

MapServer vs. Mapserver

Jáchym Čepický

Department of geoinformatics

Faculty of Forestry and Wood Technology, Mendel University of Agriculture and Forestry in
Brno

E-mail: jachym.cepicky@centrum.cz

David Procházka

Department of Informatics

Faculty of Business and Economics, Mendel University of Agriculture and Forestry in Brno

E-mail: xproch17@pef.mendelu.cz

Jitka Machalová

Department of Informatics

Faculty of Business and Economics, Mendel University of Agriculture and Forestry in Brno

E-mail: machalov@mendelu.cz

Klíčová slova: WMS, MapServer, ArcIMS

Abstrakt

Můžeme říci, že mapy prožívají díky moderním technologiím svou renesanci. Díky aplikacím, jako je [Google Maps](http://maps.google.com)¹ či [Seznam-Mapy](http://mapy.seznam.cz)² mohou uživatelé sítě náhle pracovat s geoinformacemi způsobem, na jaký doposud nebyli zvyklí - a tento nový způsob je baví.

Pro vytvoření obrázku mapy, který je buď zobrazen v GISu nebo v okně prohlížeče je potřeba mít stroj odpovídajícího výkonu a programové vybavení schopné takový obrázek vytvořit.

Tento článek se snaží pomoci nalézt odpověď na otázku "Jaký mapový server je pro má data nejvhodnější?".

Mapové servery

Mapové servery jsou programy generující požadovanou mapu jako obrázek na základě požadavků klientských programů. Výsledný obrázek předávají webovému serveru, a ten pak zpět klientskému programu. Primární funkcí mapového serveru je načíst data z různých zdrojů a jejich spojení dohromady do výsledného obrázku [1].

Na „data“ poskytovaná mapovými servery lze přistupovat buď ze specializovaných programů - Geografických informačních systémů - a nebo na příklad z klientských aplikací napsaných pro webové prohlížeče. Lze předpokládat, že uživatele obou typů těchto programů zajímá kromě vzhledu výsledného obrázku především rychlost s jakou se tento obrázek objeví u nich na obrazovce.

¹<http://maps.google.com>

²<http://mapy.seznam.cz>

Správce takových serverů zase zajímá zátěž, které je vystaven stroj, na němž jsou uložena data a na němž probíhá vykreslování výsledných obrázků, ať už z hlediska zatížení disků, tak z hlediska zátěže procesoru.

Komunikace mezi klienty a mapovými servery může probíhat buď pomocí proprietárního rozhraní a nebo přes rozhraní standardní. Standardem v případě mapových serverů je tzv. služba **WMS**³ (*Web Mapping Service*), definovaná konsorciem **OGC**⁴ (*Open Geospatial Consortium*).

Web Mapping Service

OGC je mezinárodní standardizační sdružení, zabývající se především standardy v oblasti geografických informačních systémů, jejich výměnných formátů a podobně. Mezi standardy definované touto organizací patří mimo jiné WMS (*Web Mapping Service*), WFS (*Web Feature Service*), WCS (*Web Coverage Service*), GML (*Geography Markup Language*) a další.

Web Mapping Service⁵ je definovaná v dokumentu OGC 06-042 [2], který popisuje komunikaci mezi mapovým serverem a klientskou aplikací. Mezi dotazy, které musí takový server být schopen spracovat patří mimo jiné

- **GetCapabilities**, vrátí popis dostupných dat na mapovém serveru, jejich formátů, geografické projekce a další informace
- **GetMap**, který vrátí výsledný obrázek - mapu podle zadaných vstupních parametrů.

Protože právě WMS je rozhraní, na jehož základě komunikuje většina programů stahujících z mapových serverů data (**ArcGIS**⁶, **QGIS**⁷, **Udig**⁸, **GRASS**⁹, **UMN MapServer**¹⁰, ...), zaměřili jsme se v tomto testu právě na toto rozhraní.

