

TESTOVÁNÍ STAHOVACÍCH SLUŽEB ČÚZK TĚMATU BUDOVY

Jan RŮŽIČKA, Jiří HORÁK, Tomáš INSPEKTOR

Institut geoinformatiky, HGF, VŠB-TUO, 17. listopadu 15, 708 33, Ostrava-Poruba, Česká republika
jan.ruzicka@vsb.cz, jiri.horak@vsb.cz, tomas.inspektor@vsb.cz

Abstrakt

V závěru roku 2015 byly provedeny výkonové a zátěžové testy 4 služeb tématu Budovy ČÚZK pro ověření souladu s požadavky INSPIRE a zjištění charakteristik výkonnosti, kapacity a dostupnosti při běžném i zátěžovém (stress) režimu. Byly testovány operace Get Download Service Metadata, Get Spatial Object, Get Spatial Data, Describe Spatial Data set / object. Testování bylo prováděno v režimu v režimu klienta mimo intranet služby a ke generování požadavků a vytváření logů monitorování odpovědí serveru byl použit Jmeter. Distribuce náhodných požadavků byla nastavena v souladu s Technical Guidance for the implementation of INSPIRE Download Services. Výsledky ukazují na dobré výkonové parametry řešení ČÚZK.

Abstract

Testing of download services of COSMC for theme Buildings. The download services provided by COSMC (Czech Office for Surveying, Mapping and Cadastre) for theme Buildings were tested for compliance with INSPIRE requirements at the end of 2015. The load and stress tests were focused on following operations: Get Download Service Metadata, Get Spatial Object, Get Spatial Data, Describe Spatial Data set / object. Testing was done on the client side using Jmeter software. Random service requests were distributed according to the Technical Guidance for the implementation of INSPIRE Download Services. The results show high level of quality of the services.

Klíčová slova: INSPIRE;ČÚZK;budovy;testování;WFS

Keywords: INSPIRE;COSMC;buildings;testing;WFS

ÚVOD

Testování služeb dle pravidel INSPIRE je velmi důležitou součástí infrastruktury prostorových dat. Každá služba by měla projít úvodním testováním a být také testována průběžně v průběhu svého životního cyklu. Český úřad zeměměřický a katastrální (ČÚZK) požádal naše pracoviště o testování nově spuštěné stahovací služby, která zpřístupňuje INSPIRE téma Budovy (Buildings).

Testování může být realizováno mnoha způsoby. Jedním ze základních kritérií je rozlišení testování na straně serveru a na straně klienta. V případě tohoto testování bylo využito klientské testování, přestože jsme si vědomi, že příslušná nařízení vyžadují splnění kritérií na straně serveru a ideální by bylo komplexní hodnocení zahrnující testování na obou stranách (Horák et al., 2013). Testy byly prováděny výhradně s využitím open source nástrojů (zejména Jmeter, PostgreSQL/PostGIS, GRASS GIS, MySQL, xmllint). Byly provedeny testy výkonu a dostupnosti, zátěžové testy a testy datových struktur. Výsledky testů ukazují na splnění všech požadovaných kritérií i při hodnocení na straně klienta. Drobné nedostatky byly nalezeny v případě testování datových struktur.

TESTOVÁNÍ

Rozsah testování zahrnoval testování operací Get Download Service Metadata, Get Spatial Object, Get Spatial Data, Describe Spatial Data set / object a testování souladu datových struktur metadat a dat s deklarovanými schématy.

Rozsah testování podle požadavků INSPIRE zahrnoval tyto 4 služby, resp. adresy:

<http://services.cuzk.cz/gml/inspire/bu/epsg-5514/>,

<http://services.cuzk.cz/gml/inspire/bu/epsg-4258/>,

<http://services.cuzk.cz/wfs/inspire-bu-wfs.asp>,

<http://services.cuzk.cz/xsd/wfs/BasicFeature/>.

Testování bylo prováděno v režimu klienta mimo intranet služby. Testování zahrnovalo 2 týdenní testy pro stanovení výkonových parametrů (při standardní zátěži 20 virtuálních uživatelů), 2 víkendové testy pro určení zátěžových charakteristik a provedení validace metadat a získaných dat.

Testování bylo provedeno ve 2. polovině listopadu a v prosinci 2015, začalo 16. 11. a skončilo 20. 12. 2015.

