

Maanmittauslaitoksen taustakarttasarja WMS-palveluksi

Jukka Rahkonen, <http://latuviitta.org>

Lisenssi [Creative Commons Attribution-Share Alike 3.0 Unported](http://creativecommons.org/licenses/by-sa/3.0/)

Viimeksi muokattu 25. lokakuuta 2012

Tiivistelmä

Ohjeessa selostetaan, kuinka Maanmittauslaitoksen taustakarttasarjan yhteentoista eri mittakaavatasoon kuuluvat 16962 karttatiedostoa muunnetaan sellaiseen muotoon, joka soveltuu hyvin ainakin MapServer- ja GeoServer-karttapaalvelimien lähtöaineistona käytettäväksi. Lisäksi selostetaan, kuinka muunnetuista kartoista tehdään MapServer-palvelimella WMS-taustakarttapaalvelu, joka täyttää INSPIRE-katselupalveluille asetetut suorituskyky- ja kapasiteettivaatimukset moninkertaisella marginaalilla. Käytetyt menetelmät, ohjelmistot, laitteisto ja suorituskykymittausten tulokset raportoidaan. Ohje on tehty Windows-käyttöjärjestelmää varten, mutta pienellä soveltamisella toimii myös Linuxilla.

Ohjelmistot ja muut valmistelut

Ohje on testattu seuraavilla ohjelmistoilla:

Käyttöjärjestelmä:

- Windows Vista (32-bit) ja Windows 7 (64-bit)

Ohjelmistot

- cwRsync - Rsync for Windows versio 3.0.7
<https://www.itefix.no/i2/cwrsync>
- GDAL MSVC (Win32) development, ladattu 17. lokakuuta 2012
<http://www.gisinternals.com/sdk/>
- MapServer 6.2.0-beta4 -päivitys MS4W v 3.0.6 version päälle
<http://maptools.org/ms4w/index.phtml?page=downloads.html>
- Apache JMeter -testausohjelma
<http://jmeter.apache.org/>

Levyt ja hakemistot

- MML:n taustakarttatiedostot ladataan tyhjälle USB-levylle, joka on ohjeessa levyasema D:
- Muunnetut tiedostot tallennetaan toiselle tyhjälle USB-levylle (levyasema E:)
- MapServer hakee kartat levyasemalta E:

Mikä tahansa muu hakemistojärjestys vaatii muutoksia ohjeessa esitettyihin komentojonoihin ja asetustiedostoihin. Muutokset ovat pieniä ja helppoja, mutta ne on tehtävä oikein.

Suorituskykymittauksissa käytetty testipalvelin

- Kannettava Dell
- prosessori Intel Core i7-3720QM
- 8 Gt muistia
- ulkoinen USB-kovalevy karttatiedoille

Taustakartta-aineiston lataaminen cwRsync-ohjelmalla

Taustakarttasarjan lataaminen on helpointa tehdä sivustolta <http://kartat.kapsi.fi/> Rsync-ohjelmistolla.

Ensimmäinen tehtävä on asentaa cwRsync-ohjelma. Ohjelman latausosoite on <https://www.itefix.no/i2/cwrsync>

Ladataan tiedosto ja suoritetaan asennusohjelma. Asennus vaatii järjestelmänvalvojan oikeudet.

Lataamisen jälkeen cwRsync-hakemistossa on mallikomentojenotiedosto "**cwrsync.cmd**". Avataan tämä tiedosto tekstinkäsittelyohjelmalla. Jos asennus tehtiin hakemistoon "**C:\Program Files\cwRsync**", niin tiedostoon tarvitsee vain lisätä muutama rivi. Jos asennus kuitenkin haluttiin tehdä jonnekin muualla, niin seuraavaa kohtaa on muokattava vastaamaan todellista tilannetta:

```
REM ** CUSTOMIZE ** Specify where to find rsync and related files (C:\CWRSYNC)
SET CWRSYNCHOME=%PROGRAMFILES%\CWRSYNC
```

Jos asennus tehtiin esimerkiksi hakemistoon c:\ohjelmat\cwRsync, niin tähän siis kirjoitetaan

```
SET CWRSYNCHOME=c:\ohjelmat\CWRSYNC
```

Seuraavaksi siirrytään aivan cwrsync.cmd-tiedoston loppuun. Viimeisellä rivillä lukee:

```
REM ** CUSTOMIZE ** Enter your rsync command(s) here
```

Lisätään tämän rivin perään seuraavat 11 uutta riviä riviä ja tallennetaan tiedosto.

```
rsync -arPv tiles.kartat.kapsi.fi::mml/taustakarttasarja/taustakartta_8m/100m/etrs89/ /cygdrive/d/t8000/
rsync -arPv tiles.kartat.kapsi.fi::mml/taustakarttasarja/taustakartta_4m/500m/etrs89/ /cygdrive/d/t4000/
rsync -arPv tiles.kartat.kapsi.fi::mml/taustakarttasarja/taustakartta_2m/250m/etrs89/ /cygdrive/d/t2000/
rsync -arPv tiles.kartat.kapsi.fi::mml/taustakarttasarja/taustakartta_800/100m/etrs89/ /cygdrive/d/t800/
rsync -arPv tiles.kartat.kapsi.fi::mml/taustakarttasarja/taustakartta_320/25m/etrs89/ /cygdrive/d/t320/
rsync -arPv tiles.kartat.kapsi.fi::mml/taustakarttasarja/taustakartta_160/12_5m/etrs89/ /cygdrive/d/t160/
rsync -arPv tiles.kartat.kapsi.fi::mml/taustakarttasarja/taustakartta_80/5m/etrs89/ /cygdrive/d/t80/
rsync -arPv tiles.kartat.kapsi.fi::mml/taustakarttasarja/taustakartta_40/2_5m/etrs89/ /cygdrive/d/t40/
rsync -arPv tiles.kartat.kapsi.fi::mml/taustakarttasarja/taustakartta_20/2_5m/etrs89/ /cygdrive/d/t20/
rsync -arPv tiles.kartat.kapsi.fi::mml/taustakarttasarja/taustakartta_10/1m/etrs89/ /cygdrive/d/t10/
rsync -arPv tiles.kartat.kapsi.fi::mml/taustakarttasarja/taustakartta_5/0_5m/etrs89/ /cygdrive/d/t5/
```

Kirjasinkoko yllä on tarkoituksella pieni, jotta rivit eivät katkea.