Test

K vlastnímu testu byl použit server HP Proliant ML 350T03 (<http://indica.mendelu.cz>¹¹), s nainstalovaným operačním systémem MS Windows 2003 SP1. Vybavený je dvěma 73GB SCSI disky v konfiguraci RAID-1, 2GB pamětí, procesorem Intel Xeon 3 3.06GHz (32b). Důvod pro zvolení tohoto operačního systému byl, že server se používán i pro běh licenčního manažeru pro produkty *ESRI*, který je pouze pro MS Windows. Jako webový server byl, na základě doporučení firmy *ESRI*, zvolen *Apache 2.0* s nadstavbou *Tomcat 5.0* a *JRE 1.4.2*.

³http://en.wikipedia.org/wiki/Web_Map_Service

⁴<http://opengeospatial.org>

⁵http://en.wikipedia.org/wiki/Web_Map_Service

⁶<http://www.esri.com/software/arcgis/>

⁷<http://qgis.org>

⁸<http://udig.refractor.net/confluence/display/UDIG/Home>

⁹<http://grass.itc.it>

¹⁰<http://mapserver.gis.umn.edu>

¹¹<http://indica.mendelu.cz>

Použitá data

K testu byla použita data u oblasti školního lesního podniku Křtiny - „Masarykův les“. Území, na kterém byl test prováděn je ohraničeno souřadnicemi 16d35'30.12"E 49d13'5.52"N a 16d48'44.64"E 49d21'18.72"N (WGS84). Pro test byly použity vrstvy:

- Letecké snímky - 50 leteckých snímků v infra-červeném spektru v rozlišení 0.5 m, jejichž barevná paleta byla redukována na 256 barevných odstínů. Průměrná velikost jednoho rastru je 1.89 MB (formát TIFF + TFW)
- Digitální model terénu - Jednotná rastrová mapa v rozlišení 5 m (formát TIFF + TFW)
- Mapa využití půdy - vektorová mapa ve formátu *ESRI Shapefile*, obsahující 32488 linií a 11647 ploch rozdělených do 13 kategorií. Servery byly nastaveny tak, aby se mapová vrstva vykreslovala *bez* popisků s pouze vybarvenými plochami.
- Typologická mapa - vektorová mapa ve formátu *ESRI Shapefile*, obsahující 14731 linií a 5282 ploch rozdělených do 133 kategorií. Mapová vrstva byla nastavena tak, aby se plochy vykreslily pomocí šraf (rastrový obrázek velikosti 2x2 pixely) s popiskami jednotlivých kategorií.

Od původního záměru, testovat data v zobrazení S-JTSK jsme upustili z důvodu problémů s tímto zobrazením v knihovně *PROJ.4* (kterou využívá *UMN MapServer*) na operačním systému MS Windows a problémů při konfiguraci *ArcIMS*. Data tak byla pomocí nástrojů **gdalwarp** a **ogr2ogr** převedena do zobrazení Lat/Long (referenční elipsoid WGS84).

Pro *UMN MapServer* byly ještě letecké snímky opatřeny souborem obsahujícím „mapu“ jednotlivých leteckých snímků - *tile index*.

Použité mapové servery

[University Of Minnesota MapServer](http://mapserver.gis.umn.edu/)¹² (současný název je pouhé *MapServer*) je asi nejrozšířenější Open Source a Free Software GIS program. Je vyvíjen a udržován početnou komunitou uživatelů. Lze jej spouštět buď jako CGI aplikaci nebo pomocí rozhraní *MapScript* zapracovat do různých programovacích jazyků. Pro účely testu byl použit předkompilovaný *MapServer* verze 4.8.1, stažený ze stránek projektu, který byl spouštěn jako CGI.

Druhým použitým mapovým serverem je *ArcIMS*¹³ firmy *ESRI*. Nativním formátem pro komunikaci serveru je *ArcXML*. Pro testy bylo použito *ArcIMS* verze 9.1 SP1, ve které integrován *WMS connector* – nadstavba napsaná v jazyku Java, která umožňuje publikovat data ve formátu WMS. *WMS connector* přistupuje k mapovým službám běžícím na serveru a *de facto* pro ně vytváří WMS rozhraní.

