

Vihjeitä GeoPackage-tietokannan tehokkaaseen käyttöön

Käyttäjää kehoitetaan ensisijaisesti tutustumaan käyttämiensä ohjelmistojen omiin ohjeisiin, koska ne ovat parhaiten ajan tasalla ja niissä otetaan huomioon ohjelmistokohtaisten toteutusten erikoispiirteet. Se, että GeoPackage on standardi, takaa perusasioiden toimimisen kaikissa ohjelmistoissa, mutta ei sitä, että kaikki ohjelmistot toimisivat joka suhteessa samalla tavalla ja yhtä tehokkaasti.

GeoPackage-standardin suhteen paras tietolähde on sivusto <http://www.geopackage.org> josta löytyy myös linkit OGC:n standardien GeoPackage-sivulle sekä GitHub-sivustolle, jossa tapahtuu standardin tulevien versioiden kehitystyö sekä jo julkaistujen versioiden virheiden käsittely. GeoPackagen pohjalla olevan SQLite-tietokannan ohjeistus on projektin sivustolla <https://sqlite.org/index.html>.

Seuraavat vihjeet on kirjoitettu auttamaan käyttäjiä maastotietokannan GeoPackage-julkaisun käyttöönotossa ja ne perustuvat käytännön kokemuksiin kirjoitushetkellä (huhtikuu 2020). Maanmittauslaitos ei tarjoa muuta tukea ohjelmistojen käyttöön.

Yleisiä ohjeita

1. Älä lue kartalle liian paljon kohteita

Maastotietokannan GeoPackage-tietokannan taulussa "maastokuvionreuna" on yli 10 miljoonaa viivaa, taulussa "korkeuskayra" yli 6,8 miljoonaa viivaa ja taulussa "rakennus" on yli 5,4 miljoonaa polygonia. Jos paikkatieto-ohjelma yrittää lukea ja näyttää tällaisten taulujen kaikki kohteet kartalla niin se toimii todennäköisesti hyvin hitaasti. Aseta ohjelmistoosi mittakaavaraja niin että kohteet näytetään vain silloin kun niitä osuu ruudulle kohtuullinen määrä. Sopiva mittakaavaraja riippuu taulusta. Pienillä tauluilla ei tarvita mitään rajoituksia, kun taas korkeuskäyriä tuskin kannattaa näyttää pienemmässä mittakaavassa kuin 1:50000. Järville voi olla hyvä lisätä uusi attribuutti ja tallentaa siihen pinta-alat, jolloin pienissä mittakaavoissa voidaan näyttää vain suurimmat järvet.

2. Ohjelmistosi ei ehkä osaa käyttää spatiaali-indeksiä

Kaikille maastotietokannan GeoPackage-julkaisun tauluille on luotu spatiaali-indeksit, jotka mahdollistavat kohteiden löytämisen nopeasti pieneltä rajatulta alueelta. Indeksien olemassaolo ei kuitenkaan välttämättä tee hakuja yhtään nopeammaksi, koska spatiaali-indeksi on GeoPackage-standardin laajennus, jota sovellusten ei ole pakko tukea. Jos suurimmat tietokantataulut eivät toimi sovelluksellasi nopeasti suoraan GeoPackage-tietokannasta, niin muunna tiedot johonkin sovellukselle paremmin sopivaan muotoon. Työketju on tässä tapauksessa sama kuin GML-tiedostoja käytettäessä, mutta 3000 tiedoston lataamisen ja yhdistämisen sijasta riittää, kun lataa yhden tiedoston.

Ohjelmistokohtaisia huomautuksia

ESRI:n ohjelmistot

GeoPackage-tietokantoja voidaan käyttää parhaiten siirtoformaattina ArcMap- ja ArcGIS Pro -ohjelmille. Molemmat ohjelmistot pystyvät myös lisäämään karttatasoja työtilaan suoraan GeoPackage-tietokannasta. Tämä toimii hyvin pienillä tietokantatauluilla, mutta suurimmat maastotietokannan taulut kannattaa tallentaa ensin ESRI:n omiin formaatteihin, koska ESRI:n ohjelmat toimivat niillä paljon nopeammin. ESRI:n ohjelmilla ei voi muokata GPKG-tietokannassa olevia kohteita, mutta uuden GeoPackage-tietokannan luominen on mahdollista.

Lähde: ArcGIS-blogi <https://www.esri.com/arcgis-blog/products/product/data-management/how-to-use-ogc-geopackages-in-arcgis-pro/>

MapInfo

MapInfo Pro pystyy lukemaan karttatasoja GPKG-tietokannasta. Myös kohteiden muokkaus ja uusien GeoPackage-tietokantojen luominen on mahdollista.

Lähde: <https://www.geopackage.org/implementations.html>

QGIS

GeoPackage on QGIS-ohjelman ensisijainen vektoriformaatti ohjelman versiosta 3.0 alkaen, mistä syystä ohjelma tukee GeoPackage-formaattia hyvin. Maastotietokannan suurimpienkin taulujen käyttö suoraan GPKG-tietokannasta on nopeaa, kunhan karttatasojen piirtämiselle asetetaan järkevä mittakaavarajoitus.

