A Case Study in Optimizing Computer Laboratory Resources: The High-Speed Backup and Restoration (HiSBaR) System for Computer Lab Workstations

Robert B. Sweeney
bsweeney@usouthal.edu
School of Computer and Information Sciences
University of South Alabama
Mobile, AL 36688, USA

Abstract

This paper describes HiSBaR (High-Speed Backup and Restoration) which is a system developed to allow students performing projects to perform a high-speed backup or restoration of the entire project, including an operating system installation to a server. This HiSBaR system allows for more efficient use of limited computer laboratory resources as well as providing a more flexible laboratory environment that is not difficult to maintain. In addition, students using this computer lab were able for the first time to perform the projects individually and consequently gain a greater understanding over the subject matter.

Keywords: gigabit, Ethernet, computer, laboratory, installation, backup, restoration, high-speed, multi-session, HiSBaR

1. INTRODUCTION

This paper describes the planning, design, implementation, and maintenance of a computer lab consisting of six workstations and a server. The goal of this computer lab was to be able to support more than 40 simultaneous individual student projects involving the installation of a complete operating system and other software. Each student using the lab was required to individually install and configure an operating system on one of the six workstations. The configuration of the lab made it possible for students to work on their installation project at any time the lab was open and to use any of the six available workstations. A system for High-Speed Backup and Restoration (HiSBaR) was developed to meet the needs of this lab.

2. BACKGROUND

The use of specialized hardware and software to save and distribute a computer image to campus computer labs is not new. Many Universities have implemented a system to refresh images of computers used in campus computer laboratories and faculty and staff offices to a standard state (Janz 2003). In this way, software on a computer that has been deleted or is malfunctioning can be automatically restored to the system and unauthorized software or data that has been added to the system may be deleted, depending on the configuration of the system.

An extreme example of this type of system is found at The Rochester Institute of Tech-

nology (RIT) (Weeden 2003). RIT utilizes a specially developed system of imaging for their computer labs that allows complete restoration of a computer image within minutes. They can configure a single lab in multiple ways even in the limited time period between classes to support different types of instruction using only student workers to perform the change. Their setup allows a single lab to support multiple curricula and numerous uses, resulting in an extremely flexible lab environment. Furthermore, the RIT setup supports the important ability to allow a student to start a project or work in one session and complete it at a later session, while being able to utilize the computer the student started on for other activities between these sessions.

Creating a system to allow backup of individual machines is often difficult and costly. Utilizing such systems when they are implemented can be inconvenient and may require extensive administrative effort. Various approaches to distributed backup of computer systems have been proposed including centralized backup using the Internet, and peer-to-peer backup using excess disk capacity (Cox 2002). The HiSBaR system developed for this computer lab was relatively simple to implement, not prohibitively expensive, is easy to manage, and is simple for students to use with appropriate instruction.

2. LAB CONFIGURATION

The HiSBaR system was made possible by a high speed network created in part with the installation of a relatively inexpensive Net-Gear 10/100/1000 8-port Ethernet Switch purchased for under $200 plus seven Intel Pro 1000 network adapter cards purchased for under $50 each. These network adapters were specifically selected because they had stable (not beta) DOS driver software. To act as a repository for the data students would add to each workstation, a 200 GB Maxtor hard drive (7200 RPM, 100/133 ATA) was purchased for under $150 and added to the lab server.

These purchases were made to supplement the six Dell workstations (40GB hard drives, 256MB RAM, 1.2GHz Celeron processor) that were already in the lab. In addition, a Dell workstation running Windows Server 2003

(80GB hard drive, 512 MB RAM, 2.0 GHz Celeron processor) was purchased for under $1,000 into which the 200 GB Maxtor hard drive was installed, replacing its 80GB hard drive, so that it could act as the lab file server. The lab was already wired with CAT 5 cable, however, subsequent testing of the cable setup revealed some correctable wiring problems. Finally, the lab contained a HP LaserJet 4000N networkable laser printer. See Figure 1 for a network configuration view of the HiSBaR system.
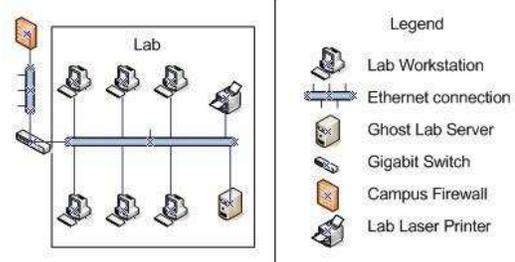
Figure 1 – Architecture of HiSBaR