Na základě analýzy nařízení a doporučení, vztahující se k těmto službám, byl připraven seznam požadavků na tyto operace. Ke generování požadavků a vytváření logů monitorování odpovědí serveru na tyto požadavky byl použit SW Jmeter (apache-jmeter 2.6 + JMeter Plugins 0.5.2, GNU/Linux, Java OpenJDK 1.8).

U každé operace byly vygenerovány náhodné požadavky pro sledované operace a sledované vrstvy a souřadnicové systémy dle zadání. V případě stahování prostorových dat byly generovány velikosti výřezů několika velikostí dle předpokládaných potřeb uživatele.

V přípravné fázi byly staženy všechny soubory ze zdroje <http://services.cuzk.cz/gml/inspire/bu/epsg-5514/> k určení referenčních bodů pro konstrukci dotazů. Důvodem je určení vhodných bodů pro vymezení obdélníku pro stahování dat, které nebudou směřovat do prázdné oblasti, ale do míst s lokalizovanými objekty. K určení referenčních bodů byla využita shluková analýza v prostředí nástroje GRASS GIS.

Testování proběhlo v souladu s Technical Guidance for the implementation of INSPIRE Download Services (INSPIRE, 2013).

Paralelně 20 klientů stahovalo data ze všech uvedených zdrojů. Požadavky byly mezi zdroje distribuovány náhodně, ve stejném poměru, což ve výsledku znamená následující poměry pro dané operace:

- cca 10 % Get Download Service Metadata
- cca 10 % Describe Spatial Data Set nebo Describe Spatial Object Type
- cca 80 % Get Spatial Data Set nebo Get Spatial Object

To odpovídá doporučením uvedeným v Technical Guidance (INSPIRE, 2013).

Pro operaci Get Spatial Object jsou dotazována místa kolem referenčních bodů získaných v přípravné fázi tak, aby byl zaručen dostatečný objem dat požadovaný specifikací (odpovědi o velikosti nejméně 1 MB). Přitom je použit parametr BBOX, jak je požadováno v Technical Guidance (INSPIRE, 2013).

Paralelně byly testovány přístupy a odezvy na 4 referenční servery z důvodu vyloučení problémů s počítačovou sítí.

Z výsledků provozního testování byly vyhodnoceny následující testy dle INSPIRE:

- výkonové testy (Performance test)
- testy kapacity (Capacity test)
- testy dostupnosti (Availability test)

K registrovaným parametrům patří:

- Datum a čas příchodu odpovědi (doručení poslední části odpovědi)
- Doba do příchodu 1 bytu (TTFB), resp. příchodu 1. paketu (dále doba odezvy, latency)
- Doba příchodu posledního bytu (TTLB) (měřeno od okamžiku odeslání požadavku z klienta)
- Velikost odpovědi v bytech (SIZE)
- Kód odpovědi serveru (identifikace chyb a jejich zdroje)

- Skupina testování (identifikace počítače a identifikace jednoho z 20 vláken pro testování, simulující jednotlivé paralelně pracující uživatele)
- Identifikace požadavku (jaká operace, jaký souřadnicový systém apod.)
- Název vrstvy
- Rozměry okna pro stažení dat

VÝSLEDKY TESTOVÁNÍ

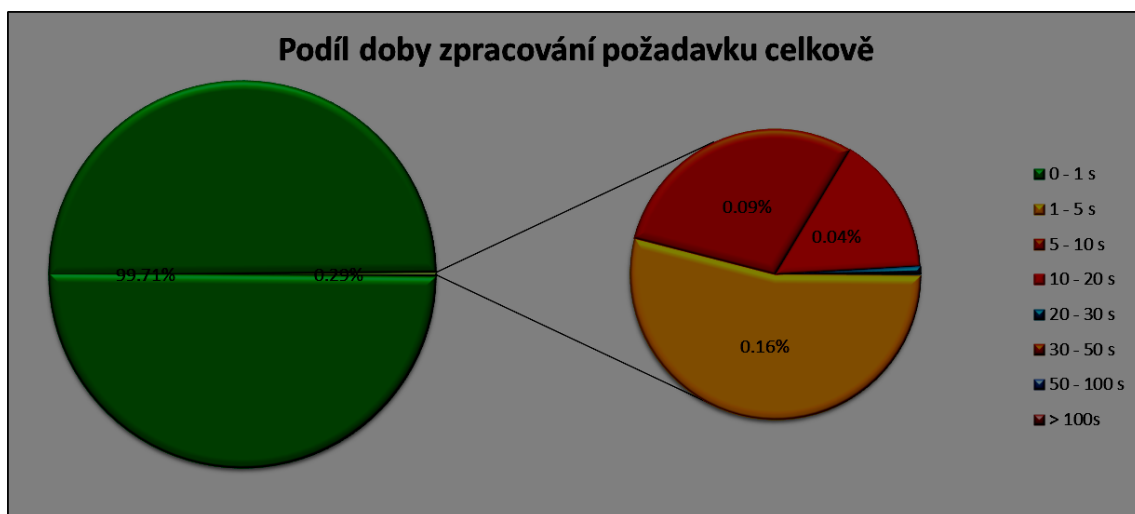
Výsledky testování jsou rozděleny do pěti skupin, dle testované operace.

