

VYSOKÁ ŠKOLA BÁŇSKÁ – TECHNICKÁ UNIVERZITA OSTRAVA

**Hornicko – geologická fakulta
Institut geoinformatiky**

**Porovnání rychlosti mapového serveru GeoServer při
přístupu k různým datovým skladům**

Bakalářská práce

Autor:

Adam Schreier

Vedoucí bakalářské práce:

Ing. Jan Růžička, Ph.D.

Ostrava 2014

Zadání bakalářské práce

Student: **Adam Schreier**

Studijní program: B3646 Geodézie a kartografie

Studijní obor: 3646R006 Geoinformatika

Téma: Porovnání rychlosti mapového serveru GeoServer při přístupu k různým datovým skladům
Comparison of Performance of GeoServer when Dealing with Different Data Warehouses

Zásady pro vypracování:

Úkoly:

1. Prostudujte problematiku publikování dat s využitím nástroje GeoServer.
2. Prostudujte nástroje PostGIS, Oracle, případně H2.
3. Prostudujte specifikaci WMS.
4. Připravte datové sklady v prostředí PostGIS, Oracle případně i H2 ve dvou různém objemových variantách: do 1 GB dat a nad 5 GB dat.
5. Připravte jednoduché WMS pro přístup k datům v PostGIS, Oracle případně i H2.
6. S využitím nástroje WMSTester (nebo JMeter) posuďte rychlost odezvy při využití různých datových skladů.
7. Volitelně můžete zkusit porovnat i varianty přístupu bez JNDI a s JNDI, případně chování na různých operačních systémech.

Rozsah grafických prací:
dle potřeby

Rozsah původní zprávy:
30 - 40 stran textu

Seznam doporučené odborné literatury:

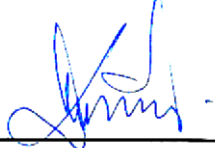
- Kropla B. Beginning MapServer. 2005. Apress. ISBN 1-59059-490-8
Růžička J., Peňáz T., Horák J., Stankovič J.: Publikování prostorových dat na internetu. Distanční text. VŠB-TU Ostrava, 2003, ISBN 80 – 248 - 0416 – 6
ISO. ISO 19128. Web map Service.

Formální náležitosti a rozsah bakalářské práce stanoví pokyny pro vypracování zveřejněné na webových stránkách fakulty.

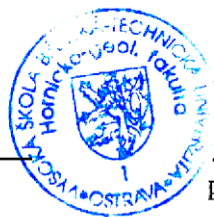
Vedoucí bakalářské práce: **Ing. Jan Růžička, Ph.D.**

Datum zadání: 31.10.2013

Datum odevzdání: 30.04.2014



prof. Ing. Zdeněk Diviš, CSc.
vedoucí institutu



prof. Ing. Vladimír Slivka, CSc., dr.h.c.
děkan fakulty

Prohlášení

- ***Celou bakalářskou práci včetně příloh, jsem vypracoval samostatně a uvedl jsem všechny použité podklady a literaturu.***
- ***Byl jsem seznámen s tím, že na moji bakalářskou práci se plně vztahuje zákon č.121/2000 Sb. - autorský zákon, zejména § 35 – využití díla v rámci občanských a náboženských obřadů, v rámci školních představení a využití díla školního a § 60 – školní dílo.***
- ***Beru na vědomí, že Vysoká škola báňská – Technická univerzita Ostrava (dále jen VŠBTUO) má právo nevýdělečně, ke své vnitřní potřebě, bakalářskou práci užít (§ 35 odst. 3).***
- ***Souhlasím s tím, že jeden výtisk bakalářské práce bude uložen v Ústřední knihovně VŠBTUO k prezenčnímu nahlédnutí a jeden výtisk bude uložen u vedoucího bakalářské práce. Souhlasím s tím, že údaje o bakalářské práci, obsažené v Záznamu o závěrečné práci, umístěném v příloze mé bakalářské práce, budou zveřejněny v informačním systému VŠBTUO.***
- ***Bylo sjednáno, že s VŠB-TUO, v případě zájmu z její strany, uzavřu licenční smlouvu s oprávněním užít dílo v rozsahu § 12 odst. 4 autorského zákona.***
- ***Bylo sjednáno, že užít své dílo – bakalářskou práci nebo poskytnout licenci k jejímu využití mohu jen se souhlasem VŠB-TUO, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly VŠB-TUO na vytvoření díla vynaloženy (až do jejich skutečné výše).***

V Ostravě dne 28. 4. 2014

Adam Schreier

Anotace

Tato bakalářská práce je věnována porovnání rychlosti GeoServeru při práci s různými datovými sklady.

První část práce popisuje nejdůležitější softwarové prostředky použité během zpracování, tedy PostgreSQL, Oracle, GeoServer a WMS Tester a dalším pomocným nástrojům. Věnuje se také specifikaci pro komunikační protokol WMS. Součástí této práce je popis instalace jednotlivých komponentů a zkompletování celého systému do funkčního celku. Následuje popis konfigurace testovacího nástroje WMS Testeru pro daný úkol, kterým je simulovaný zátěžový test, který by měl ukázat možnou regresní závislosti mezi počtem klientů a rychlostí odezvy serveru. Výsledky testů pro různé datové sklady a jejich porovnávání je cílem této bakalářské práce.

Klíčová slova: GeoServer, WMS, PostgreSQL, PostGIS, Oracle, Spatial, ShapeFile, Systém řízení báze dat, datový sklad, rychlost odezvy.

Summary

This thesis is devoted to compare the speed of GeoServer when working with different data warehouses.

The first part describes the most important software, used during processing, thus PostgreSQL, Oracle DB, GeoServer and WMS Tester. Next chapters contain a specification for a communication protocol WMS and afterwards describes how to install the individual components and complete the entire system into a functional unit. The following one describes the configuration test tool, WMS Tester for the task, which is simulated stress test, which should show a possible regression dependence between the number of clients and the server's response time. The test results for different data warehouses and comparison of them is the aim of this thesis.

Keywords: : GeoServer, WMS, PostgreSQL, PostGIS, Oracle Spatial, ShapeFile, Database management system, data warhouse, response time.

OBSAH

1	ÚVOD.....	1
2	CÍLE BAKALÁŘSKÉ PRÁCE.....	2
3	POUŽITÉ PROSTŘEDKY.....	4
3.1	Software	4
3.1.1	GeoServer	4
3.1.2	WmsTester	4
3.2	SŘBD a použité datové sklady.....	5
3.2.1	SŘBD PostgreSQL	6
3.2.2	SŘBD Oracle	7
3.2.3	ShapeFile	8
3.3	Testovací data.....	9
4	Web Map Service.....	10
4.1	OGC (Open Geospatial Consortium)U	10
4.2	WMS Specifikace.....	10
4.3	GetMap.....	11
5	WMS TESTER.....	12
5.1	Nastavení WMS Testeru	12
5.2	Výstupní soubory WMS Testeru.....	14
6	TESTOVANÁ DATA	15
6.1	Struktura dat	16
7	REALIZACE.....	17
7.1	Použitý hardware.....	17
7.2	Předinstalační příprava.....	18
7.2.1	Příprava pro Windows	18
7.2.2	Příprava pro Linux	19
7.3	PostgreSQL	21
7.3.1	Instalace PostgreSQL.....	21
7.3.2	Import a selekce dat	21
7.3.3	Pomocné nástroje.....	22
7.3.4	Export dat pro další použití.....	23
7.3.5	Opětný import pro objektivní testování	23
7.4	Oracle	24
7.4.1	Instalace	24
7.4.2	Instalace nadstavby Spatial.....	25
7.4.3	Vytvoření databáze	25
7.4.4	Import dat.....	27
7.5	GeoServer.....	28
7.5.1	Instalace GeoServer	28
7.5.2	Doporučená specifikace	28
7.5.3	Nastavení GeoServer	29
7.5.4	PŘIPOJENÍ DATABÁZÍ	30
7.5.5	Nastavení WMS	31
7.6	Publikování dat pomocí WMS	32
7.7	WMS Tester - klient.....	33
8	VÝSLEDKY A JEJICH INTERPRETACE	35
8.1	Úvod k testování	35

8.2	Sledování práce systému	35
8.3	Průběh testování	36
8.4	Malá databáze.....	37
8.5	Střední databáze	39
8.6	Kompletní databáze.....	41
8.7	Porovnání testovaných SŘBD	42
9	ZÁVĚR	43
10	SEZNAM POUŽITÉ LITERATURY:.....	45
11	SEZNAM OBRÁZKŮ A TABULEK.....	46

SEZNAM POUŽITÝCH ZKRATEK

České zkratky

ČÚZK	Český úřad zeměměřičský a katastrální
DKM	Digitální katastrální mapa
GIS	Geografický informační systém
SŘBD	System řízení báze dat

Cizojazyčné zkratky

CPU	Central Processing Unit
CRM	Customer Relationship Management
DBMS	Database Management System
ESRI	Environmental Systems Research Institute
ETRF-89	European Terrestrial Reference System 1989
EPSG	European Petroleum Survey Group
FTP	File Transfer Protocol
GIF	Graphics Interchange Format
GML	Graphic Markup Language
GNU	GNU is Not UNIX
GRASS	Geographic Resources Analysis Support System
HDD	Hard Disk Drive
HTTP	Hypertext Transfer Protocol
INSPIRE	Infrastructure for Spatial Information in Europe
ISO	International Organization for Standardization
JDK	Java Developer Kit
JPEG	Joint Photographic Experts Group
MSDN	Microsoft Developer Network
NTFS	New Technology File System
OGC	Open Geospatial Consortium
OWS	OGC Web Service
PC	Personal Computer
PNG	Portable Network Graphics
RAM	Random Access Memory
SATA	Serial Advanced Technology Attachment

SDE	Spatial Database Engine
SP	Service Pack
SQL	Structured Query Language
URL	Uniform Resource Locator
WCS	Web Coverage Service
WFS	Web Feature Service
WMS	Web Map Service
WRS	World Registry Service
WWW	World Wide Web
XML	Extensible Markup Language

1 ÚVOD

Na počátku počítačových dob, byla myšlenka šíření prostorových dat a jejich publikace široké veřejnosti myšlenkou hodně vzdálenou. Bylo nutné překonat mnohé problémy a vyřešit nelehké technické úlohy.

Dnes je však situace úplně jiná a dnešní počítače dokáží prostorová data nejen zpracovat, ale zvládnou celý proces od jejich získávání přímo v terénu až po jejich interpretaci, analýzu a publikování způsoby, jaké byly ještě v nedávné době téměř nemyslitelné.

GIS je již běžnou součástí života každého z nás, ať už jde o pouhé zobrazování map při přehledu předpovědi počasí, nebo zjišťování informací z katastru nemovitostí. A pro všechny operace, kde GIS figuruje, je nutné využívat různá prostorová data, tedy data nesoucí informaci o své poloze v prostoru.

V dnešní době, je proces získávání dat z velké části ulehčen a automatizován, sběr dat se stává stále více efektivní a možnosti využití stále širší. S tím také logicky narůstá i množství dat, které je nutné zpracovat, analyzovat, popřípadě publikovat a sdílet dále, aby bylo možné je využít v co možná největším měřítku.

Prakticky veškeré moderní mapování, vyjma leteckého snímkování je prováděno ve vektorové podobě. Ať už jsou tato data získávána při digitalizaci starých mapových děl, nebo vznikají jako součást děl nových, (např.: tvorbou digitálních katastrálních map, map pro navigační systémy, atd...) je nutné řešit i otázku dostupnosti a možnosti práce s nimi.

Tímto způsobem vzniká velké množství dat, které se stávají čím dál více nepřehledné a stále obtížněji zpracovatelné, pokud není možné je rychle a jednoduše distribuovat nejrozličnějším zájemcům z řad státních úřadů, magistrátů, soukromých společností ale i amatérských uživatelů pro další zpracování či publikování.

Základem je všechna tato data uložit do řádné struktury v databázi, kterou poté dokážeme využít pro další zpracování. Tato bakalářská práce se zabývá možnostmi a výkonem systémů řízení báze dat, které dokáží s těmito databázemi pracovat a ve spolupráci s GeoServer tato data i sdílet; A jejich porovnání s řešením, které tyto systémy nevyužívá.

Tato problematika je často opomíjena a mnoho kvalitních prací na toto téma neexistuje.

Jediná kvalitně zpracovaná studie, která byla během rešerše nalezena, porovnává PostGIS a Oracle Spatial. V této studii však byl pro každý SŘBD použit jiný operační systém a výsledky tedy byly ovlivněny i prostředím, ve kterém jednotlivé systémy pracovaly.

V této bakalářské práci jsou všechny testy prováděny v jednom operačním systému a za totožných podmínek, tak aby k ovlivnění výsledků nedošlo. Odkaz na tuto práci je uveden v seznamu zdrojů [7].

2 CÍLE BAKALÁŘSKÉ PRÁCE

Cílem této bakalářské práce je zjištění případných rozdílů v rychlosti odezvy v závislosti na použitém datovém skladu.

Jako mapový server je zde použit open-source server GeoServer. Práce porovnává tři rozdílné, datové sklady, patřící v současnosti mezi nejpoužívanější, pro uložení velkých vektorových dat v závislosti na rychlosti vyřízení požadavku klienta na server. Pro testování byly vybrány tyto datové sklady: PostgreSQL a jeho prostorová nadstavba PostGIS, Oracle a jeho prostorová nadstavba Spatial a datový formát ShapeFile.

Prvním krokem bylo seznámení se s možnostmi publikování prostorových dat prostřednictvím internetu či intranetu, a to pomocí serveru GeoServer. Dalším krokem pak bylo prostudování možností vybraných datových skladů s ohledem na jejich možnosti a volby nastavení.

Důvodem výběru těchto konkrétních datových skladů bylo především dáno tím, že PostgreSQL a Oracle jsou dva z nejpoužívanějších SŘBD a přitom se v mnohém liší, především formou licencování, kde PostgreSQL je open-source a Oracle je komerční, placený software.

Třetí, nepovinou, volbou z testovaných datových skladů, měl být původně H2. V průběhu zhotovování práce a zejména během testování bylo využito určité volnosti ve volbě 3. datového skladu a z důvodů popsaných níže byl zvolen k otestování obyčejný prostorový datový formát Shapefile.

