A secure and highly accessible examination system - *E-matura*: a case study

D. Jeske, M. Krasuski & R. Stryjek

Technical University of Łódź
Łódź, Poland

ABSTRACT: We live in an Internet world where designing a secure and highly accessible IT system is not an easy task to do. Each step of communication in complex systems is exposed to a range of types of hacker attack, wire tapping and others techniques that can compromise the system and make it unreliable. This paper describes each aspect of designing a secure and highly accessible IT system, using as an example the E-matura examination system which provides infrastructure for examination of large numbers of students simultaneously.

INTRODUCTION

The *E-matura* is a modern examination system created under the direction of Professor Sławomir Wiak at Technical University of Łódź. The system is design to handle a couple of thousands of students in real time manner and to provide maximum security for data safety. This complex and highly accessible system has many possible weaknesses, which had to be taken care of. Figure 1 shows all the steps in the communication process between the client and the server, where something could go wrong and the system was exposed to a hacker attack.

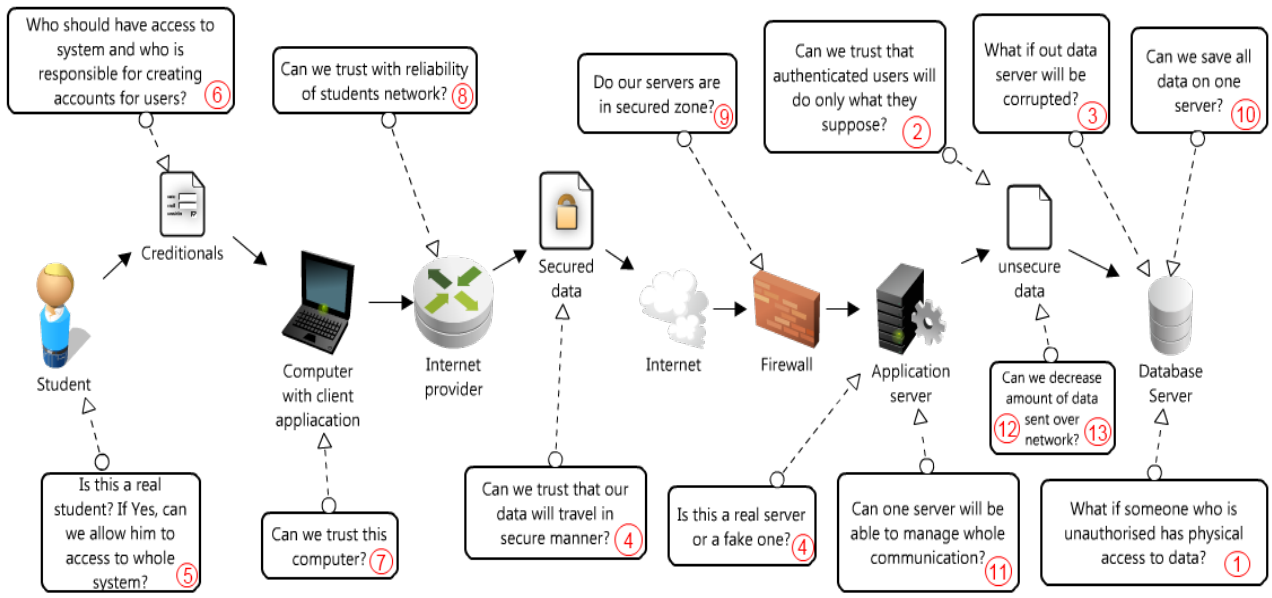

Figure 1: The communication process.

The paragraphs that follow provide answers to some of the questions that might arise from analysing this diagram and discusses them in great detail.

DATA SAFETY

Encrypting Personal Data

An IT system, which stores and processes the personal data of its users, must ensure very tight data security. In the E-matura project, this problem also appeared because the system has to store detailed information about the users (Name, Surname, Personal ID), i.e. students and teachers who participate in the project. Storing these data is required in order to identify each user of the project by the financial institution funding this project. Moreover, classified data such as examination questions and survey questions, which are asked during examination, are also stored in E-matura.

