Secure Data Management in an English Speaking Test Implemented in General-Purpose PC Classrooms

Hideo Masuda
Kyoto Institute of Technology
Matsugasaki, Sakyo
Kyoto, JAPAN 606-8585
+81 75 724 7956
h-masuda@kit.ac.jp

Masayuki Mori
Kyoto Institute of Technology
Matsugasaki, Sakyo
Kyoto, JAPAN 606-8585
morim@kit.ac.jp

Katsunori Kanzawa
Kyoto Institute of Technology
Matsugasaki, Sakyo
Kyoto, JAPAN 606-8585
kanzawa@kit.ac.jp

Yasushi Tsubota
Kyoto Institute of Technology
Matsugasaki, Sakyo
Kyoto, JAPAN 606-8585
tsubota-yasushi@kit.ac.jp

Yumi Hato
Kyoto Institute of Technology
Matsugasaki, Sakyo
Kyoto, JAPAN 606-8585
hato@kit.ac.jp

Yasuaki Kuroe
Kyoto Institute of Technology
Matsugasaki, Sakyo
Kyoto, JAPAN 606-8585
kuroe@kit.ac.jp

ABSTRACT

The Kyoto Institute of Technology Speaking Test, “English for the 21st Century” is being developed to assess the English speaking ability of undergraduate students learning English as a lingua franca. The ultimate goal of this project is to introduce the computer-based English speaking test as part of entrance examinations to graduate programs. Despite the high-stakes nature of the test, it needs to be implemented in general-purpose PC classrooms mainly due to financial constraints. A secure data sharing system needs to be established between the PCs used for the test and the servers to preserve confidentiality, integrity, accessibility for the audio data. External graders will need access to carry out online evaluation of the collected data.

Also, the computer rooms must resume normal operation soon after the test administration. We administered the first two large-scale feasibility tests (approximately 700 examinees each) in January and December 2015. In this paper we will demonstrate the Windows custom image and secure data sharing tools we have developed for the tests and also report on how they were operated in the actual administration of the tests.

Keywords: PC classroom, computer-based test, online evaluation.

1. INTRODUCTION

In October 2012, Kyoto Institute of Technology (KIT) established a project team to introduce English speaking tests into the entrance examination to graduate programs [1]. Since the onset of the project, the academic and technical staff of the Center for Information Science at KIT have been working collaboratively with the academic staff teaching English at the institute in developing and implementing the computer-based speaking test, and

examining its feasibility and practicability as a high-stakes examination.

The project team decided to develop the computer-based test (CBT) for practical reasons. The difficulty with face-to-face interview tests, which are the prevailing option for English speaking tests, is that it would require employing and training a significant number of experienced English interview graders to assess the ability of some 700 applicants in under a week. Both budget limitations and concerns about the consistency of applying the English language interview grading rubric among so many temporary employees have made a computer-based test an attractive option.

While the amounts of data transmitted in ordinary CBTs are quite limited, far larger amounts of data need to be transmitted in English speaking tests because the sound quality of recorded questions and examinees' responses must be high enough to ensure the fairness required of a high-stakes examination.

Further, highly secure data transmission must be established between the on-campus server used for the test administration and the off-campus server prepared for the external raters to mark the examinees' oral responses. However, given the low frequency of administering the entrance examinations (less than 3 times a year), a large budget cannot be allocated for the installment of equipment used exclusively for the speaking test, not to mention the PC rooms designed for the test. The project team therefore decided to use some general-purpose PC rooms (“PC Labs” in Figure 1) that are used by students mainly for programming exercises [4].

In this paper, we will demonstrate how the computer system in these PC rooms were customized so that the institute could implement the speaking test, ensuring the high level of confidentiality, integrity, and availability required for quality data transmission for the entrance examination. We will also examine the results of the first two large-scale feasibility tests administered in 2015, and discuss further challenges to address to achieve the goal.

This work is licensed under a Creative Commons Attribution International 4.0 License.

Copyright is held by the owner/author(s).
SIGUCCS '16, November 06-09, 2016, Denver, CO, USA
ACM 978-1-4503-4095-3/16/11.
<http://dx.doi.org/10.1145/2974927.2974957>

Computer Infrastructure (System9) Center for Information Science, Kyoto Institute of Technology

March 2014.

Figure 1 Overview of Our Computer System (System9)

2. PREREQUISITES

KIT updates its computer system every four years [2,3,4]. The newest system named "System 9"[4] provides virtual server services, file sharing services, Web services for off-campus users, e-mail services, LMS services, user authentication services, terminal services for computer exercises, printing services and so on (figure 1). In this section, we will explain some of the important services used for the speaking test.