Hlavním důvodem pro výběr Shapefile byla myšlenka, že pro publikaci dat prostřednictvím WMS není potřeba žádných zvláštních funkcí, nebo nastavení a je tedy možné využít přímo Shapefile, bez nutnosti instalovat jakýkoli SŘBD a následně importovat do některého z nich. Společně s touto myšlenkou je možné si položit také otázky:

Bude využití SŘBD přinášet výhodu v rychlosti zpracování? V jaké situaci je lepší použití jednoduchého řešení a kde se vyplatí využít SŘBD?

Získat odpovědi na tyto, ale i další otázky je předmětem této práce.

Dalším stěžejním bodem práce bylo zprovoznění a nastavení serverové služby WMS pro publikování tří testovacích vrstev, (oproti 2 zadaných byla doplněna ještě jedna vrstva tak, aby bylo pokryto větší spektrum možných situací) prostřednictvím služby Web Map Service. Samotné testování závislosti mezi rychlostí odezvy webového klienta a použitým datovým skladem pak obstarával WMS Tester – aplikace pro zátěžové testování WMS služeb.

Hlavním cílem této práce je však pomoci uživateli při rozhodování, které řešení je pro něj to nejlepší z hlediska rychlosti zpracování a s tím souvisejícího uživatelského komfortu.

Z tohoto důvodu je práce psaná částečně ve formě manuálu, podle kterého by měl každý alespoň trochu počítačově znalý člověk být schopný provést veškeré zde popisované testy ve vlastním prostředí, nezávisle na tom jestli používá Windows, nebo Linux a přesně tak určit které z předkládaných řešení je pro něj to ideální.

3 POUŽITÉ PROSTŘEDKY

3.1 Software

Pro účel bakalářské práce byl využit níže uvedený software. A to jak pro vlastní publikování testovacích dat, tak pro jejich otestování na rychlost publikování dat prostřednictvím GeoServeru v závislosti na daném SŘBD.

3.1.1 GeoServer

GeoServer je open source software napsaný kompletně v jazyce Java. Umožňuje uživatelům sdílet a upravovat geodata. Dokáže publikovat data z téměř libovolného zdroje prostorových dat s využitím otevřených standardů.

Jako komunitní projekt je GeoServer vyvíjen, testován a podporován různorodými skupinami jednotlivců a organizací z celého světa.

V GeoServeru je samozřejmostí implementace standardů Open Geospatial Consortium (OGC) Web Feature Service (WFS) a Web Coverage Service (WCS), stejně jako vysoký výkon a plná kompatibilita s Web Map Service (WMS). GeoServer tvoří základní součást webu pracujícím s prostorovými daty. [1]

3.1.2 WmsTester

WMS Tester je Java-aplikace sloužící k testování WMS serverů a jejím hlavním cílem je extrémní zátěžový test, který se snaží ověřit možnou regresní závislost (počet klientů x rychlost WMS serveru). Jedná se o komunitní projekt. Nástroj je volně dostupný na stránkách www.sourceforge.net.

Více v samostatné kapitole 5

3.2 SŘBD a použité datové sklady

System řízení báze dat (zkracováno na SŘBD či DBMS) je software, který zajišťuje práci s databází, tzn., tvoří rozhraní mezi aplikačními programy a uloženými daty. Občas se pojem zaměňuje s pojmem databázový systém. Databázový systém však je SŘBD dohromady s databází.

Aby mohl být nějaký programový systém označený za SŘBD, musí být jednak schopen efektivně pracovat s velkým množstvím dat, ale také musí být schopný řídit (vkládat, modifikovat, mazat) a definovat strukturu těchto perzistentních dat (čímž se liší od prostého souborového systému). [2]

Schopnosti SŘBD budou tedy srovnány se souborovým systémem. Souborový systém si můžeme představit i velice jednoduše jako adresář, nebo strukturu adresářů, ve kterých se nachází soubory s daty, které je možné číst a využít. Jak už z logiky věci vyplývá, soubor sám o sobě není schopen vkládat, modifikovat, nebo mazat vlastní obsah. Pro jednoduchou publikaci dat ale žádná z těchto schopností není nutná a ve prospěch tohoto prostého řešení také mluví fakt, že není prakticky nic, co by se muselo nastavovat, nebo instalovat, pokud se rozhodneme data publikovat tímto způsobem.

3.2.1 SŘBD PostgreSQL

PostgreSQL je výkonný open-source, objektově-relační databázový systém. Má za sebou více než 15 let aktivního vývoje a osvědčenou architekturu, kvůli které si v komunitě uživatelů zasloužil silnou pověst pro svou spolehlivost, integritu dat a korektnost. Je kompatibilní na všech hlavních operačních systémech, včetně Linuxu, UNIX (AIX, BSD, HP - UX, SGI IRIX, Mac OS X, Solaris, Tru64), a Windows. Má plnou podporu cizích klíčů a zahrnuje většinu SQL:2008 datových typů. Včetně INTEGER, BOOLEAN, CHAR, VARCHAR, DATE, INTERVAL a časového razítka. Podporuje ukládání rozsáhlých binárních objektů, včetně obrázků, zvuků nebo videa. [1]

Výhody:

- Open-source
- Velmi rozšířený
- Aktivní komunita neustále vylepšující jeho chod

Některé obecné limity PostgreSQL jsou uvedeny v následující tabulce.

Limit	Hodnota
Maximální velikost databáze	neomezeno
Maximální velikost tabulky	32 TB
Maximální velikost řádky	1.6 TB
Maximální velikost jednoho pole	1 GB
Maximální počet řádků	neomezeno
Maximální počet sloupců	250 - 1600 záleží na datovém typu

Tabulka 1 Limity PostgreSQL

PostGIS

PostGIS přidává podporu pro geografické objekty v objektově-relačním SŘBD PostgreSQL. Ve skutečnosti, PostGIS přidává prostorovou funkcionalitu pro PostgreSQL, díky čemuž může být použit jako prostorová databáze pro geografické informační systémy (GIS), stejně jako Shapefile nebo Oracle Spatial.

Vývoj PostGIS byl zahájen jako výzkumný projekt open source technologie, prostorové databáze. PostGIS je uvolněn pod GNU General Public Licence. PostGIS je nadále rozvíjen skupinou přispěvatelů vedené řídicím výborem projektu a stále jsou přidávány nové funkce. [3]

3.2.2 SŘBD Oracle

Oracle je systém řízení báze dat, společně s daty tvoří moderní multiplatformní databázový systém s velice pokročilými možnostmi jejich zpracování. Tento nástroj je komerční ekvivalencí pro PostgreSQL.

Oracle Corporation je název firmy, oficiální název databázové platformy je Oracle Database. Oracle Corporation je jedna z hlavních společností vyvíjejících relační databáze, nástroje pro vývoj a správu databází či customer relationship management (zkráceně CRM) systémů. Byla založena v roce 1977 a v roce 2005 zaměstnávala 50 000 lidí. [4]

Výhody:

- Funkce automatické optimalizace dat efektivně spravují více dat, snižují náklady na úložiště a zvyšují výkon databáze.
- Profesionální podpora
- Velké zkušenosti firmy s řízením objemných dat

Některé obecné limity Oracle jsou uvedeny v následující tabulce.

Limit	Hodnota
Maximální velikost databáze	neomezeno
Maximální velikost tabulky	32 TB
Maximální velikost řádky	1.6 TB
Maximální velikost jednoho pole	4 GB
Maximální počet řádků	neomezeno
Maximální počet sloupců	1000

Tabulka 2 Limity Oracle

Oracle Spatial

Oracle Spatial (dále jen Spatial) je integrovaná množina funkcí a procedur, které umožňují uložení, přístup a analýzu prostorových dat rychlým a efektivním způsobem v databázi Oracle. Spatial je dostupný pouze v edici Enterprise.

Spatial podporuje mimo jiné objektově-relační model uložení prostorových dat. Pro uložení prostorových dat využívá objektového datového typu SDO_GEOMETRY. V jedné tabulce jsou tak pro prvek popisná data uložena společně s prostorovými. [5] Výhody nabízené objektově-relačním modelem jsou především:

- Podpora mnoha geometrických typů, např. oblouků, kružnic, liniových řetězců či polygonů.
- Snadné použití při vytváření a údržbě prostorových indexů a dotazů.
- Uložení geometrických popisů prvků v jednom řádku a jednom sloupci tabulky.

3.2.3 ShapeFile

ShapeFile je vektorový datový formát pro ukládání prostorových dat a jejich atributů.

Prostorová informace je ukládána jako jednoduchý zápis souřadnic.

Popisuje hlavní geometrické prvky, jako jsou body, linie, polygony,...

Protože ShapeFile neukládá žádnou topologickou informaci, je jeho zpracování oproštěno od dalších procesů počítající i topologickou strukturu dat. Tento fakt může znamenat rychlejší zpracování a také menší nároky na prostor na disku.

Shapefile jako takový není pouze jediným souborem, ale sadou spolupracujících souborů, kde každý z nich má přesně specifikovaný úkol.

Celý ShapeFile „komplet“ se sestává z více souborů. Jedním je samotný ShapeFile. Další jsou soubory pomocné. [6]

Struktura ShapeFile souborů

Koncovka	Funkce
.sup	Ukládá samotné prostorové informace ve formě vektorových souřadnic ve vlastním formátu
.dbf	Ukládá atributy, atributy a geometrie jsou propojeny
.shx	Obsahuje „prostorové“ indexy
.prj	Nese informaci o kartografickém systému, ve kterém jsou prostorová data uložena (spatial bin) rozděluje prostor ShapeFile do obdélníkových výřezů (bins).
.sbn	Každý z nich obsahuje určitý počet Features a je zaznamenána oblast, kterou pokrývá. Při prostorových dotazech jsou tak do výpočtů zahrnuta je n ty „biny“, které spadají do hledané oblasti. Toto způsobuje velký nárůst výkonu při výpočtech.
.sbn	Pomocný soubor pro .sbn. Obsahuje identifikátory jednotlivých řádek a jejich velikost.

Tabulka 3 Struktura ShapeFile

Využití Shapefile

ShapeFile je prakticky nejpoužívanější formát pro ukládání prostorových dat.

Byl vyvinut společností ESRI (Environmental Systems Research Institute) začátkem 90. let a od té doby se těší velké popularitě. Tuto popularitu si zasloužil zejména pro svou jednoduchost a zároveň schopnost ukládat data kvalitně tak, že je možné je využít na širokou škálu dalších operací.

Tento formát se stal téměř standardem a prakticky každý GIS program umí s tímto formátem pracovat. Navíc existuje velká šance, že jestli máte někde uložena prostorová data, bude to právě Shapefile, kde jsou uložena, nebo alespoň se do Shapefile dají jednoduše převést.

3.3 Testovací data

Pro samotné testování a práci na této bakalářské práci je nutné zvolit vyhovující data. Předmětem bakalářské práce je vlastní testování vybraných dat o různé datové velikosti publikovaných prostřednictvím GeoServer.

Data použitá v této bakalářské práci byla získána z ČUZK a je to harmonizovaná část digitální katastrální mapy podle INSPIRE. Tato data jsou ve vektorové podobě.

Vzhledem k povaze zadání a účelům testu bylo nutné získat objemná data v rozmezí minimálně 5 GB. Z vektorových dat DKM o celkové velikosti 8GB byly tedy exportovány 3 soubory ve formátu Shapefile.

Více v samostatné kapitole 6

Data jsem obdržel 18. 10. 2013.

Použitý souřadnicový systém je ETRF-89

Geografické souřadnice zobrazené oblasti jsou:

Min X = 12,7790813446045

Min Y = 48,9878387451172

Max X = 18,5905704498291

Max Y = 51,0215263366699

4 Web Map Service

Je světový standard, který byl vyvinutý OGC a je i nadále rozšiřován

Patří do skupiny standardů OWS (OGS Web Service),

Mezi OWG patří také WCS (Web Coverage Service), WFS (Web Feature Service) a WRS (Web Registry Service).

4.1 OGC (Open Geospatial Consortium)

Open Geospatial Consortium (OGC) je nevládní mezinárodní nezisková standardizační organizace, zahrnující celosvětově více než 400 komerčních, vládních, nevýdělečných a vědeckých organizací za účelem spolupráce na procesu otevřené shody podporující vývoj a implementaci standardů pro geoprostorová data a služby, GIS, zpracování dat a jejich výměnu. Založena byla roku 1994 jako náhrada za OGF (Open GRASS Foundation).

4.2 WMS Specifikace

WMS je postavena na principu Klient-Server a slouží jako přístupové rozhraní pro požadavky a dotazy přenášené skrze protokol HTTP na prostorová data prostřednictvím rastrových výstupů ve zvoleném formátu.

ISO (International Organization for Standardization) přijala WMS jako mezinárodní standard pod označením ISO 19128 Geographic Information: Web Map Service v roce 2005 [8].

Přehled základních typů dotazů [9]:

GetCapabilities - Pokud daný klient nezná možnosti a vlastnosti spravovaných dat, což se skoro ve všech případech děje, potřebuje tyto možnosti zjistit. Proto klient při první komunikaci se serverem sestaví GetCapabilities dotaz pro získání metadat WMS serveru ve formátu XML.

GetMap - Tento typ dotazu lze považovat za hlavní, klientovi zpřístupní mapu ve formě obrazových dat v určitém formátu. Viz níže....

GetFeatureInfo - Tento typ dotazu vrací klientovi informaci zapsanou v XML s atributy daného prvku na mapě o určitých souřadnicích.

4.3 *GetMap*

Je hlavním požadavkem protokolu WMS a skládá se z následujících prvků:

FORMAT - specifikuje výstupní rastrový formát dat. Zahrnuje běžné obrazové formáty (GIF, PNG a JPEG).

SRS - (Spatial Reference System) určuje referenční systém pomocí prefixu EPSG kódu rozdělených dvojtečkou.

BBOX - (Bounding Box) definuje obdélníkový výřez mapy pomocí čtyř rohových hodnot: minimální délka, minimální šířka (levý spodní roh), maximální délka a maximální šířka (pravý horní roh) a to v jednotkách podle zvoleného souřadnicového systému

LAYERS – Seznam požadovaných vrstev, které se mohou překrývat. Pořadí vrstev je první-nejníže, poslední nejvýše.

STYLES – Vyvolá požadovaný, předem definovaný styl, podle kterého je mapa vykreslována. "STYLES="

WIDTH a **HEIGHT** - definuje šířku a výšku výsledného rastrového obrazu. Při nedodržení poměrů stran zadaných v BBOX je obraz deformován

TRANSPARENT - určuje průhlednost a nabývá hodnot "TRUE" nebo "FALSE" (implicitní je FALSE)

BGCOLOR - barva pozadí, definována hexadecimálně, implicitní je bílá (0xFFFFFFFF). Zapisuje se ve formátu 0xRRGGBB.