Kun tiedosto on tallennettu, avataan tietokoneella Windows-komentoikkuna, siirrytään cwRsync-asennushakemistoon eli siihen hakemistoon, mihin cwrsync.cmd -tiedosto on tallennettu ja annetaan komento "cwrsync.cmd"



Ohjelma alkaa ladata karttatiedostoja Kapsin palvelimelta ja noin puolentoista vuorokauden päästä D: -asemalla on seuraavat tiedostohakemistot:

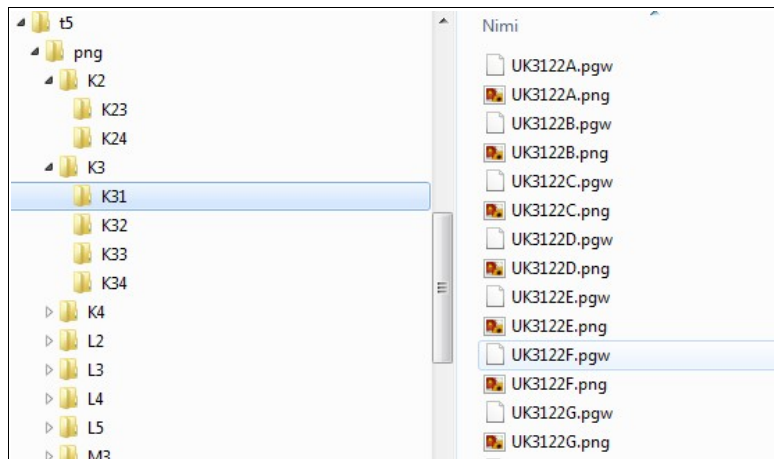
Hakemisto	Tiedostoja	Hakemiston koko
t8000	1	163 kt
t4000	1	474 kt
t2000	1	1,1 Mt
t800	1	5,0 Mt
t320	1	49,3 Mt
t160	65	191 Mt
t80	65	710 Mt
t40	216	2,9 Gt
t20	1519	3,5 Gt
t10	2940	14,2 Gt
t5	12152	35,7 Gt
Yhteensä	16962	58,0 Gt

Yllä olevien numeroiden perusteella voi tarkistaa, onko hakemistoissa oikea määrä tiedostoja. Jos ei ole, niin cwrsync.cmd -tiedostosta voi pyyhkiä pois ne rivit, joiden lataus on onnistunut, ja suorittaa komennon uudestaan. RSync on viisas ohjelma, eikä se lataa uudestaan jo ladattuja tiedostoja, joten mahdollinen täydennyslataus on nopea.

Kun lataus on kokonaan valmis, niin kaikki Maanmittauslaitoksen taustakarttasarjan karttatiedostot ovat D- levyllä png-muodossa. Koska tämä tiedostomuoto ei sovi alkuunkaan tehokkaaseen paikkatietokäyttöön, niin ne seuraavassa vaiheessa ne muunnetaan sisäisesti laatoitetuiksi tiff-tiedostoiksi, eli kuvatiedostot tallennetaan siten, että tiff-kuvan sisältä voidaan poimia kuvadataa halutuista paikoista pieninä 256x256 pikselin suuruisina paloina ilman, että koko suurta kuvaa tarvitsisi avata. Vaikka png-tiedostomuodosta luovutaankin, niin png-tiedostojen pätevä pakkausmenetelmä, nimeltään "DEFLATE", voidaan kuitenkin säilyttää. Tiff-muunnoksen jälkeen uusien kuvatiedostojen sisään lasketaan vielä 5 yleistettyä kuvatasoa. Tämä tarkoittaa sitä, että jos MML:n kartan pikselikoko on esimerkiksi 1 metri, niin tiedoston sisään tallennetaan yleistetyt versiot, joiden pikselikoot ovat 2, 4, 8, 16 ja 32 metriä. Näistä yleistetyistä versioista tulee olemaan paljon iloa, jos koko karttalehti halutaan näyttää vain pienenä postimerkinä kuvaruudulla. Tehokkaista ja tehottomista paikkatietojen kuvatiedostomuodoista voi lukea lisää täältä: http://latuviitta.org/documents/Paikkatietojen_nopeat_ja_hitaat_kuvatiedostomuodot.pdf

Tiff-muunnoksen yhteydessä yksinkertaistetaan samalla kertaa karttatiedostojen hakemistorakennetta niin, että kaikki saman mittakaavatason kartat tallennetaan yhteen hakemistoon. Maanmittauslaitoksen tapanaan on tallentaa karttojen metatiedon hakemistopuuhuun. Tätä menetelmää selventää alla oleva kuvaruutukaappaus ja tässä esimerkkinä annettava kokonainen polku, joka johtaa yhteen karttatiedostoon.

```
/mml/taustakarttasarja/taustakartta_5/0_5m/etrs89/png/L3/L31/UL3111A.png
```



Komentojonojen rakentaminen

Tiedostolistausten tekeminen mittakaavatasoittain

Annetaan D: -aseman juurihakemistossa seuraavat 11 komentoa:

```
dir /s /b t8000\*.png >t8000.bat
dir /s /b t4000\*.png >t4000.bat
dir /s /b t2000\*.png >t2000.bat
dir /s /b t800\*.png >t800.bat
dir /s /b t320\*.png >t320.bat
dir /s /b t160\*.png >t160.bat
dir /s /b t80\*.png >t80.bat
dir /s /b t40\*.png >t40.bat
dir /s /b t20\*.png >t20.bat
dir /s /b t10\*.png >t10.bat
dir /s /b t5\*.png >t5.bat
```

Avataan nämä komentojonojen pohjat tekstinkäsittelyohjelmaan. Alla esimerkkinä kaksi ensimmäistä riviä **t5.bat** -tiedostosta

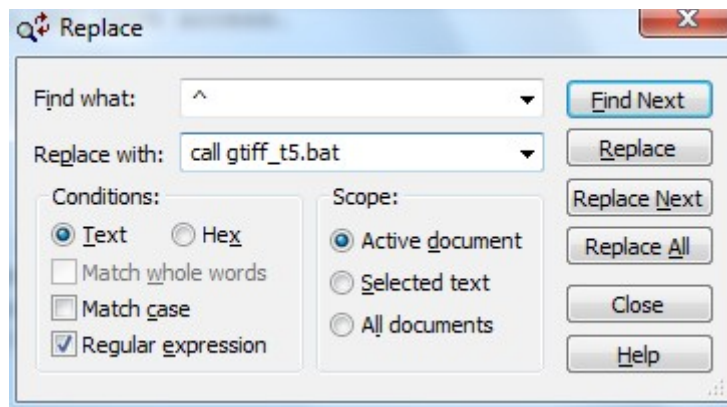
```
d:\t5\png\K2\K23\UK2344A.png
d:\t5\png\K2\K23\UK2344B.png
```