Vyhodnocení operace „získat metadata služby stahování dat“ (Get Download Service Metadata) (GetCapabilities)

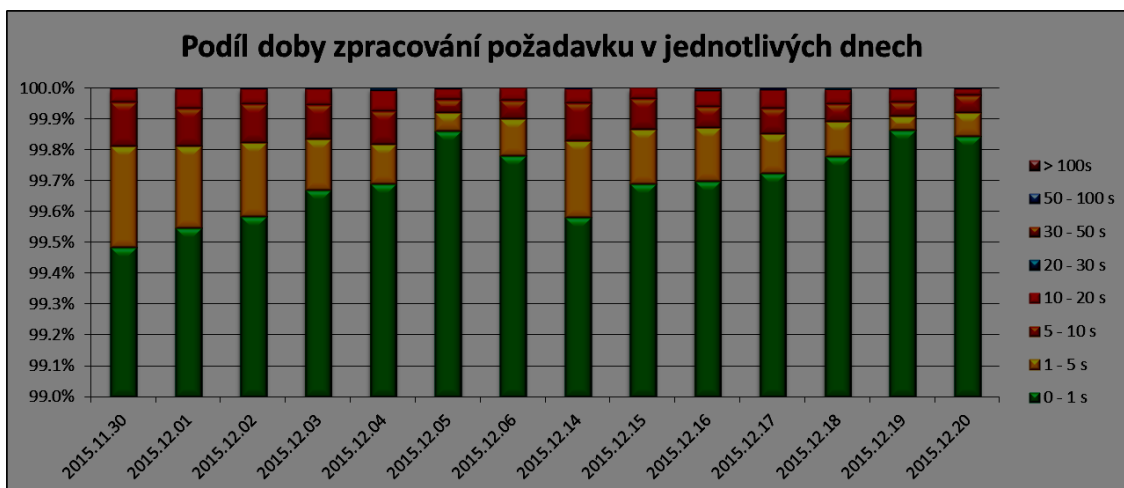
Při hodnocení výkonnosti se kontroluje, zda za normální situace je doba odezvy k odeslání prvotní odpovědi nejvýše 10 sekund. Normální situace má představovat nejméně 90% času, tj. podíl odpovědí s dobou odezvy horší než 10 sekund nesmí být vyšší než 10 %.

Výsledky jsou vynikající, podíl požadavků s dobou odezvy k odeslání prvotní odpovědi na GetCapabilities větších než 10 s je pouze 0.047%, což zdaleka nedosahuje hranice 10%. V žádném ze zkoumaných dnů se nedosahuje hodnot, které by ukazovaly na možné ohrožení splnění této podmínky.

Pokud srovnáme výsledky s hodnocením testování stahovacích služeb pro témata Parcel, Boundary, Zoning ČUZK v roce 2012, byly výsledky o řád lepší (0,004%) (Horák et al., 2013).



