

# Measured Performance of an Information System

**Szikora Péter**

Budapest Tech, Hungary  
szikora.peter@kgk.bmf.hu

*Abstract: I know some problems about our scholar information systems (SIS), my colleague made a little student questionnaire about it. The most frequently mentioned problems were aborting connections, short timeout and slowness.*

*I think these problems are because of the quality of conceptual level data model, and the tools used for application development. Most SISs are written in programming languages and operate in a software environment consists of also commercial elements, and need a lot of resources. We use open source softwares to solve this problems, like apache, php, mysql. The test environment consists of PC computers and free softwares. The database server has an Intel processor of four cores and 8 GB RAM. It runs Linux, Apache with PHP and MySQL as an RDBMS.*

*The results were better than I had expected before. All the registrations of all the pupils were successful. The total time needed for the 32.768 registrations of the 8.192 students was 3 minutes and 7 seconds, so the average time needed for a test student 0.023 second.*

## 1 Scholar Information Systems (SISs)

I know some problems about our scholar information systems, my colleague made a little student questionnaire about it. The most frequently mentioned problems were aborting connections, short timeout and slowness.

I think these problems are because of the quality of conceptual level data model, and the tools used for application development. Most SISs are written in programming languages and operate in a software environment consists of also commercial elements, and need a lot of resources.

I would like to prove, that we can develop a better data model, better application so I can make a faster system, than general SIS. I chose Linux Operating system, because it is free and responsible. Linux is a free Unix-type operating system originally created by Linus Torvalds with the assistance of developers around the world. Developed under the GNU General Public License, the source code for Linux is freely available to everyone. Click on the link below to find out more about the operating system that is causing a revolution in the world of computers.

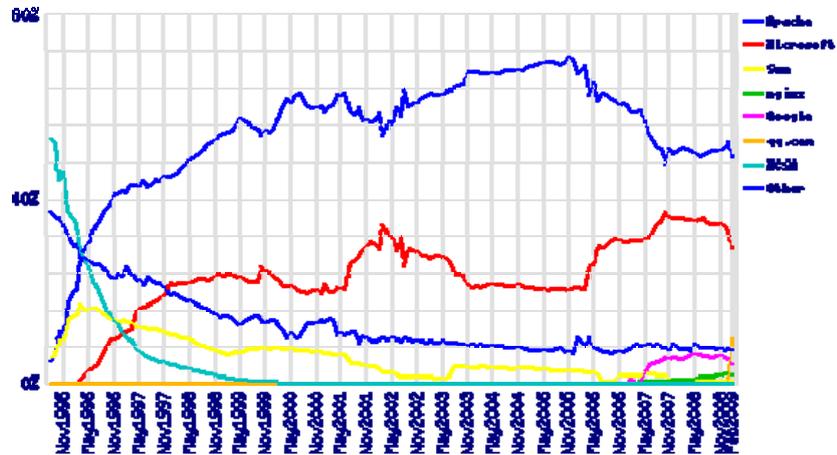


Figure 1

Market Share for Top Servers Across All Domains August 1995 - February 2009 [1]

The reasons for choosing Apache as the web server are numerous. The primary reasons, however, boil down the licensing, and the active support and development. Apache, on top of Linux, is used in most cases in all over the world so it can be considered the most accurately tested http server. See Figures 1 and 2.

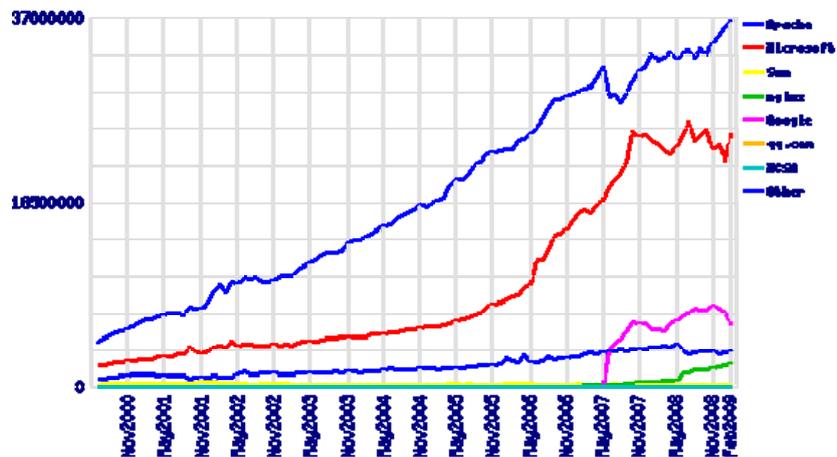


Figure 2

Totals for Active Servers Across All Domains June 2000 - February 2009 [1]

The MySQL database has become the world's most popular open source database because of its consistent fast performance, high reliability and ease of use. MySQL runs on more than 20 platforms including Linux, Windows.

