

View Service Quality Testing according to INSPIRE Implementing Rules

J. Ardielli, J. Horak, J. Ruzicka

Institute of Geoinformatics, Faculty of Mining and Geology, VSB – Technical University of Ostrava, 17. listopadu 15, 70830, Ostrava-Poruba, Czech Republic, phone: +420596993236, e-mail: jiri.ardielli@vsb.cz

crossref <http://dx.doi.org/10.5755/j01.eee.119.3.1367>

Introduction

A spatial data infrastructure is based on network services. A quantitative evaluation of the service quality should contain both server-side testing and client-side testing. Directive 2007/2/EC of the Council and the European Parliament establishes the legal framework for setting up and operating an Infrastructure for Spatial Information in Europe (INSPIRE) based on infrastructures for spatial information established and operated by the member states. Quality of services (QoS) according to the INSPIRE directive follows the three aspects - performance, capacity and availability of services [1].

For the purpose to meet QoS demands, even in the case of accesses as many users as possible, respectively sending the maximum number of requests, it is necessary to use monitoring tools to observe network status for the subsequent evaluation, performance analysis of network behavior at certain time and based on it to provide feedback to a management system [2].

Besides evaluation of the content and function capabilities it is necessary to measure availability of the service and its performance. The availability is given by quality of data packet transmission, which is primarily a function of data packet loss [3]. Reliable data transmissions are of vital importance of today's systems. The main cause of data loss in systems is a network failure [4].

Performance means the minimal level by which an objective is considered to be attained representing the fact how fast a request can be completed within an INSPIRE Network Service. This includes on one side the ability of the server to quickly response to request, on other side the quality of data channel to transfer data with minimal latency. The quality of data channel depends on type of used technology of transfer, hence on the type of terminal equipment and his network characteristics. In wireless technologies the theoretical maximum connection speed is only achievable on laptops where high-performance

components are used (in comparison to mobile devices). Other mobile devices like family PDAs or smart phones have low-performance components with a limited connection speed [5]. Personal computers usually have a high-speed wire connection to the network and the speed of data transmission is usually determined by the performance of elements in the network.

When presenting the technology to end-users the most commonly used network characteristic is the maximum bandwidth. The bandwidth is an average rate of successful message delivery through a communication channel. Any decreasing of the given bandwidth means a substantial degradation of QoS [6]. Bandwidth is usually measured in bits per second (b/s). The maximum theoretical bandwidth is closely related to the channel capacity [7]. The overall performance (pages/s) according to INSPIRE covers delays at the server and the bandwidth of the network.

The first testing of QoS for web map services (WMS) according to the INSPIRE directive was performed in the Czech Republic in 2008 and 2009 [8].

In this paper we propose an innovative way how to test and measure the availability of network view services of WMS for end users. The improved methodology includes testing on the bases of layer definition, scale definition, network status verification by parallel querying services of control servers and a complex evaluation on the client and server side.

New tests of WMS services have been prepared in collaboration with COSMC (Czech Office for Surveying, Mapping and Cadastre, one of main provider of Czech spatial network services). These tests were focused on two main aspects:

a) Monitoring the availability of WMS related with records of errors, and continuous monitoring of response quickness and other performance parameters;

b) Stress tests, which were carried out during weekends. These tests measure the period of the reaction of the web service on the demand of the users depending

on the number of the simultaneous accesses [9]. It simulates the high server load and monitors its behavior under such conditions. Tests can help determine physical limits of the system performance.

The testing helps to understand how the system handles the load caused by many concurrent users. Any provider is able to test a latency of application, but for end user it is important to measure the overall latency (the response time) and also other characteristics like error occurrence, availability and performance.

The results allow to assess the situation of users in real network traffic. Results of the client-side testing are evaluated together with analysis of server's logs and compared with the parameters required by the implementing rules of the INSPIRE directive.

Measurements of QoS

Analysis of web systems can be focused on server-side or client-side evaluations.

Server-side analyses usually explore web server log files, including i.e. click stream analysis [10]. One of the main analytical objectives is to explore a dependency between a content of rendered image and time for its rendering. Results of the server-side tests can be used for optimization of the service based on several techniques (i.e. geoweb caching or load balancing).

Hicks et al. [11] specifies following 5 types of client-side tests: Test precondition data, Functional testing, Stress test functionality, Capacity testing and Performance testing.

Performance testing is the most well known form of testing. Tests are based on software emulating common users' behavior or uses some random pattern to access the server.