EXCEPTIONS - uvádí předdefinována chybová hlášení pro klienta. Vracejí chybovou hlášku zapsanou v XML.

5 WMS TESTER

WMS Tester je aplikace, která slouží k simulaci zátěžových testů WMS serverů a nepřímo jejich databázových systémů. Tato aplikace je napsána kompletně v programovacím jazyce Java, její fungování napříč platformami tak není prakticky nijak omezeno. Výstupem jsou 4 soubory s naměřenými hodnotami, viz. níže. Hlavním účelem tohoto testování je ověřit závislosti mezi počtem klientů přistupujícím k testované službě a časem odezvy a zpracování požadavku pro jednotlivé dotazy. Konkrétní verze použitá v této práci je verze 1.2 r1.

5.1 Nastavení WMS Testeru

WMS Tester jako většina programů napsaných v Java je spustitelný pomocí příkazů z příkazové řádky, nebo přímým spuštěním inicializačního souboru.

V případě spuštění z příkazové řádky stačí v adresáři instancí WMS Testeru spustit příkaz „java -jar wms tester.jar“

Jakmile je WMS Tester spuštěn, inicializuje se příkazová řádka ve které se WMS Tester obsluhuje a zobrazí se nápověda a přehled příkazů.

Jejich přepis vidíte níže:

Konkrétně **WmsTester-1.2 r1** umožňuje uživateli nastavit 5 základních parametrů testu:

-e	--endpoint	URL adresa testovaného WMS serveru.
-l	--layer	Seznam požadovaných WMS vrstev oddělených „&&“.
-v	--vendor	Seznam zprostředkovatelů WMS
-u	--user	Autorizační údaje pro chráněný server.
-p	--passwd	Heslo pro chráněný server.

Pro další upřesnění a nastavení testu je v této verzi připraveno dalších 11 nastavovacích a 2 aplikační parametry:

-t	--test-name	Možnost zvolit si název výstupních souborů
-d	--read-delay	Pauza mezi dotazy (default: „0 s“).
-n	--read-num	Požadovaný počet rastrů na klienta (default: „10“).
-f	--format	Formát výstupního rastru (default: „image/png“).
-t	--threads	Počet vláken (klientů) (default: „1“).
-i	--increase	Interval připojování dalších klientů (default: „10“).

-dd	<code>--debug</code>	Úroveň debugování
-log	<code>--log</code>	Ukládá obrázky, které vznikají při dotazech klientů
-tt	<code>--total-time</code>	Časový limit délky testu (default: „300 s“).
-eh	<code>--extend-height</code>	Volitelná výška rastru (default: „800 px“).
-ew	<code>--extend-width</code>	Volitelná šířka rastru (default: „1200 px“).
	<code>--help</code>	Nápověda k nastavení aplikace.
	<code>--version</code>	Výpis verze aplikace.

```

Microsoft Windows [Version 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. All rights reserved.

C:\Users\Test>java -jar wmsTester\WmsTester-1.2-r1-jar-with-dependencies.jar
Dub 05, 2014 4:04:27 ODP. org.apache.grape.commons.PropertiesConfig reload
INFO: Loaded from class path ==> project.properties
Dub 05, 2014 4:04:28 ODP. org.apache.grape.commons.PropertiesConfig reload
INFO: Missing in full path ==> project.properties
Dub 05, 2014 4:04:28 ODP. org.apache.grape.commons.PropertiesConfig reload
INFO: Config path is not defined
Usage: -m mode -e <URL> [-l <layer-name>] [-v <vendors-params>] [-d <delay-sec>]
[-i <increment-time>] [-c <read-count>] [-t <threads-count>] [-eh <height>]
[-ew <width>] [-u <user> -p <password>] [--help] [--version] [-log <0|1>] [--dd
<0-3>]

This application is intended to test WMS services in diferent way. Primary goal
is to test service under extreme user use and check if there is any performace r
egression.

Options:
  -m, --mode           Test modes:
  -m seq              Reads random sequence.
                     simulation of real user (presudo random ZoomIn,
                     Move, ZoomOut)
  -m rer              Reads random extends from
                     whole layer
  -m zer              ZoomIn sequence in the
                     middle of the layer
  -m rsr              Reads random
                     image size from a whole layer (default:'seq')
  -T, --test-name     Name used for stats and logged files
  -e, --endpoint      WMS URL
  -l, --layer         List of WMS layer names separated by '&&'
  -v, --vendor        List of WMS vendor specific params separated by
                     '&&', name and value are separated by '='
  -u, --user          User name to pass authentication
  -p, --passwd        Users password
  -d, --read-delay    Delay before the next requests is send (s)
                     (default:'0')
  -n, --read-num      Maximum number of requests sent by one thread
                     (client) (-1 infinity) (default:'10')
  -f, --format        Format (default:'image/png')
  -t, --threads       number of parallel running threads (clients)
                     (default:'1')
  -i, --increase      time between the increasing number of active
                     clients (s) (default:'10')
  -tt, --total-time   Maximal test time (s) (default:'300')
  -log, --log         log downloaded images (default:'0')
  -dd, --debug        debug level (default:'0')
  -eh, --extend-height extend height (px) (default:'600')
  -ew, --extend-width extend width (px) (default:'1024')
  --help             Implementation of WMS Tester Client
  --version          Print application version

C:\Users\Test>_

```

Obrázek 1 Výpis parametrů WMS-Tester

Nejdůležitějším parametrem je parametr „-e“, který při zadání URL adresy kde se nachází WMS server spustí základní předdefinovaný test, jehož chování můžeme měnit o ovlivňovat tak celý test pomocí nastavovacích parametrů. Výsledky testů jsou uloženy do souborů, jejichž popisem se zabývá následující podkapitola.

5.2 Výstupní soubory WMS Testeru

WMS Tester ukládá naměřené hodnoty do 4 souborů. Tyto soubory jsou v datovém formátu CSV a je možné s nimi pracovat a dále analyzovat tyto výsledky.

Program, který byl při analýze těchto výsledků použit, byl Microsoft Excel s balíku Microsoft Office

Výstupní soubory jsou tyto:

RandomExtentRead-mt.csv

Zaznamenává a hodnoty všech dotazů zaslaných na server za použití požadavku „GetMap“.

RandomExtentRead-mt-avg.csv

Zaznamenává hodnoty všech parametrů v předchozím souboru jako průměry.

RandomExtentRead-mt-slate.csv

Zaznamenává sečtené hodnoty pro jednotlivé bloky testu. Jeden blok je vždy mezi zapojením dalšího klienta do testu po intervalech -i.

RandomExtentRead-mt-thread.csv

Zaznamenává sečtené hodnoty pro jednotlivé bloky testu. Jeden blok je vždy mezi zapojením dalšího vlákna do testu.

Tyto soubory jsou vytvářeny při spuštění testu v WMS Testeru a můžeme je nalézt vždy ve stejném adresáři jako WMS Tester samotný. Problém může ale nastat při spuštění více testů v řadě. Pokud není nastaveno jinak WMS Tester nevytváří soubory nové, ale přepisuje soubory stávající, proto je nutné před započítím nového testu změnit nastavení testu, konkrétně parametr zodpovědný za pojmenování výstupních souborů, aby nedošlo k jejich ztrátě.

Struktura jednotlivých souborů a jejich popis je součástí přílohy č. 1

6 TESTOVANÁ DATA

Jako testovací, byla vybrána data, jenž pochází z ČUZK, konkrétně se jedná o digitální katastrální mapu v harmonizované podobě podle směrnice INSPIRE. Data jsou ve vektorovém formátu a již v databázi pro PostgreSQL verze 9.1 a novější.

Tato data jsou přibližně 8GB velká, jedná se tedy o poměrně rozsáhlou databázi, u které bude možné nasimulovat vícero testů, tedy test na malou střední a kompletní databázi.

Po importu do PostgreSQL byla měla databáze 5,29GB dat, což splňuje podmínku ze zadání, otestovat tedy databázi větší, než 5GB dat. Dále byla pro účely testování databáze ještě zmenšena dle zadání na menší než 1GB dat. Vzhledem k výsledkům test, které jsou uváděny později a účelu testování byla data rozdělena a otestována nad rámec zadání ještě na malou databázi (cca 10MB). Stření testovaná databáze měla cca 150MB.

Tato data byla vybrána zejména pro jejich dostupnost, kvalitu zpracování a pro přiblížení se co nejvíce reálné situaci využití této bakalářské práce jako pomůcky při rozhodování o použití určitého SŘBD. K bakalářské práci je v elektronické příloze z důvodů objemu dat přiložena pouze „malá“ a „střední“ databáze, která pro otestování a nastavení celé struktury plně postačuje. Kompletní data je vždy možné stáhnou z oficiálních stránek ČUZK, nebo budou dostupná u vedoucího práce.

Nakonec byly takto nově získaná data rozdělena do samostatných tabulek, tvořící zmenšeniny původní databáze. Tyto byly následně podle postupu uvedeného v kapitole 7.3.6 exportovány do formátu Shapefile, se kterým bylo následně pracováno.

Dělení databází	Velikost Shapefile	Velikost .shp	Počet řádků	Název databáze
Malá	18.2MB	11 734KB	24 999	Vrozek2
Střední	375MB	127 111KB	499 999	Vzorek1
Kompletní	10.4GB	3 803 915KB	15 309 000	CP

Tabulka 4 Rozdělení testovaných dat

6.1 Struktura dat

Následující obrázek nastiňuje strukturu dat v databázi a jejich popis

	gid [PK] serial	beginlifes character varying(80)	areavalue character vai	localid character varying	label character vai	nationalca character varyim	referencp character varying(80)	the_geom geometry(MultiPolygon,4258)
1	15031071	2012-06-13T06:56:27Z	2151	CP.3032193712	25/14	991902-25/14	16.945852 49.489293	0106000020A210000001000000010300000001000
2	15031070	2011-10-27T12:17:19Z	8343	CP.3032192712	25/13	991902-25/13	16.945983 49.489682	0106000020A210000001000000010300000001000
3	15031069	2012-06-13T06:56:27Z	3075	CP.3032191712	25/12	991902-25/12	16.946273 49.49007	0106000020A210000001000000010300000001000
4	15031068	2011-10-27T12:17:19Z	2979	CP.3032190712	25/11	991902-25/11	16.94628 49.490293	0106000020A210000001000000010300000001000
5	15031067	2011-10-27T12:17:19Z	3057	CP.3032189712	25/10	991902-25/10	16.94631 49.490493	0106000020A210000001000000010300000001000
6	15031066	2011-10-27T12:17:19Z	3117	CP.3032188712	25/9	991902-25/9	16.946378 49.490719	0106000020A210000001000000010300000001000
7	15031065	2011-10-27T12:17:19Z	3185	CP.3032187712	25/8	991902-25/8	16.946386 49.490935	0106000020A210000001000000010300000001000
8	15031064	2012-06-13T06:56:27Z	73	CP.3032256712	163/13	991902-163/13	16.947142 49.485734	0106000020A210000001000000010300000001000
9	15031063	2011-10-27T12:17:19Z	135	CP.3032255712	163/12	991902-163/12	16.946879 49.48587	0106000020A210000001000000010300000001000
10	15031062	2012-06-12T08:56:28Z	63	CP.3032254712	163/11	991902-163/11	16.946641 49.486032	0106000020A210000001000000010300000001000
11	15031061	2011-10-27T12:17:19Z	185	CP.3032253712	163/10	991902-163/10	16.946355 49.486296	0106000020A210000001000000010300000001000
12	15031060	2012-06-12T08:56:28Z	51	CP.3032252712	163/9	991902-163/9	16.946179 49.486462	0106000020A210000001000000010300000001000
13	15031059	2011-10-27T12:17:19Z	3	CP.3032251712	163/8	991902-163/8	16.945431 49.488278	0106000020A210000001000000010300000001000
14	15031058	2012-06-13T08:56:29Z	35	CP.3032250712	163/7	991902-163/7	16.945315 49.488457	0106000020A210000001000000010300000001000
15	15031057	2012-06-13T06:56:27Z	14	CP.3032249712	163/6	991902-163/6	16.945112 49.488828	0106000020A210000001000000010300000001000
16	15031056	2012-07-03T12:56:30Z	1	CP.3032248712	163/5	991902-163/5	16.945006 49.489032	0106000020A210000001000000010300000001000
17	15031055	2012-06-13T06:56:27Z	4649	CP.3032172712	23/29	991902-23/29	16.953766 49.481748	0106000020A210000001000000010300000001000
18	15031054	2011-10-27T12:17:19Z	4294	CP.3032171712	23/28	991902-23/28	16.95406 49.481649	0106000020A210000001000000010300000001000
19	15031053	2011-10-27T12:17:19Z	4767	CP.3032170712	23/27	991902-23/27	16.954388 49.481583	0106000020A210000001000000010300000001000
20	15031052	2011-10-27T12:17:19Z	4885	CP.3032169712	23/26	991902-23/26	16.954747 49.480462	0106000020A210000001000000010300000001000
21	15031051	2012-06-13T08:56:29Z	35	CP.3032168712	23/25	991902-23/25	16.955341 49.480882	0106000020A210000001000000010300000001000
22	15031050	2012-06-13T08:56:29Z	3929	CP.3032167712	23/24	991902-23/24	16.956027 49.481437	0106000020A210000001000000010300000001000
23	15031049	2012-06-12T08:56:28Z	3896	CP.3032166712	23/23	991902-23/23	16.956627 49.482472	0106000020A210000001000000010300000001000
24	15031048	2011-10-27T12:17:19Z	3810	CP.3032165712	23/22	991902-23/22	16.956278 49.482644	0106000020A210000001000000010300000001000
25	15031047	2012-06-13T06:56:27Z	4532	CP.3032164712	23/21	991902-23/21	16.955884 49.482787	0106000020A210000001000000010300000001000
26	15031046	2011-10-27T12:17:19Z	4597	CP.3032163712	23/20	991902-23/20	16.95485 49.483204	0106000020A210000001000000010300000001000
27	15031045	2011-10-27T12:17:19Z	3206	CP.3032186712	25/7	991902-25/7	16.946325 49.491121	0106000020A210000001000000010300000001000
28	15031044	2012-06-13T06:56:27Z	3288	CP.3032185712	25/6	991902-25/6	16.94611 49.491915	0106000020A210000001000000010300000001000
29	15031043	2012-06-12T08:56:28Z	3323	CP.3032184712	25/5	991902-25/5	16.946064 49.492032	0106000020A210000001000000010300000001000
30	15031042	2011-10-27T12:17:19Z	255	CP.3032181712	23/37	991902-23/37	16.951917 49.480711	0106000020A210000001000000010300000001000
31	15031041	2012-07-03T12:56:30Z	50	CP.3032178712	23/35	991902-23/35	16.940674 49.482074	0106000020A210000001000000010300000001000
32	15031040	2012-07-03T12:56:30Z	49	CP.3032177712	23/34	991902-23/34	16.940641 49.482072	0106000020A210000001000000010300000001000

Obrázek 2 Struktura testovaných dat v prostředí PostgreSQL

Popis důležitých sloupců:

- gid jednoznačný identifikátor (primární klíč)
- referencp souřadnice rohů jednotlivých polygonů reprezentující parcely
- The_geom geometrie zapsaná ve formátu pro prostorové dotazy



Scale = 1 : 10K

14.06622, 49.96444

Obrázek 3 Ukázka zobrazení dat v prostředí OpenLayers

7 REALIZACE

7.1 Použitý hardware

Veškeré testování probíhalo a bylo zpracováváno vždy na jednom počítači. Prakticky to tedy znamená, že roli serveru i klienta zastával tentýž stroj. Vzhledem k velmi nízkým výkonovým požadavkům ze strany WMS Testeru, který prakticky jen vysílá požadavky směrem k serveru, není toto řešení nikterak omezující a na výsledky nemá vliv.

