The Evolution of a Virtualized Laboratory Environment

Bill Stackpole
Department of Networking, Security and Systems Administration
Rochester Institute of Technology
152 Lomb Memorial Drive, Rochester, NY 14623
(585) 475-5351
bill.stackpole@rit.edu

ABSTRACT
Distance learning techniques for lecture-based courses have been employed with success for quite some time. Creating a hands-on laboratory environment to serve the same remote student population has proven to be difficult. This paper addresses the evolution of a remote laboratory system designed to provide faculty and students with a functional and accessible solution that offers an experience equivalent to that of an on-site laboratory. It addresses successes and failures. The designs for each of the iterations are discussed. Future integration with additional curriculum, as well as plans for community outreach, are proposed.

Categories and Subject Descriptors
K.3.2 [Computer and Information Science Education]: Distance Education, Human-Computer Interface, Lab environments

General Terms
Design, Experimentation.

Keywords
Curricular Issues, Distance Learning, Laboratory Environments, Networking, IT Education, System Sizing and Benchmarking, Virtualization.

1. INTRODUCTION
Distance learning has been a popular mechanism for the delivery of lecture-based material for over ten years. Courses developed for in-person stand-up lecture have been modified for delivery via computer-based technologies. The Nation Center for Education Statistics has documented exponential increases in participation in distance education programs (NCES, 1999-2002) [2]. The deployment of Course Management Software (CMS) and Learning Management Systems (LMS) has helped support both students and faculty to this end. Advancements in virtualization technologies are beginning to increase the success of delivering laboratory instruction in remotely accessible environments.

Many of the courses in the Networking, Security, and Systems Administration Department at Rochester Institute of Technology have strong, lab-based, “hands-on” components designed to help students anchor concepts presented during the lecture portion of the class. Successful completion of courses in the areas of systems- and network-administration are dependent upon the student demonstrating their skills in the lab environment. This lab-centric approach has proven to be an effective environment for both teaching and learning. Until 2005, this had been the only method employed to provide a laboratory experience to students in the systems- and network-administration fields.

2. THE PHYSICAL LAB
There are a number of physical computing laboratories currently in use at RIT. One example is the Systems Administration Lab. The physical infrastructure of this lab environment includes eighty student machines (twenty benches with four machines per bench, each with its own keyboard, mouse, and monitor). The student machines are periodically replaced through a “cascade program” that allows the re-use of older machines through a “hand-me-down” process. This helps to ensure that equipment purchased is used to the end of its life.

In addition, the lab includes other items providing support for the infrastructure. A number of servers provide network services such as DHCP, DNS, WINS, and storage for student images. Physical wiring to support the machines is in place. Each student machine has two network adapters that connect to the lab infrastructure. One network adapter is always connected to an “imaging network”. The other adapter is wired to a patch panel so students can manipulate network connectivity and perform packet inspection via the bench hub. This constitutes a very large investment in equipment, infrastructure, and space.

This lab requires operational support to keep it running smoothly. A full-time lab manager as well as student workers are employed for these tasks. During “open hours”, staff is present to ensure the security and function of the lab.

3. LAB AVAILABILITY
The labs are scheduled in roughly two-hour increments to support different courses. When the labs are scheduled to support the classes in session, the lab is available only to students who are enrolled in that course section. When classes are not in session, the lab has “open hours,” at which time students are free to use the lab to complete their assignments and work on personal projects. Labs are typically unlocked fourteen hours a day, every day of the week.
While the labs are well-equipped with physical hardware, the utilization of the labs varies over the course of an academic quarter. Early in the quarter, the labs are not highly utilized. Near the end of the quarter however, there are more students competing for lab space than there is space available (demand exceeds capacity).

Even in a well-equipped lab with modern equipment, hardware failures will occur. Such failures take time to identify and repair. During both of these periods, the machines are unavailable for student use. This can interrupt and inconvenience those attempting to utilize the lab.

4. REASONS FOR A VIRTUAL LAB
The current physical teaching laboratory has been proven effective. It provides an environment in which the student can learn network- and system-administration tasks. However, there are still a number of opportunities for improvement, some of which are outlined in this section. Many of these perceived improvement opportunities were helpful in driving the development of a remote laboratory system.

4.1 Cost
Maintaining a physical laboratory environment is expensive. Some commonly overlooked costs associated with the lab include: allocation of physical space to host the lab, maintenance of the space and equipment, heating and cooling of the space, lighting, furniture, electricity, cabling, and security. The “cascade program” allows some cost savings by providing an outlet to continue to use machines until they truly reach end-of-life. However, this does not eliminate the need to have workers on hand at all times to support hardware failures. The cost of labor alone is expensive, independent of other factors. While a remote lab environment cannot eliminate all of these costs, it can reduce many of them. For example, the use of virtualized CD-rom drives eliminates issues caused by failed physical drives.

