ANALYSIS AND MODELING OF THE BEHAVIORS OF AN ASTERISK SERVER USING THE METHOD OF LEAST SQUARES

Ravonimanantsoa Ndaohialy Manda-Vy, Rakotondraina Tahina Ezechiel, Randriamitantsoa Andry Auguste
University of Antananarivo
ESPA BP 1500 Antananarivo/ Madagascar
ndaohialy@blueline.mg

ABSTRACT

With the convergence of new technologies over IP and the emergence of new VoIP services, what are the most efficient methods for estimating the use of a VoIP server? This paper is oriented towards the study of the behaviors of an Asterisk server using the method of least squares. The results of the simulations show the limitations of server usage based on the number of clients.

KEY WORDS

VoIP, Asterisk, Least squares, SIP.

1. Introduction

The study of Asterisk servers and the use of hardware resources based on the number of calls are made using interpolation techniques. This allows an examination of the relationships between various and complex server physical parameters. In addition, despite research conducted on VoIP [1] [2] [3] [4] [5], both in the theoretical and experimental fields, VoIP servers lack analytical models predicting their behaviors, especially in relation to number of users. The least square method is often used in complex models. A large number of measures were unachievable, and underlying laws were unknown. The least square method seems to meet the required specifications of our server simulation.

Our primary goal is to design a model as a parameter based on number of calls and use of resources. This will allow us to have a set of data that will be grouped according to the criteria of least squares [8] to better model this server. The article [9] shows that it is possible to get an estimate of resources material using the method of least squares estimates.

First, we describe the state of the research on Asterisk servers. Next, we describe the use of the least square method in term of analysis and modeling. Finally, we build a model out of our analysis using MATLAB environment.

2. State of art

2.1 Comparative studies

In the article "a comparative study of standard VoIP with asterisk" [1], the authors perform different research-based VoIP protocols. The experiment is as follows: The block diagram of the system deployed for the study comprises three components: an Asterisk server (the unit under observation), a call generator (for calls to be sent to asterisk), and a Server Monitor (that will monitor the Asterisk server). The scenario for the interconnection of these three elements consists of the call generator and the Asterisk server connected by two different Ethernet interfaces (eth0 and eth1). The calls will initially be transmitted via eth1 link supporting the VoIP flow. The monitoring equipment is connected to the same switch as the eth0 link to the other two units. For the Asterisk server, they installed a Debian GNU Linux system on a Pentium 4 (2’4 GHz) with 1GB of RAM memory. Asterisk 1.4 has been installed from the Debian packages, using standard configurations.

2.2 The results

<table>
<thead>
<tr>
<th>Number of calls</th>
<th>Real value</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>2,36</td>
</tr>
<tr>
<td>20</td>
<td>4,64</td>
</tr>
<tr>
<td>30</td>
<td>7,12</td>
</tr>
</tbody>
</table>

Table 1

<table>
<thead>
<tr>
<th>Number of calls</th>
<th>Real value</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>33,38</td>
</tr>
<tr>
<td>20</td>
<td>38,23</td>
</tr>
<tr>
<td>30</td>
<td>42,09</td>
</tr>
</tbody>
</table>

Table 2
If the relation R (x, y) ≈ ± 1, there exists a relationship between x and y. If the relation R (x, y) = 0, the relationship does not exist. We will consider the following variables:

- X = number of calls
- Y = CPU usage
- Z = use of memory
- T = the bandwidth of the use

### 3 Modeling and analysis of a VoIP server

#### 3.1 VoIP server analysis

Based on the foregoing, we used the least squares method to determine whether there is a correlation between two variables predicting the use of server hardware resources:

- Forecast for the processor frequency useful for a number of users,
- Prediction of the size of RAM occupied by a given number of users,
- Prediction of the bandwidth occupation
- And forecast number of users based on specific hardware resources.

We will define a second mathematical formulation and it will be used with the least square method. There will be an adjustment in the form of:

\[
 y = ax + b.
\]

Thus, we obtained

\[
0238X - Y = 0, 05,333
\]
\[
Z = 0,4355 X 29,19
\]
\[
T = -0,0025 X 86,066667
\]

Since Y, Z and T are not integral equations, given above are the number of calls spent on an Asterisk server using as parameter the material, the ideal model to determine the number of potential customers is to set Y and / or Z and / or T.

The memory size required, the clock frequency required for the processor, and the bandwidth of the server.