Na počítači byl nainstalován operační systém Windows 7, jenž byl získán se školní licenci z programu MSDN. Aby byla tato práce nápomocná při testování pro co největší počet uživatelů, byla přidána další složka: „instalace a realizace v operačním systému Linux“, která se liší zejména v předinstalační fázi. Při krocích, kde se postup neliší, nebo je velmi podobný se instalace nerozděluje.

Parametry použitého PC

Operační systém	Windows 7, 64bit / Oracle Linux Server 6, 64bit
Procesor	Intel(R) Core™2 Duo CPU E450 @ 2.2GHz
Operační paměť	4 GB RAM (3GB dostupných)
Pevný disk	1 x HDD 100 GB, SATA, NTFS

7.2 Předinstalační příprava

Pro bezproblémový průběh těchto testů je doporučena čistá instalace jednoho z operačních systému, pro které je následující postup popsán.

Tedy konkrétně **Windows 7 64bit SP1**, nebo **Oracle Linux 6**.

Oracle Linux 6 byl vybrán z důvod nejvyšší kompatibility ze strany Oracle, co se podporovaných OS týče.

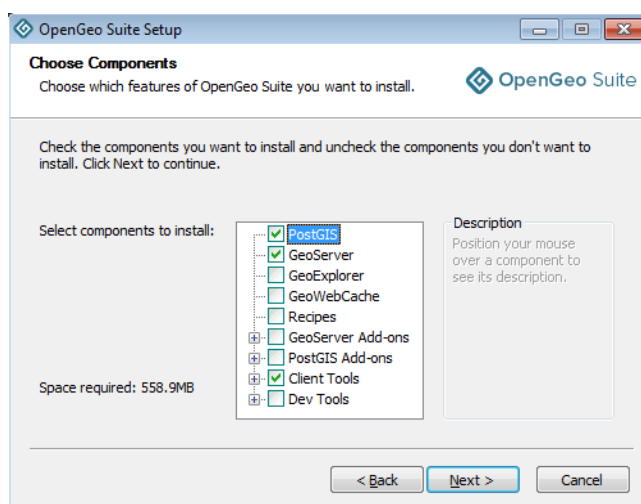
7.2.1 Příprava pro Windows

V případě nainstalovaného a připraveného systému, je možné na něj nainstalovat kompletní balíček OpenGeo Suite, který nám značně usnadní práci a pomůže se vyhnout mnoha chybám, které by mohli nastat.

Pro instalaci tohoto balíčku je ale nejprve nutné nainstalovat .Net framework 4, nebo novější. Pro tuto instalaci doporučuji udělat pomocí webové instalační služby, která vše dokončí automaticky, stačí ji jen spustit z oficiálních stránek Microsoftu.(Odkaz č. 1)

Dále už nám nic nebrání instalaci samotného balíčku OpenGeo suite, který je možné po rychlé registraci stáhnout ze Boundless (Odkaz č. 2). OpenGeo suite je také součástí elektronické přílohy.

Samotná instalace je dalším z nikterak složitých kroků, protože se zde setkáváme s klasickým instalátorem, který nás celou instalací provede a nastaví tak, aby plně vyhovovala potřebám této práce. Jediným bodem instalace, kterému je nutné věnovat větší pozornost je volba jednotlivých komponent balíčku, které se mají instalovat. S tímto pomůže následující obrázek:



Obrázek 4 Volba komponent balíčku OpenGeo suite

Tento balíček tedy nainstaluje kompletní SŘBD PostgreSQL i s prostorovou nadstavbou PostGIS.

Pro další postup budou využívány také komponenty „clinet tools“ tedy klientské nástroje na správu databáze. Budou dále využívány pro export dat a kontrolu struktury a funkčnosti.

Klientské nástroje obsahují pgAdmin a pgShapefileLoader.

Více o těchto pomocných nástrojích v samostatné podkapitole 7.3.4

Posledním krokem je instalace podpory pro Java aplikace, Java SE Runtime Environment. Aktuální verzi je možné stáhnout a nainstalovat přímo ze stránek výrobce. (Odkaz č. 3)

Seznam odkazů použitých v kapitole

(1) <http://www.microsoft.com/cs-cz/download/details.aspx?id=17851>

(2) <http://boundlessgeo.com/solutions/opengeo-suite/download/windows-download>

(3) <http://www.oracle.com/technetwork/java/javase/downloads/index.html>

7.2.2 Příprava pro Linux

U linuxové distribuce je potřeba připravit na instalaci Oracle, také instalaci balíčku OpenGeo z výše zmíněných stránek Boundless a instalaci programu Java, další přípravy nejsou nutné, tyto jsou však pro další bezproblémový postup nezbytné.

Jedná se o několik jednoduchých příkazů v Terminálu a editací konfiguračních souborů. Následující řádky tyto příkazy obsahují, společně z vysvětlením co příkazy reprezentují.

1. Stáhnutí a nastavení konfiguračních yum souborů, sloužících pro další nastavování a instalace

```
# cd /etc/yum.repos.d
# wget http://public-yum.oracle.com/public-yum-ol6.repo
```

2. Automatická instalace všech potřebných balíčků a jejich následná aktualizace se spustí příkazem

```
# yum install oracle-validated
```

3. V souboru "/etc/selinux/config" změňte hodnotu u položky SELINUX

```
SELINUX=permissive
```

4. Další příkaz vytvoří adresáře, do kterých se bude program instalovat a přidělí uživateli potřebná práva.

```
mkdir -p /u01/app/oracle/product/11.2.0/db_1
chmod -R 775 /u01
```

5. Tento příkaz přidá počítač do seznamu povolených serverů

```
xhost +<machine-name>
```

6. Poslední úpravou bude přidání následujícího do souboru ".bash_profile"

```
# Oracle Settings
TMP=/tmp; export TMP
TMPDIR=$TMP; export TMPDIR
ORACLE_HOSTNAME=ol5-112.localdomain; export ORACLE_HOSTNAME
ORACLE_UNQNAME=DB11G; export ORACLE_UNQNAME
ORACLE_BASE=/u01/app/oracle; export ORACLE_BASE
ORACLE_HOME=$ORACLE_BASE/product/11.2.0/db_1; export ORACLE_HOME
ORACLE_SID=DB11G; export ORACLE_SID
PATH=/usr/sbin:$PATH; export PATH
PATH=$ORACLE_HOME/bin:$PATH; export PATH
LD_LIBRARY_PATH=$ORACLE_HOME/lib:/lib:/usr/lib; export LD_LIBRARY_PATH
CLASSPATH=$ORACLE_HOME/jlib:$ORACLE_HOME/rdbms/jlib; export CLASSPATH
```

7. Následuje už jen instalace Java, bez čehož bychom nemohli spouštět Java programy

```
$ su -c "yum install java-1.7.0-openjdk"
```

8. Přidáme adresář, ze kterého se bude instalovat balíček OpenGeo suite

```
wget http://yum.opengeo.org/suite/v3/rhel/6/x86_64/OpenGeo.repo
```

9. Instalace OpenGeo suite

```
yum install opengeo-suite
```

7.3 PostgreSQL

7.3.1 Instalace PostgreSQL

Tento krok je již zahrnut v instalaci balíčku OpenGeo Suite, je možné ho tedy přeskočit.

7.3.2 Import a selekce dat

Import dat pro PostgreSQL se nese v duchu příkazů v příkazové řádce nebo Terminálu a celý proces je více méně automatizovaný, jedinou podmínkou je správné sestavení příkazu (Příkaz č. 1). Import dat byl proveden s kompletní databází získanou od vedoucího práce a až následně byla provedena selekce na výše zmiňovanou „malou“ a „střední“ databázi. K tomuto účelu posloužil příkaz zadávaný již přímo do konzole „psql“, tedy do konzole PostgreSQL (Příkaz č. 2). Vzhledem k velikosti databáze tyto operace mohou trvat i několik minut. Pro přihlašování do této konzole slouží předdefinovaný administrátorský účet „postgres“ s heslem „postgres“. Tento účet je využíván v celé práci. Tento uživatel má přidělenou i volnou databázi, která se vytvoří rovněž při instalaci a je využívána v případě, pokud je prováděn import dat a žádná jiná databáze není specifikována.

Tímto byly vytvořeny další 2 tabulky, obsahující data těchto dalších 2 vrstev sloužících pro testování.

Seznam odkazů použitých v kapitole:

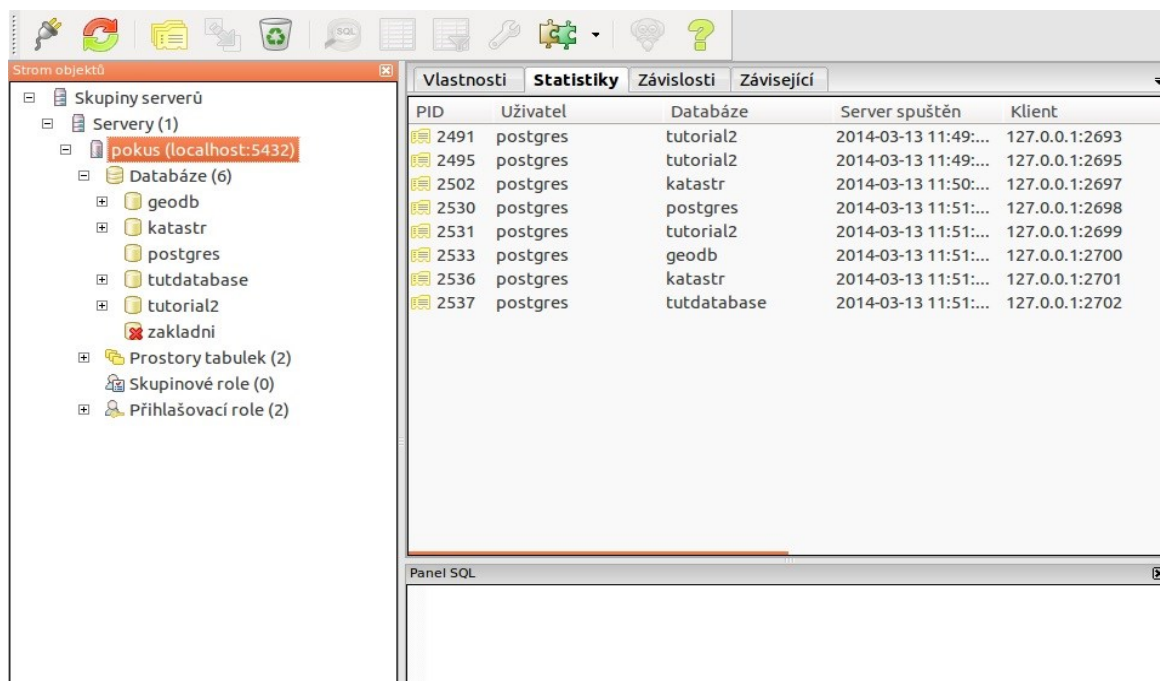
(1) > psql -f cuzk.sql

(2) psql> SELECT * INTO vzorek2 FROM public.cp WHERE cp.gid < 25000;

7.3.3 Pomocné nástroje

pgAdmin

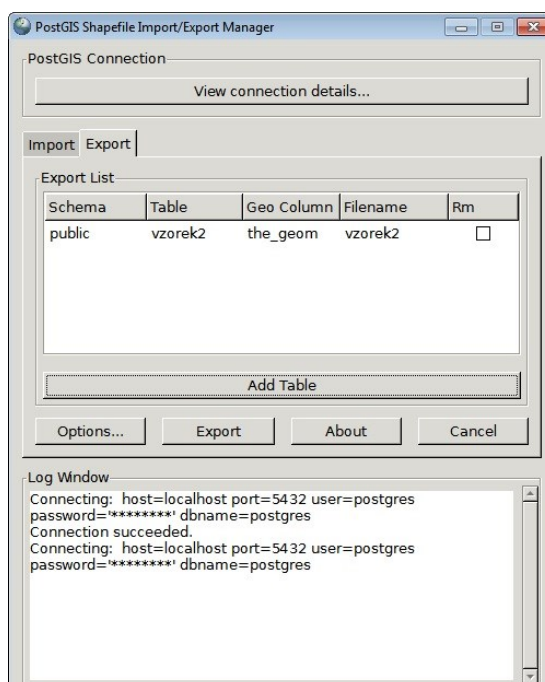
pgAdmin je grafické rozhraní pro SŘBD PostgreSQL, jeho hlavní výhodou je, že dokáže zastat ve většině případů konzoli „psql“ a je možné v něm vytvářet nové databáze, tabulky a dotazy. V této práci sloužil hlavně pro kontrolu a pro přehled struktury dat.