In the E-matura project, an SQL Server 2008 R2 Microsoft database was used, which provided several possibilities for protection of data against unauthorised access. Taking into account two basic features (the rate of application performance and security of classified data security), it was decided to encrypt data at a line level using the AES algorithm. Only table columns that store sensitive data were encrypted to avoid slowing down the whole database. By using encryption, even people with physical access to data are unable to use them without the proper decryption key.

Data Access Policy

When developing a Web-based project, which involves users who are very determined to pass the examination, other safety-related factors must be taken into account. One known problem, which often occurs in these types of systems, are SQL injection attacks. These attacks are very common in systems where the business logic used to retrieve data from the database is located within the application. An example of a malfunction can be the creation of a mechanism to login in, which queries about the user's existence are built into the application and then run as a dynamic SQL in the database. In this case, an unauthorised person can affect on the appearance of a built query by typing additional SQL commands. In this way, the system will check, if the password is correct and then execute the hacker's command.

To avoid this problem and to prevent access to data directly from the application, all communication between the Web application and the database must be done through stored procedures. In this way, the user cannot get to the tables and data, because the stored procedure which provides such access, is not available.

Database Backup and Log-shipping

Another aspect in the construction of a Web-based service the aim of which is to collect examination and personal data is protection against losing such data.

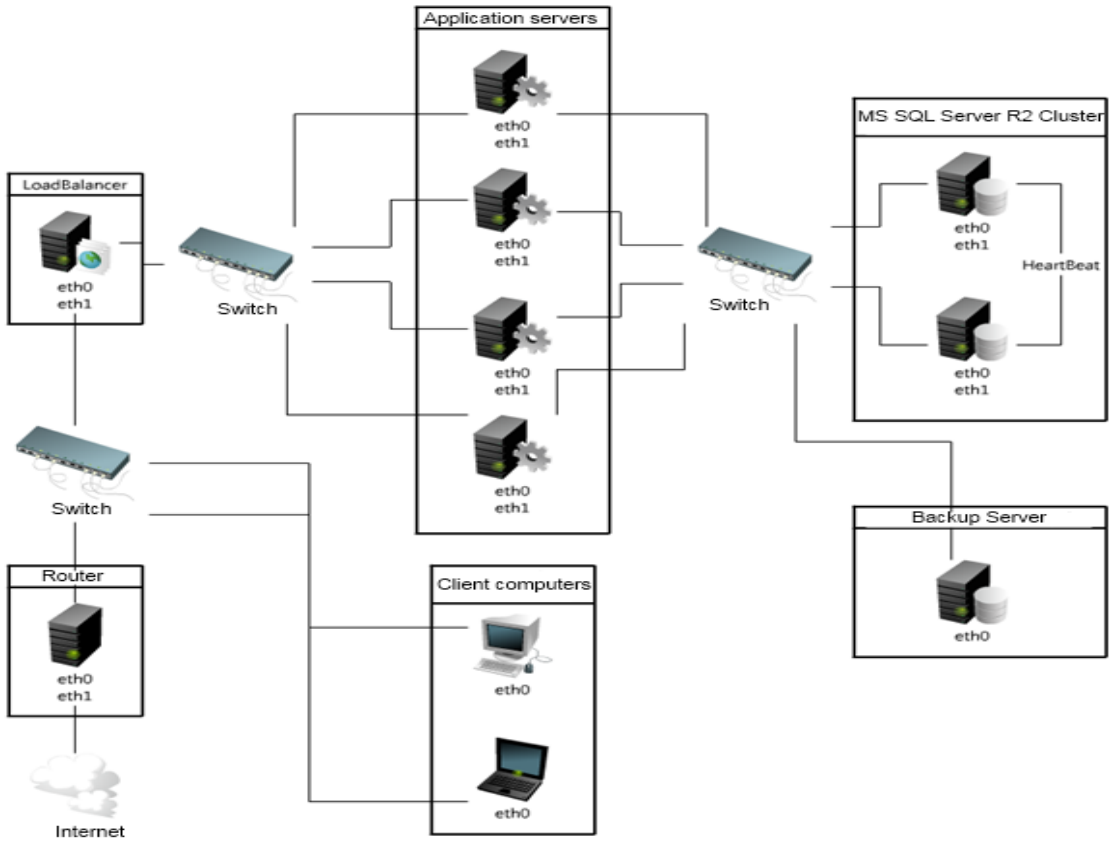

Figure 2: Database server logic separation.

To meet all the requirements, an infrastructure was built in which the entire database is backed up every night onto a separate server, which is located on a different subnet and connected directly to the database server. This server is not visible to application servers so data security is increased. To provide even greater data security and the possibility of restoring the data from a specified period, e.g. reaching the database of a user who could delete data, log shipping is used, which allows to perform transaction log backups every 15 minutes. Fifteen-minute spaces between transaction log backups provide with sufficient granularity of backups, while the database server loads low. Figure 2 shows the server infrastructure, which has been built in the E-matura project.

The above diagram shows the infrastructure of the E-matura project. To ensure greater security, most systems are duplicated and divided into separated subnets. Users who connect to the E-matura are unable to connect to database servers, because these servers are connected to a separate subnet to which only application servers have access.

Creation of such separations required using separated NICs in application servers to handle traffic from the Internet and separated network cards to communicate with the database. Due to this approach, an unauthorised person, who would like to have access to the data will have to deal with several servers on the way and, thus, such a person will have to spend much more time on the break-in. This situation will give the system management more time for emergency reaction.

Encrypted Connection between the Application Server and the Client Application