Obr. 1. Podíl doby zpracování požadavku celkově



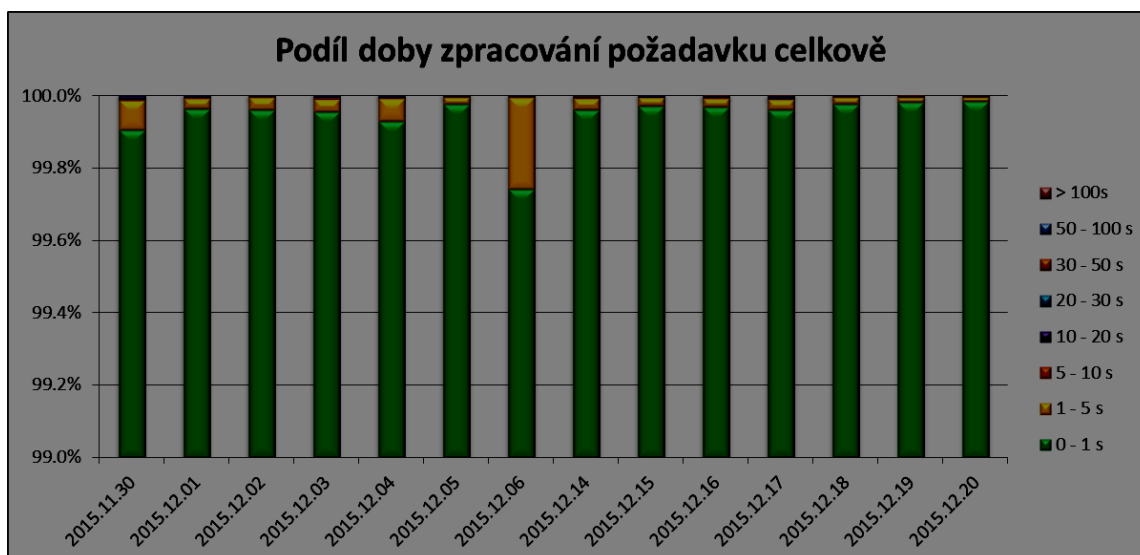
Obr. 2. Podíl doby zpracování požadavku v jednotlivých dnech

Vyhodnocení operace „získat sadu prostorových dat“ (Get Spatial Data Set)

Požadavky „získat sadu prostorových dat“ jsou realizovány jako požadavky na stažení příslušných komprimovaných souborů, které příslušnou sadu dat obsahují (např. <http://services.cuzk.cz/gml/inspire/bu/epsg-5514/id.zip>). Rozlišoval se souřadnicový systém u sad, tedy S-JTSK (EPSG:5514) a ETRS89 (EPSG:4258).

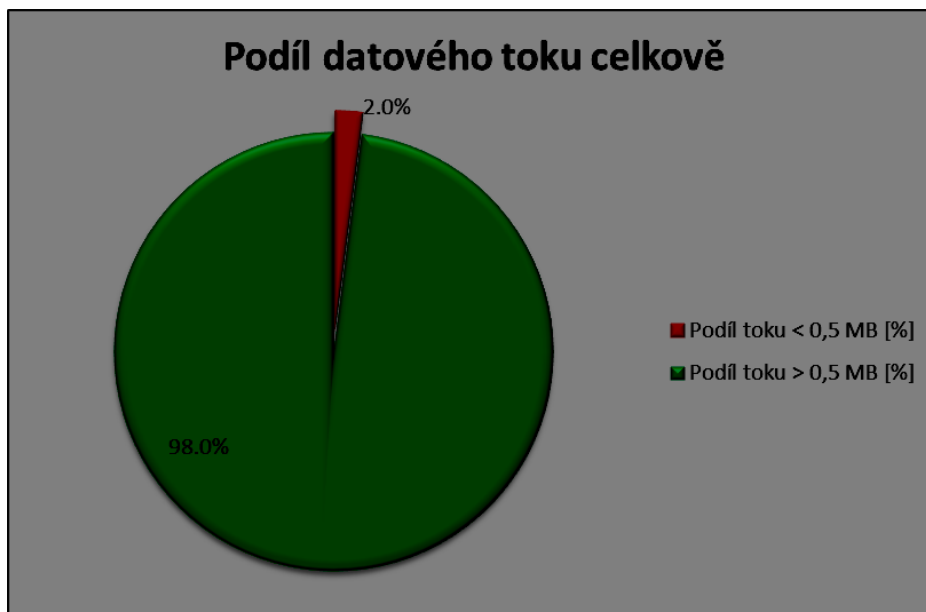
Kontroluje se, zda za normální situace je doba odezvy k odeslání prvotní odpovědi nejvýše 30 sekund. Dále se kontroluje, zda služba stahování dat udržuje za normální situace odezvu vyšší než 0,5 megabytu za sekundu nebo vyšší než 500 prostorových objektů za sekundu.

Výsledky jsou vynikající, z celkového počtu více než 5.6 mil. požadavků na Get Spatial Data Set pouze 20 bylo delších než 30 s. Proto je podíl požadavků s dobou odezvy k odeslání prvotní odpovědi větších než 30 s pouze 0.000%. V žádném ze zkoumaných dnů se nedosahuje hodnot, které by ukazovaly na možné ohrožení splnění podmínky 10%.



