A Web Service Monitoring Indicator and Model System and Performance

Zhang Yinsheng, Qiao Xiaodong, Han Feng, Wang Jitian, Liang Jian, Li Peng

1 Institute of Scientific & Technical Information of China, 15 Fuxing Road, Beijing China, 100038 zhangyinsheng@istic.ac.cn
2 Open sources Qizhi (Beijing) Ltd, Inc., Beijing, China, 100125, wjt@qzsoft.cn

Abstract

This paper analyses the current technical limitations in the monitoring web services which mainly focus on the lower layers rather than application layer in TCP/IP stacks. So, the paper puts forward an indicator and model system, proposes some technologies to monitor web services. The indicators add users’ behavior and features in comparing with the current technologies, and integrates some services concerning browser activities in the layer higher than application layer protocol in TCP/IP stacks. For realizing the monitoring indicators, key technologies are given which are to insert tags in the HTML files indicating the services and link browser to monitoring center sending the browser information for analyzing users’ behaviors.

1. Introduction

Generally speaking, web services refer to the serial processes of operation of software codes in Internet based on HTTP or other Internet standard protocols to satisfy another serial processes of operation of software codes. A foundational feature of web services lies in that it can be described in XML.

Many technologies are developed to monitor web services. Table 1 enumerates the basic functions of the current web services monitoring technologies.

Table 1 shows that the mainstream technologies monitoring web services narrow on the lower layer segments in TCP/IP which posses the following limitations:

1. Web services mainly described in XML belong to the events in the layers of application and presentation in TCP/IP. It means that every service mostly perform all the protocols in the stack of TCP/IP, so the technologies monitoring events in the lower layer segments in TCP/IP are not enough for the users who hope knowing a web service through all the protocols of TCP/IP.

2. The above mentioned technologies lacks of capabilities monitoring the users in Internet who browse the servers, in the other words, the most performance information of web services are lost using the above technologies.

3. The clear indicators and models of general web services, which can describe the main functional characters of web services, are still short.

So, the present paper aims at establishing a system of the indicators with their models to realize monitoring web services for a common website, knowing what do the web services do, and how is the performance of web services.

2. The Relevant Studies on the Indicators of Web Services

To describe the main functions and attributes of web service resources, we need to integrate a serial of indicators. Some relevant studies on the indicators of web services mainly enumerated as following:

International Telecommunication Union (ITU) constituted some standards regulating the technology specifications in TCP and IP layers such as "IP Data Communications Business-IP Packet Transmission Performance and Availability Parameters" (ITU-T Y.1540) 10, "Internet Protocol Communications Services - IP Packet Transmission and Distribution and Availability Parameters " (ITU-T Y.1541) 11, "IP Network Technical Requirements - the Network Performance Parameters and Indicators" (YD / T 1171-2001) 12 These documents mainly regulate the

\[ \text{All the authors are co-first authors and Qiao Xiaodong is the corresponding author.} \]
indicators in IP or lower layers. SOAP Version 1.2 regulates the specifications in higher layers. As the functions of web services are complex, the general indicators describing web services has not yet formed. The achieved reaches mainly focus on security, reliability and security.

3. A Monitoring Indicator and Model System

The proposed indicators with models describing the main functions and attributes of web service are about as following.

I. Indicators depicting users

I-1. The Number of Users
The Number of Users = the number of computers accessed the monitored website

I-2. The Number of Users in Different Business (or Resources)
The Number of Users in Different Business (or Resources) = the number of computers accessed the different pages indicating the business (or resources)

I-3. The Number of Users in Different Nations or Regions
The Number of Users in Different Nations or Regions = the number of computers accessed the monitored website in the different nations or regions

I-4. The Number of Users with Different IP Addresses
The Number of Users with Different IP Addresses = the number of computers accessed the website with the different IP addresses

II. Indicators depicting web service features

II-1. The Number of Accessing a Website
The Number of Accessing a Website = The number of different computers accessed on the monitored website

II-2. The Number of Accessing Important Pages
The Number of Accessing Important Pages = The number of different computers accessed on the Important Pages

II-3. The Number of Accessing Important Pages in Distribution of Users’ IP addresses
The Number of Accessing Important Pages in Distribution of Users’ IP Addresses = The number of computers accessing important pages in distribution of users’ IP addresses

II-4. The Number of Accessing Important Pages in Distribution of Users’ Nation or Region
The Number of Accessing Important Pages in Distribution of Users’ Nation or Region = The Number of computers accessing important pages in distribution of users’ nation or region

II-5. Most Popular Services
Most Popular Services = The ranking of the pages denoting the different services

II-6. Web Services Amount with Different Types
Web Services Amount with Different Types = the number of accessing the page combines represented a service

II-7. Utility Efficiency of Different Information Resources
Utility Efficiency of Different Information Resources = the number of accessing the page combines represented the resource by order

III. Indicators depicting the web services under application protocol layer

III-1 Host Monitoring Indicators
III-1-1. Management Information Server
III-1-2. Server Basic Parameters
III-1-3. Server Trouble-Free Run-Time
III-1-4. The Basic Information of Operating System
III-1-5. Operating System in Trouble-Free Time