The client application communicates with the server part via a network service based on Windows Communication Foundation technology, version 4.0. This technology, using open standards such as HTTP and SOAP, provides the functionality to client applications. In order to ensure secure communication between the client and the server, an SSL, which allows the security of the system in two dimensions, has been used.

- Verification of Application Server;
- Encryption of transmitted data.

Verification of the server is based on a system of certificates through which it is possible to check an entity identified by a given certificate. The certificate is issued by a special certification centre at the request of an applicant. Each certificate is assigned to a particular Internet domain and is linked with a company or with an individual user. Such certificates can be issued only by an authorised centre so it cannot be issued by a fraudster or a third person having no such rights.

Each centre is further verified by the parent centre for further verification and preservation of the so-called certification path, consisting of a chain-level review of each issue of the certificate. Each certificate contains information about parent certification organisations so certificates can always be verified by the end user.

Each certificate includes a pair of keys: private and public, so that it is possible for an asymmetric encryption of data transferred between client and server. During the communication between the application server and the application installed on an end user's computer, keys are used to encrypt and decrypt transmitted information. By using this solution the data sent between the client and the server cannot be overheard by a third party.

Authentication Client Application by Using the Token with a Limited Lifetime

Encryption of the connection and server identification through a certificate issued by a trusted certification centre does not provide a total protection system. The user of the system knows that he/she is communicating with the original server and the data he/she enters will not get into the hands of third parties, but the same client must be verified to determine the resources that should be accessible. In order to identify a user properly, both authentication and authorisation checks need to be performed.

Authentication involves checking whether the person is who he/she claims to be. By the authorisation process, the system checks which resources/functionality the user is authorised to access. In the E-matura system, authentication is based on the username and password, which are checked when the user logs into the system. If the user gives the correct data, as a result of this operation, he/she will receive a specially-called token-generated number that is assigned to the current logon session.

This token is used to authenticate all Web service methods that provide a unique communication layer between the client and the server. By using a token, the username and its password are not sent with each request to a Web service, and this increases safety by reducing to a minimum the transmission of confidential data to the user. Additional protection is the life span of the token counter. Each token has a set life span that is incremented each time the site is accessed. If the service using the generated token is barred because of its invalidity, any subsequent attempt to gain access to the site returns an error and redirection to the login page. With this approach, a token captured on a victim cannot be used on another computer or the same computer in another session.

