International Journal of Advances in Computer Science and its Applications – IJCSIA
Volume 4 : Issue 1 [ISSN 2250 – 3765]
Publication Date : 09 January 2014

Random or Pseudo-random Testing of Download Services for Geodata

[Jan Ruzicka, Jiri Ardielli]

***Abstract* — The paper discuss one of problems that appear when testing download services (web feature service, WFS) for geodata. The paper follows results of performance tests of download services of Czech Office of Surveying, Mapping and Cadaster (COSMC) according to INSPIRE requirements. Methodology of testing used for performance tests of download services of COSMC was based on expected user behavior, but generally the performance tests of geoweb services are based on randomly generated inputs. The paper compares these two approaches for download services.**

***Keywords*—performance testing, spatial data, download service, INSPIRE**

I. Introduction

The development of Spatial Data Infrastructure is based on integration of global, European, national and local spatial initiatives. 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). INSPIRE requires to establish and operate following types of network services for the spatial data sets and services: discovery services, view services, download services, transformation services and services allowing spatial data services to be invoked.

Download services are services that allow download geodata with geometry representation in a vector format. There are two types of services: first is based on static geodata usually stored in files that represents administrative or other spatial units, second is based on dynamically selected geodata according to spatial (or other) filter conditions, for example extent specified with two points in geographical space.

Basic indicators of quality of services (QoS) and limit values are given by implementing of regulations according to INSPIRE [1, 2]. Full satisfaction of end users usually require to implement higher standards (than those given by INSPIRE) and provide better performance [3]. Novel approaches emphasize the central role of users and the importance of elaborated testing of the final user satisfaction [4].

Jan Ruzicka
Institute of Geoinformatics / VSB Technical University of Ostrava
Czech Republic
jan.ruzicka@vsb.cz

Jiri Ardielli
TietoCzech, s. r. o,
Czech Republic
jiri.ardielli@tieto.com

The aim of our testing was to verify the fulfillment of obligatory parameters required by above mentioned regulations and to evaluate the capacity to meet also higher requirements.

A quantitative evaluation of the service quality should contain both server-side testing and client-side testing.

II. Methodology

The download services of COSMC were tested during two weeks in May and June 2012. The methodology used for Get Spatial Object operation was based on pseudo randomly generating algorithm. The requests were created randomly close to centroids of cadastral areas where is available digital cadastral map in a vector format. Area covered by individual request was either 1 square kilometre or 250 square meters.

According to recommendations for testing that were published after the testing in 2012 year the testing should be based on requests randomly distributed in the extent of whole tested service [5]. Technical guidelines [6, 7] does not specify any algorithm that should be used for testing.

The research presented in this paper was focused to compare a testing based on methodology used in 2012 for testing of COSMC download services and a testing based on recommendations for randomly generated requests.

The testing was not done directly on COSMC services due to administrative issues. The testing was done on two services published in local intranet.

A. *Cadastral Map Service*

Digital cadastral map service has been created upon the same data that are used by COSMC to build a service for serving information about plots in the Czech Republic. Whole available geodata were downloaded from COSMC to create local service.

Downloaded data were imported to PostgreSQL/PostGIS database.

As coordinate system there was used EPSG:4258.

B. *Buildings Service*

Vector data for basic topographical map of Czech Republic from Open Street Map project (OSM) [8] were used to build this service. For testing was used layer that consist of buildings. The area of Czech Republic is not fully covered with buildings so we expect that the random testing should give better results than testing based on selected areas.

Downloaded data were used in file format (ESRI Shapefile format) without any transformation.

Coordinate system was used EPSG:4326.

C. Tested operation

This testing was focused only on dynamically selected objects according to specified conditions (in this testing only predefined extent of area was used). According to INSPIRE terminology was tested Get Spatial Object operation of dynamic download service. According to OGC (Open Geospatial Consortium) was tested GetFeature operation of WFS.