Obr. 3. Podíl doby zpracování požadavku v jednotlivých dnech

Sledování datového toku ukazuje splnění požadavků INSPIRE na udržení odezvy nad 0.5 MB/s. Celkově je požadavků s vyšší rychlostí datového toku než 0.5 MB/s celých 98%. Pro srovnání v roce 2012 byl podíl nadlimitních rychlostí stahování ještě o něco vyšší (99.7%).



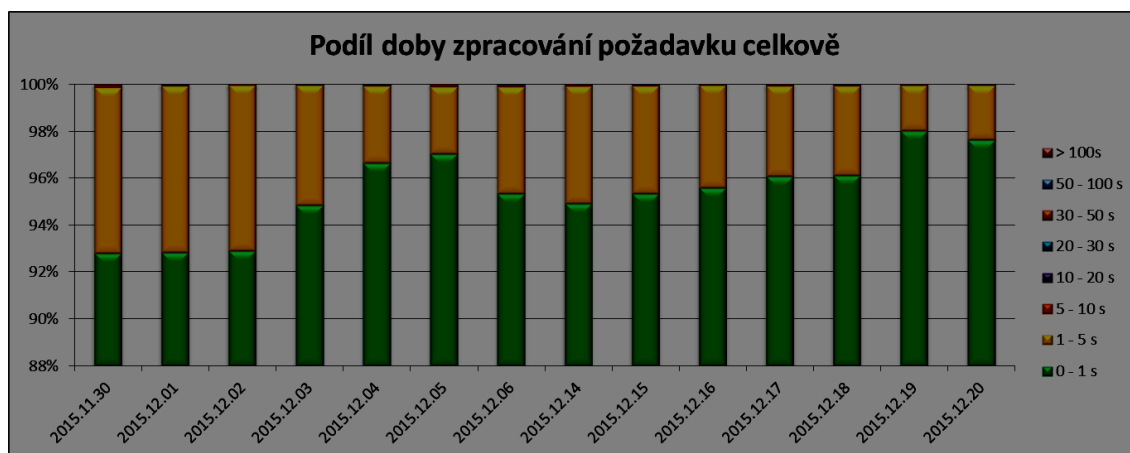
Obr. 4. Podíl datového toku celkově

Vyhodnocení operace „získat prostorový objekt“ (Get Spatial Object) (GetFeature)

Kontroluje se, zda za normální situace je doba odezvy k odeslání prvotní odpovědi nejvýše 30 sekund.

Dále se kontroluje, zda služba stahování dat udržuje za normální situace odezvu vyšší než 0,5 megabytu za sekundu nebo vyšší než 500 prostorových objektů za sekundu.

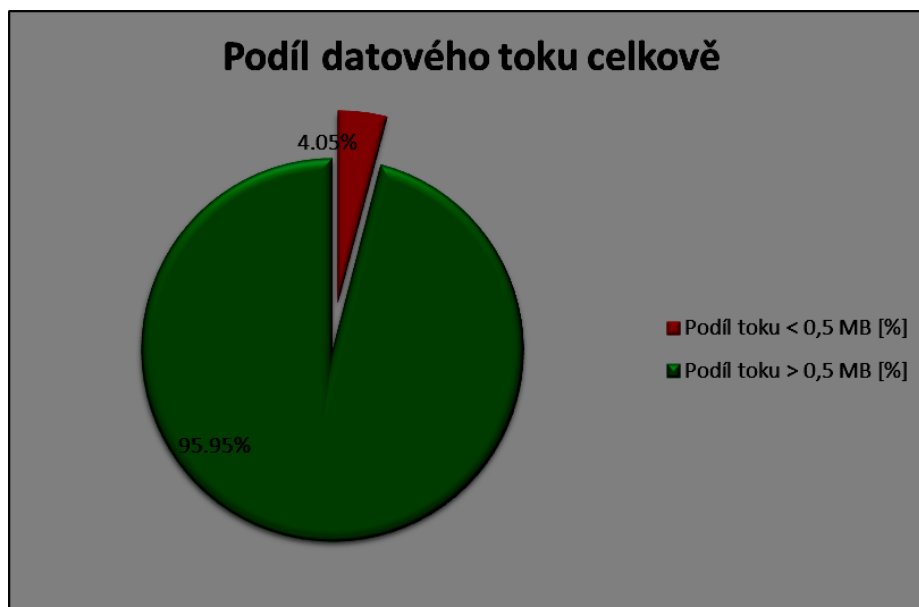
Výsledky jsou vynikající, podíl požadavků s dobou odezvy k odeslání prvotní odpovědi na Get Spatial Object větších než 30 s je 0.000%. V žádném ze zkoumaných dnů se nedosahuje hodnot, které by ukazovaly na možné ohrožení splnění této podmínky. Za celou dobu testování bylo zjištěno pouze 25 požadavků překračujících sledovaný čas. Nejvyšší podíl horších reakčních časů byl registrován 14. 12. 2015, kdy bylo zaregistrováno 7 požadavků (pouhých 7 z 376 tisíc) s časy dokonce vyššími než 100 s a 2 požadavky nad 50 s.