Muokataan komentojonoja niin, että jokaisen rivin alussa kutsutaan toista, varsinaisen työn tekevää komentojonoa.

```
call gtiff_t5.bat d:\t5\png\K2\K23\UK2344A.png
pause
call gtiff_t5.bat d:\t5\png\K2\K23\UK2344B.png
```

t5.bat komentojono laitetaan kutsumaan **gtiff_t5.bat** -komentojonoa; vastaavasti **t10.bat** pannaan kutsumaan **gtiff_t10.bat** -komentojonoa ja niin edelleen. Ensimmäisen rivin jälkeen on pantu **pause**-komento, joka pysäyttää komentojonon suorituksen, kun ensimmäinen tiedosto on käsitelty. Katkerien kokemusten perusteella suosittelen tarkistamaan, näyttääkö ensimmäisen muunnoksen tulos oikealta, ennen kuin antaa luvan jatkaa komentojonon suorittamista. Ei ole täysin mahdotonta, että jossain on joku pieni virhe, ja silloin on parempi tuhota yksi tiedosto tai tehdä yksi väärä muunnos kuin tuhota kokonainen hakemistollinen tiedostoja tai hukata vuorokausi aikaa vääriin muunnoksiin

Tiedostojen muokkaamiseen kannattaa käyttää "Regular Expression" -määrittelyä ymmärtävää teksturia, esimerkiksi TextPad- tai Notepad++ -ohjelmia, sillä niiden avulla voidaan kätevästi lisätä haluttu teksti rivin alkuun tekemällä kuten tämä kuva näyttää.



Komento **gtiff_t5.bat**, jota edellinen komentojono kutsuu, sisältää kaksi erillistä komentoa ja näyttää tältä:

```
gdal_translate -of GTiff -a_srs epsg:3067 -co TILED=YES -co compress=DEFLATE
%1 E:\t5\%~n1.tif

gdaladdo -r average --config COMPRESS_OVERVIEW DEFLATE
E:\t5\%~n1.tif 2 4 8 16 32
```

Selvyiden vuoksi molemmat komennot on pitänyt jakaa kahdelle riville. Rivinvaihdot on poistettava, jos komennot kopioidaan komentojonotiedostoksi yllä olevasta esimerkistä .

Komentojonon toiminta

Komento 1, muunnos gdal_translate-ohjelmalla

- muuntaa png-tiedostot geotiff-muotoon (-of GTiff)
- liittää geotiffiin tiedon karttojen koordinaattijärjestelmästä, joka on suomalainen ETRS-TM35FIN, ja jonka koodi on EPSG:3067
- kirjoittaa tiffin laatoitettuna (-co TILED=YES)
- käyttää pakkaamiseen samaa "deflate" -pakkausmenetelmää, mitä myös png käyttää (-co compress=DEFLATE)
- ottaa muokattavaksi t5.bat-komentojonon syöttämän tiedoston (%1)
- antaa tulostiedoston nimen niin, että se tallenetaan E: -asemalle hakemistoon \t5, ja tiedostonimenä käytetään lähtötiedoston nimen alkuosaa ja tarkentimeksi annetaan .tif (E:\t5\%~n1.tif)

Huomaa, että levyasemalla E: täytyy olla valmiiksi olemassa hakemisto \t5. Komennossa on eräs kenties oudolta vaikuttava kohta, eli %~n1. Se merkitsee kutsuvan komentojonon välittämää ensimmäistä parametria %1, mutta niin, että kun kutsun mukaan tulee tiedoston täydellinen nimi, jossa on mukana levyaseman tunnus, polku, tiedostonimi ja tiedostonimen tarkenne, niin %:n ja 1:n väliin kirjoitetulla osalla "~n" (tilde ja n) poimitaankin pelkkä tiedostonimi. Tämän muistaa jatkossa, kun ajettelee, että se on "än niin kuin name".

Komento 2, yleistettyjen tasojen lisääminen gdaladdo-ohjelmalla

- käyttää yleistyksessä ympäröivien pikseleiden keskiarvoa (-r average). Tämä johtaa sen verran paremman näköiseen laatuun kuin oletuksena käytettävä "-r nearest", että parametria kannattaa ehdottomasti käyttää; erityisesti kartan tekstit säilyvät näin paljon paremmin luettavina
- pakataan myös yleistetyt tasot deflate-menetelmällä (--config COMPRESS_OVERVIEW DEFLATE)
- lasketaan yleistetyt tasot samalle tiedostolle, joka edellä kirjoitettiin E: -levylle (E:\t5\%~n1.tif)
- lasketaan viisi yleistystasoa (2 4 8 16 32)

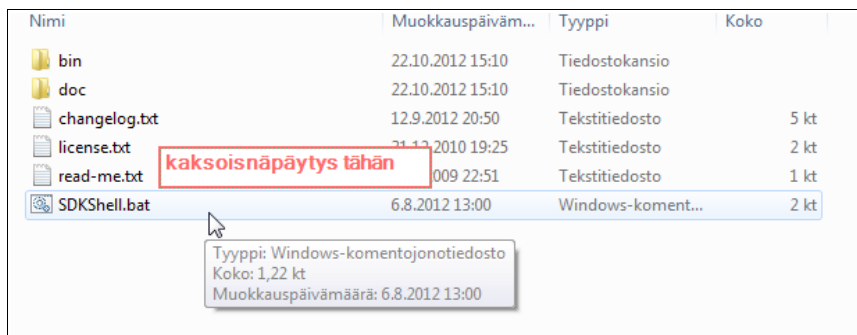
Muut komentojonot (gtiff_t10.bat jne.) ovat muuten täsmälleen samanlaisia, mutta niissä käytetään muita mittakaavatasoja vastaavia hakemistopolkuja.

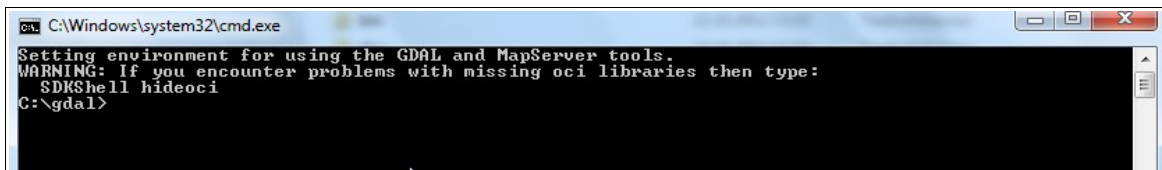
Komentojonojen suoritus

Komentojonot käyttävät kahda GDAL-ohjelmaa "gdal_translate" ja "gdaladdo", joten koneelta täytyy löytyä GDAL asennettuna. Windowsille GDAL-ohjelmat saa esimerkiksi täältä <http://gisinternals.com/sdk/>

Esimerkiksi 32-bittisille Windows-käyttöjärjestelmille asennus onnistuu ilman järjestelmänvalvojan oikeuksiakin lataamalla tämä zip-paketti <http://gisinternals.com/sdk/Download.aspx?file=release-1600-gdal-mapserver.zip>

