, H., Valacich, J. S., Wright, R. T., Kaiser, K., Nunamaker, J. F., Jr., Sipior, J. C., & de de Vreede, G. J. (2010). IS 2010: Curriculum Guidelines for Undergraduate Degree Programs in Information Systems, *Communications of the Association for Information Systems*, 26, article 18. Available at <http://aisel.aisnet.org/cais/vol26/iss1/18>.