D. Monitored parameters

There were monitored several parameters: date and time of response arrival, time spent till the arrival of a first byte of the response (TTFB), time spent till the arrival of a last byte of the response (TTLB), size of the response in bytes (SIZE), the HTTP response code of the server (to identify errors and their sources), group of tests (identification of one from ten used threads for testing, simulating individual users) and extent. One of the required parameter was also bandwidth (Bps), which could be obtained by usage of one of monitoring tool [9]. Because used testing tool collects parameters like response time and size of response, so the bandwidth parameter was calculated on the bases of it.

According to INSPIRE directive there were compared time spent till the arrival of a first byte of the response that must be less than 30s and download speed, that must be more than 0.5 MBps. Both criteria must be fulfilled for more than 90% of requests.

E. Pseudo randomly generated requests for cadastral map service

The requests were created randomly close to centroids of cadastral areas where is available digital cadastral map in a vector format. Area covered by individual request was either about 1 square kilometre or about 250 square meters. There were used about 6000 points that should represent centroids. The requests were generated in the same way as in a case of methodology used for testing of COSMC services in 2012 year. There was only difference between the requests: the coordinate system used to define bounding boxes.

F. Randomly generated requests for cadastral map service

There were generated about 250000 of requests randomly in a geographical space based on two limits. First limit was that extent of the request must be inside polygon of Czech Republic. Second limit was that request must not cover area bigger than 1 square kilometre and not smaller than 250 square meters.

G. Pseudo randomly generated requests for building service

There were randomly selected 50000 buildings from tested layer. Two requests were generated according to the location of building. First request covers area of 64 square kilometres and the second request covers area of 16 square kilometres. This should simulate similar conditions as in a case of methodology used for testing of COSMC services in 2012 year.

H. Randomly generated requests for building service

There were generated 100000 requests randomly in a geographical space based on two limits. First limit was that extent of the request must be inside polygon of Czech Republic. Second limit was that request must not cover area bigger than 64 square kilometres and not smaller than 16 square kilometres.

I. Tests for cadastral map service

There was run one test for each type of generated requests. Each test takes 12 hour that is enough for load testing. The test was done with 10 independent clients.

J. Tests for building service

There was run one test for each type of generated requests. Each test takes 60 minutes that is, in this case, enough for load testing, because of short response times. The test was done with 10 independent clients.

III. Used technology

For performance testing of web based applications there can be used wide area of testing software. One of easy and cost effective solution is usage of Apache Jmeter [10, 11]. Jmeter (apache-jmeter 2.6 + JMeter Plugins 0.5.2, GNU/Linux, Java OpenJDK 1.7) software was used for generating of requests and for logging service's responses.

GeoTools and XSLT scripts were used to generate extents for Jmeter requests.

GeoServer 2.2.4 was used to publish download services with Open Street Map and Cadastral Maps.

PostgreSQL 9.1 and PostGIS 2.0 were used for storing digital cadastral map in database.

IV. Results

Results show summary of parameters of responses. There are two columns. First column shows results for all responses except errors. Second column shows values only of results that contains at least one feature, so the empty collections of geodata are filtered out. In a case of buildings service were no responses longer than 30 s, so for purposes of comparison has been added a new limit to 50 ms.

A. Pseudo randomly generated requests for cadastral map service

TABLE I. PSEUDO RANDOMLY GENERATED REQUESTS FOR CADASTRAL MAP SERVICE

Parameters		
Parameter	*All responses*	*Responses without empty collection*
Average time to first byte (ms)	16666	16683
Maximum time to first byte (ms)	46942	46942
Percent of responses > 30 s	0.37	0.52
Minimum download speed (Bps)	826	50697
Percent of responses < 0.5 MBps	1.47	1.09

B. Randomly generated requests for cadastral map service

TABLE II. RANDOMLY GENERATED REQUESTS FOR CADASTRAL MAP SERVICE

Parameters		
Parameter	*All responses*	*Responses without empty collection*
Average time to first byte (ms)	17553	17560
Maximum time to first byte (ms)	47724	47724
Percent of responses > 30 s	0.22	0.24
Minimum download speed (Bps)	62923	62923
Percent of responses < 0.5 MBps	0.88	0.04