Top reasons to use MySQL: scalability, flexibility, high performance and availability, strong data protection, open source freedom and 24 x 7 support and lowest total cost of ownership [2].

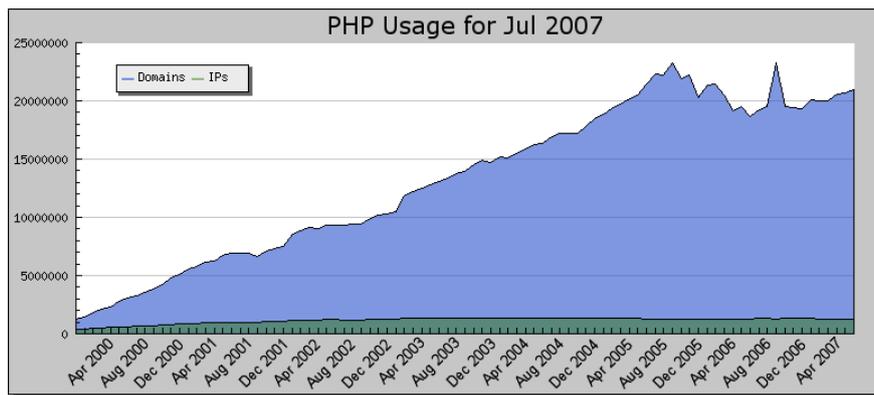


Figure 3  
Usage Stats for April 2007 [3]

On Figure 3 shows that PHP broke the 20 million domains, that means 20 million sites running PHP. PHP has the best documentations, and it is free too, and very easy to use.

The Open Source Definition is used by the Open Source Initiative to determine whether or not a software license can be considered open source. The definition was based on the Debian Free Software Guidelines, written and adapted primarily by Bruce Perens. Perens in turn based his writing on the "four freedoms" of Free Software from the Free Software Foundation [4].

## 2 The Background

The test environment consists of PC computers and free softwares. The database server has an Intel processor of four cores and eight GB RAM. It runs Debian Linux, Apache 2.2.9 with PHP 5.2.6 and MySQL 5.1 as an Relational DataBase Management System (RDBMS).

My colleague developed a little, clear but consistent data model for testing a SIS module. This data model is very minimal. This is a very important issue, redundant database need much more storage capacity. The test database contains 8192 students, 4 examinations for each of them and the appropriate number of dates for them. The registration's method is we called a php script via in an offline browser.

We need to make parallel connections to web and database server. We chose offline browser wget to call PHP script with get method, because when we using the GET variable all variable names and values are displayed in the URL. Wget is a simple computer program that retrieves content from web servers, and is part of the GNU Project. [5] If we didn't use an offline browser we ought to more than 8000 "test people" to click in the same minute – which would need the same number of computers.

In the test database we have 11 tables. It's size is less than 2 Megabyte. There are 8192 students with many data of them, in table 'HALLGATOK', like name, UserID, birthday, birthplace and mother name. There are more then 1500 exam occasions and subjects in 'VIZSGAALKALOM' and in 'TARGY' tables.

### 3 PHP Script – Applying for Exam

The script is written in PHP[6]. PHP, which stands for "*PHP: Hypertext Preprocessor*" is a widely-used Open Source general-purpose scripting language that is especially suited for Web development and can be embedded into HTML. Its syntax draws upon C, Java, and Perl, and is easy to learn. The main goal of the language is to allow web developers to write dynamically generated web pages quickly, but you can do much more with PHP. The PHP script needs to be integrated in short blocks, because of the developing, and manageability.

We use sessions in code, because with it, we can distinguish students. When we use session, the visitor who is accessing the web site is assigned a unique id, the so-called session id. This is either stored in a cookie on the user side or is propagated in the URL [7].

We started it from a shell script with the student ID as a parameter. First of all, we need to see the blocks. in the first block, user logs in to system. After it, he can see all of his exams. The script chooses one exam occasion for all of his subject, and applying for them. If there is no space on that occasion, chooses another exam. The choice is made by a random function. Only that occasions are listed which are available for the student. Time at the beginning of the connection and when the response was saved to a local file by the browser was logged.