The software that made the HiSBaR system possible was Symantec Ghost 7.5 and luckily our School already had a license for this version. In addition, custom Ghost scripts were written for backing up and restoring data and which also provide instructions to the students to walk them through the backup and restoration processes. These Ghost scripts are available from the author upon request.

To allow each student to perform a backup or recovery of their entire installation, a custom DOS boot diskette was created for each student. To customize each diskette, the autoexec.bat, network.ini and protocol.ini files had to be edited. This was done with a script written on the author's Linux workstation which created custom versions of each of these files. In addition, a hex editor was used to disable most of the DOS commands on the command.com file used for this boot diskette so that students would not be able to list or erase the contents of the shared drive using that diskette. These scripts are also available upon request from the author.

Testing the HiSBaR System

Although the lab was created relatively quickly, it was tested extensively before the live student projects were started. A single test machine was setup which contained an

Intel Pro1000 network adapter and Ghost scripts were written and tested using various configuration settings to see the effect on transfer speed. This network adapter was selected because it had stable DOS driver software and worked well, providing the necessary transfer speeds

Like all projects, there were unexpected problems to be overcome. The Intel Pro 1000 network adapter cards that had been tested were not ordered. Instead, 3-COM Gigabit Network Adapters were purchased for all lab machines which had only a beta DOS driver. Extensive testing revealed that the 3-COM drivers were incapable of achieving the necessary data transfer speeds. Consequently, a second purchase of the appropriate Intel Pro 1000 network adapter cards was made. Once they were installed, adequate transfer speeds were achieved. The relevance is that maintaining control over the purchasing of the network components is sometimes difficult in a University environment but is nevertheless critical to project success.

One aspect of testing was the speed with which the backup and restore process could be completed. When this project was envisioned, a total time of 10 minutes to complete either a backup or restore was considered the absolute maximum acceptable. This time was not however optimal and a goal was to cut this time in half, which hopefully would increase student usage of the lab.

Testing in the lab revealed that the use of a well-written DOS driver for the gigabit network adapter was of critical importance. Testing determined that the Intel Pro 1000 network adapters were able to complete the backup or restore in under 5 minutes for the a 3GB workstation. This includes the time required for the DOS boot up to complete, the time for a disk check to be run, and the time required by the Ghost server to create a backup of the previous backup, in addition to the actual data transfer time.

A breakdown of the time required to complete the backup process time shows that it consists of 1 minute to boot the DOS diskette and login to Ghost server, 1 minute to check the file system and create the image backup, plus 1.5 minutes/gigabyte of data to be transferred. For example, for a typical 3 GB installation, the time would be 1 minute for boot up/login plus 1 minute for file system check plus 4.5 minutes transfer time, for a total of 6.5 minutes. The restore process is nearly identical in terms of time required.

These time measures are for one transfer occurring at a time. When two simultaneous transfers occurred there was a slight reduction in the transfer rate to perhaps 2 minutes/GB. However, when three simultaneous transfers were attempted there was degradation to the extent that transfers took about 4-5 minutes/GB. On the other hand, since doing a single transfer only took between 4-5 minutes to complete, it was rare to have even two transfers going simultaneously because students worked on their projects asynchronously.

In order to keep the situation of multiple simultaneous transfers from occurring, warning signs were placed in the lab and an announcement was made in class. These explained that if a backup or restoration was in progress, starting a subsequent transfer would slow both transfers more than if the second student waited until the previous transfer completed before starting their own. An informal survey of students at the end of the semester suggested that three simultaneous transfers occurring was the most any had ever experienced.

3. USING HISBAR

Students used HiSBaR for two projects during its initial semester of implementation. The first project was performed individually by each student and required the installation and configuration of the Linux RedHat9 operation system plus the Apache web server with support for PHP, MySQL, and SSL among other things. The second project was performed as a two-person team and involved the installation and configuration of the Windows 2003 Server – Web in addition to the IIS6 web server with support for PHP, MySQL and SSL.