Obr. 5. Podíl doby zpracování požadavku v jednotlivých dnech

Z hlediska požadované rychlosti stahování je průměrný výsledek vyhovující (95.9%), rovněž většina dnů prokázala splnění podmínky s dostatečnou rezervou, avšak 30. 11. a 4. 12. 2015 nebyla dosažen požadovaný podíl menších toků do 10%. 30. 11. bylo zjištěno jen 86.3% požadavků s rychlostí stahování nad 0.5 MB/s a 4. 12. 2015 jen 88.1%.

Ze srovnání s výsledky z roku 2012 (Horák et al., 2013) pro jiné služby vyplývá, že rychlosti doby zpracování odpovědí jsou prakticky velmi podobné a podíl nadlimitní rychlosti stahování je o něco menší (byl 99.7%), což však může být ovlivněno novým požadavkem na minimální velikost stahované odpovědi 1 MB.



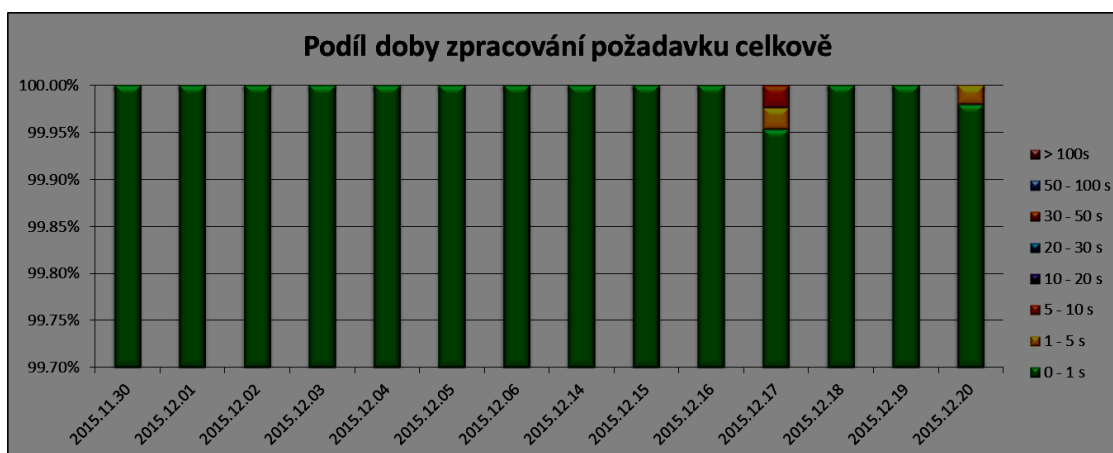
Obr. 6. Podíl datového toku celkově

Vyhodnocení operace „popsat sadu prostorových dat“ (Describe Spatial Data Set)

Kontroluje se, zda za normální situace je doba odezvy k odeslání prvotní odpovědi nejvýše 10 sekund.

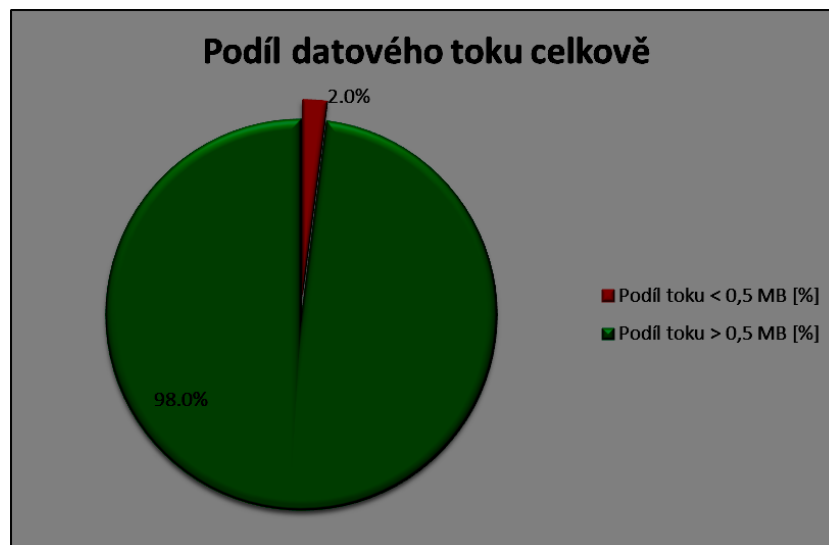
Dále se kontroluje, zda služba stahování dat udržuje za normální situace odezvu vyšší než 0,5 megabytu za sekundu.