Obrázek 5 pgAdmin

pgShapefileLoader

Tento nástroj je velmi jednoduchým grafickým rozhráním pro nástroj „shp2psql“ který slouží pro převod souborů ve formátu .SQL, který se používá pro import dat do PostgreSQL. Tímto způsobem je možné rovnou importovat, ale také exportovat data z PostgreSQL.



Obrázek 6 pgShapefileLoader

7.3.4 Export dat pro další použití

Vzhledem k tomu, že data poskytnutá vedoucím práce, byla ve formátu .SQL, bylo tedy nutné tato data exportovat do formátu, který je možné dále použít. Pro tyto účely byl vybrán formát ESRI Shapefile pro jeho rozšířenost a širokou podporu. Pro export do tohoto formátu byl použit zmiňovaný pgShapefileLoader.

Na předchozím obrázku je možné vidět již připravený program, který je připojen na databázi a pro spuštění exportu stačí pouze použít tlačítko „export“.

7.3.5 Opětný import pro objektivní testování

Data exportovaná z PostgreSQL ve formátu Shapefile byla dále využita jako data pro samotné testování GeoServer bez použití SŘBD jako samotný datový sklad, tak i jako vstupní data do SŘBD Oracle. Proto v rámci zamezení možnosti ovlivnění výkonu a výsledků případnými skrytými chybami při exportu, byly získané soubory využity jako vstupní data i pro PostgreSQL a původní data byla nahrazena těmito novými.

7.4 Oracle

System řízení báze dat od společnosti Oracle se shodným názvem je distribuován na oficiálních stránkách jako free-to-use software, který je možné stáhnout a se základními funkcemi používat. Nutno podotknout, že základními funkcemi je myšleno opravdu široké portfolio možností a nastavení, takže i náročnější uživatel si může vystačit se základní licenci, v případě použití pro osobní potřebu.

7.4.1 Instalace

Pro systém Windows i Linux je připraven instalační balíček, který je rozdělen na 2 komprimované soubory a je možné je stáhnout přímo z oficiálních stránek společnosti Oracle. Soubory je možné stáhnout a využívat, jedinou podmínkou, kladenou ze strany Oracle je registrace členského bezplatného účtu.

Instalační soubory jsou připraveny pro celou řadu operačních systémů a také možnost výběru 32, nebo 64bitové varianty.

Oracle Database 11g Release 2

Standard Edition, Standard Edition One, and Enterprise Edition

7/13: Patch Set 11.2.0.4 for Linux and Solaris is now available on support.oracle.com. Note: it is a full installation (you do not need to download 11.2.0.1 first). See the README for more info (login to My Oracle Support required).

(11.2.0.2.0)

zLinux64 File 1, File 2 (2GB) See All

(11.2.0.1.0)

Microsoft Windows (32-bit) File 1, File 2 (2GB) See All

Microsoft Windows (x64) File 1, File 2 (2GB) See All

Linux x86 File 1, File 2 (2GB) See All

Linux x86-64 File 1, File 2 (2GB) See All

Solaris (SPARC) (64-bit) File 1, File 2 (2GB) See All

Solaris (x86-64) File 1, File 2 (2GB) See All

HP-UX Itanium File 1, File 2 (2GB) See All

HP-UX PA-RISC (64-bit) File 1, File 2 (2GB) See All

AIX (PPC64) File 1, File 2 (2GB) See All

Obrázek 7 Stahovací obrazovka z oficiálních stránek Oracle

7.4.2 Instalace nadstavby Spatial

Spatial nadstavba je již v instalaci integrována a je nutné ji pouze aktivovat příkazem z příkazové řádky, nebo Terminálu.

Po spuštění příkazové řádky následuje přihlášení do Oracle příkazem „sqlplus sys as sysdba“.

Po přihlášení je již možné zadávat systémové příkazy a aktivovat spatial nadstavbu příkazem.

Windows: @%ORACLE_HOME%\md\admin\mddins.sql

Linux: @\$ORACLE_HOME/md/admin/mddins.sql

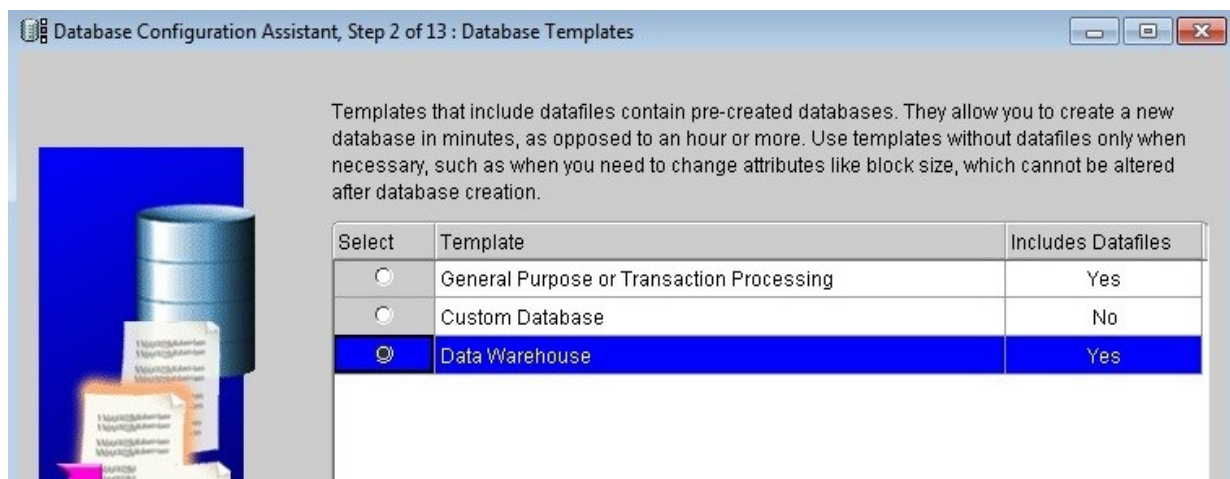
Aktivace nadstavby proběhla v pořádku a nyní již můžeme přistoupit ke tvorbě databáze a importu dat.

7.4.3 Vytvoření databáze

Vytvoření databáze je nutné pro import a správu dat. K vytvoření databáze, která bude schopna korektně zobrazit prostorová data, můžeme využít vestavěný grafický nástroj „Database configuration assistant“, který se nainstaloval společně se samotným systémem řízení báze dat.

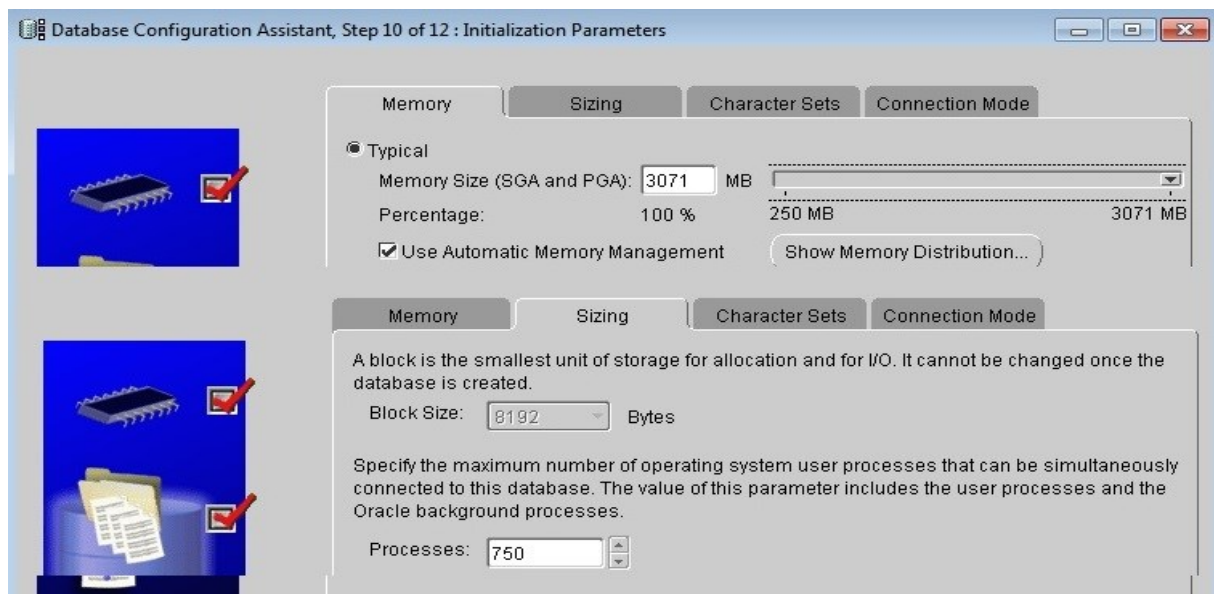
Proces vytváření databáze v Database configuration assistant je opatřen grafickým rozhraním, a tak není vytvoření první databáze nic složitého. Každopádně určité nastavení, především z důvodu maximalizace výkonu při testování, je dobré udělat. Následující obrázky by s tímto nastavením měly pomoci a vysvětlit také z jakého důvodu byla tato nastavení provedena.

Nastavení typu databáze je důležitým krokem. Pro velkou databázi, jejíž účelem je pouze uchovávat data, ze kterých bude uživatel hlavně číst je na doporučení Oracle zvolen typ „Data Warehouse“ tedy „datový sklad“.



Obrázek 8 Volba typu datového skladu Oracle

Následně se dostaneme až téměř na samotný závěr tvorby databáze a máme možnost upravit parametry, které ovlivní výkon testování. A to maximální použitou paměť RAM a počet procesů, které mohou najednou běžet v databázi. Paměť RAM byla nastavena na maximum a počet procesů na 750, tedy tak, aby i při testování počet spuštěných procesů nepředstavoval omezení, což by způsobilo umělé zkreslení a zhoršení výsledků testů.



Obrázek 9 Nastavení tvorby databáze Oracle

Po vytvoření databáze, následuje již poslední krok při přípravě tohoto SRĚBD, tedy import dat.

7.4.4 Import dat

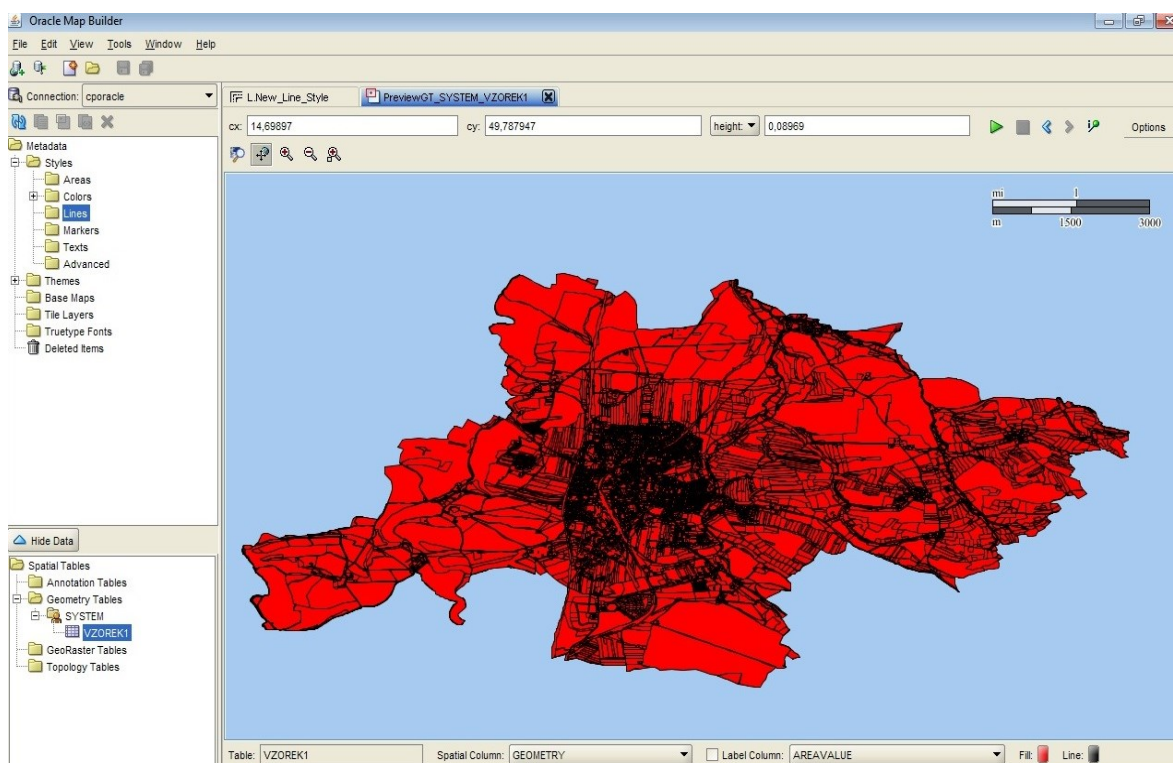
Import dat je možné provést několika způsoby, vzhledem k účelům, pro které by měla tato práce sloužit, zde byla snaha o nalezení co nejjednoduššího způsobu.

Nakonec bylo zvoleno řešení, které v sobě zahrnuje využití podpůrného nástroje „MapBuilder“ který byl vytvořen přímo k těmto účelům samotnými výrobci tohoto SŘBD, tedy společností Oracle.

MapBuilder má jednoduché grafické rozhraní a vzhledem k tomu, že je napsán v jazyce Java, tak pro jeho spuštění je jediný předpoklad, tedy instalace SE JDK JAVA. Poté již je možné spustit jej na jakékoli platformě.

Spuštění probíhá na všech operačních systémech totožně a to zadáním příkazu

„java -jar mapbuilder.jar“



Obrázek 10 MapBuilder

Spojení s databází navážeme pomocí tlačítka 

Dále můžeme vyplnit potřebné údaje pro vytvoření spojení.

Posledním krokem je import souborů shapefile, jejichž vytvořením se zabývá kapitola 7.3.5.

Ten provedeme přes nabídku tools a zvolíme „import shapefile“.

Importem dat končí přímá práce s tímto SŘBD a je možné jej propojit s GeoServerem a zahájit testování.

7.5 GeoServer

7.5.1 Instalace GeoServer

GeoServer je součástí balíčku OpenGeo Suite, není tedy nutné ho instalovat separátně a tento krok, je možné podobně jako u PostgreSQL a PostGIS přeskočit a rovnou zadat domovskou adresu serveru, protože server se spouští automaticky se startem PC.

„<http://localhost:8080/geoserver/web>“

7.5.2 Doporučená specifikace

Vzhledem ke skutečnosti, že GeoServer je napsaný kompletně v Java, tak je jeho použití velice všestranné a nesetkává se s prakticky žádným omezením.