Vlastní test

Vlastní test byl proveden skriptem k tomuto účelu vytvořeným. Skript byl napsán v programovacím jazyce Python. K měření času potřebného k vytvoření obrázku s mapou byl využit

¹²<http://mapserver.gis.umn.edu/>

¹³<http://www.esri.com/software/arcgis/arcims/index.html>

modul `timeit` s funkcí `repeat`.

Tato funkce¹⁴ potřebuje ke svému běhu dva parametry:

- Počet opakování tohoto testu
- Počet volání testované funkce vrámci jednoho testu.

Obě hodnoty byly nastaveny na 10. Funkce vrací číselné pole, obsahující počty sekund potřebné k provedení každého testu (v našem případě pole o deseti prvcích).

Výsledná čísla jsou tedy počty sekund potřebných k vytvoření, stažení a uložení 10 obrázků s mapami.

Nelze říci, že se jedná o počet sekund potřebných k vytvoření obrázku. Celkový čas je kromě parametrů na straně serveru závislý na dostupnosti sítě, rychlosti ukládání souboru na straně klienta a dobou potřebnou k vykonání samostatné funkce.

Funkce `save_file` dostane jako svůj parametry vždy URI, ze kterého má stahovat potřebná data. Ještě před tím, než započne se stahováním dat, jsou hraniční souřadnice upraveny o náhodnou hodnotu. Cílem této úpravy bylo zamezit načítání odpovědí z cache na straně serveru.

```
[...]
t = timeit.Timer("""mapserverVsMapserver.save_file(
                    "%s",
                    verbose=%d,
                    bbox="%s",
                    mapserv="%s"
                    )"" % \
                    (URI, verbose, bbox, mapserv),
                "import mapserverVsMapserver")
times = t.repeat(10,10)
[...]
```

```
number = 0
def save_file(uri, verbose=0, bbox=None, mapserv=""):
```

```
    # bbox randomization
    newBox = ""
    rand = random.random()/100-0.005
    for cord in bbox.split(","):
        newBox += str(float(cord)+rand)+", "
    newBox = newBox[:-1]
    uri += "&BBOX=%s" % (newBox)
    global number
    number += 1
    if verbose == 3:
        print number
```

¹⁴http://diveintopython.org/performance_tuning/timeit.html

```
if verbose > 3:
    print number, uri
map = urllib.urlopen(uri)
input = open('map-%s-%03d.png' % (mapserv,number), 'wb')
input.write(map.read())
input.close()
return
```

Testovány byly následující varianty:

- „Jednoduchý“ rastrový soubor celého území (digitální model terénu)
- „Jednoduchý“ rastrový soubor celého území (digitální model terénu) - výřez území
- „Náročná“ vektorová mapa celého území, včetně rastrových textur a popisků (typologická mapa)
- „Náročná“ vektorová mapa celého území, včetně rastrových textur a popisků (typologická mapa) - výřez území
- „Náročný“ rastrový soubor celého území - letecké snímky
- „Náročný“ rastrový soubor celého území - letecké snímky - výřez území
- Kombinace „náročného“ rastru s „náročnou“ vektorovou mapou na celém území - letecké snímky + typologická mapa
- Kombinace „náročného“ rastru s „náročnou“ vektorovou mapou na celém území - letecké snímky + typologická mapa - výřez území
- „Jednoduchá“ vektorová mapa celého území (mapa využití půdy)
- „Jednoduchá“ vektorová mapa celého území (mapa využití půdy) - výřez území.

Celé území bylo ohraničeno (výchozími) souřadnicemi 16d35'30.12"E 49d13'5.52"N a 16d48'44.64"E 49d21'18.72"N. Výřez pak 16d42'7.128"E 49d17'14.388"N a 16d42'39.996"E 49d17'35.412"N.

Obrázky byly stahovány ve formátu PNG o velikosti 400×400 pixelů.