Usually applied metrics follow:

- Response time;
- Error frequency;
- Availability.

The response time represents a difference between the time when a user sends a request to a server and the time when the user receives a complete response from the server. The number of concurrent users of the server has to be taken into account for any evaluation of this metric.

The error frequency is usually expressed as a percentage of errors occurred during the test. Alternatively an error occurrence evaluates an average of time between failures.

The availability usually refers to a percentage of time when a user can access the service. The interpretation depends on the application type (according to requirements). INSPIRE based QoS does not distinguish if returned results are correct or not. Even an error message returned by the monitored service is considered as an evidence for the availability of the service.

Testing methodology and the test environment

The aim of the study is to evaluate quality parameters of main Czech network services for end users and check the implementation of QoS requirements of INSPIRE. For

measurements of service conditions for end users we selected the client-side black-box testing. Even though various drawbacks of such testing exist [8], it still represents the main way how to verify real service conditions for end users. The independency of such testing from the service provider is welcome.

The reliable evaluation of service conditions is based on a long-term testing. It is necessary to set up appropriate number of virtual users, establish an appropriate temporal schema (time schedule) for sending (generating) of their requests and define a content of requests. Any request consists of pre-prepared sequences of GetMap requests. Basic map layers were selected according to Annex I and II of INSPIRE specification. All parameters and time schedule should reflect behavior of existing users. Analysis of server logs provides important information about usage patterns (average time delays, scales, zooming, panning etc.). A final evaluation of testing results is focused on three main metrics: performance, error occurrence and response time. These aspects are usually highlighted for utilization and acceptance of network services by end users [12].

Settings for performance testing (long-term period):

- Version: WMS 1.3.0;
- Layers: 11 (in each request just one layer);
- Format: image/png;
- Width: 800 px;
- Height: 600 px;
- Bbox: coordinates defined according to selected scales (1:500, 1:750, 1:1000, 1:2000, 1:50000, 1:350000, 1:750000) and positions;
- Crs: coordinate reference system EPSG102067;
- Cache-control: no-cache;
- Duration: 2 months (four weeks for performance tests with 20 virtual users and four weekend stress tests up to 500 virtual users);

For testing of web systems it is possible to develop new, fully customized software or to select an appropriate system from the large offer of existing tools. One of specific tool designated especially for purpose of INSPIRE compliant remote web service testing is described in [13]. A wide range of different tools for testing of web systems are documented at websites <http://www.softwareqatest.com>. We decided to use a WAPT software v.3.0 (SoftLogica LLC). The WAPT software is focused on testing common web systems. It can be configured to provide adequate tests of services based on WMS specification. It allows setting of general parameters for testing such as number of users, number of iterations, interval between runs and delay between users. Sets of requests may create scenarios. The results of tests are available in reports and graphs. The reports and graphs are based on logs that can be used for further processing. Applied metrics include:

- Average request processing time;
- Maximum request processing time;
- Minimum request processing time;
- Overall performance;
- Percentage of errors;
- Average bandwidth.

Outputs of the client-side testing are influenced by the current status of network (i.e. high traffic during peak hours), hardware and software conditions on the client side. For unbiased evaluation of QoS it is needed to eliminate measurements recorded during unusual conditions when some problems occur (no matter of the reason and location of problems). A detection of such period is based on evaluation of outputs from continual monitoring of WMS services of selected reliable servers. Six servers with high availability were continuously tested with low loading. Requests were generated for one user with delay of one second. The purpose of parallel tests was to monitor connectivity in the network and detect any issues with outages or significant delays in transmission. The time slots when more than half of controlled services returns bad results (typically the service is unavailable or it provides significant delays) are recorded. Afterwards these periods are excluded from processing of WMS service testing results.

Tests were carried out on an office PC (connection using LAN 100 Mbps), which corresponds to today's standards. The server solution integrates Microsoft IIS 7.5, Oracle DBMS and Geomedia Web Map.

To evaluate the activity on the server side COSMC provides two types of logs describing the situation at the time of the tests carried out by the client.

Laboratory testbed

The test period was 2 months. For each layer we specified a scale and a location in the map and other parameters defined in requirements generated by the client. The table 1 provides a list of all requested layers.