2.1 Network Boot for PC Labs

The terminal system for programming exercises at KIT consists of about 300 PC terminals, and students and faculty staff have free access to the system for self-study and programming exercises. Windows 7 and Linux (CentOS 6) are installed on the PC terminals and users can select either of them depending on their purpose of use. Although the OS in each PC terminal needs to be kept up-to-date for security reasons, it is difficult to treat some 300 terminals separately. Our system therefore utilizes a network boot method to realize updating work in an integrated way. For Windows 7, we have adopted Citrix PVS plus CO-CONV Readcache method, and for Linux, we use NFS root plus unionfs. This enables us to change OS without accessing each client PC's hard drive as long as we prepare necessary OS images and keep them updated on the server. CO-CONV Readcache is a mechanism to cache data blocks in local hard drives and reduces the data traffic at the time of the network boot. This alleviates the

traffic concentration in the case of simultaneous executions of the same application, which often occurs when computer exercise classes are in session.

2.2 File Sharing Server for Individual Users

System 9 uses a moving profile so that clients can use the same settings and access their files at every terminal. As a result, each user's file is not saved on local hard drives, but in the shared file server. The shared file server gives access only to the appropriate files for which the authenticated user has read and write permissions. We also set up a one gigabyte quota restriction to prevent users from creating files without a size limit and consequently suppress file space available for other users. The access protocols to the shared file server are CIFS for Windows and NFS for Linux. The Disk space accessible from Windows 7 is limited to \$Home/windows, which is a part of \$HOME that is accessible from Linux. This enables users to refer to their files, regardless of the OS they use.

Figure 2 Overview of the Speaking Test backend

2.3 User Management

System 9 has a master authentication database which has the account information of all the students and staff members at KIT. The database is updated automatically during the night. With the help of the master authentication database, the common authentication method can be used and the permission process can be executed only with specific user groups for each service. For the terminals used for programming exercises, there is an active directory server that has the information of students and academic staff only, and the use for administrative purposes is prohibited.

3. REQUIREMENTS

The followings are required for administering the English speaking test in the entrance examination.

[R0] In taking the test, the same conditions are ensured for all the examinees.

[R1] Test items are kept confidential until the commencement of the test.

[R2] No one but those specified in advance can take the test.

[R3] The test is not interrupted or suspended due to network or PC troubles.

[R4] Examinees' oral responses are securely recorded and saved on the server.

[R5] Examinees' oral responses are securely delivered to each external rater.

4. IMPLEMENTATION

Approximately 220 terminals installed for programming exercises in three separate PC rooms are used for the English speaking test. Using these terminals addresses [R0].

In order to prevent access by non-test-taking users, the test administration system is called after each examinee logs on to the Windows system using his or her own account. By registering the examinees' accounts in advance, access by those who are not registered is denied on the test day. Preventing access to the test by users not registered to take the examination addresses [R2].

If access to the external link were required for obtaining the test questions, there would be a greater possibility of network or PC troubles that inevitably lead to the suspension of the test. Our system utilizes a network boot method, and when a PC cannot access the Netboot server, the PC does not work. As a result, the situation where the PC is working but the test questions cannot be loaded is more likely to be avoided. Disabling terminals with network difficulties addresses [R3].

The test application is embedded in the Windows custom image, and therefore, access to the test application is ensured once the OS image loads. Having the test application integrated into the Windows images addresses [R3].

However, there arises a concern for the leakage of the test questions when the test application is installed on the OS image in advance. To prevent this, the OS image is changed/modified from a general-purpose image to a test image on the test administration day, utilizing the Network boot function. Swapping out the image on the testing day helps address [R1].

In most cases, three or more candidates take the speaking test, using the same PC terminals on the same day. By ensuring that the PC is shut down when each candidate finishes taking the test, his/her data files in the temporary file stage are deleted and cannot be manipulated.

Each candidate's oral responses are saved on the file sharing server. The backup is also stored on USB flash memory connected to each PC terminal before the test to be retrieved after the test. The storage area could be prepared on the local hard drive for the same purpose, but higher security can be ensured by saving the backup in USB flash memory. This helps address [R4].

All the oral responses of each examinee are saved in one folder prepared for him/her on the file sharing server, and, employing the rsync plus SSH method, a copy is made in the server used for the rating of the test. When the rating is completed the speech responses on the file sharing server are deleted. This addresses [R5].