Zip-arkisto puretaan johonkin hakemistoon, jonka jälkeen tiedostonhallinnan avulla etsitään käynnistystiedosto "sdkshell.bat" kuten alla olevista kuvista selviää.





```
ca: C:\Windows\system32\cmd.exe
Setting environment for using the GDAL and MapServer tools.
WARNING: If you encounter problems with missing oci libraries then type:
  SDKShell hideoci
G:\gdal>
```

Kuinka saat GDAL-ohjelmilla täydet tehot irti moniprosessorisesta tietokoneesta

GDAL-ohjelmat eivät yleensä käytä monisäikeistä suoritusta, joten useampiprosessoriset koneet tekevän muunnoksia osateholla. Tätä ei kannata surra, vaan avata useita komentoikkunoita ja ajaa niistä jokaisessa omaa komentojonoaan. Tällä tavalla kokonaisteho on jopa parempi kuin ajattaessa monisäikeistä suoritusta tukevia ohjelmia yhtenä prosessina, koska yksisäikeisen ohjelman ei tarvitse tuhlaata voimiaan säikeiden pitämiseen oikeassa tahdissa. Tätä ohjetta testattaessa muunnokset ajettiin neljässä eri ikkunassa tietokoneella, jossa on kahdeksanytiminen prosessori. Tai niin sitä mainostetaan, vaikka käytännössä ”hyperthreading” tuntuu ajattaessa samalta kuin vanhemmat neliprosessoriset. Testissä ei muutenkaan ollut mitään hyötyä ajaa muunnoksia neljää useammassa ikkunassa, koska kovalevy ei enää pysynyt vauhdissa mukana, ja kokonaisteho alkoi heikentyä. Komentojonon jakaminen eri ikkkunoihin tehtiin raa’asti mittakaavatason mukaan, mikä johti hyvään alkuvauhtiin mutta laiskaan lopetukseen, koska 1:5000 -hakemistossa oli niin paljon enemmän tiedostoja kuin muissa hakemistoissa, että viimeisen vuorokauden ajan tietokone teki muunnoksia vain yhdellä prosessorilla, koska kaikki muut työt olivat jo valmistuneet. Hätähousuille tässä on helppo paikka tehostaa kokonaisprosessia jakamalla 1:5000 -komentoiono esimerkiksi kolmeen tai neljään eri ikkunoissa ajattavaan osaan.

Koko taustakarttasarjan kaikkien tiedostojen muunnosten tekemiseksi, sitten kun komentojonot ovat valmiina, täytyy siis vain ajaa nämä 11 komentoa oman mieltymyksen ja tietokoneen suorituskyvyn mukaan joko yhdessä tai useammassa ikkunassa

```
t8000.bat
t4000.bat
t2000.bat
t800.bat
t320.bat
t160.bat
t80.bat
t40.bat
t20.bat
t10.bat
t5.bat
```

Mallisuoritus hyvällä kannettavalla kesti 32,5 tuntia, mutta tästä ajasta koko viimeinen vuorokausi muunnettiin pelkkiä 1:5000 -tason tiedostoja yhdessä ainoassa ikkunassa ja vain yhtä prosessoria käyttäen.

Valmistelut MapServer-taustakarttaa varten.

Tässä vaiheessa on saatu aikaan E: -levylle 11-hakemistoa ja niihin MML:n taustakarttasarjan karttatiedostot hyvin optimoituun muotoon MapServer-karttapalvelinta varten. Samat tiedostot toimivat kyllä hyvin myös GeoServer-palvelimella ja kenties muillakin merkeillä. Nopeuden periaatteet ovat kuitenkin kaikilla samat, eli karttapalvelimen on päästävä iskemään välittömästi kiinni mahdollisimman oikeaan resoluutioon ja mahdollisimman oikean kokoiseen alueeseen.

Muunnoksen varjopuolena näyttää olevan se, että tiedostojen vaatima levytila on kasvanut Maanmittauslaitoksen alkuperäiseen png-jakeluun verrattuna huomattavasti. Jos tavoitteena on nopea WMS-palvelu, niin tämä hinta kannattaa ehdottomasti maksaa, sillä suuret png-tiedostot eivät kertakaikkiaan sovi lähtöaineistoksi WMS-käyttöön.

Hakemisto	Tiedostoja	Hakemiston koko
t8000	1	310 kt
t4000	1	942 kt
t2000	1	2,3 Mt
t800	1	9,6 Mt
t320	1	105 Mt
t160	65	397 Mt
t80	65	1,4 Gt
t40	216	5,6 Gt
t20	1519	7,1 Gt
t10	2940	25,7 Gt
t5	12152	63,0 Gt
Yhteensä	16962	104 Gt

Seuraavat toimenpiteet ovat tarpeen erityisesti MapServeriä varten, mutta koska niihin tarvittavat ohjelmat ovat myös GDAL-apuohjelmia niin ne on tässä vaiheessa jo käyttövalmiina.

Tehtäviä on kaksi: Ensiksi luodaan jokaista mittakaavatasoa varten shapefile-muotoinen hakemistotiedosto, johon tallennetaan kuvatiedostojen peittämät maantieteelliset alueet ja tiedostojen hakemistopolku. Näiden hakemistotiedostojen eli tileindex-tiedostojen perusteella Mapserver osaa etsiä ja avata oikeat kuvatiedostot. Toisessa vaiheessa luodaan vielä näille hakemistotiedostoille spatiaali-indeksit. Niiden avulla saadaan kuvahakuihin hieman lisää vauhtia.

Mittakaavatasoilla 1:8000000 - 1:32000 hakemistotiedoston tekeminen on oikeastaan turhaa, koska noilla tasoilla ei ole kuin yksi ainoa karttatiedosto, ja MapServerille voitaisiin antaa suorat ohjeet lukea juuri noita tiedostoja, mikä mikä olisi myös hieman nopeampaa kuin tutkia ensin shapefile-hakemistoa. Ainoa syy shapefile-indeksin käyttämiseen jokaisella mittakaavatasolla on laiskuus, sillä sen ansiosta voidaan käyttää enemmän leikkaa-liimaa -menetelmää ja käsityön määrä vähenee.