Table 1. List of explored layers and testing scales

Name	No. of scenes	Testing scales
rst_kn	10	1:2000, 1:1000
rst_kmd	10	1:2000, 1:1000
next_p_maps	10	1:2000, 1:1000
border_parcel	10	1:2000, 1:1000
parcel_numbers	10	1:750, 1:500
def_parcel	10	1:2000, 1:1000
def_buildings	10	1:2000, 1:1000
rst_pk	10	1:2000, 1:1000
overview_region_lines	5	1:750000
overview_cadaster_centers	5	1:350000
overview_cadaster_territory	5	1:50000

For each scale five scenes were defined and repeatedly accessed. The evaluation includes especially average, maximum and minimum response time to execute requirements, the overall performance of the tested application, the error rate of approaches and the average bandwidth in the access.

- Performance testing.

The challenge of test was to monitor long-term behavior of the server with minimum load of 20 users. Tests were performed in working days during four weeks.

- Load and stress testing.

The stress testing shows the system behavioral in case the data is accessed by more users simultaneously [14]. It was a short-term testing, which was designed to determine the border of the number of simultaneously user accession, where the server is still able to handle the requirements. Simultaneously it was necessary to monitor server performance and afterwards to evaluate errors in the extremely high load. Tests were performed during holidays when minimal utilization of WMS is expected. The limit was set to the 500 users.

Experiments and results

The INSPIRE Directive for assessing the quality of network service sets the minimum number of requirements simultaneously viewing service to 20 requests per second. Tests were conducted in full operation and therefore the viewing services were accessed also by standard users. In this sense, the tests were more stringent than required by the Directive.

First, the records of parallel testing (six control servers) are processed and evaluated, especially error occurrence and response time.

Further the evaluation of the time when control servers (or their services) are unavailable is provided. When more than three problems in access to these servers (either errors or response time over 20 seconds) occur, we conclude there is an unspecified problem on the client side or in the network. Testing of COSMC server has no meaning at this time. The respective test results at that time are marked as invalid and excluded from processing.

During *performance tests* it was generated over 49 millions of requests.

The average response time for all tested layers is depicted in fig. 1. It ranges between 200 ms and 1 s. The lowest time is recorded for a simple vector layer with labels (parcel_numbers). Commonly used vector cadastral maps (like def_parcel, borders of parcels) provide about 400 ms. Raster layers (with rst prefix) show significantly slow response time about 700 ms. Surprisingly administrative layers with low scales show very large response time.

The significant increase of response time between scales 1:1000 and 1:2000 was recognized. This increase ranges between 18 % and 64 % and highly depends on the layer type (Fig. 2).

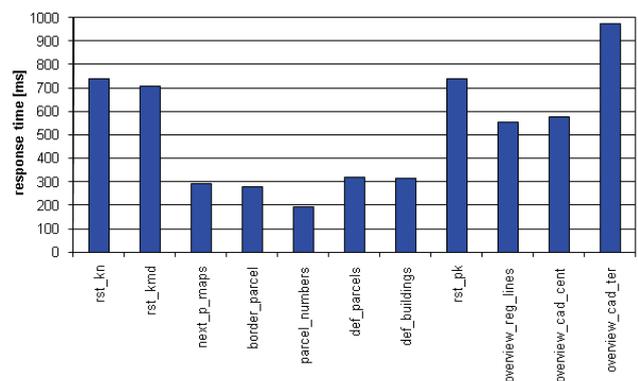


Fig. 1. The response time for each layer

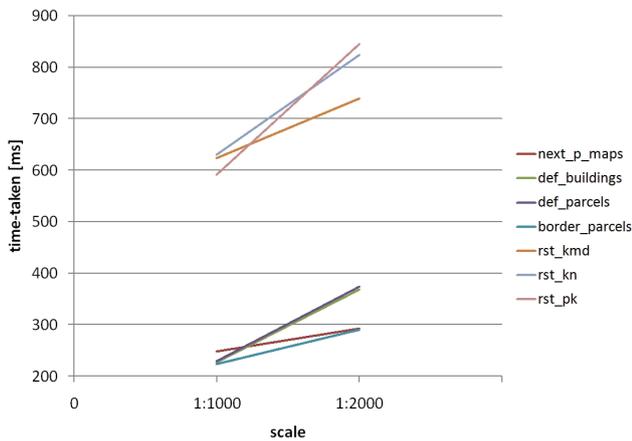


Fig. 2. Time increase due to the change of scale

In the graph of linear regression relationship (Fig. 3) two raster layers show the worst results (extremely high response time): `overview_cadaster_territory` and `cadastral maps` at scale 1:2000. On the other hand the best ratio is declared for the vector layer of lots (parcels) at scale 1:2000.