Výrobce neudává žádné minimální konfigurace a při testování bylo dokázáno, že GeoServer běží i na starších méně výkonných strojích.

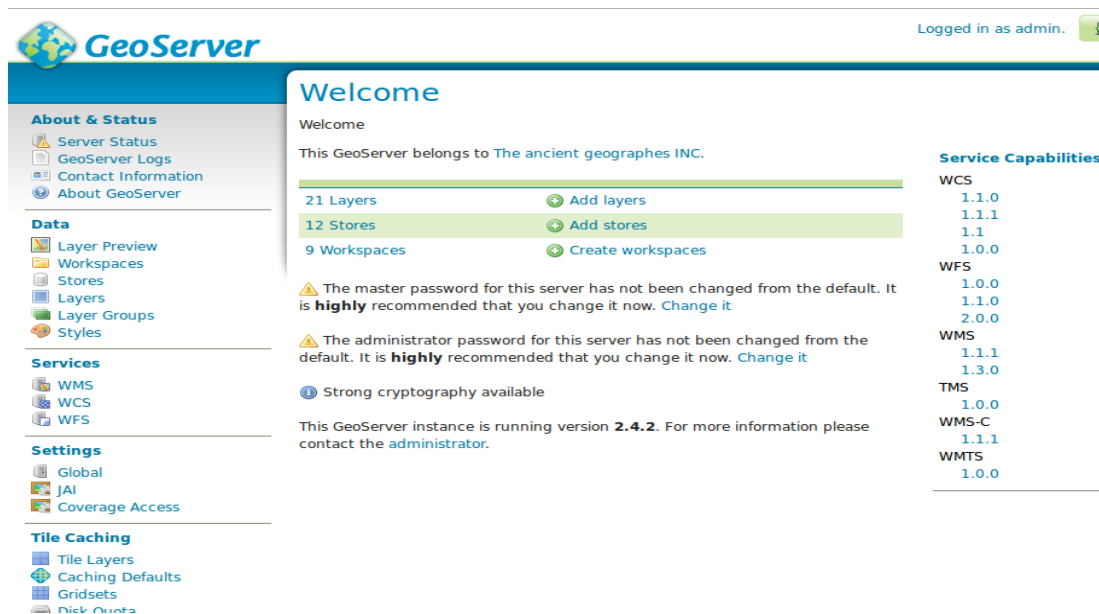
Pro jeho správu prakticky postačuje jakýkoli operační systém a podpora Java, konkrétně Java verze 6 a vyšší na daném stroji, na kterém chceme GeoServer uvést do chodu.

Nastavování v aplikaci již není nutné, vývojáři provedli veškerá standartní nastavení již při tvorbě a proto je možné ihned po spuštění započít s prací a dalšími kroky.

7.5.3 Nastavení GeoServer

Po zadání výše zmíněné adresy GeoServeru se uživateli zobrazí již grafická podoba GeoServeru a je možné ho dále nastavovat a upravovat.

Předdefinované přihlašovací údaje pro administrátora jsou „Admin“ a heslo „geoserver“.



Obrázek 11 GeoServer

Po přihlášení je již možné provádět veškeré modifikační zásahy podle potřeb.

Prvním krokem bylo vytvoření nového pracovního prostoru „workspace“, který slouží pro sdružení vrstev, které mají respektovat jednotné nastavení, které je možné pro daný „workspace“ definovat. Do tohoto pracovním prostoru budou později připojeny databáze, resp. vrstvy vytvořené z tabulek databází v jednotlivých SRBD.



Obrázek 12 Přidávání nového pracovního prostoru

7.5.4 PŘIPOJENÍ DATABÁZÍ

Dalším nutným krokem je import samotných databází, tedy spojení se SRBD.

V GeoServer je pro tento případ tato funkce přednastavena a k jejímu požití slouží záložka „Stores“. V nabídce, které se pod touto položkou skrývá, najdeme seznam podporovaných databázových systémů a úložišť. Po vybrání odpovídajícího z nich následuje rychlá konfigurace, propojení s vybraným datovým skladem a výběr požadované vrstvy z načtené databáze.

New data source

Choose the type of data source you wish to configure

Vector Data Sources

- CSV - Comma delimited text file
- Directory of spatial files (shapefiles) - Takes a directory of shapefiles
- H2 - H2 Embedded Database
- H2 (JNDI) - H2 Embedded Database (JNDI)
- Microsoft SQL Server (JNDI) - Microsoft SQL Server (JNDI)
- Oracle NG - Oracle Database
- Oracle NG (JNDI) - Oracle Database (JNDI)
- Oracle NG (OCI) - Oracle Database (OCI)
- PostGIS - PostGIS Database**
- PostGIS (JNDI) - PostGIS Database (JNDI)
- Properties - Allows access to Java Property files containing FeatureInfo
- Shapefile - ESRI(tm) Shapefiles (*.shp)
- Web Feature Server - The WFSDataStore represents a connection to the server, and the ability to perform transactions on the server (when available)

Raster Data Sources

- ArcGrid - Arc Grid Coverage Format
- GeoTIFF - Tagged Image File Format with Geographic information
- Gtopo30 - Gtopo30 Coverage Format
- ImageMosaic - Image mosaicking plugin
- WorldImage - A raster file accompanied by a spatial data file

Other Data Sources

- WMS - Cascades a remote Web Map Service

Obrázek 13 Napojení datového skladu krok 1

Formulář na připojení databáze požaduje jen minimum informací a jeho vyplnění by nemělo dělat potíže. Host a port jsou vždy přednastaveny a uživateli stačí vypsát pouze název nově vytvářeného spojení, název databáze a své přihlašovací údaje k této databázi.

New Vector Data Source

Add a new vector data source

PostGIS
PostGIS Database

Basic Store Info

Workspace *
PostGIS

Data Source Name *

Description

Enabled

Connection Parameters

host *
localhost

port *
5432

database

schema
public

user *

passwd

Namespace *
PostGIS.cz

Expose primary keys

max connections
10

min connections
1

Obrázek 14 Napojení datového skladu krok 2

<< < | > >> Results 1 to 8 (out of 8 items)

Published	Layer name	Action
	cp	<input checked="" type="button" value="Publish"/>

Obrázek 15 Napojení datového skladu krok 3

Při připojování ShapeFile není logicky vyžadováno umístění serveru, port ani přihlašovací údaje, ale pouze cesta k adresáři, kde se daný ShapeFile nachází.

Po úspěšném připojení můžeme zkontrolovat přítomnost vrstvy v záložce vrstvy „Layers“ a následně i vizualizovat výsledek pomocí „Layer preview“, tedy náhledu vrstvy.

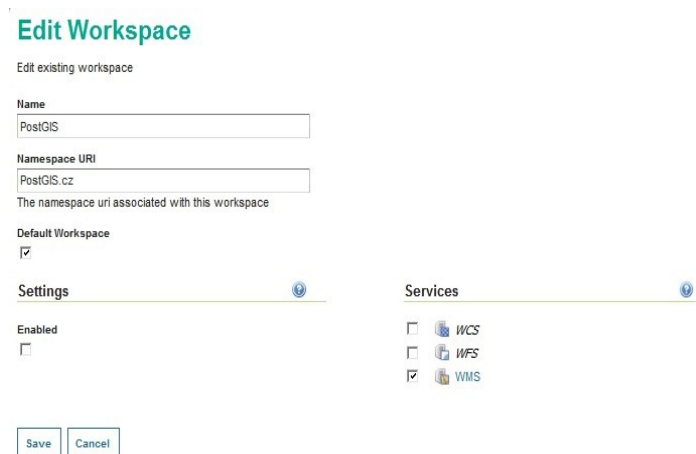
7.5.5 Nastavení WMS

Vzhledem k tomu, že WMS je standard a jeho nastavení je již předem definované, odpadá nám jakékoli složité nastavování.

Veškeré potřebné nastavení jsou již provedena v základním kódu programu a není nutné další nastavování, veškeré nahrané vrstvy jsou zároveň dostupné pro WMS

Uživateli může, volitelně označit

„aktivaci WMS“, tak jak ukazuje obrázek č. 15 a atributy (např.: informace o majiteli, organizace, další informace) WMS také libovolně nastavit.



Obrázek 16 Aktivace WMS

Tato „aktivace WMS“ se provádí pomocí nastavení „workspace“ o kterém bylo psáno dříve. Dále jsou již vrstvy dostupné klasickou cestou a je možné je dále využívat a vše je již připraveno na samotné testování.

7.6 Publikování dat pomocí WMS

Po nastavení WMS přímo v GeoServeru, již je možné tuto službu libovolně využívat.

Spuštění a korektní nastavení služby můžeme otestovat základním požadavkem na server.

„*http://localhost:8080/geoserver/wms?version=1.1.1&request=GetCapabilities*“

„localhost“ jako název domény byl v tomto případě použit proto, že server i klient se nachází na stejném PC a je mu přiřazen port 8080, jde tedy o lokální instalaci. V případě instalace síťové je nutné použít adresu, která je registrována a dostupná, nejčastěji ve formě IP adresy.

V tomto případě by obecný předpis zněl takto:

„*http://<IP-adresa-serveru>/geoserver/wms?version=1.1.1&request=GetCapabilities*“

Výpis GetCapabilities ve formátu XML je očekávaný výsledek. Pokud je odpověď v tomto formátu a samotný soubor obsahuje výpis vrstev, služeb a podporovaných souřadnicových systémů, můžeme prohlásit WMS za připravenou.

Soubor s výpisem GetCapabilities je součástí elektronické přílohy.

7.7 WMS Tester - klient

WMS Tester, program popsany již v kapitole 5, poslouží pro testování rychlosti GeoServer. První testy se odehrávali zcela podle pokynů skromného návodu přiloženém k programu při stahování. Tento návod, společně s celou aplikací, je také umístěn v elektronické příloze k této práci.

Pro první zkušební spuštění byl zadán z příkazové řádky operačního systému příkaz:

```
„java -jar wms tester.jar -e http://localhost:8080/geoserver/wms“
```

Tého příkaz, bez specifikace konkrétní vrstvy, je kontrolním příkazem funkčnosti programu a odezvy serveru. V tomto případě program vypíše seznam získaných z *GetCapabilities* a upozornění „*Layer not specified*“.

Dalším krokem je spustit první jednoduchý test:

```
„java -jar wms tester.jar -e http://localhost:8080/geoserver/wms -l  
PostGIS:Vzorek2“
```

Test takto spuštěný, bez bližší specifikace atributů, se chová podle přednastavených požadavků. Vytvoří se jeden klient, který vysílá požadavky na server, dokud nezíská 10 náhodných rastrů ze zvolené vrstvy. V příkazové řádce/Terminálu, můžeme pozorovat průběh testu a jeho zakončení hláškou „*Background thread finish*“.

Po dokončení testu jsou již k dispozici 4 již zmiňované soubory, jejichž strukturu a vysvětlení najdete v kapitole 5.2 a příloze č. 1.

Pro optimální výsledky je ale možné test různě konkrétněji specifikovat.

Při testování bylo využito vícero nastavení testů a ke každé testované situaci je přesně zadání testu uvedeno v následující kapitole a příloze č. 2, která se samotným testováním a především zhodnocením výsledků zabývá.

Příklad takového kompletně sestaveného testu si však můžeme ukázat již nyní.

```
„java -jar wms tester.jar -e http://localhost:8080/geoserver/wms -l “vzorek2“  
-d 0 -t 50 -i 10 -tt 600 -n 100“
```

Takto specifikovaný test vytvoří 50 klientů-vláken, (-t 50) která budou odesílat požadavky na server bez časové prodlevy mezi vyřízením dotazu a opětovným zadáním (-d 0) dokud

každý z nich nezíská 100 rastrů (-n 100), nebo dokud nevyprší maximální čas určený k testování, tedy 600 sekund (-tt 600).

Mezi zapojením jednotlivých klientů do testu je vždy 10sekund prodleva (-i 10).

Jak test probíhá, každých 10 sekund přibývá další klient, který sebou nese další požadavky na server a tím postupně navyšuje celkový čas odpovědi serveru, z důvodů vyšší zatíženosti serveru a datového skaldu.

8 VÝSLEDKY A JEJICH INTERPRETACE

8.1 Úvod k testování

Testování a celá práce byla zprvu vypracovávána na Linuxové distribuci, která byla spuštěna přes Virtual Box, nástroji na tvorbu virtuálních PC. Toto řešení se ale ukázalo jako nevhodné, protože i když bylo možné data publikovat, výsledky testů byly nepoužitelné. Kvůli nízkého výkonu aplikace soustavně kolabovaly a bylo nutné celou práci zopakovat na výkonnějším stroji.

Z tohoto důvodu byl použit stroj, jehož specifikace je možné najít v předchozí kapitole. Pro maximalizaci využitelnosti této práce jako případný zdroj informací při testování ve vlastních podmínkách, nebo při řešení obdobného problému, byl na tomto PC zvolen operační systém Windows.

Pro ověření funkcionality a kvality zápisu postupu, byla celá práce realizována kompletně znovu. Navíc by tak bylo možno odhalit případné chyby. Vzhledem k tomu, že žádné chyby objeveny nebyly a prakticky veškeré postupy zde uvedené jsou znalosti získané z oficiálních dokumentací, je tedy možné tvrdit, že výsledky nejsou ovlivněny chybnou instalací, nebo špatným postupem.

8.2 Sledování práce systému

Testy probíhaly postupně, finální testování proběhlo 3.4.2014

Výkon systému byl během testu monitorován a zapsán. K tomuto účelu posloužil vestavěný nástroj systému Windows „Monitorování zdrojů“.

		Oracle	PostgreSQL	ShapeFile
Server	Využití CPU	89%	90%	64%
	Využití RAM	2286 MB	2427 MB	2611 MB
	Vytížení CPU Databází	18%	13%	0%

Tabulka 5 Zatížení serveru při testování malé databáze

		Oracle	PostgreSQL	ShapeFile
Server	Využití CPU	93%	95%	75%
	Využití RAM	2425 MB	2623 MB	2733 MB
	Vytížení CPU Databází	43%	30%	0%

Tabulka 6 Zatížení serveru při testování střední databáze

		Oracle	PostgreSQL	ShapeFile
Server	Využití CPU	95%	98%	82%
	Využití RAM	2535 MB	2678 MB	2702 MB
	Vytížení CPU Databází	47%	36%	0%

Tabulka 7 Zatížení serveru při testování kompletní databáze

8.3 Průběh testování

Celkově bylo provedeno více než 60 testů aby bylo možné určit parametry testů s výsledky, které nejvíce vypovídají o chování GeoServeru ve spojení s jednotlivými datovými sklady. Nakonec byly zvoleny parametry individuálně pro každou testovanou variantu, tak, aby bylo možné posoudit výsledky mezi sebou a výsledky objektivně vyhodnotit.