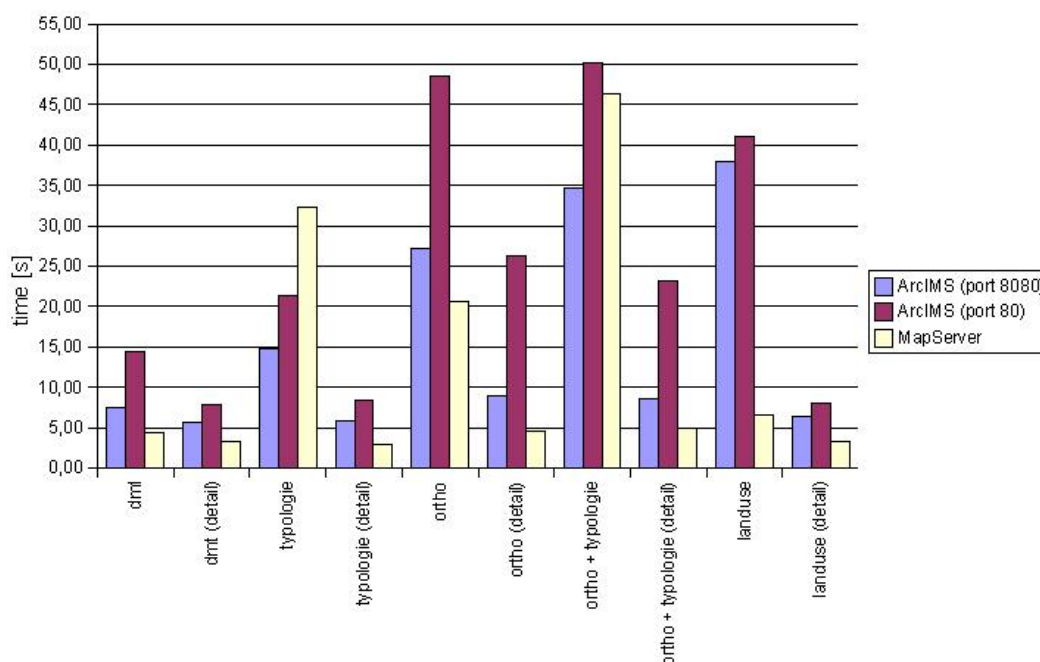
Výsledky

Sekvenční dotazy z jednoho klienta na server

Tyto testy popisují řady sekvenčních dotazů z jednoho klienta na mapový server. Pro kladení dotazů na mapový server *ArcIMS* jsou použity porty 80 a 8080. Pokud je dotaz položen přes port 80, převezme jej webový server *Apache*, ten jej předá *Tomcatu* a následně se zavolá samotný *WMS connector* (který získá data z běžící mapové služby *ArcIMS*). Pokud je dotaz položen přes port 8080, probíhá komunikace přímo s nadsstavbou *Tomcat* a je tedy ušetřen čas komunikace s web serverem.

	dmt	dmt (detail)	typologie	typologie (detail)	ortho	ortho (detail)	ortho + typologie	ortho + typologie (detail)	land-use	land-use (detail)
ArcIMS (port 8080)	7,52	5,61	14,73	5,82	27,23	8,88	34,59	8,62	37,90	6,38
ArcIMS (port 80)	14,51	7,89	21,37	8,38	48,61	26,27	50,15	23,17	41,03	8,10
MapServer	4,38	3,28	32,38	3,00	20,65	4,66	46,37	4,85	6,55	3,28