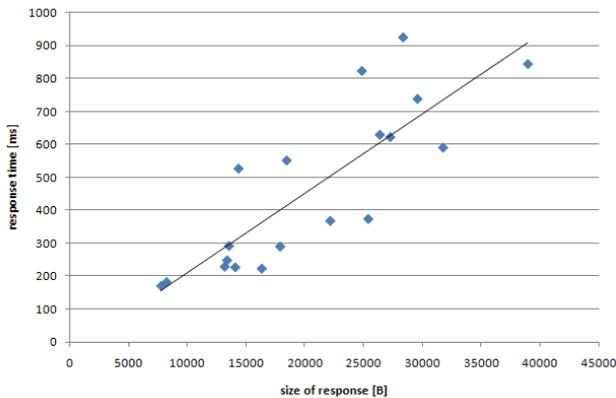


Fig. 3. The relationship between the average size of the response and the response time

Table 2 describes the distribution of response times. The number of requests processed over 5 seconds is only 0.28% (with capacity requirements fulfillment), which fully meets the reference criteria.

Table 2. Summary values for all tests

Time-taken [s]	Number of requests	Number of requests [%]	Average response time [ms]	Average size of the response [B]
< 1	46003408	94,06	386,57	19770
1 -5	2768316	5,66	1450,33	35038
>= 5	134561	0,28	9373,26	27056

The response time highly depends on time during a day (Fig. 4). In the night the average response time is close to 420 ms, after 6 AM the recorded time continuously grows up to almost 600 ms at 11 AM. The peak hours are recognized between 11 AM and 3 PM. Later the average response time goes down, but the quiet night status is reached only after 10 PM.

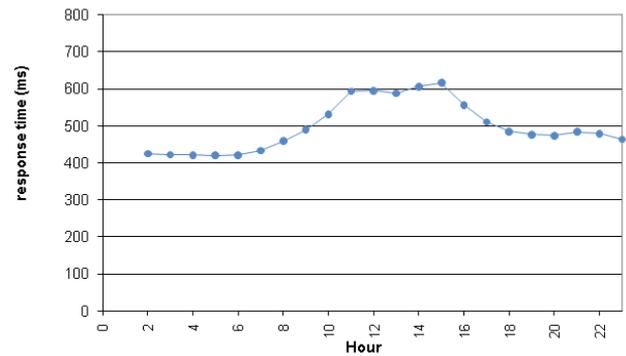


Fig. 4. Development of the average response time during a day

Over 17 millions of requests were generated during **stress tests**. By increase of the maximum number of virtual users from 150 to 250 the average value of time-taken (response time) is about half bigger (the value of 1033 ms raises up to 1548 ms). The number of requests completed over 1 second and over 5 seconds is also increased. In both cases, the number of requests completed over 5 seconds do not exceeded the 10% of all requests, thus the results are in compliance with the evaluation performance criteria of the INSPIRE directive.

In the case of limit values currently set for accession user 500 is the average value of the attribute Time-Taken in 2679 ms and the number of requests handled in a time of 5 seconds is close to 16% (Fig. 5).

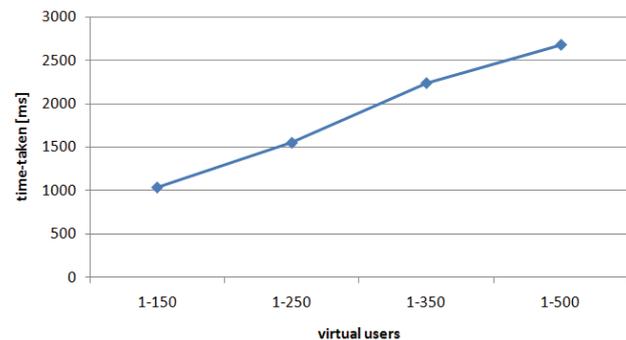


Fig. 5. Average time-taken with various numbers of users

Fig. 6 represents an overall performance of WMS server with growing number of virtual users. As we can see the limit of server is 90 pages/second. At this level the system is stable and has constant behavior except of the range for 240-260 users, when network errors occur.