Zpracování těchto dat probíhalo pomocí programu Microsoft Excel z Balíku Microsoft Office.

Data byla zpracována do tabulek a grafů.

Tabulky zobrazují data rozdělená do tříd a průměrů za celý test. Tabulky byly vytvořeny pro větší přehlednost výsledku a zahrnují zajímavá data z průběhu testu. Grafy již zobrazují průběh testu jako takového a zejména slouží jako grafický pomocník pro odhad regresní závislosti mezi počtem klientů a odezvou GeoServeru. Zobrazena jsou pouze data relevantní pro posouzení testu, kompletní sady souborů s výsledky jsou součástí elektronické přílohy.

Přesné parametry výsledných „ostrých“ testů jsou zaznamenány v příloze číslo 2.

8.4 Malá databáze

Všechny 3 datové sklady byly nejprve otestovány s nejmenší sadou dat.

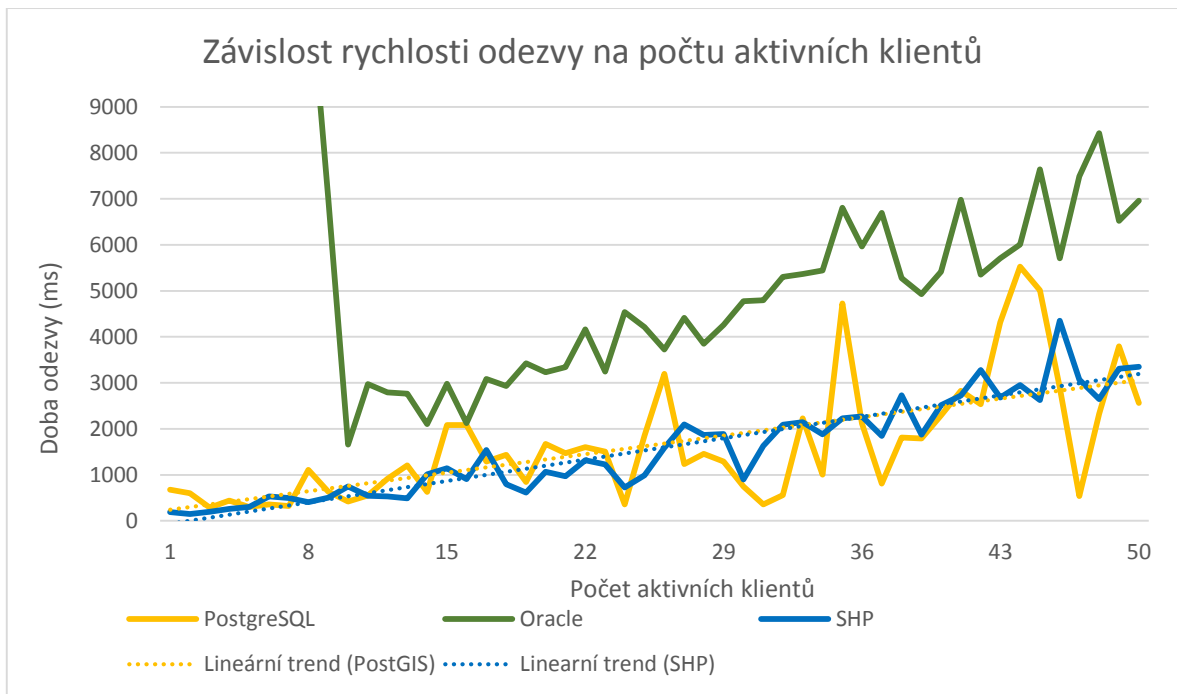
Výsledky uvedené níže byly zapsány formou 2 tabulek a stejného počtu grafů. Tabulky nabízí pohled na nastavení a výsledky jednotlivých testů (tabulka č 8), které byly navíc rozděleny do tříd (tabulka č. 9), aby lépe vynikla závislost počtu aktivních klientů na rychlosti s jakou GeoServer poskytoval odpovědi na vyslané požadavky. Grafy výsledky přehledně zobrazují v grafické podobě jako průměrné časy odpovědi při zapojení určitého množství klientů (graf č. 1) a také ukazují stabilitu a průměrnou rychlost odpovědi při plném zatížení (graf č. 2). Při tomto testování Oracle test nedokončil, z důvodu přílišného zatížení přestal reagovat a test tak byl ukončen při plném zatížení při 50 klientech.

Nastavení testu	PostgreSQL	Oracle	ShapeFile
Počet klientů	50	50	50
Počet dotazů na klienta	100	100	100
Rozestup aktivace klientů	1	5	1
Hodnoty testu	PostgreSQL	Oracle	ShapeFile
Celkový čas testu (s)	571	304	356
Zodpovězeno dotazů	5000	1611	5000
Maximální doba vyřízení	18,58	44,960	16,789
Minimální doba vyřízení	0,1	0,109	0,078
Modus vyřízení dotazu	0,187	4,836	2,137
Průměrný počet dotazů/s	8,757	5,516	14,045
Průměrné vyřízení dotazu	2,365	5,395	2,754

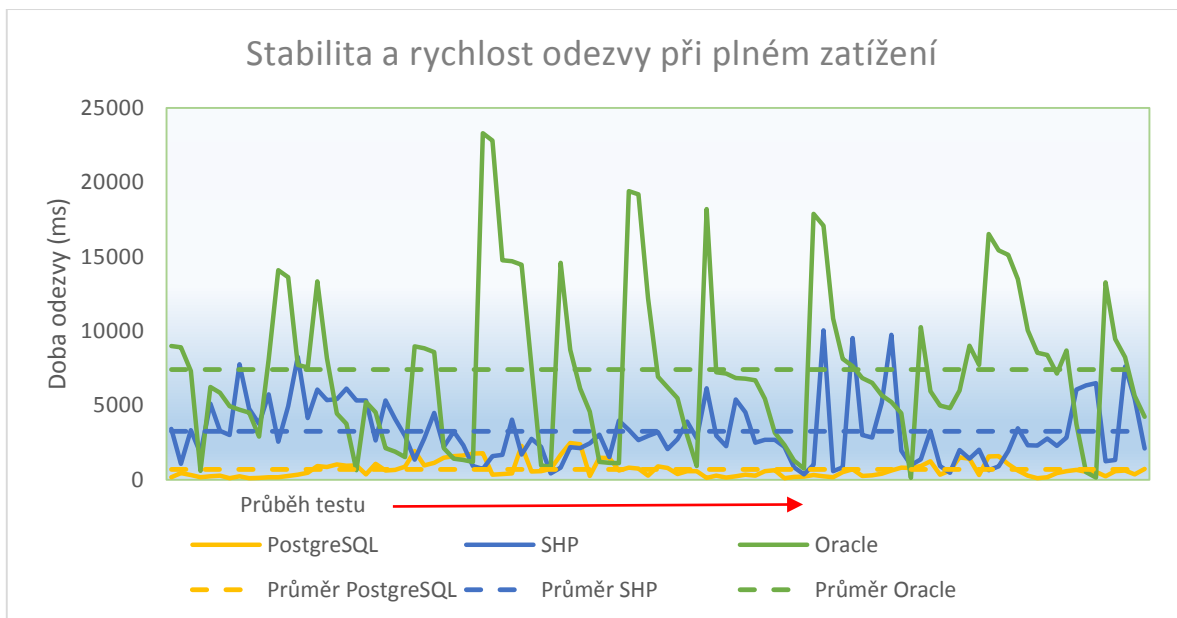
Tabulka 8 Nastavení a výsledky jednotlivých testů Malá databáze

Klientů	Závislost rychlosti vyřízení dotazu na počtu dotazů za sekundu					
	PostgreSQL		ShapeFile		Oracle	
	počet dotazů	Průměrné vyřízení (s)	počet dotazů	Průměrné vyřízení (s)	počet dotazů	Průměrné vyřízení (s)
1-4	5	0,940	61	0,208	4	39,113
5-14	146	0,728	231	0,588	127	3,068
15-24	152	1,510	281	1,015	29	3,299
25-34	314	1,400	442	1,784	334	4,607
35-44	260	2,560	924	2,633	336	5,928
45-49	341	1,933	974	3,207	174	7,102
50	3718	2,564	2087	3,350	338	7,089

Tabulka 9 Výsledky testů rozdělených do tříd Malá databáze



Graf 1 Závislost rychlosti odezvy na počtu aktivních klientů Malá databáze



Graf 2 Stabilita a rychlost odezvy při plném zatížení Malá databáze

8.5 Střední databáze

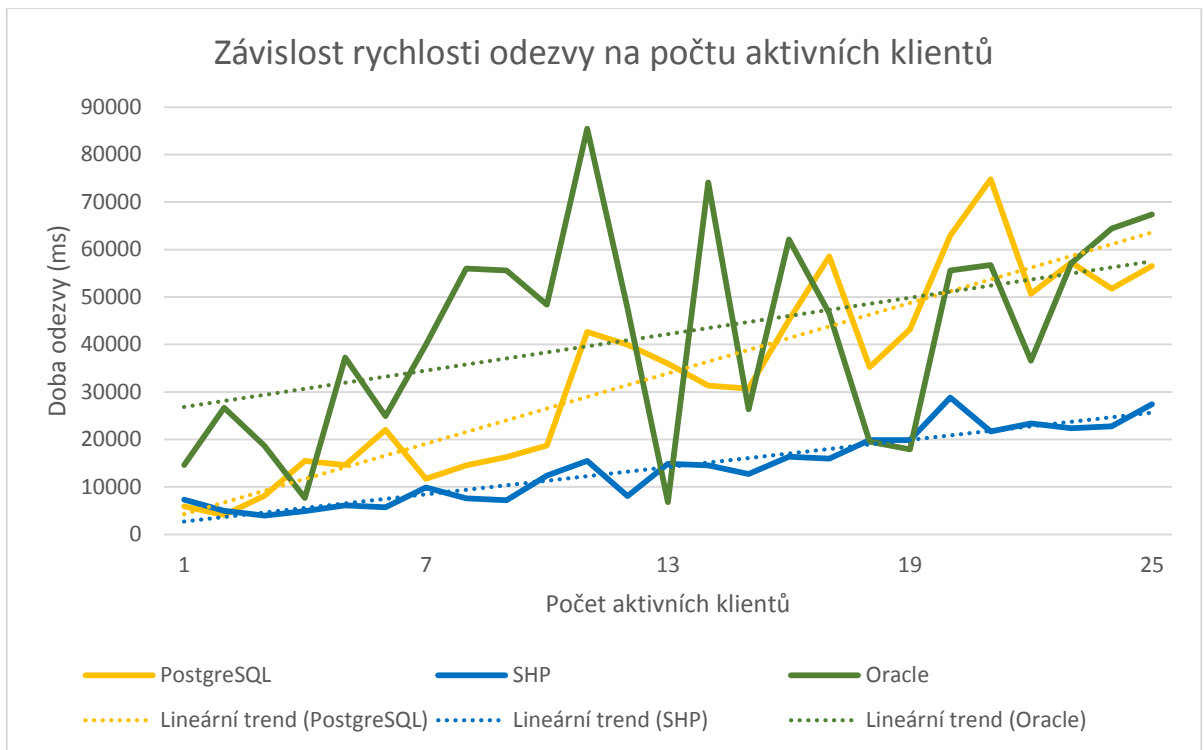
Testování této sady dat je interpretováno totožným způsobem, jako testování v podkapitole předešlé. Za zmínku však stojí, že SŘBD Oracle při testech nezvládnul zapojení všech klientů při delším testování a již po několika zaslaných dotazech při plném zatížení opět přestal odpovídat a test tak byl automaticky ukončen. Výsledky celou situaci objasňují.

Nastavení testu	PostgreSQL	Oracle	ShapeFile
Počet klientů	25	25	25
Počet dotazů na klienta	50	50	50
Rozestup aktivace klientů	10	10	10
Hodnoty testu	PostgreSQL	Oracle	ShapeFile
Celkový čas testu (s)	2883	472	603
Zodpovězeno dotazů	1250	75	1250
Maximální doba vyřízení	176,701	376,601	91,026
Minimální doba vyřízení	0,842	0,124	0,078
Modus vyřízení dotazu	3,774	17,94	0,437
Průměrný počet dotazů/s	0,434	0,159	2,073
Průměrné vyřízení dotazu	51,190	90,489	23,018

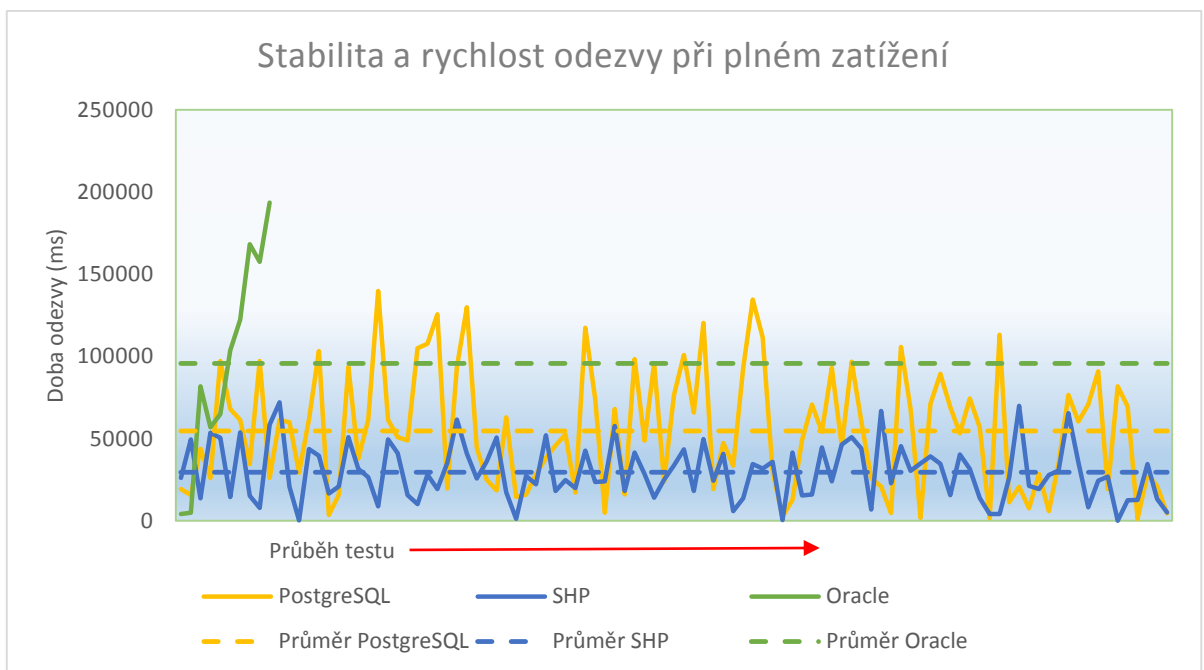
Tabulka 10 Nastavení a výsledky jednotlivých testů Střední databáze