In addition to providing these cost savings to the institute, the system is even more attractive if one factors fuel savings into the equation. This was not a consideration when the labs were being built, but turns out to be an added benefit to both students and faculty.

4.2 Availability
The lab environment is not always available. While the labs are open approximately fourteen hours a day, between six and eight of those hours are reserved for course sections to be taught in the physical lab. This leaves students to juggle their schedules to ensure they can get into the lab to complete their assignments. When the labs are closed, the students have no access or ability to complete the machine-dependant portions of their assignments.

Inclement weather can also pose a problem when students need to travel to the lab. A remote lab environment would allow access to the labs at all hours of the day or night -- as well as from locations across- or even off-campus. Students could share the lab environment at the same time as other classes are being held.

Given that the lab is utilized near capacity toward the end of the academic quarter, making the lab available as a virtual lab is an attractive option to alleviate demands on space and equipment.

4.3 Performance
The machines in the lab are always a year or two older than current state-of-the-art. While this does not appear to be a major issue when one considers the type of load being put on these machines, newer, faster hardware is beneficial in terms of getting the job done more quickly. A virtualization platform can offer the capability to upgrade the entire infrastructure by adding to rather than replacing the machines in the infrastructure.

4.4 Community Outreach
The current lab infrastructure limits the ability to share the facilities with other institutions. Having a remotely accessible environment can allow virtual access to the labs to be offered to others. By sharing the system, institutions that lack such an infrastructure can have an opportunity to expand their curricular offerings.

For the reasons cited above it was clear that a virtual lab environment could offer alternatives that the current lab infrastructure simply could not support.

5. THE FIRST RUN AT A VIRTUAL LAB – ASSEMBLING THE SYSTEM
In the fall quarter of 2005, the decision was made to pilot a remotely accessible laboratory environment. A number of elements were taken into consideration for this system.

5.1 Funding and The Hardware Platform
No funding had been allocated for this system (for hardware, software, personnel, or support) so a search for unused equipment was begun. It happened that a number of high-end workstation machines were available in the fall quarter. Ten of these machines were commandeered and put into service as remote-access hosts.

The approximate hardware specs for the machines included:
- 2.5Ghz CPU
- 2gig of system RAM
- 80G hdd
- 100B Ethernet connection

5.2 The Virtualization Platform
Vollrath and Jenkins (2004) [3] used Microsoft Virtual PC in their implementation of a remote laboratory to teach systems administration. After careful consideration, it was decided that the Virtual PC environment did not provide sufficient control over the network to allow a direct translation of the lab exercises targeted for this project. The features available in the VMware product offered a better fit for this application.

The Workstation product from VMware was selected. A number of licenses for this product were available for use at the institute. Negotiations for the use of these licenses allowed them to be allocated to this project.

The plan provisioned one machine per student and allowed them administrative access to the host operating system. From the console of the machine, the student would use VMware to install and test services on a group of virtual machines. A series of pre-built images were accessible via a network share and would be used as the base upon which the students would perform their exercises.
5.3 The Host Operating System
A host operating system to support the VMware platform needed to be selected. Since there was very little time to deploy the system, the decision was made to install Microsoft Windows XP Professional on all ten machines and enable the Remote Desktop Protocol (RDP) for remote access. All ten machines were placed on the network behind a router providing Network Address Translation. An Access Control List on the router blocked all but ten non-standard consecutive TCP ports from external access to the network. Each of these ten ports were forwarded to one of the ten machines in use. Unique port numbers allowed allocation of a machine to a single user while sharing a single outside-facing IP address and DNS name.

5.4 Imaging and Disaster Recovery
The existing physical lab infrastructure was built to facilitate disk imaging with Norton Ghost. The virtual host machines had identical specifications and had access to the physical lab infrastructure. Thus it was possible for a disk image to be created. This provided restoration of each machine to a default state should it become unusable due to drive failure, the presence of malicious code, etc. Default accounts were created and administrative privileges provided to the student account.

5.5 Selection of Participants
The course selected as the pilot for this program had a typical lecture population of between 36 and 48 students. Labs were scheduled with either 14 or 28 students, limited by the instructors and hardware available in the lab.