Výsledky jsou vynikající, podíl požadavků s dobou odezvy k odeslání prvotní odpovědi na Describe Spatial Data Set větších než 10s je 0, protože za celou dobu testování nebylo zjištěno žádné překročení sledovaného času. Nejvyšší podíl horších reakčních časů byl registrován 17.12.2015, kdy bylo zaregistrováno 0,023% požadavků s časy odezvy 1 - 5 s a 0,023% požadavků s časem odezvy v intervalu 5 - 10 s.



Obr. 7. Podíl doby zpracování požadavku v jednotlivých dnech

Hodnoty datového toku ukazují na splnění požadavků INSPIRE na udržení odezvy nad 0.5 MB/s. Celkově je 98,7% požadavků s vyšší rychlostí datového toku než 0.5 MB/s, což plně vyhovuje. Tyto výsledky jsou totožné s výsledky testování v roce 2012.



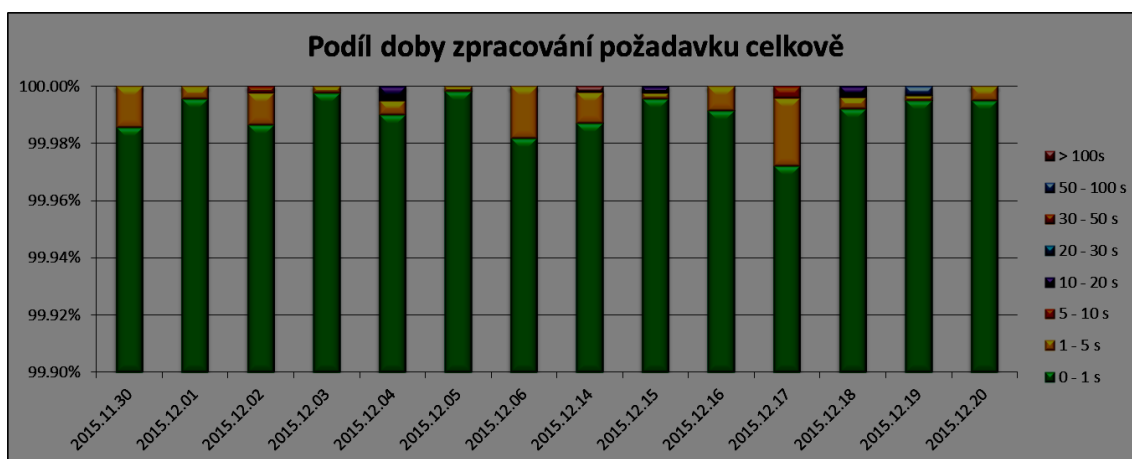
Obr. 8. Podíl datového toku celkově

Vyhodnocení operace „popsat typ prostorového objektu“ (Describe Spatial Object Type)

Kontroluje se, zda za normální situace je doba odezvy k odeslání prvotní odpovědi nejvýše 10 sekund.

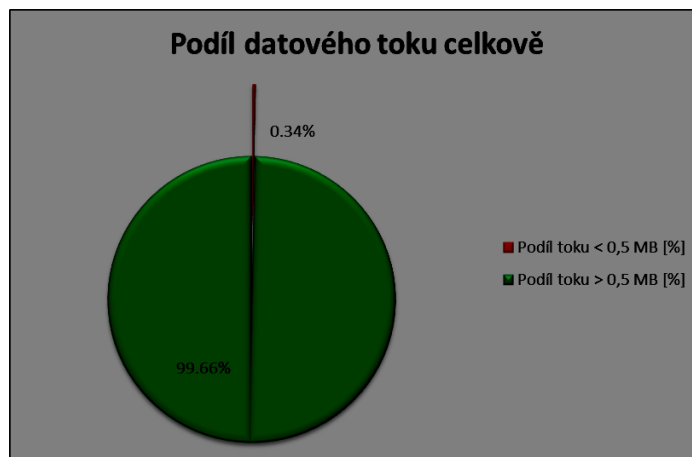
Dále se kontroluje, zda služba stahování dat udržuje za normální situace odezvu vyšší než 0,5 megabytu za sekundu.