Tileindex -hakemistotiedostojen luominen

Tehdään E: -levyaseman juureen uusi hakemisto "tindex" ja annetaan GDAL-komentoikkunasta tässä hakemistossa seuraavat 11 komentoa:

```
gdaltindex -write_absolute_path t8000.shp \t8000\*.tif
gdaltindex -write_absolute_path t4000.shp \t4000\*.tif
gdaltindex -write_absolute_path t2000.shp \t2000\*.tif
gdaltindex -write_absolute_path t800.shp \t800\*.tif
gdaltindex -write_absolute_path t320.shp \t320\*.tif
gdaltindex -write_absolute_path t160.shp \t160\*.tif
```



```
gdaltindex -write_absolute_path t80.shp \t80\*.tif
gdaltindex -write_absolute_path t40.shp \t40\*.tif
gdaltindex -write_absolute_path t20.shp \t20\*.tif
gdaltindex -write_absolute_path t10.shp \t10\*.tif
gdaltindex -write_absolute_path t5.shp \t5\*.tif
```

Tämän jälkeen luodaan näille shapefile-tiedostoille spatiaali-indeksit shptree-komennoilla.

```
shptree t8000
shptree t4000
shptree t2000
shptree t800
shptree t320
shptree t160
shptree t80
shptree t40
shptree t20
shptree t10
shptree t5
```

MapServer-WMS-palvelun koeajot

Alla olevassa taulukossa esitetään JMeter-testausohjelmalla saadut tulokset WMS-palvelimen suorituskyvystä. Mittaukset on tehty INSPIRE-latauspalveluiden testaamisesta annettujen teknisten ohjeiden mukaisesti, eli yhtäaikaisten käyttäjien lukumäärä oli 20, 8-bittisen tuloskuvat koko 600x800 pikseliä, ja BBOX-parametri on jokaisella pyynnöllä eri alueelta ja sen maantieteellinen koko vaihtelee satunnaisesti. Tarkempia tietoja testauksesta saa toisesta latuviitan ohjeesta.

Vertailtavat asiat testissä olivat MapServer-ohjelman käyttötapa (cgi tai fast-cgi), tuloskuvan näytteistämiseen käytetty menetelmä (nearest, bilinear tai average) sekä laskennassa käytetty ylinäytteistysuhde (ei ylinäytteistystä tai 2,5-kertainen ylinäytteistys).

näytteistys- menetelmä	näytteitä ms	keskiarvo ms	mediaani ms	90% ms	minimi ms	maksimi ms	kuvaa/ sek.	kt/ sek.
cgi-moodi								
Oversample =1,0								
nearest	2000	550	475	961	121	2060	34,3	1852,3
bilinear	2000	933	860	1420	233	3332	20,7	3063,7
average	2000	1031	942	1636	260	3573	18,4	1947,4
Oversample =2,5								
nearest	2000	750	607	1404	142	3073	25,2	961,0
bilinear	2000	1118	981	1914	243	5499	17,3	2250,6
average	2000	1106	1011	1792	254	3315	17,4	2266,2

fast-cgi-moodi

Oversample =1,0

nearest	2000	273	153	406	49	8308	63,7	3439,7
bilinear	2000	573	292	658	137	24751	27,9	4128,9
average	2000	719	494	1187	139	15763	23,6	2500,9

Oversample =2,5

nearest	2000	438	304	969	50	6549	40,5	1542,4
bilinear	2000	792	666	1389	152	8949	23,3	3031,6
average	2000	1010	755	1926	150	8129	17,8	1763,1

INSPIRE-vaatimus suorituskvyn suhteen on korkeintaan 5 sekunnin viive pyynnön vastaanottamisen ja vastauksen alkamishetken välillä normaalitilanteessa. Normaalitilanne puolestaan tarkoittaa joukkoa, johon kuuluu vasteaikonen nopeimmat 90 %. Tämän vaatimuksen testikone selvitti kirukkaasti, sillä 90 % vasteaika oli alle 2 sekuntia kaikilla testatuilla asetuksilla.

INSPIRE:n kapasiteettivaatimuksen mukaan suorituskvyyvaatimusten on täytyttävä 20 yhtäaikaisella käyttäjällä, eli alun 5 sekunnin mittaisen kiihdytysajan jälkeen palvelusta täytyy saada kestävästi ulos 20 kuvaa sekunnissa. Testipalvelimen kapasiteetti vaihteli eri asetuksilla välillä 17,3 kuvaa sekunnissa – 63,7 kuvaa sekunnissa. Nopeimmilla asetuksilla INSPIRE-vaatimukset siis ylitettiin kolminkertaisesti.

Prosessointiasetusten vaikutukset palvelun nopeuteen ja laatuun

Cgi-toimintamuoto on kertakäyttöinen: prosessi luodaan, sillä tehdään tilattu työ ja prosessi lopetetaan. WMS:n yhteydessä tämä merkitsee yhtä prosessia jokaista WMS-pyyntöä kohti. Fast-cgi-toimintamuodossa sen sijaan perustetaan pitkäkestoinen prosessi, joka vastaa toistuvasti sille tehtäviin pyyntöihin. Fast-cgi-prosessin perustaminen on kalliimpaa kuin cgi-prosessin perustaminen, mutta kun fast-cgi-prosessi on valmis vastaanottamaan työmääräyksiä, niin se toteuttaa niitä nopeammin. Taulukon numeroista huomataan selvästi, että mitä enemmän tietokoneen on tehtävä raskasta laskentaa ja mitä kauemmin tehtävän suoritus kestää, sitä pienemmäksi tulee cgi:n ja fast-cgi:n nopeusero. Cgi toimii myös paljon tasaisemmin ja ero nopeimpien ja hitaimpien suoritusten välillä on pieni ja hitaimmat yksittäiset suoritukset saadaankin, ehkä yllättäen, fast-cgi:stä.

Näytteistysmenetelmien laskennan raskaudessa on eroa ja menetelmät vaikuttavat laatuun. Ei kuitenkaan ole ollenkaan selvää, kuinka suuria ja millaisia laatuvaikutukset ovat, sillä lähtöaineiston ominaisuuksilla on suuri vaikutus eikä sanonta ”hitaasti hyvä tulee” sovi ainoaksi ohjeeksi. Parasta on tehdä vertailutestejä todellisilla aineistoilla ja todellisessa käyttötilanteessa.

Ylinäytteistämällä tarkoitetaan sitä, että jos esimerkiksi ylinäytteistyssuhde on 2,5, niin 10x10 pikselin kuvapala prosessoidaan 25x25 pikselin tarkkuudella. Tämä parantaa usein laatua erityisesti jos muunnos johtaa kuvan kiertämiseen, koska mm. kirjaimiin suoriin viivoihin ei synny yhtä selvästi erottuvia portaita. Ylinäytteistyssuhde 1 tarkoittaa, että kuvaa ei ylinäytteistetä ja luvuista nähdään, että se nopeuttaa prosessointia. Ylinäytteistämisen ei pitäisi koskaan huonontaa laatua,

mutta se, kuinka paljon se parantaa laatua, selviää vain kokeilemalla.

Kuvien prosessoinnissa on paljon samaa, tehtiinpä sitä minkä merkkisellä ohjelmalla tahansa, joten yllä olevat kommentit eivät koske pelkästään MapServer-palvelinta.