Test registrations were started almost simultaneously with a two second pause after every one hunderd starts. The settings of the offline browser wget were: max 4 retries, 30 seconds timeout, 10 to 30 seconds between to retries.

## 4 The Results – Average Load

The results were better than I had expected before. All the registrations of all the pupils were successful. The total time needed for the 32.768 registrations of the 8.192 students was 3 minutes and 7 seconds, so the average time needed for a test student 0.023 second.[8] On figure 4 we can see the average load of the server, in every minute . We can clearly see in the first half of the figure, how the load grows. There are increases when the script starts new exam-applications. At the end the load decreases as the rest of the exam-applications are recorded into the system. [9]

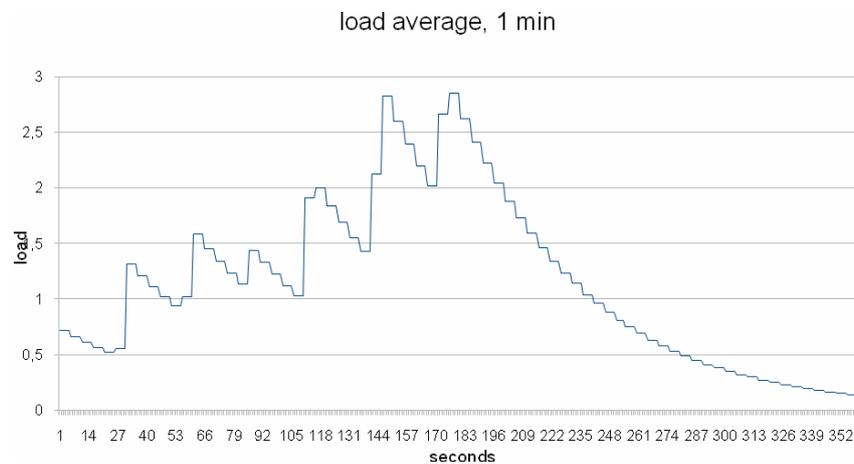


Figure 4  
Load average, 1 min

The load average tries to measure the number of active processes at any time. As a measure of CPU utilization, the load average is simplistic, poorly defined, but far from useless. The load average is the sum of the run queue length and the number of jobs currently running on the CPUs. High load averages usually mean that the system is being used heavily and the response time is correspondingly slow. [10]

### Conclusions

Summarizing the above I can state that (much) better results can be achieved even in proper hardware environment if we choose the appropriate software tools not really for developing but for the operating environment as well. [11]

### Acknowledgement

First of all, I would like to thank my colleague Keszthelyi András for all of his help in this work. Also I would like express my thanks to Kadocsa György who played a decisive role. Several people have expressed their willingness in this subject.

## **References**

- [1] Apache, [http://news.netcraft.com/archives/2009/02/18/february\\_2009\\_web\\_server\\_survey.html](http://news.netcraft.com/archives/2009/02/18/february_2009_web_server_survey.html) 2009
- [2] Mysql Documentation, <http://dev.mysql.com/doc/>, 2009
- [3] PHP usage, <http://www.php.net/usage.php>, 2009
- [4] Open Source definition, [http://en.wikipedia.org/wiki/Open\\_source](http://en.wikipedia.org/wiki/Open_source), 2009
- [5] Peter Moulding: PHP Black Book, Perfect-Pro Kft., 2002
- [6] Wget, <http://en.wikipedia.org/wiki/Wget>, 2009
- [7] PHP Documentation, <http://www.php.net/manual/en/>, 2009
- [8] András Keszthelyi: The Role of Data Modeling in Information System, 2<sup>nd</sup> International Conference for Theory and Practice in Education, Budapest, Hungary, May 29, 2009
- [9] Péter Szikora: The Role of the Tools and Methods of Implementation in Information System Efficiency, 2<sup>nd</sup> International Conference for Theory and Practice in Education, Budapest, Hungary, May 29, 2009
- [10] Load Average, <http://www.teamquest.com/resources/gunther/display/5/>, 2009
- [11] András Keszthelyi: How to Measure an Information System's Efficiency?, 7<sup>th</sup> International Conference on Management, Enterprise and Benchmarking, Budapest, Hungary, June 5-6, 2009