Closed Registration Procedure

Because the E-matura is not an open system in which every user, who knows the server address, can have access to the system, an additional layer of security could be introduced. The registration process begins with a personal or telephone contact with a representative of the recruiter who wants to obtain their own account in the system. The person concerned will later receive a registration form, which is a request to participate in the program. By using the traditional method of verification at this stage, all lead to the elimination of those who would like to get an unauthorised system account. After verifying the data, the person concerned receives via e-mail the information needed to log into the system. A person receiving such information has the authority to set up accounts for users within his/her unit.

Sharing the Examination Only to Verified Computers

In addition to verification of persons participating in examinations, it should be remembered that the account that receives the user can be stolen, or disseminated by this user. In order to provide additional protection against such an eventuality, a special verification system, which can connect computers to the system, has been created.

The person responsible for students' accounts must also perform the registration procedure for his/her computer that will be used during the examination. All computers that will participate in the examination must be verified before the examination and the number cannot exceed the number of the users in a given unit. (Precisely speaking, this number may be slightly higher than the number of registered computers because there must be a computer backup in case of failure of the main unit.) With this approach, access to the examination will only have verified and recorded CPUs.

Working Offline Application (Working on Questions Collected in Advance)

Safety is not only to ensure confidentiality of data and verification of the participants. For complete safety, the continuity of conducting the examination in all conditions should be ensured. The E-matura is designed to work with unexpected interruptions in the Internet access and in the power supply on both the client and server-side applications. After starting the examination, the system retrieves all questions for the client and stores them locally in an encrypted form. While providing each response, the system first saves it locally in an encrypted form and, then, tries to send the reply to the server.

If the answer cannot be sent immediately, it goes to a response pending queue. The queue is periodically checked and in the case of contact with the server, all replies are sent. If the failure continues, the system shuts down without giving out the examination results; however, all responses are stored on a local disk in a secure form and can be sent as soon as the fault is removed.

Firewall

The last element in ensuring security is already a standard at this point but one cannot forget about it when describing a comprehensive solution to a secure information system. This refers to the firewall, separating servers that are within the internal network from the so-called external world or the global Internet. The firewall allows for locking of the computer's ports on which the services are not operating, in order to increase safety and to not leave loopholes for malicious software that could get through such ports to the internal network.

ENSURING HIGH AVAILABILITY

Database Server Cluster

In systems for students' examinations, in addition to providing high security for data, database availability, while writing the examination, should also be ensured. The E-matura project's aim is to make examinations fully accessible to all candidates during the designated hour. With these assumptions, a huge number of visitors is handled, who begin an examination at exactly the same time. To ensure that the database is highly accessible, a cluster database based on Microsoft operating systems Windows 2008 R2 was built. The infrastructure of the database servers is identical and consists of the following elements:

- IBM HS22 Blade Server;
- 2x Xeon Processor E5640;
- 32 GB RAM.

Both servers use a shared disk array attached via a SAN. The cluster is configured as Active-Passive, allowing only one server to be running at a time, while the second one takes over the role immediately after the occurrence of any problem, without losing any data. All data (including transaction log files) are stored on a shared matrix thus the access to them has got to the server, which currently operates in an Active. Figure 3 shows how to connect the servers in a cluster.

A Cluster of Application Servers (Load Balancing)

The classic application running in the on-line environment is activated on a single Web server, which handles all the traffic generated by this and other applications that are installed on the computer. This solution works in most cases because the average numbers of people who use the Web server at the time are not able to overload the server's resources. A situation where several tens of thousands of people are referred at the same time to a single server refer is unacceptable. It is because to any connection to the server must be allocated a certain amount of memory and CPU time. To meet the requirements of high availability of the examination, the solution based on so-called Load Balancing was applied.