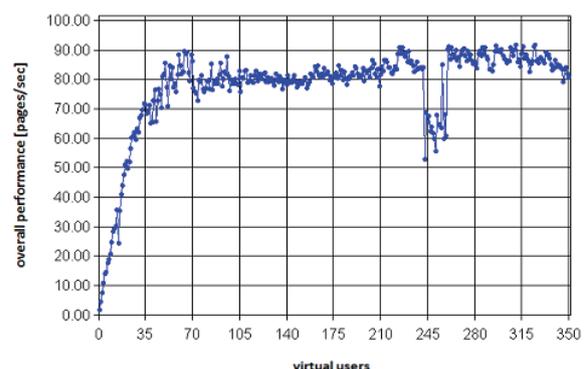


Fig. 6. Overall performance of WMS server

A final part of evaluation explores internal server errors frequency. HTTP status codes delivered from the server were analyzed and the status code 500 (internal server error) was selected for further processing. The number of internal errors occurred in 1 minute according time during a day is depicted in Fig. 7. The purpose was to discover if there is some dependency.

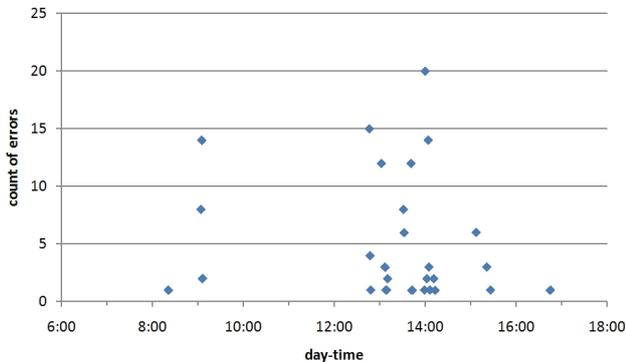


Fig. 7. Count of internal server errors in one minute against time during a day

The distribution of errors is not even but it is surprisingly concentrated to several time slots during a day. The higher probability of internal server errors refers to intervals between 8 and 9 a.m. and between 13 and 14 p.m. These periods do not fully correspond to peak hours indicated in fig. 4.

Conclusions

Results of performance and stress tests confirm that management capacity and performance criteria of "Regulation for Network Services" INSPIRE directive (INSPIRE 2009) are fulfilled with no critical comments.

Response times vary as expected for vector and raster - bitmaps provide a longer response, depending highly on the scale and complexity of layers. On the server side the range is between 180 and 850 ms; on the client side we recognize an increase of about 2-13% (192-975 ms). Number of requests handled in a time over 5 seconds on the client side is only 0.34%.

The capacity criterion was also fulfilled. The stress test shows the ability to handle up to 90 requests per second which exceeds the requirements of the implementation of INSPIRE.

In terms of operational reliability there was observed errors. Internal server error is at a very low level, client-side error rates are continually recorded from 3 to 5%. This aspect requires further investigation.

References

- INSPIRE (2009) Commission Regulation (EC) No 976/2009 of 19 October 2009 implementing Directive 2007/2/EC of the