Nopeammin, paremmin, halvemmalla eli tulosten tarkastelu

Tässä ohjeessa on käsitelty varsin isoa kokonaisuutta, johon kuuluu monia erilaisia paikkatietojen käsittelyvaiheita. Jokaisessa vaiheessa on varmasti parantamisen varaa, mutta ohjeeseen ei kuitenkaan ole otettu mukaan mitään ihan toivottoman huonoja menetelmiä. Siitä syystä yhden yksityiskohtaan optimoiminen voi muuttaa jotain toista yksityiskohtaa huonompaan suuntaan. Kirjoittajan mielestä ohje kokonaisuudessaan on varsin hyvässä tasapainossa.

Nopeammin

Jos aloitetaan lähtöaineiston hankkimisesta, niin Rsync on hieno menetelmä, jos joku vain tarjoaa sekä palvelun ylläpidon että verkkoyhteydet. Toinen hyvin toimiva tapa näin suurten aineistojen hankkimiseen on Bit Torrent, mutta nopea se on vain silloin, kun torrent on virkeä eli verkossa on paljon jakajia ja lataajia. Rsyncin hienous on mahdollisuus hakemistojen synkronointiin ja sitä kautta osittain päivitettyjen aineistojen ajan tasalla pitämiseen.

Tiedostojen muuntaminen GDAL-ohjelmilla on nopeaa, mutta silti yhdellä hyvälläkin tietokoneella voi tehdä tiukkaa muuntaa kaikki 16962 tiedostoa yhden vuorokauden aikana. Muunnos olisi tarpeeton, jos lähtöaineisto sopisi ilman muunnoksia WMS-palvelun tekemiseen, mutta suuremman tiedostokoon takia sitten olisi enemmän ladattavaa. Yleismaailmallista, kaikki käyttäjät tyytyväiseksi tekevää tiedostomuotoa, ei kuitenkaan ole olemassa, joten muunnosten tekemistä voi olla mahdoton välttää. Kovassa kiireessä kannattaa harkita ladattujen png-kuvien jakamista esimerkiksi esimerkiksi tuhannen kuvan eriin omille USB-levyilleen ja aloittaa niiden muuntaminen toisilla tietokoneilla vielä rsync-latauksen kestäessä. Kotoa tai konttorista luultavasti löytyy muutama kone, jotka eivät ole yötä päivää varattuja.

Ohjeessa kuvattuun prosessiin kuuluu useita MapServeriä varten tehtyjä optimointeja. Muunnos geotiffeiksi ja yleistettyjen tasojen laskeminen ei periaatteessa ole välttämätöntä, koska MapServer kyllä osaa lukea myös png-kuvia. Tehonhukka on kuitenkin niin suuri, että ei siinä kyllä ole mitään järkeä.

Jonkin verran järkeä voisi olla siinä, että ensimmäisessä vaiheessa tehtäisiin vain muunnos geotiff-muotoon ja WMS-palvelu käynnistettäisiin niiden varassa heti kun kuvat ovat kasassa. Gdaladdo-osuus voitaisiin sitten tehdä tuotannon aikana. Tällä tavalla WMS-palvelu voitaisiin avata käyttöön muutamaa tuntia aikaisemmin. Muutamia kymmeniä prosessoriminuutteja säästyisi myös, jos gdaladdo-ohjelmalla käytettäisiin nopeinta nearest-menetelmällä, mutta laatu kyllä huononisi silmännähdän.

MapServerin tuotannonaikaisen nopeuden parantamiseksi voidaan tämän mallisuorituksen ja pikaisen testauksen perusteella esittää lähinnä arveluita.

- Testikone oli tavallinen, vaikkakin tehokas toimistokannettava, jossa oli tavallinen 64-bittinen Windows 7 -käyttöjärjestelmä. Erityisesti palvelinkäyttöön suunniteltu tietokone ja käyttöjärjestelmä saattaisi olla nopeampi.
- MapServer:in on monissa testeissä todettu olevan Linuxilla selvästi nopeampi kuin Windowsilla.

- Lähtökuvat olivat ulkoisella USB-levyllä, sisäinen kovalevy saattaisi olla nopeampi, SSD-levystä puhumattakaan. Tosin INSPIRE-testauksen aikana prosessorikuorma pysytteli 100 prosentin tasolla, joten levy ei ilmeisesti ollut pullonkaulana.
- Lähtökuvat olivat pakattuja, pakkaamisen purkaminen vaatii prosessorivoimaa ja testin aikana koneen kaikki ytimet kävivät 100 % teholla. Pakkaamattomien kuvien käyttö voisi olla nopeampaa. Varjopuolena olisi se, että 100 gigatavun sijasta tarvittaisiin yli teratavu levytilaa.
- INSPIRE-testissä käytettyä JMeter-testiohjelmaa, jota ajettiin samalla koneella kuin MapServeriä, rasitti epäilemättä sekin tietokonetta. Hidastumisen huomasi selvästi, jos testi suoritettiin niin, että JMeter piirsi koko ajan graafista kuvaajaa testin aikana.

Paremmiin

- Suurin laadun huononeminen syntyy todennäköisesti siitä, että jos toimitaan ohjeen mukaan, niin gdaladdo ohjelman ”average” -menetelmällä tehdyt yleistetyt kuvaversiot tarjotaan ulos MapServeristä toiseen kertaan uudelleennäyteistettynä. Yhteen kertaan tehty uudelleennäyteistys voisi johtaa näkyvästikin parempaan lopputulokseen. Tämän voi halutessaan selvittää helposti jättämällä gdaladdo-vaiheen tekemättä ja tuottamalla WMS-palvelun pelkästään laatoitetuista tiff-kuvista. Hidastuminen ei välttämättä olisi kovin suurta, koska taustakarttasarjan 11 eri mittakaavaa toimivat jo itsessään pyramiditasoina.
- MapServer pystyy kirjoittamaan tuloskuvat useisiin eri tiedostomuotoihin ja niiden versioihin. Toiset niistä ovat varmasti parempilaatuisia kuin jotkin muut, mutta tässä testissä kokeiltiin vain INSPIRE-testausohjeiden mukaista 8-bittistä png-tiedostomuotoa.

Halvemmallalla

- Kaikki käytetyt ohjelmistot ovat avointa lähdekoodia joten taustakartta-WMS:n pystyttämiseen ei silkkää rahaa kulu ollenkaan. Koska muunnokset ovat nopeita ja komentojonojen avulla automatisoitavissa niin ihmistyötäkään ei paljon tarvita. Laitteistokustannukset ovat myös pienet, koska kaikki työvaiheet onnistuvat melkein millä tahansa tietokoneella. Rasteritaustakartta-aineiston WMS-palvelu saatiin kevyesti täyteen INSPIRE-vauhtiin noin 1800 euron hintaisella kannettavalla (vuoden 2012 hintataso). Muutamaa satasta halvemmalla voisi selvittää hankkimalla pöytä- tai palvelinkone, jossa on hyvä prosessori ja kovalevy mutta halpa näytönohjain. Todennäköisesti joillakin muilla avoimen lähdekoodin ohjelmistoilla voidaan päästä suunnilleen samassa ajassa yhtä hyvään suorituskykyyn. Kaupalliset muunnos- ja palvelinohjelmistot tuskin pärjäävät hintakilpailussa, vaikka myös työaika- ja laitteistokustannukset otettaisiin huomioon, joten niiltä odotetaan parempaa vauhtia ja laatua. Laadun voi määritellä myös niin, että se sisältää tukipalvelut ja mielikuvat luotettavuudesta ja jatkuvuudesta.