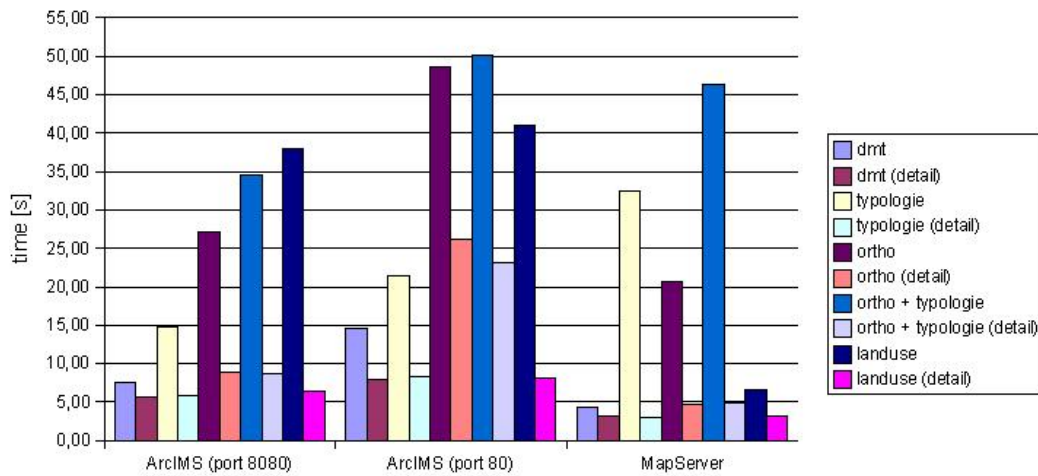
Tabulka 1: Srovnání rychlosti vykreslování různých typů vrstev servery



Obrázek 1: Jak rychle je daná vrstva zpracována různými servery (větší čas znamená horší výsledek)

Sekvenční dotazy ze serveru na server

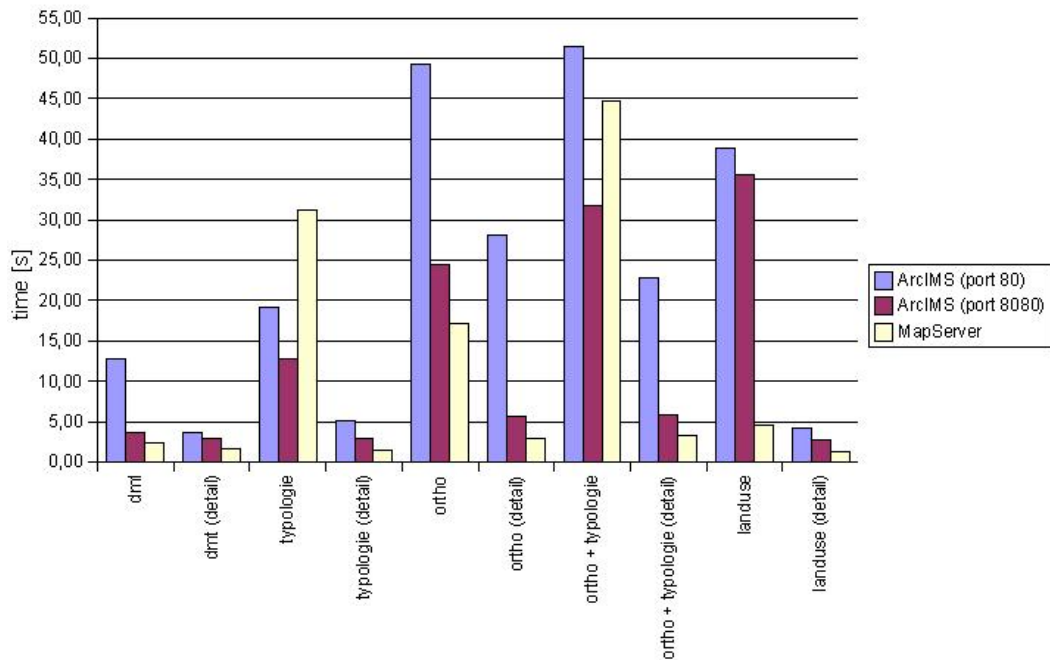
Tyto hodnoty popisují dotazy pokládané ze stejného počítače, jako je ten, na kterém je instalován mapový server. Porovnáním s předchozími hodnotami získáme vliv komunikace po síti na výsledné časy.