- European Parliament and of the Council as regards the Network Services. – 10 p.
- Jankūnienė R., Gvergždys J., Budnikas A. Monitoring the Quality of Heterogeneous MAN Performance // *Electronics and Electrical Engineering*. – Kaunas: Technologija, 2007. – No. 2(74). – P. 69–74.
- Rindzevičius R., Augustaitis M., Tervydis P., Pilkauskas V. Performance Measures of Priority Queuing Data Network Node with Unreliable Transmission Channel // *Electronics and Electrical Engineering*. – Kaunas: Technologija, 2008. – No. 3(83). – P. 37–42.
- Stankus M., Penhaker M., Srovnal V., Cerny M., Kasik V. Security and Reliability of Data Transmissions in Biotelemetric System // *IFBME Proceedings 29/1*. – Berlin: Springer, 2010. – No. 29(2). – P. 216–219.
- Krejcar O. Problem Solving of Low Data Throughput on Mobile Devices by Artefacts Prebuffering // *EURASIP Journal on Wireless Communications and Networking*. – Hindawi Publishing Corporation, 2009. – P. 1–8.
- Bridova I., Vaculik M., Brida P. Impact of Background Traffic on VoIP QoS Parameters in GPON Upstream Link // *Electronics and Electrical Engineering*. – Kaunas: Technologija, 2011. – No. 1(107). – P. 111–114.
- Batkauskas V., Kajackas A. Quality of Heterogeneous Mobile Data Services: Capabilities and End-user Achievements // *Electronics and Electrical Engineering*. – Kaunas: Technologija, 2010. – No. 5(101). – P. 43–46.
- Horak J., Ardielli J., Horakova B. Testing of Web Map Services // *International Journal of Spatial Data Infrastructures Research*. – JRC, 2009. – P. 1–19.
- Krejcar O., Janckulik D., Motalova L. Architecture, Development and Testing of Home Care Biomedical System // *In Proceedings of Vehicular Technology Conference Fall (VTC 2009–Fall)*. – IEEE Conference Publishing Services, 2009. – P. 2108–2113.
- Sun J., Xie H. Mining Sequential Patterns from Web Click-Streams Based on Position Linked List // *In proceedings of Asian-Pacific Youth Conference On Communication Technology*. – Scientific Research Publishing, 2010. – P. 466–470.
- Hicks G., South J., Oshisanwo A. O. Automated testing as an aid to systems integration // *BT Technology Journal*. – Kluwer Academic Publishers Hingham, 1997. – No. 15. – P. 26–36.
- Amberg M., Fischer S., Schörder M. Evaluation of User Acceptance for Web-based Aptitude Tests // *Communications of the IIMA*. – The International Information Management Association, 2006. – Vol. 6. – Iss. 3. – P. 121–128.
- Kliment T., Cibulka D. Testovanie vyhladavacich a zobrazovacich sluzieb podla INSPIRE poziadavek // *Proceedings of GIS Ostrava'2011*. – VSB-TU Ostrava, 2011. – P. 1–9.
- Motalova L., Krejcar O. Stress Testing Data Access via a Web Service for Determination of Adequate Server Hardware for Developed Software Solution // *In Proceedings of Second International Conference on Computer Engineering and Applications ICCEA'2010*. – Vol. 01. – IEEE Conference Publishing Services, 2010. – P. 329–333.

Received 2011 04 01
Accepted after revision 2011 09 09

J. Ardielli, J. Horak, J. Ruzicka. View Service Quality Testing according to INSPIRE Implementing Rules // Electronics and Electrical Engineering. – Kaunas: Technologija, 2012. – No. 3(119). – P. 69–74.

In collaboration with COSMC it has been prepared new tests of WMS services which focus on two main aspects: a) Monitoring the availability of WMS related with records of errors, and continuous monitoring of response quickness and other performance parameters. Test period was two months. For each layer are specified scale and location in the map and other parameters in requirements generated by the client. The evaluation utilizes average, maximum and minimum execution requirement times, the overall performance of the tested application, the error rate of approaches and the average bandwidth in the access. b) Stress tests, which were carried out during the holidays to simulate the high server performance and its behavior under such conditions. Tests help determine the physical limits of system performance.

The aim of both versions of tests is to evaluate the availability, performance characteristics and stress behavior of this service from the client's perspective. Results of the client-side testing are evaluated together with analysis of server's logs. Results of performance and stress tests confirm that management capacity and performance criteria of the INSPIRE directive. Internal server error is at a very low level, client-side error rates are continually recorded from 3 to 5%. Ill. 7, bibl. 14, tabl. 2 (in English; abstracts in English and Lithuanian).

J. Ardielli, J. Horak, J. Ruzicka. Paslaugos kokybės testavimas pagal INSPIRE įgyvendinimo taisykles // Elektronika ir elektrotechnika. – Kaunas: Technologija, 2012. – Nr. 3(119). – P. 69–74.

Bendradarbiaujant su COSMC buvo parengti nauji WMS testai, kurie remiasi dviem pagrindiniais aspektais: a) WMS prieinamumo, susijusio su klaidų įrašais, monitoringu; nenutrūkstamu atsako greičio monitoringu bei kitų parametrų monitoringu; b) stresiniais testais, kurie buvo vykdomi poilsio dienomis imituojant didelį serverio našumą ir jo veikimą tokiomis sąlygomis. Testai leidžia nustatyti fizines sistemos veikimo ribas. Abiejų testų tikslas – įvertinti prieinamumą, veikimo charakteristikas ir elgseną stresinėmis sąlygomis žvelgiant iš kliento perspektyvos. Testų rezultatai parodė, kad vidinių serverio klaidų lygis yra labai žemas, kliento pusėje klaidų lygis svyruoja nuo 3 iki 5 %. Il. 7, bibl. 14, lent. 2 (anglų kalba; santraukos anglų ir lietuvių k.).