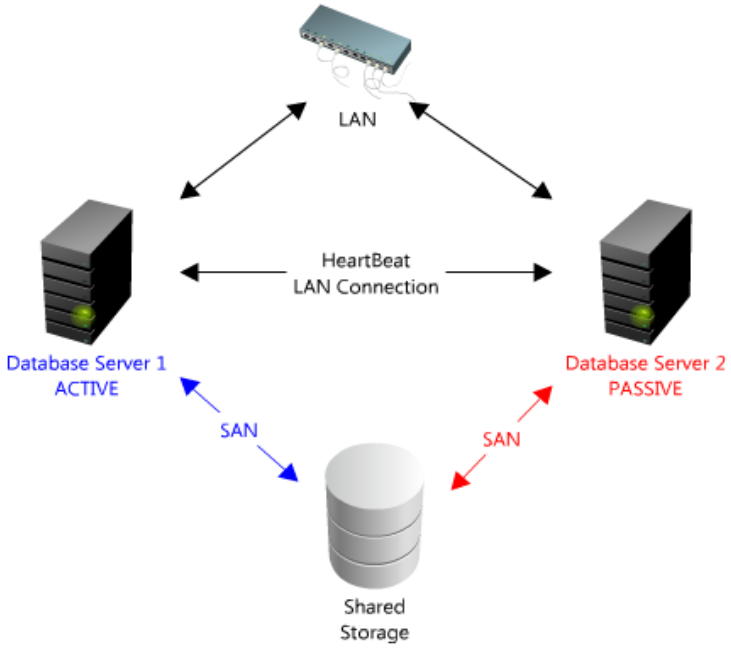

Figure 3: Connection of the servers in a cluster.

The solution was to build a cluster of servers in which two main parts can be distinguished. The first one is a computer constituting of the AP to the examination. All connections are directed to this computer, but they are not directly supported but only transmitted to computers in the cluster.

Based on the selected load management algorithm, this server redirects traffic to the least loaded server in the cluster. This solution is highly scalable and allows for virtually unlimited expansion of the cluster - the only limitation is the Internet bandwidth at which the communication takes place.

An additional feature that may satisfy this server is decoding an encrypted SSL message and passing it along as a decrypted message. This causes a reduction of the burden on the target servers; however, with a l arge number of connections it may cause an overload of the server balancing traffic.

Caching Data Retrieved From the Database

Sometimes data retrieved from the database server and sent to the application server duplicate the data relating to each customer. For example, the examination questions for the group of examinees are the same. Therefore, one can optimise access to the database by buffering data on server side applications. Data are collected only once during the first query to the application and, then, stored in an associative array, where there is a key that uniquely identifies buffering data.

For each data retrieval operation, this board is checked and when it appears that data have already been buffered, the database query can be replaced by fast data downloading from the buffer. It is worth remembering that stored data cannot occupy too much space in the computer's memory, because it can cause it to slow down or be immobilised.

A Limited Amount of Information Exchange

The last step of the application performance optimisation is to minimise the number and size of information exchanged. In order to ensure an adequate volume of exchanged data, all data sent between the client and servers are encoded in binary form. Unlike the classic implementation of the SOAP protocol in which data are sent in plain text, this is the solution that helps to reduce the message size by up to 40%.

CONCLUSIONS

Ensuring full security and high accessibility of data in the E-matura was not an easy task due to the nature of this project, which requires simultaneous access to data by all users participating in an examination. By constructing this solution, the developers have tried to ensure maximum scalability. If the need to support a greater number of users occurs, the solution could be adding just one more server without rebuilding the entire application.

The aim of this project is to provide an examination for about 30,000 graduates from the Łódź region in 2012. About 3,000 students attended the first test examination on 28 April 2011. They used server resources at the level of 8-10%. Taking into account the nonlinearity in respect of the load equipment, increasing the number of users and conducting synthetic tests, one can conclude that the purchased equipment should meet the requirements.