C. Pseudo randomly generated requests for building service

TABLE III. PSEUDO RANDOMLY GENERATED REQUESTS FOR BUILDING SERVICE

Parameters		
Parameter	*All responses*	*Responses without empty collection*
Average time to first byte (ms)	23	23
Maximum time to first byte (ms)	468	468
Percent of responses > 30 s	0	0
Percent of responses > 50 ms	3.29	3.29
Minimum download speed (Bps)	127818	127818
Percent of responses < 0.5 MBps	0.01	0.01

D. Randomly generated requests for building service

TABLE IV. PSEUDO RANDOMLY GENERATED REQUESTS FOR BUILDING SERVICE

Parameters		
Parameter	*All responses*	*Responses without empty collection*
Average time to first byte (ms)	18	19
Maximum time to first byte (ms)	1616	1616
Percent of responses > 30 s	0	0
Percent of responses > 50 ms	0.64	0.67
Minimum download speed (Bps)	11242	43968
Percent of responses < 0.5 MBps	2.79	0.34

Fig. 1 and 2 shows number of responses depending on TTFB, which is relative converted to 1 byte of response. Fig. 2 represents a more detailed interval between 0 and 1 ms from the first figure. It's seen, that generating speed of one byte is different for random and pseudo-random requests.

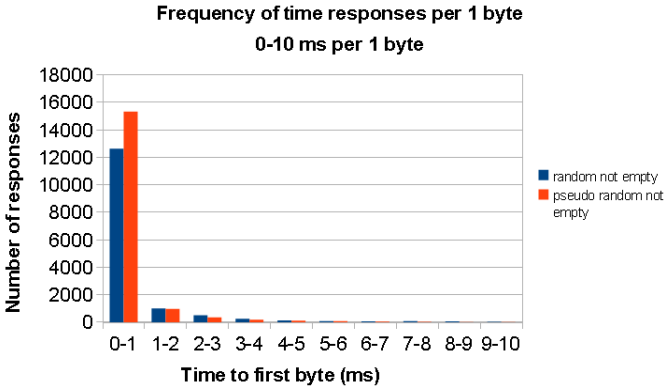

Figure 1. Nr. of responses depending on TTFB per 1 B in range 0-10 ms

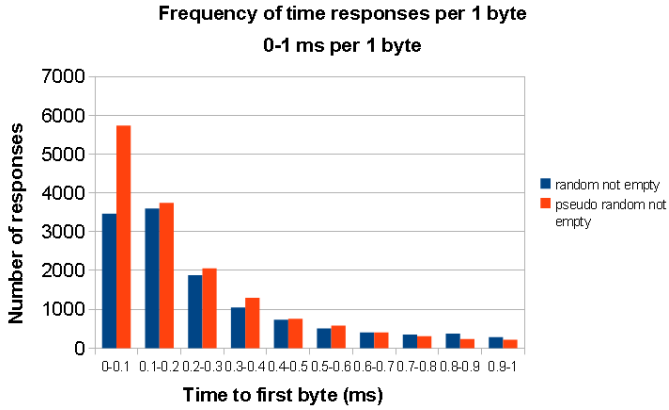

Figure 2. Nr. of responses depending on TTFB per 1 B in range 0-1 ms

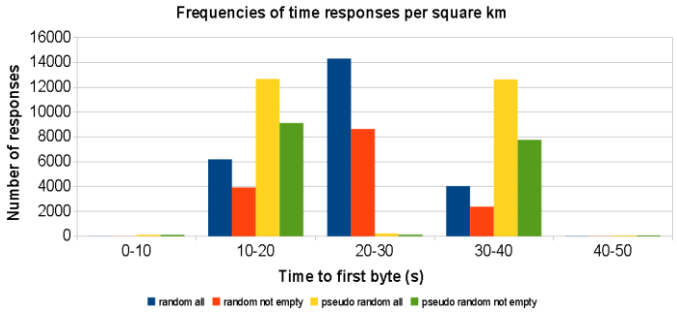

Figure 3. Nr. of responses depending on TTFB per 1 square km

On fig. 3 is shown number of responses of random and pseudo random requests depending on classes defined by TTFB. This numbers are relative converted to 1 square kilometer. There are big differences in results for random and pseudorandom requests. Randomly generated requests have normal distribution, while pseudo random have low values in interval 20-30 s. Randomly generated requests had area in interval 250 square meters to 1 square kilometer, while pseudo randomly generated requests had two areas, 250 square meters and 1 square kilometer.

v. Conclusion

We can conclude that there is significant difference when dealing with percent of responses higher than specified time limit in a case of random or pseudo random requests in both tested services.

Randomly generated requests produce smaller amount of responses over the time limit than pseudo randomly generated requests.

Monitored number of responses and measured TTFB unified per 1 byte also show differences between random and pseudo random requests.

Results also shows that it depends on how big area the extent covers and that usage of small amount of predefined areas gets other results than usage of values interval for areas generating of very different acreage.