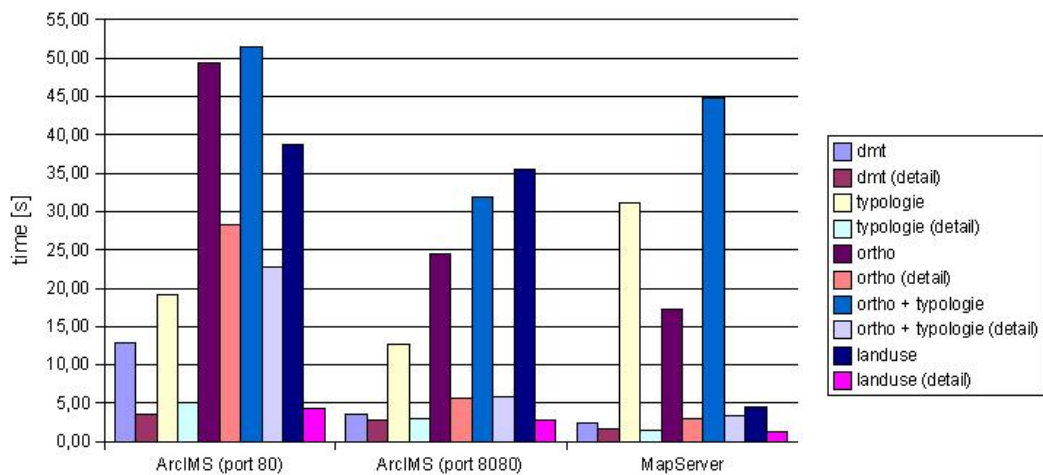
Obrázek 2: Jak rychle dokáží servery zpracovávat jednotlivé vrstvy

	dmt	dmt (detail)	typologie	typologie (detail)	ortho	ortho (detail)	ortho + typologie	ortho + typologie (detail)	landuse	landuse (detail)
ArcIMS (port 8080)	3,65	2,89	12,74	2,95	24,43	5,63	31,80	5,78	35,52	2,83
ArcIMS (port 80)	12,84	3,64	19,10	5,17	49,37	28,20	51,50	22,85	38,82	4,27
MapServer	2,39	1,67	31,20	1,53	17,17	2,98	44,76	3,39	4,50	1,21

Tabulka 2: Srovnání časů nutných odpovědi při kladení dotazů pouze v rámci serveru (neprobíhá komunikace po síti)



Obrázek 3: Jak rychle je daná vrstva zpracována různými servery



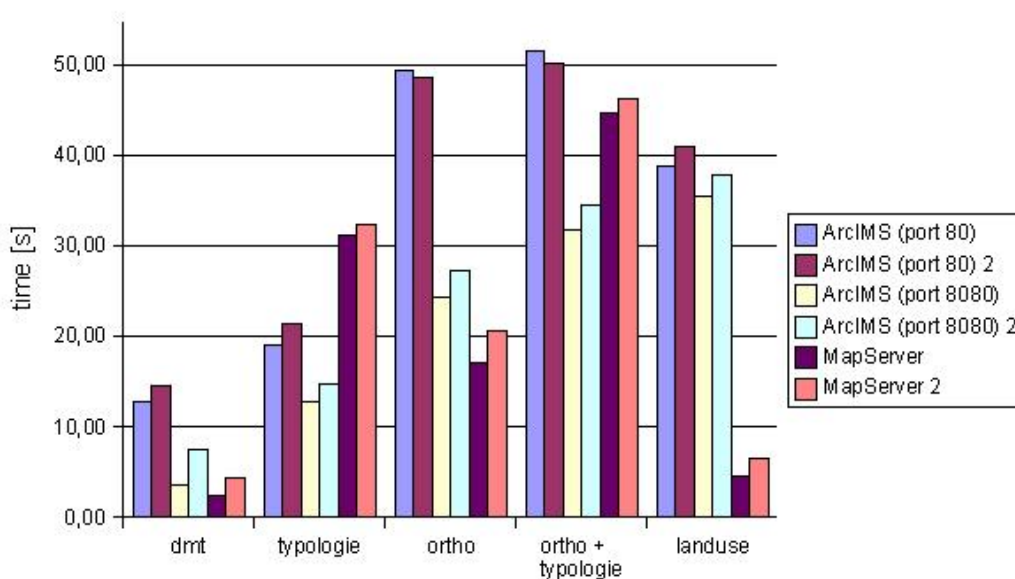
Obrázek 4: Jak rychle dokáží servery zpracovávat jednotlivé vrstvy