The selection of students to participate in the project was performed randomly from the lecture population. A spreadsheet containing all of the student names had a random number field generated. The list was then sorted by the random number field. The class was provided a description of the “new” lab environment and an explanation of how it might differ from the physical environment with which some students were already familiar. The students were offered the chance to participate according to the order their name appeared in the spreadsheet after sorting. An opportunity for those randomly selected students to “opt out” was provided, at which time another student was offered the chance to participate.

Selected students who opted to participate in the project were provided remote access and login credentials for one of the ten machines. An initial on-campus training session was provided to familiarize the students with the remote platform and the virtualization software. Students participating in the project had access to an electronic forum through which they would interact with their instructor and their peers, obtain and submit lab assignments, and provide feedback to the instructor on the pros and cons of the system throughout the quarter. Participants in the pilot lab program would not attend the scheduled physical lab environment, but would continue to have face-to-face contact with their instructor during regularly scheduled on-campus lectures. Their interaction with the lab instructor was facilitated through the use of the FirstClass email system, which has the capacity to support conferencing, course drop-boxes, and instant-messaging services.

5.6 Communication and Forums
A mechanism was required to provide both communications for the students and forums for discussion. Two clear options existed. One was a new Learning Management System (LMS) that the Institute had just introduced. The alternative was a separate FirstClass mail and conferencing system with which the faculty and the students were already familiar. Rather than struggle with the new and unproven LMS, the decision was made to use the well-known FirstClass system.

5.7 Migration of Lab Exercises
The virtual lab environment, designed to mirror the experience a student would have in the physical lab environment, used assignments that were functionally identical. The exercises and deliverables remained the same for both student populations. In order to support the network ranges, file shares, and other features unique to the virtual environment, the lab assignments from the physical lab environment needed to be rewritten. Each lab was modified to substitute elements specific to the VMware environment in place of those specific to the physical lab environment.

5.8 Results of the First Trial
Data collected during this first pilot indicated that student performance and satisfaction in the virtual lab was at least equal to that experienced by students in the physical lab environment [1]. This pilot provided an advantage by offering remote access to a lab environment but did not afford any economies of scale or hardware cost savings. While every attempt was made to address issues specific to VMware in the lab rewrites, some items were overlooked and needed to be modified “on-the-fly” during the course of the quarter.

Students submitted improvement suggestions throughout the pilot project. Many of their suggestions were specific to performance improvements in the virtualized environment. Through trial and error, the students had discovered that guest operating systems could run acceptably with less than two-hundred fifty-six megabytes of memory allocated to them. This gave rise to the idea of sharing the workstation machines with more than one simultaneous user.

While no major issues had occurred, the ability to prevent or mitigate potential problems would have been advantageous. Therefore, more administrative control was identified as an area that needed attention. The ability to use the environment for multiple courses was also desired.

6. THE SECOND RUN – MODIFICATIONS TO THE SYSTEM
Between the fall and winter quarters, faculty worked to modify the project to implement improvements based on lessons learned during the operation of the pilot program, from student feedback, and from other discussions. The selection of students was performed in the same random fashion with each student retaining the option to decline participation.

Changes included:
• provisions for administrative control over the systems
• sharing of hardware
• merging two other courses into the remote lab infrastructure
• migration to the Institute Learning Management System (LMS) platform

6.1 Operating System Change
The first change moved the OS from Windows XP and Remote Desktop to an Active Directory environment built on Windows Server 2003 with Microsoft Terminal Services. This provided a central point of administration and facilitated implementation of Group Policy Objects. No longer would the participants be granted administrative control over the host machines, rather their interaction with the environment would be limited through the Remote Desktop interface.

6.2 Administrative Control
Secondly, the user load on each individual workstation was increased. The same ten machines were available but the number of students being supported was doubled and two machines were left unused in reserve. The students were given user accounts rather than administrator accounts. Six machines were assigned two students each, while two machines were assigned four students each, for a total of twenty students being supported over eight machines. The intention was to provide a fallback position should this deployment prove unworkable. If the system could not support the four student load, user load on all machines could be balanced by moving users to the two standby machines.

6.3 Additional Workstations
Additional workstation machines were merged into the infrastructure. This additional group was physically on the same network and a part of the same domain; however, it was allocated for use to support a graduate laboratory environment and additional courses. Storage for students continued to remain local to the machines to which they were assigned but the authentication for the users moved from the local machine to the Active Directory Domain Controller.

6.4 A New Learning Management System
The transition was made to the new campus Learning Management System. This provided a single point of interaction for all classes. The forums for the students in the virtual lab would no longer be separate from those in the physical lab. This allowed virtual lab participants and physical lab participants to interact online, as well as in the lecture.