The procedure used by the students during these projects went as follows. Each student was given a custom DOS boot disk with which they could connect to the server and perform a backup or restore of the data on

their workstation. Students were then supplied with their custom DOS boot diskette, instructions for using this diskette, the CDs to install the operating system, a copy of a diskette containing the driver for the Intel Pro 1000 network adapter, and instructions for installing the operating system and network adapter driver.

In order to test and learn how to use HiSBaR the students were instructed first to test their DOS boot diskette to be sure it worked by making a connection to the server and downloading a small (100 MB) test file. This was done to keep the students from using a bad DOS diskette. For example, if a student performed their operating system installation first, then attempted to back up that installation using a faulty DOS boot diskette they would not be able to complete the backup. In order to keep this from occurring they were required to first test their DOS boot diskette to be sure it was working.

The students would then proceed with the installation of the operating system and when they were at a stopping point, would shut it down safely, then boot up with the DOS boot diskette, and select BACKUP. This would run a Ghost backup script on the server which would first backup the previous image saved by the student if there was one, and then create a copy of their entire workstation installation on the server share. The names of the backups on the server were easily identified by student as they were related to the IP address each student was assigned.

The backup process was more potentially dangerous to a student's work than the restore process because for example a student might make configuration changes that would result in a non-bootable operating system then back them up on top of a working operating system. Consequently two measures were taken to hopefully minimize risk when backing up. First, as discussed previously, the Ghost backup script contained a DOS copy command that created a backup of the student's previously saved image before the current backup was overwritten. Second, after the students choose BACKUP when they had booted from their DOS boot diskette, they were shown a warning screen explaining exactly what backing

up did, and giving them the option of discontinuing the backup process.

4. ADVANTAGES OF HISBAR

Among the advantages obtained by the use of HiSBaR are that loss of a single lab machine was not catastrophic as students were free to use any available lab machine. In the past, when teams of students were assigned an individual machine, if that machine failed, then a huge problem of finding a suitable replacement and hopefully transferring work previously done to this new machine had to be undertaken. HiSBaR allows a student to restore a previously saved image to any available lab machine regardless of which lab machine was used to originally create the image.

Another advantage of HiSBaR is that it allows each student to perform an installation project individually, thereby allowing the author to ascertain that a student actually knew how to perform the tasks involved in the project. Prior to the introduction of HiSBaR, the projects were performed as groups and as a result some group members where not performing all the tasks of the project or learning all the concepts associated with project tasks. Having the students individually perform an Linux/Apache installation project, for example, greatly increased the general level of student understanding about this operation system and software which many of them had never previously used.

5. DISADVANTAGES OF HISBAR

Some of the unexpected disadvantages to this lab setup were the result of the additional number of projects being performed. For example, prior to the implementation of the HiSBaR system, one of the installation projects had been performed by groups of three to four students. The HiSBaR system allowed individual students to perform the installation project which multiplied the number of projects being performed by 3 to 4 times. Accordingly, this resulted in 3 to 4 times more projects to grade and also an increase in the number of project-related support questions to answer. To address this second issue, an online web forum for the projects was created to allow students to post questions and answers to project-related issues or problems. In fact, part of

the student's assessment for the project was related to their use of the support forum.

A related problem that must be faced by an instructor using the HiSBaR system is the ability to distinguish between real problems experienced by students and a student's inability to follow directions. For example, the process to install the gigabit Ethernet driver for the Intel Pro1000 network adapter under Linux was somewhat complicated and required an entire page of instructions since this network adapter was not automatically recognized by the Linux RedHat9 operating system used in the project. The author fielded a large number of student questions related to this procedure because it was required before the student could establish network and Internet connectivity. However, almost all of the problems students experienced with establishing Internet connectivity were not related to this complicated procedure but rather to their entering an incorrect Primary DNS address for their workstation.

Another issue related to the HiSBaR system that had to be overcome was that Ghost account used by the Ghost software requires full write access to drive where backups are stored. This would potentially allow someone to delete image files backed up on the server. In order to partially control this potential problem the boot disks given to students had the DOS commands that might be misused, such as DIR, DEL and ERASE, disabled using a hex editor.