When testing the quality of download services for geodata (WFS), the placement of the query in the space and size of extent greatly affects the results of these measurements.

Acknowledgment

We are appreciative employees of Czech Office for Surveying, Mapping and Cadaster for their cooperation and support of the analysis. We also would like to thank to all open source developers and open data collectors.

References

[1] INSPIRE, “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”, 2009, p.10.

[2] INSPIRE, “Commission Regulation (EU) No 1088/2010 of 23 November 2010 amending Regulation (EC) No 976/2009 as regards download services and transformation services”, L323/1, 2010, p.10.

[3] T. Mildorf, V. Cada, “Reference Data as a Basis for National Spatial Data Infrastructure”. Geoinformatics FCE CTU 9, 2012, pp. 51-61.

[4] P. Voldan, “Developing web map application based on user centered design”. Geoinformatics FCE CTU 7, 2011, pp.131-141.

[5] X. He, “Testing Download Service for INSPIRE Compliance”, Available at: http://sdi-testbed.eu/ [Accessed 1 August 2013].

[6] Technical Guidance for the implementation of INSPIRE View Services, version 3.1. 114 p. Available at: http://inspire.jrc.ec.europa.eu/ [Accessed 25 April 2012].

[7] Technical Guidance for the implementation of INSPIRE Download Services, version 3.0. 82 p. Available at: http://inspire.jrc.ec.europa.eu/ [Accessed 1 August 2013].

[8] Open Street Map Community, Wiki.[online] Available at: <http://wiki.openstreetmap.org/wiki/Main_Page> [Accessed 12 July 2012].

[9] A. Sinha, N. Kumar, “Bandwidth Monitoring Tools”, UACEE International Journal of Advances in Computer Networks and its Security, vol. 3, 2013, pp.116-120.

[10] K. Tirghoda, “Web Services Performance Testing Using Open Source Apache Jmeter”, International Journal of Scientific & Engineering Research, vol. 3, 2012, pp.318-320.

[11] G. Khare, D. Pathy, A. Rajan, A. Jain and A. Rawat, “Performance Analysis of Web based Applications on Single and Multi Core Servers”, UACEE International Journal of Computer Science and its Applications, vol. 2, 2012, pp. 34-38.

About Author (s):

[Jan Ruzicka has gained his Ph.D. degree at Institute of Geoinformatics at VSB Technical university of Ostrava, where he works now. He is mainly focused on GeoWeb services research, GeoWeb services orchestration and metadata for spatial data.]

[Jiri Ardielli formerly worked at Institute of Geoinformatics at VSB where cooperates now. His research is focused on GeoWeb services. Actually he works at Tieto Company on IT services managing according to ITIL.]