5. LESSONS LEARNED

With the system described in the previous sections, the computer-based English speaking test was administered in January and December, 2015. 551 first-year students in the undergraduate program at KIT took the first test, together with 37 anchor examinees who took all three versions of the test used on that occasion. Anchor examinees took the test to help equate the versions. The second test was administered to 575 first-year students and 69 anchor examinees in the following academic year.

On both occasions, the test application was supplied by the test developer we work with one day before the test administration. We installed the test into the base image of Windows 7 and changed the setting of the boot server to boot the new image to the terminals on the test day. Just after the test, we transferred sound files and summary files to the vendor server. In addition, we uninstalled the test application and deployed the base image with out the test to the terminals.

Basically, the feasibility tests were successfully completed, but the followings are the problems for us to solve:

1. Since we deployed the new image for the test just before the test administration, at the first boot, ReadCache was not effective. We should be able to avoid this by booting all the terminals before the actual test.
2. There were a few errors related to writing speech responses to the file sharing system. Although this may have been due to the test application, we need to investigate the causes rigorously.

With regard to the first problem, we are planning to develop a slimmer OS image that only has the modules necessary for the test application and thus reduce the time to create the cache data. The second problem may be solved by further sophisticating the measures the test application has for avoiding temporary writing errors.

6. CONCLUSIONS

In this paper, we have demonstrated the system we developed for administering the English speaking component of the entrance examination in the general-purpose PC rooms and also reported on how the system was operated in the actual administration of the test. We have successfully completed the first two large-scale feasibility tests where some 700 first-year students took the speaking test. One challenge we now need to address concerns user authentication. At present, user authentication is operated as an account database of students and staff members of KIT. However, in order to introduce

the speaking test into the entrance examination, user ID and passwords need to be issued for each candidate. We should therefore decide on the way of giving user names to each candidate. At KIT, the user name each student gets issued is based on his/her student number, and s/he uses the user name until s/he left the institute. It is therefore necessary to seek a way of giving individual candidates the user names which they can continue to use after they matriculate to the institute.

7. ACKNOWLEDGEMENTS

This work was supported by JPRS KAKENHI Grant Number 16H03448. Special thanks to e-communications, Inc.

8. REFERENCES

- [1] Yumi, H., Katsunori, K.: Development and Execution of CBT English Speaking Test: Evaluation of Trial for Entrance Examination, KOUHOU of Center for Information Science, KIT, No.34, 30-48 (2015, In Japanese).
- [2] Hideo, M., Seigo, Y., Michio, N. and Akinori, S.: Using coLinux to Provide a Linux Environment on Windows PC in Public Computer Labs, In Proceedings of the 34th annual ACM SIGUCCS fall Conference, 221-224 (2006). DOI= <http://dx.doi.org/10.1145/1181216.1181266>.
- [3] Hideo, M., Kazuyoshi, M., Yu, S., Kouichiro, W. and Yasuaki, K.: KIT's Campus Computer System by Virtual Machine Technology and Integrated Identity Service, In Proceedings of the 38th annual ACM SIGUCCS fall Conference, 251-256 (2010). DOI= <http://dx.doi.org/10.1145/1878335.1878398>.
- [4] Hideo, M., Kazuyoshi, M., Yu, S., Kouichiro, W. and Yasuaki, K.: Distributed Campus Computer Infrastructure - Integrate Education, Research, Library and Office Activities, In Proceedings of the 42nd annual ACM SIGUCCS conference on User services, 93-96 (2014). DOI= <http://dx.doi.org/10.1145/2661172.2668055>.
- [5] Hideo, M., Kazuyoshi, M., Yu, S.: Low TCO and High-Speed Network Infrastructure with Virtual Technology. In *Proceedings of the 37th annual ACM SIGUCCS fall Conference*, 321-324 (2009). DOI= <http://dx.doi.org/10.1145/1629501.1629563>.
- [6] Hideo, M., Kazuyoshi, M., Yu, S. and Yasuaki, K.: High-Speed Network Infrastructure between KIT's Campuses for Computer System Redundancy, In Proceedings of the 40th annual ACM SIGUCCS Service & Support Conference, 109-110 (2013). DOI= <http://dx.doi.org/10.1145/2504776.2504818>.
- [7] Hideo, M., Kazuyoshi, M., Yuki, S., Yu, S. and Yasuaki, K.: Moodle Integration of an Automated Account Enabling System and a User Status Collection System, In Proceedings of the 39th annual ACM SIGUCCS fall Conference, 207-210 (2011). DOI= <http://dx.doi.org/10.1145/2070364.2070418>.