Finally, the fact that there were only 6 machines available led to some excessive demand over supply right before the project due dates. Even though many were warned to complete the project early and were given 6 weeks to complete the project the lab was completely full for most hours in the 48 hours before the projects were due. Because of some charges that 'camping' on terminals had occurred in this period, a policy of 2 hours maximum time per machine was instituted for the lab and a sign in/sign out sheet was created for each lab machine.

6. FUTURE DEVELOPMENTS

Version 7.5 of the Ghost software which was used in this lab has a limitation when backing up a Linux system in that it will allow the system to be backed up even if the system was not shut down cleanly. When this happens it is possible that upon recovery of the backed-up system the user will experience a 'Kernel Panic' error and the system will not boot. This, in fact, happened to a student once during the semester even though all students had explicit instructions never to backup unless they had performed a clean software system shutdown. These instructions were also taped to the front of every monitor in the lab. Unfortunately for this student, he ignored all instructions and was forced to start the installation process over.

Another concern when using HiSBaR is that you must warn students to be careful when deciding whether to backup or restore. Two students after working on their projects for several hours and then deciding to quit and backup their work, actually selected the restore option and completely overwrote their latest work with previously saved work. Both the backup and restore scripts display warning screens and information on how to stop the script prior to a backup or restore actually taking place and students are warned verbally as well as in the written project installation instructions. In addition, a note taped to each monitor in the lab warned students to think carefully before selecting the backup or restore option.

One thing that would have greatly increased the productivity of the lab is to have a lab assistant available for simple maintenance tasks and to help students with some of the commonly asked project-related questions. At the present time funding does not allow hiring a lab assistant and consequently all responsibilities for the planning, implementation, and maintenance for the lab fell into the hands of the author.

Ghost version 8 actually contains some enhancements over version 7.5 which allows a disk check to be run on a Linux installation prior to its being backed up and terminates the backup process if an error is found. In addition, Ghost 8 also supports the Linux GRUB bootloader program, which version 7.5 does not, which constrained the projects to the use of the LILO bootloader. Consequently our plans are to purchase Ghost 8 for the lab in order to take advantage of this new feature.

One issue that should be addressed in future implementations of HiSBaR is to provide some means of automatically backing up the entire share drive containing all workstation images that was contained on the Ghost server using perhaps a tape drive, a second hard drive, or a network drive. The HiSBaR system performed an internal backup using the primary server hard drive of the last two images saved by each student; however, no backup was performed of the entire server hard drive. The simplest solution would probably be the installation of a second hard drive with some type of mirroring software and this is the approach planned as a future HiSBaR enhancement.

Finally, due to space constraints and accessibility issues, the server used for the backup and recovery was located in the same room as the lab workstations. Even though no problems were experienced in the initial semester of use, moving the server to a secure location or locking it in some kind of case within the lab is planned for the future in order to provide greater physical security.

7. CONCLUSIONS

This paper describes the planning, design, implementation, and utilization of a computer lab which allows high-speed backup and recovery of complete student workstation installations. The HiSBaR system provides many advantages and is not difficult to implement in addition to being relatively inexpensive. Issues and problems encountered in the development and use of HiSBaR were discussed and solutions, if found, were outlined. The HiSBaR system is flexible enough to meet the needs of both students and faculty and can provide an enhanced computer laboratory learning environment.

8. REFERENCES

Cox, Landon P., Murray, Christopher D. and Brian D. Noble, 2002, "Pastiche: making backup cheap and easy." OSDI '02: Proceedings of the 5th symposium on Operating systems design and implementation, December 2002, pp. 285–298.

Janz, Kenneth and Pei-Yi Hu, 2003, "Liberating lab computing: building a stable yet flexible computing environment for students and faculty." Proceedings of the 31st annual ACM SIGUCCS conference on User services, September 21-24, 2003, pp. 125–128.

Weeden, Elissa M., Scarborough, Gary R., and Dianne P. Bills, 2003, "Lab management strategies for IT database curriculum." Proceeding of the 4th conference on Information technology curriculum, October 16-19, 2003, pp. 62–66.