Klientů	Závislost rychlosti vyřízení dotazu na počtu dotazů za sekundu					
	PostgreSQL		ShapeFile		Oracle	
	počet dotazů	Průměrné vyřízení (s)	počet dotazů	Průměrné vyřízení (s)	počet dotazů	Průměrné vyřízení (s)
1-4	27	8,398	22	6,894	10	18,468
5-9	52	15,123	66	7,488	14	62,611
10-14	31	31,774	102	12,722	14	126,627
15-20	125	41,449	192	17,968	16	120,744
21-24	160	53,273	79	22,942	10	97,273
25	855	56,529	786	27,379	10	95,753

Tabulka 11 Výsledky testů rozdělených do tříd Střední databáze



Graf 3 Závislost rychlosti odezvy na počtu aktivních klientů Střední databáze



Graf 4 Stabilita a rychlost odezvy při plném zatížení Střední databáze

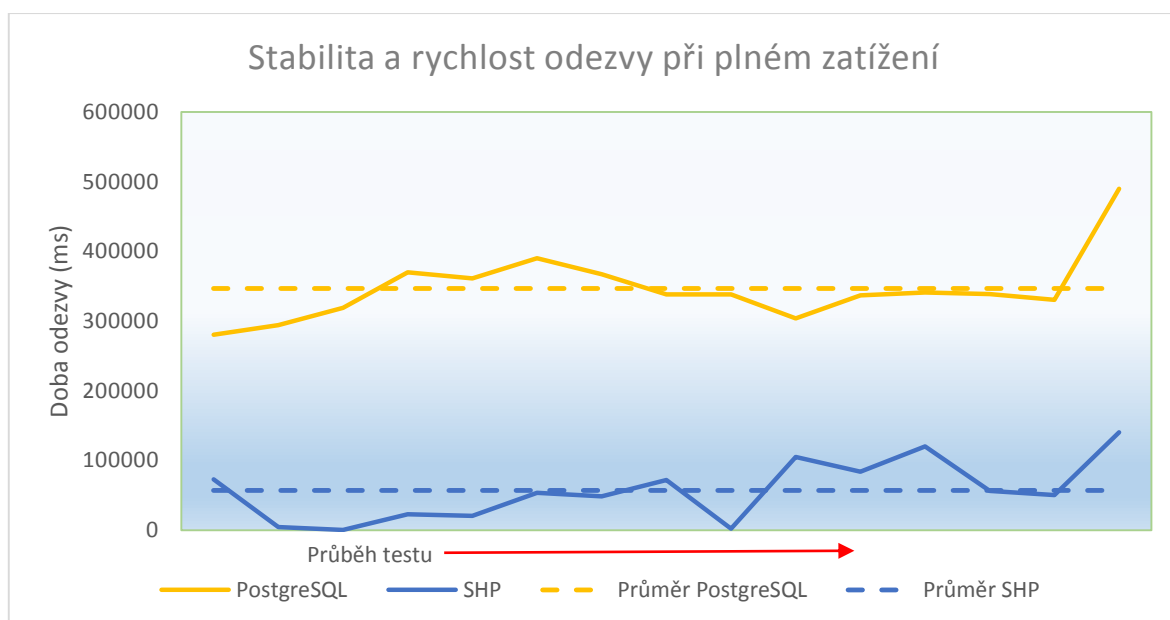
8.6 Kompletní databáze

Testování kompletní databáze bylo už od počátku problematické. A to ze zřejmých důvodů, tedy z důvodu nedostatečné výpočetní kapacity stroje, na kterém bylo testování prováděno. Z tohoto důvodu se podařilo kompletně dokončit test pouze pro ShapeFile.

PostgreSQL z 25 dotazů stihl v 9 minutovém limitu, určeném pro jeden dotaz, zpracovat 16. Oracle pouze první z 6 pokusů a na další rezignoval, přestal odpovídat a test byl stejně jako v předchozích případech předčasně ukončen. Z těchto důvodů je také prezentace výsledků tohoto testování o něco chudší, poněvadž není tolik dat, které by bylo zajímavé prezentovat.

Nastavení testu	PostgreSQL	Oracle	ShapeFile
Počet klientů	5	5	5
Počet dotazů na klienta	5	5	5
Rozestup aktivace klientů	10	10	10
Hodnoty testu	PostgreSQL	Oracle	ShapeFile
Celkový čas testu (s)	2258	822	310
Zodpovězeno dotazů	16	1	25
Maximální doba vyřízení	511,181	265,996	140,525
Minimální doba vyřízení	280,379	265,996	0,218
Průměrné vyřízení dotazu	356,907	265,996	51,350
Průměrné vyřízení při plném zatížení	346,622	----	53,360

Tabulka 12 Nastavení a výsledky jednotlivých testů Kompletní databáze



Graf 5 Stabilita a rychlost odezvy při plném zatížení Kompletní databáze

8.7 Porovnání testovaných SŘBD

Při porovnávání jednotlivých testů, je možné vidět, že PostgreSQL i ShapeFile si vedli poměrně dobře a jejich výkon u menších dat je porovnatelný. Zatímco řešení od společnosti Oracle za svými konkurenty zaostává, což je nejvíce zřejmé na výsledcích testů, které se ani nepodařilo dokončit.

Ačkoli v testech s malým objemem dat byly výsledky pro PostgreSQL a ShapeFile velmi podobné, se zvyšujícím se objemem testovaných dat se zvyšoval i výkonnostní rozdíl, který byl nakonec již velmi značný. Zejména na největším testovaném vzorku dat kde byl rozdíl markantní.

Oracle při posledním testu dopadl nejhůře, kdy dokázal správně dokončit v limitu 9 minut pouze 1 dotaz z 6, když dalších 19 úplně vypustil a na další práci rezignoval.

Nejlepších hodnot při testování tedy dosáhnul ShapeFile, datový sklad, který měl původně sloužit pouze pro porovnání výkonu testovaných SŘBD. Druhé místo i přes částečný neúspěch při posledním testování obsadil PostgreSQL, který se osvědčil zejména při testování méně objemných dat, kde dokázal ShapeFile dokonce i mírně překonat.

Nejhůře dopadl Oracle, který se se svou nadstavbou Spatial nedokázal vypořádat s nižším výkonem PC a nedokončil tak své testy.

Výsledky pomůže podtrhnout finální porovnávací tabulka.

		ShapeFile	PostgreSQL	Oracle
Mala data	Průměrná rychlost odezvy při plném zatížení (s)	3,350	2,564	7,089
	Délka trvání testu (s)	356	571	304 N
Střední data	Průměrná rychlost odezvy při plném zatížení (s)	27,379	56,529	95,753
	Délka trvání testu (s)	603	2883	472 N
Kompletní data	Průměrná rychlost odezvy při plném zatížení (s)	53,360	346,622	----
	Délka trvání testu (s)	310	2258	822 N

Tabulka 13 Finální porovnání

N – nedokončil test

9 ZÁVĚR

Tato bakalářská práce byla zaměřena na porovnání rychlosti GeoServeru při přístupu k různým datovým skladům. Datové sklady použité v práci byly rozděleny na povinné a volitelné. Konkrétně povinné byly PostgreSQL a Oracle, vybrány vedoucím práce. Třetí, volitelný datový sklad, původně navrhovaný H2 byl v průběhu práce s ohledem na informace získané během zpracovávání a výsledky předběžných testů zaměněn za Shapefile z důvodů vyšší informační hodnoty. Jako podklad pro testování posloužila data z ČUZK a to digitální katastrální mapa ve vektorové podobě harmonizovaná podle směrnic INSPIRE. Tato data byla exportována do formátu ShapeFile, a následně importována do vybraných SŘBD. K testování bylo potřeba podle zadání vytvořit z těchto dat i objemově zmenšenou verzi pod 1GB, aby mohlo dojít k pohodlnému načtení dat do paměti RAM. Z důvodů nízké výpočetní hodnoty při testování kompletní databáze vinou nízkého výkonu PC byla nad rámec zadání selektována ještě další „malá databáze“, sloužící k dalšímu testování a získání dalších výsledkových dat.

Data byla publikována GeoServerem pomocí WMS a nástrojem WMS Tester také otestována rychlost odezvy při zapojení jednotlivých datových skladů. Veškeré programy použité při této práci byly získány z volně dostupných zdrojů na internetu. Ke splnění zadaných úkolů bylo nutné prostudovat jejich dokumentaci a také specifikaci komunikačního protokolu OGC WMS.

Výsledky jsou prezentovány pomocí tabulek a grafů vytvořených v programu Microsoft Excel.

Testování ukázalo, že SŘBD, které využívají vlastních výpočtů, jejichž výsledky následně odesílají na server, nejsou při pouhém publikování statických dat tou nejlepší volbou.

Shapefile opakovaně dokázal, že je možné na něm provádět výpočty i bez pomoci SŘBD a vyhnout se tak nadměrnému zatížení PC. Nutno podotknout, že oproti klasickým serverovým řešením měl stroj, použitý při této práci, určité rezervy. Právě tyto „výkonové rezervy“ mohli zapříčinit nespolehlivost Oracle a negativně tak ovlivnit výsledky testů tohoto SŘBD.

Nejlépeších výsledků tedy dosahoval ShapeFile a pro publikování dat, bez větších nároků na obměny a správu, je to tedy velmi dobrá volba. Ovšem toto řešení postrádá ostatní výhody, které sebou použití SŘBD přináší. Tedy zejména správu dat na jiném stroji, menší nároky

na prostor na disku a další služby a operace s daty. ShapeFile s kompletními daty měl více než 10 GB, stejná data nahraná do PostgreSQL byla objemově téměř poloviční.

Toto vše dokazuje, že volba řešení je individuální a vždy je dobré podobné testování uskutečnit ve vlastních podmínkách. Tato práce by mohla být jednou z pomůcek, která pomůže ono ideální řešení najít.

10 SEZNAM POUŽITÉ LITERATURY:

1

OpenPlans. *GeoServer Overview* [online]. c2013, poslední revize 18.9.2013.

URL: <<http://docs.geoserver.org/2.4.0/user/introduction/overview.html>> [cit. 2014-01-10]

2

Wikipedie. *SŘBD (Data management system)* [online]. c2012. Poslední revize 27.3.2014.

URL: <http://cs.wikipedia.org/wiki/Systém_řzení_báze_dat> [cit. 2014-02-09]

3

PostGIS project Steering Committee. *PostGIS 2.1.3dev Manual* [online]. c2014.

URL: <<http://postgis.net/docs/manual-2.1>> [cit. 2014-01-18]

4

ORACLE. *Overview Oracle database 11g* [online]. c2013.

URL: <<http://www.oracle.com/cz/products/database/overview/index.html>> [cit. 2014-2-03]

5

Wikipedie. *Oracle Spatial* [online]. c2012. Poslední revize 20.4.2013

URL: <http://cs.wikipedia.org/wiki/Oracle_Spatial> [cit. 2014-02-15]

6

ESRI. *ESRI ShapeFile technical documentation* [online]. c1998

URL: <<http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf>> [cit. 2014-03-19]

7

Indian Institutes of Technology. *Comparing PostGIS to Oracle Spatial* [online]. c2011

URL:<<http://www.gise.cse.iitb.ac.in/wiki/images/c/c4/Finalreport.pdf>> [cit. 2013-11-24]

8

Open Geospatial Consortium. *WMS standard definition* [online]. c2010

URL: <<http://www.opengeospatial.org/standards/wms>> [cit. 2013-12-13]

9

Kropla B. *Beginning MapServer*. c2005. Apress. ISBN 1-59059-490-8 [cit. 2013-12-15]

11 SEZNAM OBRÁZKŮ A TABULEK

Obrázek 1 Výpis parametrů WMS-Tester	13
Obrázek 2 Struktura testovaných dat v prostředí PostgreSQL	16
Obrázek 3 Ukázka zobrazení dat v prostředí OpenLayers	16
Obrázek 4 Volba komponent balíčku OpenGeo suite	18
Obrázek 5 pgAdmin	22
Obrázek 6 pgShapefileLoader	22
Obrázek 7 Stahovací obrazovka z oficiálních stránek Oracle	24
Obrázek 8 Volba typu datového skladu Oracle	26
Obrázek 9 Nastavení tvorby databáze Oracle	26
Obrázek 10 MapBuilder	27
Obrázek 11 GeoServer	29
Obrázek 12 Přidávání nového pracovního prostoru	29
Obrázek 13 Napojení datového skladu krok 1	30
Obrázek 14 Napojení datového skladu krok 2	30
Obrázek 15 Napojení datového skladu krok 3	30
Obrázek 16 Aktivace WMS	31
Tabulka 1 Limity PostgreSQL	6
Tabulka 2 Limity Oracle	7
Tabulka 3 Struktura ShapeFile	8
Tabulka 4 Rozdělení testovaných dat	15
Tabulka 5 Zatížení serveru při testování malé databáze	35
Tabulka 6 Zatížení serveru při testování střední databáze	35
Tabulka 7 Zatížení serveru při testování kompletní databáze	36
Tabulka 8 Nastavení a výsledky jednotlivých testů Malá databáze	37
Tabulka 9 Výsledky testů rozdělených do tříd Malá databáze	37
Tabulka 10 Nastavení a výsledky jednotlivých testů Střední databáze	39
Tabulka 11 Výsledky testů rozdělených do tříd Střední databáze	39
Tabulka 12 Nastavení a výsledky jednotlivých testů Kompletní databáze	41
Tabulka 13 Finální porovnání	42