6.5 Results and Issues
In terms of student performance and satisfaction, similar results to those obtained during the pilot project were experienced. Student performance in the virtual section, measured by grades earned for the lab assignments, was comparable to the performance of students in the physical lab sections. Further, the virtual students were satisfied that their lab experience was appropriate to deliver the concepts under discussion.

One difficulty during this trial involved memory allocation to the virtual machines. While the hosting machines had sufficient memory to support multiple students simultaneously using virtual images, no control was in place to prevent one student from allocating all available memory to their single session. This caused issues with machine availability to the other students sharing access to that physical machine. Additionally, students would login to the machine, begin a session, and then leave it unattended indefinitely. While this was not a problem when the machines were allocated to one or two students, the machines with a four-student load were rendered unavailable to the other three students as all of the system resources had been consumed.

7. THE THIRD RUN – FURTHER MODIFICATIONS

7.1 Hardware Change
Two used rack-mount dual-processor machines had become available for the lab to borrow for the spring quarter because of an upgrade elsewhere on campus. Two new rack mount servers were purchased. One was designated to share the hosting of student infrastructure while the second provided Network Attached Storage (NAS) functions. Having the NAS in place allowed quotas to be applied to the remote storage of virtual images while the servers kept the operating system on their local drives. Should these new rack-mount servers be unable to support the virtual lab environment, they could be moved into the physical lab to upgrade the existing infrastructure.

7.2 Auto-Logoff Policies
Changes were made to the group policy objects to implement an auto-logoff policy. This was put into place to address the unattended sessions experienced during winter quarter.

7.3 Results and Issues
Twenty students were, again, supported on the three rack mount servers. The results were similar to those experienced the previous two quarters in terms of student performance and satisfaction. No major issues were experienced.

8. THE FUTURE – WHAT’S ON THE HORIZON
A partially populated IBM blade server was obtained by the department. Additional hardware was purchased to fully populate the server and a number of the blades in the server were allocated to support the project. The server also had a Storage Area Network (SAN) attached to it which was intended to replace or augment the existing NAS. The environment was expected to continue using VMWare Workstation and a Windows infrastructure. As it turned out, the blade server was not as economical a solution in terms of the number of VMs that could be supported.

A number of other possibilities were explored when considering how the infrastructure might continue to evolve. The labs in the department, currently populated with workstations, require maintenance and do not provide students with access other than during open lab hours. One option under consideration is to remove all workstations from the lab, replacing them with thin clients. Such a sweeping change would provide significant savings on hardware maintenance. This could also provide students twenty-four hour access to the lab and resolve lab scheduling issues. This would reduce the man-hours required to maintain the labs, and eliminate hardware changes during the annual cascade.
Another possibility is to implement a community outreach program. Many smaller institutions would greatly appreciate having access to such an environment but they have neither the funds nor the capability to build such an infrastructure. RIT is currently working together with other colleges to help them develop virtual labs for Information Assurance education and to provide support for their coursework.

Through the use of higher performance platforms the infrastructure can be made to scale more effectively. Scalability of the infrastructure would allow additional lab-based classes to immediately be added to the environment. Other coursework could be supported with the addition of hardware to support virtual switches and virtual consoles.

As of August 2008, the first-generation platform has been decommissioned. It has been replaced with a much higher performance system comprised of four SunFire servers and a NetApp SAN[4]. A graduate research project was recently completed that created a mechanism by which the instructors can compare student performance to “time-on-task” through the use of a scripting system and a back end database. This project was designed to track both the time a student is logged into the host, as well as time spent in each of the virtual machines. A student-driven scheduling system has also been proposed.

This project started as a proof of concept. In only four years it has evolved into a platform that has proven useful and exciting to both faculty and students. Some major employers in the area have expressed interest in collaborating with us on future development. Many different course are having their lab exercises performed on this remotely accessible system. The system has been well-received by both faculty and students.

It is expected that coursework in the department will continue to migrate to this platform. This infrastructure has the capacity to significantly change the way that lab-based exercises are experienced and delivered.

9. ACKNOWLEDGMENTS

This project would not be where it is today without the ideas, hard work, and support of the faculty and students in the department. Faculty and staff directly involved in this endeavor include: Luther Troell, Yin Pan, Peter Lutz, Sharon Mason, Daryl Johnson, Bruce Hartpence Bo Yuan, and Charles Border. Students to whom gratitude is extended include Charles Consagra, Matthew Lidestri, Jason Koppe, and Suraaj Guar.

10. REFERENCES