Porovnání sekvenčních dotazů

Níže uvedená tabulka a graf shrnují rozdíly mezi dotazy klient-server a server-server. Je patrné, že zpoždění při dotazech přes počítačovou síť je ve většině případů téměř konstantní. Tímto testem se navíc vzájemně ověřili i naměřené hodnoty. Je patrné, že při opakovaném měření dosahujeme obdobných výsledků. Hodnoty bez čísla 2 jsou naměřené při kladení ze samotného serveru. Hodnoty s číslem 2 (např. ArcIMS (port 80) 2) jsou naměřeny při kladení dotazů z klienta na server.

	dmt	typologie	ortho	ortho + typologie	landuse
ArcIMS (port 80)	12,84	19,10	49,37	51,50	38,82
ArcIMS (port 80) 2	14,51	21,37	48,61	50,15	41,03
ArcIMS (port 8080)	3,65	12,74	24,43	31,80	35,52
ArcIMS (port 8080) 2	7,52	14,73	27,23	34,59	37,90
MapServer	2,39	31,20	17,17	44,76	4,50
MapServer 2	4,38	32,38	20,65	46,37	6,55

Tabulka 3: Srovnání vlivu komunikace po síti na dobu nutnou k odpovědi



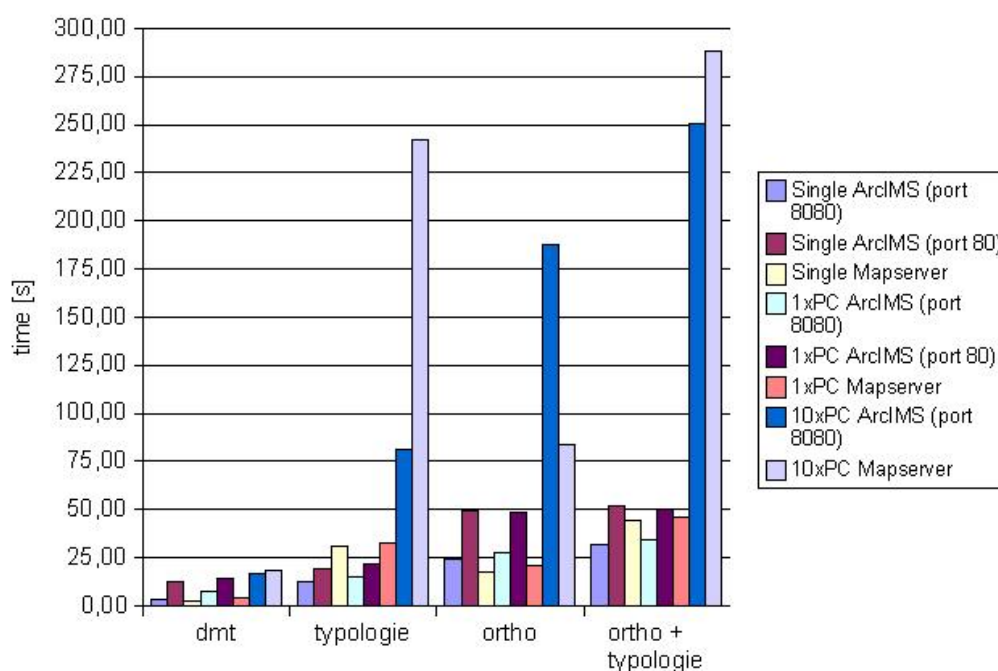
Obrázek 5: Porovnání dotazu z klient-server a ze server-server

Porovnání sekvenčních a paralelních dotazů

Následující tabulka srovnává délky odpovědí v případě kladení dotazů z jednoho PC, samotného serveru a paralelně z 10 různých PC. Zobrazená naměřená hodnota pro 10 PC je průměrem naměřených hodnot ze všech PC (rozdíly mezi hodnotami z PC byly ve všech uvedených případech pod 10%). Položky označené “Single” popisují dotazy server-server. Od testování *ArcIMS* přes port 80 bylo pro paralelní zátěž bylo upuštěno, protože je z předchozích výsledků zcela evidentní, že se pro takovéto nasazení nehodí.