III-2 IP Protocol Layer Events Monitoring Indicators
III-2-1. Network Equipment Information
III-2-2. Network Equipment Management Information
III-2-3. Packet Delay (package speed)
Packet Delay = go-in time - departure time = $t_2-t_1$

III-2-4. IP Packet Throughput
IP Packet Throughput = successful packages/duration

III-2-5. Packet Loss Ratio
Packet Loss Ratio = lost packages transiting outcomes/total packages

4. The Key Technologies Realizing the Monitoring

To realize the indicators above mentioned, the proposal technology is to gather user’s information
from browsers which is short for the current technologies of web service monitoring. For gathering user’s information, the current technologies mainly do it through the server-side access logs (apache log files, for example, IIS log files). However, this method has its limitations: it cannot get real-time information; log collection is difficult, especially in clusters and content delivery network (CDN), it is unable to collect the information about user’s environment such as access path, the number of user sessions, etc., and it cannot distinguish the computers using one IP address. Some studies explored to gather user’s information from the user’s side. Xiang Jianchi enumerated the relevant technologies including Java Applet, JavaScript, Plug-in and Frame. Wu Qi suggested recording browse behavior in user’s side using access attribute list and IE plug-in. These studies has not solved the problem how to send the gathered information to a monitoring center, and the technologies lack of detailed methods for application.

The proposed technology is to embed script codes as the tags of visited pages in the monitored page and to link the browser to the monitoring center, meanwhile, gather the information of user’s behavior. This technology can overcome the limitations in the current technologies and realize the indicators of I and II in §3.

For example, the embedded codes in a page “istic.html” are about as following:

```
<html>
<head></head>
<body></body>
</html>
<script src='http://68.160.12.53/js/istic.js'></script>
```

“istic.js” is about as following:

```
SCRIPT_URL = 'http://168.160.12.53/cgi-bin/';
WEB_ID=1;
var userDate = new Date();
document.write('<img src="'+SCRIPT_URL+'/g.fcgi?id='+WEB_ID+'&ut='+userDate.getTime()+'" width="0" height="0">');
```

In the program,

“168.160.12.53”is the IP address of monitoring center

“WEB_ID” is the number of monitored website

“userDate” is record users’ access time

“SCRIPT_URL” is the URL of the visited page embedded codes. These paths indicate the pages represent the web services as in the indicators.

The information gathering works are performed by the file “g.fcgi”, and the information will be saved in the database for computing indicators according to the models.

Table 2 and Fig.1 are the samples showing the gathering outcomes of “istic.js” for outputting indicator II-1(The Number of Accessing a Website).

5. Conclusion

Users’ information from browsers can be used widely for monitoring web services. The users’ information gathering technology from browsers is the major technology for realizing the indicators which needs to link to the monitoring center server and to analyze the page access information with the embedded the tags in the pages by programs according to the models. The outcomes can overcome the limitations of the present technologies in lacking of information about user’s browse behavior.

6. References

[12] IP network technical requirements - the network performance parameters and indicators” (YD / T 1171-2001)

The monitored objects

<table>
<thead>
<tr>
<th>The monitored objects</th>
<th>Windows Net monitoring</th>
<th>HP OVO</th>
<th>IBM ITCAM</th>
<th>SNMP</th>
<th>MRTG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internet Browser</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XML files</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Base</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JDBC/ODBC</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TELNET, HTTP, FTP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TCP/UDP object and events</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>IP object and events</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Link layer protocol object and events</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Server: operation system</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Server: CPU</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Server: disk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Server: memory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: The functions of mainstream technologies in monitoring web services

<table>
<thead>
<tr>
<th>No.</th>
<th>fields</th>
<th>type</th>
<th>index</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>web_id</td>
<td>integer</td>
<td>primary key</td>
<td>Website ID embedded the codes</td>
</tr>
<tr>
<td>2</td>
<td>date</td>
<td>integer</td>
<td></td>
<td>Date (year/month/date)</td>
</tr>
<tr>
<td>3</td>
<td>uernum</td>
<td>integer</td>
<td></td>
<td>Number of users per day</td>
</tr>
<tr>
<td>4</td>
<td>pvnum</td>
<td>integer</td>
<td></td>
<td>Number of visit per day</td>
</tr>
<tr>
<td>5</td>
<td>houru0</td>
<td>integer</td>
<td></td>
<td>Users during 00:00—00:59</td>
</tr>
<tr>
<td>6</td>
<td>hourp0</td>
<td>integer</td>
<td></td>
<td>Visit during 00:00—00:59</td>
</tr>
<tr>
<td>52</td>
<td>houru23</td>
<td>integer</td>
<td></td>
<td>Users during 23:00—23:59</td>
</tr>
<tr>
<td>53</td>
<td>hourp23</td>
<td>integer</td>
<td></td>
<td>Visit during 23:00—23:59</td>
</tr>
</tbody>
</table>

Table 2: The table structure for getting the indicator II-1 (The Number of Accessing a Website)

Fig. 1 The outcomes of indicator II-3: The Number of Accessing a Website