Hyödyllisiä linkkejä

Maanmittauslaitoksen taustakarttasarja

Aineiston kuvaus Maanmittauslaitoksen sivustolla

<http://www.maanmittauslaitos.fi/digituotteet/taustakarttasarja-rasteri>

INSPIRE-katselupalvelun suorituskykyvaatimukset

http://inspire.jrc.ec.europa.eu/documents/Network_Services/TechnicalGuidance_ViewServices_v3.1.pdf

Mapserverin mapfile-ohjaustiedoston teko

MapServerin käytön opettaminen ei kuulu tämän ohjeen aiheisiin, joten valmis ja toimivaksi testattu mapfile annetaan liitteenä. Seuraavassa on kuitenkin muutamia kommentteja MapServerissä, mapfilestä ja sillä aikaansaataavasta WMS-taustakarttapalvelu.

- Tämä rivi määrää, mihin MapServer kirjoittaa palvelun virhelokin
CONFIG "MS_ERRORFILE" "c:\ms4w\apps\log\taustakartta.txt"
- Lokitiedoston hakemisto on oltava olemassa, joten joko luo hakemisto **c:\ms4w\apps\log** tai muokkaa MS_ERRORFILE-parametria niin, että se viittaa johonkin olemassa olevaan hakemistoon
- DEBUG on oleellinen asia, jos jokin ei toimi. Ota kommenttimerkki "#" pois, ja katso lokitiedostosta, mikä menee pieleen.
- Tämän rivin kuuluu vastata todellista tilannetta:
"wms_onlinerresource" "http://localhost:8060/cgi-bin/mapserv.exe?
map=c:\ms4w\apps\maps\taustakartta_wms.map"
- Muokkaa onlinerresource-riviä, jos Mapserver (eli Apache-palvelin) on laitettu kuuntelemaan jotain muuta porttia kuin 8060, tai jos mapfile on tallennettu johonkin toiseen hakemistoon tai toisella nimellä.
- Taustakartan 11 eri mittakaavatasoa on yhdistetty GROUP-tasoksi "taustakartta". WMS-ohjelmat voivat pyytää aina tuota ryhmätasoa. Mapfile on tehty niin, että ryhmätaso käyttää kaikissa mittakaavoissa vain yhtä lähtöaineistoa. Tätä säädetään asetuksilla MINSCALEDENOM ja MAXSCALEDENOM, esimerkiksi
MINSCALEDENOM 1400000
MAXSCALEDENOM 3000000
- PROCESSING "RESAMPLE=BILINEAR" parantaa pikakokeen perusteella laatua näillä lähtöaineistoilla. Bilinear-näytteistäminen on hieman hitaampaa kuin oletuksena käytettävä "NEAREST".

Liite 1.

MapServer mapfile, joka tuottaa Maanmittauslaitoksen taustakarttasarjasta WMS-palvelun.
Optimoitu vauhtia varten seuraavilla parametreillä:

```
PROCESSING "RESAMPLE=NEAREST"  
PROCESSING "CLOSE_CONNECTION=DEFER"  
PROCESSING "OVERSAMPLE_RATIO=1.0"
```

```
=====  
MAP  
  NAME ""  
  STATUS ON  
  MAXSIZE 5000  
  #Huom! Lokitiedoston tallennushakemisto on oltava olemassa  
  CONFIG "MS_ERRORFILE" "c:\ms4w\apps\log\taustakartta.txt"  
  EXTENT 19970 6597278 749676 7785300  
  SIZE 600 600  
  SHAPEPATH "c:/ms4w/tmp/ms_tmp"  
  IMAGECOLOR 255 255 255  
  UNITS METERS  
# DEBUG 5  
WEB  
  IMAGEPATH "c:/ms4w/tmp/ms_tmp"  
  IMAGEURL "/ms_tmp/"  
  METADATA  
    "ows_enable_request" "*"   
    "wms_srs" "EPSG:3067 EPSG:2393 EPSG:4326"  
    "wms_onlineresource" "http://localhost:8060/cgi-bin/mapserv.exe?"  
map=\ms4w\apps\maps\taustakartta_wms.map"  
    "wms_title" "Latuviitan MML-taustakarttademo-WMS"  
    "wms_bbox_extended" "true"  
  END  
END  
PROJECTION  
  "init=epsg:3067"  
END  
OUTPUTFORMAT  
  NAME 'AGG_Q'  
  MIMETYPE "image/png"  
  DRIVER "AGG/PNG"  
  EXTENSION "png"  
  IMAGEMODE RGB  
  FORMATOPTION "QUANTIZE_FORCE=ON"  
  FORMATOPTION "QUANTIZE_DITHER=OFF"  
  FORMATOPTION "QUANTIZE_COLORS=256"  
END  
OUTPUTFORMAT  
  NAME "GEOTIFF"  
  DRIVER "GDAL/GTiff"  
  MIMETYPE "image/tiff"  
  IMAGEMODE "BYTE"  
  EXTENSION "tif"  
END  
LAYER  
  NAME "taustakartta_1"  
  GROUP "taustakartta"  
  STATUS ON  
  PROCESSING "RESAMPLE=NEAREST"  
  PROCESSING "CLOSE_CONNECTION=DEFER"  
  PROCESSING "OVERSAMPLE_RATIO=1.0"  
  TILEINDEX "e:\tindex\t8000"
```

```

# tai käytä suoraa viittausta kuvaan kuten seuraavalla rivillä
#DATA "e:\t8000\8000000.tif"
TYPE RASTER
#DEBUG 5
UNITS METERS
SIZEUNITS PIXELS
MINSCALEDENOM 3000000
METADATA
    "wms_title" "taustakartta_1"
    "wms_srs" "EPSG:3067 EPSG:2393 EPSG:4326"
END
PROJECTION
    "init=epsg:3067"
END
END
LAYER
    NAME "taustakartta_2"
    GROUP "taustakartta"
    STATUS ON
    PROCESSING "RESAMPLE=NEAREST"
    PROCESSING "CLOSE_CONNECTION=DEFER"
    PROCESSING "OVERSAMPLE_RATIO=1.0"
    TILEINDEX "e:\tindex\t4000"
    TYPE RASTER
    #DEBUG 5
    UNITS METERS
    SIZEUNITS PIXELS
    MINSCALEDENOM 1400000
    MAXSCALEDENOM 3000000
    METADATA
        "wms_title" "taustakartta_2"
        "wms_srs" "EPSG:3067 EPSG:2393 EPSG:4326"
    END
    PROJECTION
        "init=epsg:3067"
    END
END
LAYER
    NAME "taustakartta_3"
    GROUP "taustakartta"
    STATUS ON
    PROCESSING "RESAMPLE=NEAREST"
    PROCESSING "CLOSE_CONNECTION=DEFER"
    PROCESSING "OVERSAMPLE_RATIO=1.0"
    TILEINDEX "e:\tindex\t2000"
    TYPE RASTER
    #DEBUG 5
    UNITS METERS
    SIZEUNITS PIXELS
    MINSCALEDENOM 530000
    MAXSCALEDENOM 1400000
    METADATA
        "wms_title" "taustakartta_3"
        "wms_srs" "EPSG:3067 EPSG:2393 EPSG:4326"
    END
    PROJECTION
        "init=epsg:3067"
    END
END
LAYER
    NAME "taustakartta_4"
    GROUP "taustakartta"
    STATUS ON
    PROCESSING "RESAMPLE=NEAREST"