	dmt	typologie	ortho	ortho + typologie
Single ArcIMS (port 8080)	3,65	12,74	24,43	31,80
Single ArcIMS (port 80)	12,84	19,10	49,37	51,50
Single Mapserver	2,39	31,20	17,17	44,76
1xPC ArcIMS (port 8080)	7,52	14,73	27,23	34,59
1xPC ArcIMS (port 80)	14,51	21,37	48,61	50,15
1xPC Mapserver	4,38	32,38	20,65	46,37
10xPC ArcIMS (port 8080)	16,98	80,90	187,68	250,05
10xPC Mapserver	18,31	242,06	83,68	287,84

Tabulka 4: Srovnání vlivu sekvenčního a paralelního kladení dotazů



Obrázek 6: Porovnání sériového a paralelního kladení dotazů

Závěr

Vyvozovat závěry z naměřených výsledků je vždy problematické. Do značné míry záleží na úhlu pohledu a interpretaci. Navíc každý test může být napadnut z hlediska ne zcela objektivní metodiky měření, použitých prostředků nebo konfiguraci daného produktu. Primárním cílem tohoto článku proto bylo poskytnout výtah z námi získaných dat a popsat podmínky, za kterých byla tato data naměřena. Na základě zde zmíněných výsledků lze bez újmy na objektivitě říci:

- Jak se dalo předpokládat, je markantní rozdíl mezi rychlostí zpracování požadavku u produktu *ArcIMS* při kladení dotazů přes port 80 a port 8080. V případech, kdy je webový server *Apache* vypuštěn a dotaz je položen přímo *Tomcatu* (standardně port 8080), je požadavek vyřízen často i v poloviční době. Pro nasazení s vysokou zátěží serveru je tedy velmi vhodné zvolit tuto variantu.
- Otevřený projekt *MapServeru* je pro poskytování mapových služeb minimálně konkurenceschopných řešením komerčnímu produktu *ArcIMS*. Jak je patrné z výsledků slabým místem *MapServeru* je mapování v souborech uložených textur do vektorové mapy (vrstva typologie). Při práci s rastry i vektorovými podklady však dosahuje velmi zajímavých výsledků. Relativně nízký výkon *ArcIMS* je pochopitelný, pokud vezmeme v potaz fakt, že *WMS connector* je napsán v Javě a nemůže tedy dosáhnout takové efektivity, jako kdyby byl předkompilován pro určitou platformu. Tedy výkon *WMS connectoru* je do značné míry dán i výkonem *Tomcatu*.
- Horší výsledky při zpracování rastrových map u obou řešení by mohly být teoreticky vylepšeny vybudováním pyramid (u produktu *ArcIMS* např. uložením rastru do *ArcSDE*¹⁵).
- Sledování zátěže serveru v průběhu testů potvrdilo předpoklad, že zatímco při zpracování rozsáhlých rastrových dat je limitním faktorem rychlost disku, při zpracování vektorových dat je limitním faktorem rychlost procesoru.

Bylo by bezesporu zajímavé, porovnat i další známé a v praxi často používané mapové servery, zejména *GeoServer*¹⁶, *DeeGree*¹⁷, nově uvolněný produkt *MapGuide*¹⁸ či český *TopoL Internet Server*¹⁹. Při dalším testování hodláme rovněž porovnat rychlosti zpracování dat v nativních formátech jednotlivých řešení.

Reference

1. Mitchell, Tyler (2005): *Web Mapping Illustrated*, O'Reilly Media, Inc., Sebastopol.
2. Beaujardiere, Jeff (2006): *OpenGIS(r) Web Map Server Implementation Specification*, Open Geospatial Consortium Inc., OGC(r) 06-042, Version 1.3.0, <http://www.opengeospatial.org>

¹⁵<http://www.esri.com/software/arcgis/arcscde/>

¹⁶<http://docs.codehaus.org/display/GEOS/Home>

¹⁷<http://www.deegree.org/>

¹⁸<http://www.autodesk.com/mapguide>

¹⁹<http://topol.cz/?doc=2400>