Výsledky jsou vynikající, podíl požadavků s dobou odezvy k odeslání prvotní odpovědi větších než 10 s je pouze 0.001%, což zdaleka nedosahuje hranice 10%. V žádném ze zkoumaných dnů se nedosahuje hodnot, které by ukazovaly na možné ohrožení splnění této podmínky. Nejvyšší podíl horších reakčních časů byl registrován 4. 12. 2015 (0.005%) a 18. 12. 2015 (0,004%).



Obr. 9. Podíl doby zpracování požadavku v jednotlivých dnech

Sledování datového toku ukazuje bezproblémové splnění požadavků INSPIRE na udržení odezvy nad 0.5 MB/s. Celkově je 99,7% požadavků s vyšší rychlostí datového toku než 0.5 MB/s (což odpovídá výsledkům testování z roku 2012). Rovněž výsledky v jednotlivých dnech jsou velmi vyrovnané a nenaznačují existenci nějakých problémů v průběhu kalendáře.



Obr. 10. Podíl datového toku celkově

ZÁVĚR

Doba odezvy k odeslání prvotní odpovědi je splněna s dostatečnou rezervou, počty požadavků s dobou odezvy nad sledovaný limit jsou zcela minimální. Ani u žádné z 5 operací nedošlo k přiblížení ke sledovanému limitu.

Požadavek na dostatečnou rychlost stahování dat je u sledovaných operací splněn. Nejlepší výsledky z hlediska rychlosti stahování dat vykazuje operace Describe Spatial Object Type, kdy celkově 99,7% požadavků s vyšší rychlostí datového toku než 0.5 MB/s. V případě operace Get Spatial Data Set bylo požadované rychlosti dosaženo v 98% a žádný den testování nebyl limit ohrožen. Podobně je to v případě operace Describe Spatial Data Set (98.7%), kde se objevily horší (ale vyhovující) výsledky ve 2 dnech testování (30. 11. a 4. 12. 2015). V případě operace GetFeature byl sice celkový výsledek vyhovující (celkově 95.9% bylo staženo s vyšší než limitní velikostí), rovněž většina dnů prokázala splnění podmínky s dostatečnou rezervou, avšak 30. 11. a 4. 12. 2015 nebyla dosažen požadovaný podíl menších toků do 10%. 30. 11. bylo zjištěno jen 86.3% požadavků s rychlostí stahování nad 0.5 MB/s a 4. 12. 2015 jen 88.1%.

Hodnocení dostupnosti vycházelo ze zjištění počtu chyb při testování. Průměrné počty chyb u jednotlivých služeb kolísají od 0.085% do 0.119%, tj. zhruba 10x méně než je stanovený limit 1% dostupnosti služby. Nicméně v některých dnech (zejména 14. 12.) byla výrazně vyšší chybovost (14.12 často dosahovalo kolem 0,5% za celý den), a to naznačuje, že by v některých dnech mohlo být ohroženo splnění podmínky dostupnosti služby. Chybovost při testování jiných služeb ČUZK v roce 2012 byla pouze 0,021%. Proto doporučujeme dostupnosti služeb věnovat zvýšenou pozornost.

Celkově dobré výsledky byly potvrzeny i provedenými zátěžovými (stress) testy při simulovaném přístupu 250, resp. 500 klientů. Splnění požadavku doby odezvy je zajištěno bez potíží i při vysoké zátěži. Podíl odpovědí s dostatečnou rychlostí datového toku při zátěžových testech sice nebyla splněna, přesto jsou výsledky podobně dobré jako v roce 2012.

Testy datových struktur potvrdily souladu získaných dat s XML schématy, na které se data odvolávají v hlavičce XML souboru (vyjma drobných chyb u gml:id). Pokud jde o testovaná metadata, testy nevykázaly žádnou chybu a můžeme tedy tvrdit, že metadata jsou v souladu s deklarovanou strukturou.

LITERATURA

Horák J., Růžička J., Ardielli J. 2013. Výkonové a zátěžové testy stahovacích služeb ČUZK dle požadavků INSPIRE. In sborník GIS Ostrava 2013, Ostrava, 21-23.1.2013.

INSPIRE 2013. Technical Guidance Download Services v 3.1. 09-08-2013 EC. Dostupné na: <http://inspire.ec.europa.eu/>