```

```

PROCESSING "CLOSE_CONNECTION=DEFER"
PROCESSING "OVERSAMPLE_RATIO=1.0"
TILEINDEX "e:\tindex\t800"
TYPE RASTER
#DEBUG 5
UNITS METERS
SIZEUNITS PIXELS
MINSCALEDENOM 216000
MAXSCALEDENOM 530000
METADATA
    "wms_title" "taustakartta_4"
    "wms_srs" "EPSG:3067 EPSG:2393 EPSG:4326"
END
PROJECTION
    "init=epsg:3067"
END
END
LAYER
    NAME "taustakartta_5"
    GROUP "taustakartta"
    STATUS ON
    PROCESSING "RESAMPLE=NEAREST"
    PROCESSING "CLOSE_CONNECTION=DEFER"
    PROCESSING "OVERSAMPLE_RATIO=1.0"
    TILEINDEX "e:\tindex\t320"
    TYPE RASTER
    #DEBUG 5
    UNITS METERS
    SIZEUNITS PIXELS
    MINSCALEDENOM 80000
    MAXSCALEDENOM 216000
    METADATA
        "wms_title" "taustakartta_5"
        "wms_srs" "EPSG:3067 EPSG:2393 EPSG:4326"
    END
    PROJECTION
        "init=epsg:3067"
    END
END
LAYER
    NAME "taustakartta_6"
    GROUP "taustakartta"
    STATUS ON
    PROCESSING "RESAMPLE=NEAREST"
    PROCESSING "CLOSE_CONNECTION=DEFER"
    PROCESSING "OVERSAMPLE_RATIO=1.0"
    TILEINDEX "e:\tindex\t160"
    TYPE RASTER
    #DEBUG 5
    UNITS METERS
    SIZEUNITS PIXELS
    MINSCALEDENOM 49600
    MAXSCALEDENOM 80000
    METADATA
        "wms_title" "taustakartta_6"
        "wms_srs" "EPSG:3067 EPSG:2393 EPSG:4326"
    END
    PROJECTION
        "init=epsg:3067"
    END
END
LAYER
    NAME "taustakartta_7"
    GROUP "taustakartta"

```



```

STATUS ON
PROCESSING "RESAMPLE=NEAREST"
PROCESSING "CLOSE_CONNECTION=DEFER"
PROCESSING "OVERSAMPLE_RATIO=1.0"
TILEINDEX "e:\tindex\t80"
TYPE RASTER
#DEBUG 5
UNITS METERS
SIZEUNITS PIXELS
MINSCALEDENOM 28800
MAXSCALEDENOM 49600
METADATA
    "wms_title" "taustakartta_7"
    "wms_srs" "EPSG:3067 EPSG:2393 EPSG:4326"
END
PROJECTION
    "init=epsg:3067"
END
END
LAYER
    NAME "taustakartta_8"
    GROUP "taustakartta"
    STATUS ON
    PROCESSING "RESAMPLE=NEAREST"
    PROCESSING "CLOSE_CONNECTION=DEFER"
    PROCESSING "OVERSAMPLE_RATIO=1.0"
    TILEINDEX "e:\tindex\t40"
    TYPE RASTER
    #DEBUG 5
    UNITS METERS
    SIZEUNITS PIXELS
    MINSCALEDENOM 16800
    MAXSCALEDENOM 28800
    METADATA
        "wms_title" "taustakartta_8"
        "wms_srs" "EPSG:3067 EPSG:2393 EPSG:4326"
    END
    PROJECTION
        "init=epsg:3067"
    END
END
LAYER
    NAME "taustakartta_9"
    GROUP "taustakartta"
    STATUS ON
    PROCESSING "RESAMPLE=NEAREST"
    PROCESSING "CLOSE_CONNECTION=DEFER"
    PROCESSING "OVERSAMPLE_RATIO=1.0"
    TILEINDEX "e:\tindex\t20"
    TYPE RASTER
    #DEBUG 5
    UNITS METERS
    SIZEUNITS PIXELS
    MINSCALEDENOM 10400
    MAXSCALEDENOM 16800
    METADATA
        "wms_title" "taustakartta_9"
        "wms_srs" "EPSG:3067 EPSG:2393 EPSG:4326"
    END
    PROJECTION
        "init=epsg:3067"
    END
END
LAYER

```

```
NAME "taustakartta_10"
GROUP "taustakartta"
STATUS ON
PROCESSING "RESAMPLE=NEAREST"
PROCESSING "CLOSE_CONNECTION=DEFER"
PROCESSING "OVERSAMPLE_RATIO=1.0"
TILEINDEX "e:\tindex\t10"
TYPE RASTER
#DEBUG 5
UNITS METERS
SIZEUNITS PIXELS
MINSCALEDENOM 3500
MAXSCALEDENOM 10400
METADATA
    "wms_title" "taustakartta_10"
    "wms_srs" "EPSG:3067 EPSG:2393 EPSG:4326"
END
PROJECTION
    "init=epsg:3067"
END
```

END

LAYER

```
NAME "taustakartta_11"
GROUP "taustakartta"
STATUS ON
PROCESSING "RESAMPLE=NEAREST"
PROCESSING "CLOSE_CONNECTION=DEFER"
PROCESSING "OVERSAMPLE_RATIO=1.0"
TILEINDEX "e:\tindex\t5"
TYPE RASTER
#DEBUG 5
UNITS METERS
SIZEUNITS PIXELS
MAXSCALEDENOM 3500
METADATA
    "wms_title" "taustakartta_11"
    "wms_srs" "EPSG:3067 EPSG:2393 EPSG:4326"
END
PROJECTION
    "init=epsg:3067"
END
```

END

END