So we get the equation of our VoIP server by adding the previous equations depending on the number of calls

The mathematical model of an Asterisk server compared to the number of calls is given by the following equation:

\[
X = \frac{Y + Z + T - 115,2027}{0,671}
\]

Modeling with numerical values of an Asterisk server is then represented by the equation:

\[
\begin{pmatrix}
  \frac{x_1}{x_n} \\
  \vdots \\
  \frac{x_n}{x_n}
\end{pmatrix} = 1 \frac{1}{0,671} \begin{pmatrix}
  \frac{y_1}{y_n} \\
  \vdots \\
  \frac{y_n}{y_n}
\end{pmatrix} + \begin{pmatrix}
  \frac{z_1}{z_n} \\
  \vdots \\
  \frac{z_n}{z_n}
\end{pmatrix} + \begin{pmatrix}
  \frac{t_1}{t_n} \\
  \vdots \\
  \frac{t_n}{t_n}
\end{pmatrix} - 115,2027
\]

### 2.3 Method of least squares

A common situation in statistics is to have two data sets of size n, n, \{y_1, y_2, ..., y_n\} and \{x_1, x_2, ..., x_n\}. They are obtained experimentally or measured using a population. The regression problem is to find a relationship existing between x and y as in the form y = f (x). When the desired relationship is affine, of the form of y = ax + b, it is called a linear regression. Even though such relationship exists, the measured data do not satisfy exactly the relationship.

Taking into account of the mathematical model of observed errors, we consider the data \{y_1, y_2, ..., y_n\} as a set of achievements of random variable Y and the data \{x_1, x_2, ..., x_n\} as another set of random variable X. Variable Y is identified as a dependent and variable X is an explanatory variable.

### 2.4 The line of least squares

Data \{(x_i, y_i), i = 1, ..., n\} can be represented by a cloud of scattered plot of n points in the (x, y) plane. The center of gravity of the cloud can be easily calculated. It is the point of coordinates \((\overline{x}, \overline{y}) = \left(\frac{1}{n} \sum_{i=1}^{n} x_i, \frac{1}{n} \sum_{i=1}^{n} y_i\right)\). Finding an affine relationship between Y and X variables returns a line that fits to scatter. a is the slope of the line and b is the origin.

Calculation shows that these values, denoted a and b are equal to

\[
a = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n} (x_i - \overline{x})^2}
\]

\[
b = \overline{y} - a\overline{x}
\]

### 2.5 The correlation

The correlation between two or more random variables or numerical statistics is the measure of intensity of connection existing between the variables. In this case, we have four variables: the number of calls, CPU utilization memory usage, and bandwidth usage.

The intensity of the correlation is obtained by calculating the linear correlation coefficient. This coefficient is the ratio of covariance and nonzero product of their standard deviations.

The correlation coefficient is between -1 and 1.

### Table 3

RAM usage in kbps

<table>
<thead>
<tr>
<th>Number of calls</th>
<th>Real value</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>86,05</td>
</tr>
<tr>
<td>20</td>
<td>86,00</td>
</tr>
<tr>
<td>30</td>
<td>86,00</td>
</tr>
</tbody>
</table>

### 3.1 VoIP server analysis

Based on the foregoing, we used the least squares method to determine whether there is a correlation between two variables predicting the use of server hardware resources:

- Forecast for the processor frequency useful for a number of users,
- Prediction of the size of RAM occupied by a given number of users,
- Prediction of the bandwidth occupation
- And forecast number of users based on specific hardware resources.

We will define a second mathematical formulation and it will be used with the least square method. There will be an adjustment in the form of:

\[
y = ax + b.
\]

Thus, we obtained

\[
0238X - Y = 0, 05,333
\]
\[
Z = 0,4355 X 29,19
\]
\[
T = -0,0025 X 86,066667
\]

Since Y, Z and T are not integral equations, given above are the number of calls spent on an Asterisk server using as parameter the material, the ideal model to determine the number of potential customers is to set Y and / or Z and / or T.

The memory size required, the clock frequency required for the processor, and the bandwidth of the server.

So we get the equation of our VoIP server by adding the previous equations depending on the number of calls

The mathematical model of an Asterisk server compared to the number of calls is given by the following equation:

\[
X = \frac{Y + Z + T - 115,2027}{0,671}
\]

Modeling with numerical values of an Asterisk server is then represented by the equation:

\[
\begin{pmatrix}
  \frac{x_1}{x_n} \\
  \vdots \\
  \frac{x_n}{x_n}
\end{pmatrix} = 1 \frac{1}{0,671} \begin{pmatrix}
  \frac{y_1}{y_n} \\
  \vdots \\
  \frac{y_n}{y_n}
\end{pmatrix} + \begin{pmatrix}
  \frac{z_1}{z_n} \\
  \vdots \\
  \frac{z_n}{z_n}
\end{pmatrix} + \begin{pmatrix}
  \frac{t_1}{t_n} \\
  \vdots \\
  \frac{t_n}{t_n}
\end{pmatrix} - 115,2027
\]
And after verification, we obtained the results shown in the following tables:

### Table 4
**CPU usage in %**

<table>
<thead>
<tr>
<th>Number of calls</th>
<th>Real value</th>
<th>Value obtained by the least squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>2,36</td>
<td>2,32</td>
</tr>
<tr>
<td>20</td>
<td>4,64</td>
<td>4,7</td>
</tr>
<tr>
<td>30</td>
<td>7,12</td>
<td>7,08</td>
</tr>
</tbody>
</table>

### Table 5
**RAM usage in Megabytes**

<table>
<thead>
<tr>
<th>Number of calls</th>
<th>Real value</th>
<th>Value obtained by the least squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>33,38</td>
<td>33,54</td>
</tr>
<tr>
<td>20</td>
<td>38,23</td>
<td>37,9</td>
</tr>
<tr>
<td>30</td>
<td>42,09</td>
<td>42,255</td>
</tr>
</tbody>
</table>

### Table 6
**Bandwidth usage in kbps**

<table>
<thead>
<tr>
<th>Number of calls</th>
<th>Real value</th>
<th>Value obtained by the least squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>86,05</td>
<td>86,04</td>
</tr>
<tr>
<td>20</td>
<td>86,00</td>
<td>86,01</td>
</tr>
<tr>
<td>30</td>
<td>86,00</td>
<td>85,99</td>
</tr>
</tbody>
</table>

The mathematical model of a VoIP server is defined by:

\[
\begin{pmatrix}
  x_1 \\
  \vdots \\
  x_n
\end{pmatrix} = C_1 \begin{pmatrix}
  y_1 \\
  \vdots \\
  y_n
\end{pmatrix} + \begin{pmatrix}
  z_1 \\
  \vdots \\
  z_n
\end{pmatrix} + \begin{pmatrix}
  t_1 \\
  \vdots \\
  t_n
\end{pmatrix} - C_2
\]

All headings should be in Times New Roman and bolded. Headings may be fully or left-justified; please choose one style and use it throughout your paper. Also, ensure that the spacing between headings and text are consistent. If the heading goes over one line, ensure with:

- **x** a matrix column represents the number of calls \( \begin{pmatrix}
  x_1 \\
  \vdots \\
  x_n
\end{pmatrix} \)
- **y** a matrix columns representing CPU usage \( \begin{pmatrix}
  y_1 \\
  \vdots \\
  y_n
\end{pmatrix} \)
- **Z** a matrix columns representing the consumption of the RAM \( \begin{pmatrix}
  z_1 \\
  \vdots \\
  z_n
\end{pmatrix} \)
- **T** a matrix columns representing the occupation of the bandwidth \( \begin{pmatrix}
  t_1 \\
  \vdots \\
  t_n
\end{pmatrix} \)

\( C_1 \) and \( C_2 \): constants that define the physical limitations of the server

#### 3.2 Result of the model

- From the tables, we could see that the results obtained with our model look about the same as the results of the experiment.
- From the calculation of the size of the memory used, CPU usage and bandwidth usage on an Asterisk server, we can deduce the number of clients that we can assign to this server.

#### 4 Interpretation and discussion

We are able to propose a mathematical model of a VoIP server based on the use of resources characterizing the server.

After calculating the correlation, we found that there was a relationship between the number of calls and CPU usage. Additionally, there is a relationship between memory usage and the number of calls and enter the latter and bandwidth. Using hardware with limited capacity, the number of users on an Asterisk server is limited. An operator intending to achieve maximum number of customers has to use powerful hardware resources.

Furthermore, we found that variation in terms of bandwidth is almost insignificant compared to the change in the number of calls. This proves that VoIP does not necessarily fully depend on the bandwidth variable. Consequently, VoIP and SIP in particular partially interfere with other protocols across the Internet.

#### 5. Conclusion

In this article, we presented a model of an Asterisk server depending on the number of calls made by the server's hardware resources and the server's behaviors. The method of least squares demonstrated that there is a relationship between the number of calls and the use of resources. Furthermore, the method of least squares presented approximate results to the actual results. In term of prediction, it can be used to determine the boundary of a server. Based on these results, we developed a model using Asterisk server hardware resources. This model can be an object of further research.
Figure 1. CPU Utilisation

Figure 2. RAM Utilisation
Figure 3. Bandwidth occupation

References