Lähde: <https://qgis.org/fin/docs/index.html>

FME

FME pystyy sekä lukemaan että kirjoittamaan GeoPackage-tietokantoja. FME:stä tarvitaan vähintään "Desktop Professional" -lisenssiversio. FME:n versioon 2020.0 on tehty hyödyllisiä parannuksia siihen, miten uudet GeoPackage-tietokannat luodaan, mutta myös vanhemmat versiot tuottavat standardin mukaisia tietokantoja.

Lähde: https://docs.safe.com/fme/html/FME_Desktop_Documentation/FME_ReadersWriters/ogcgeopackage/ogcgeopackage.htm

GDAL

GDAL-ohjelmakirjastoon kuuluvan ogr2ogr-ohjelman avulla voidaan lukea, kirjoittaa ja muokata GeoPackage-tietokannassa olevaa vektoridataa. Tulos on standardin mukainen, mutta lisäksi GDAL luo GPKG-tietokantoihin standardin ulkopuolisen metatietotaulun, josta on hyötyä erityisesti QGIS-ohjelmalle. Tästä taulusta ei ole haittaa muille ohjelmistoille. Ohjelmallisesti samat toiminnot on mahdollista toteuttaa käyttämällä GDAL-kirjastoa sekä sen C++, C#, Python- ja Java-ohjelmointirajapintoja.

Lähde: <https://gdal.org/drivers/vector/gpkg.html>

GeoServer

GeoPackage sopii GeoServerin WMS- ja WFS-palveluiden tietolähteeksi.

Lähde: <https://docs.geoserver.org/latest/en/user/data/vector/geopkg.html>

MapServer

GeoPackage sopii MapServerin WMS- ja WFS-palveluiden tietolähteeksi. Yhteys tietokantaan muodostetaan OGR-vektoritasona.

Lähde: <https://www.mapserver.org/input/vector/ogr.html>

Huomioita GeoPackage-tietokannan nopeudesta

GeoPackage on pohjimmiltaan SQLite-tietokanta, joka on hyvin yleisesti käytetty tiedostopohjainen tietokanta. SQLite on nopea, mutta kuten tietokannoissa yleensäkin, suurista tauluista tehtävien kyselyiden suoritusajaksi vaikuttaa ratkaisevasti se, hyödyntääkö kysely indeksejä vai ei. Muutamia esimerkkejä mtkmaasto.gpkg-tietokannasta tehdyillä kyselyillä

Esimerkki 1. Yhden kohteen valinta indeksoidun kentän vs. indeksoimattoman kentän perusteella. Nopeusero yli 250-kertainen.

```
SELECT * FROM rakennus WHERE fid=5000001;
```

Suoritus aika: 0,032 sekuntia

```
SELECT * FROM rakennus WHERE mtk_id=1150058247;
```

Suoritus aika: 8,505 sekuntia

Esimerkki 2. Kohteiden valinta 1 km x 1 km suuruisen alueen sisältä (kohteita 284 kpl, koko taulussa kohteita 5462918 kpl) joko niin, että spatiaali-indeksiä (R*Tree-indeksiä) ei käytetä hyväksi, tai niin, että R*Tree-indeksiä käytetään alikyselyssä. Nopeusero yli 60-kertainen.

```
SELECT * FROM rakennus
```

```
WHERE ST_MaxX(geom)>=633070 AND ST_MinX(geom)<=634070
```

```
AND ST_MaxY(geom)>=6826718 AND ST_MinY(geom)<=6827718;
```

Suoritus aika: 8,468 sekuntia

```
SELECT * FROM rakennus
```

```
WHERE fid IN
```

```
(SELECT id FROM rtree_rakennus_geom
```

```
WHERE maxX>=633070 AND minX<=634070
```

```
AND maxY>=6826718 AND minY<=6827718);
```

Suoritus aika: 0,135 sekuntia

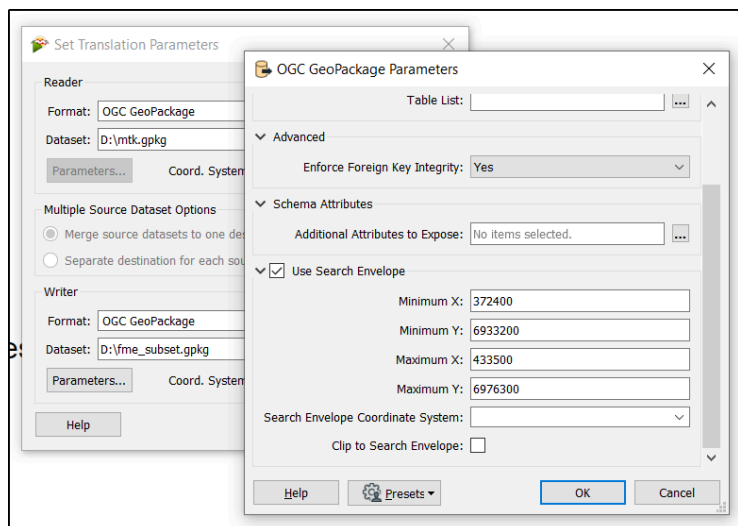
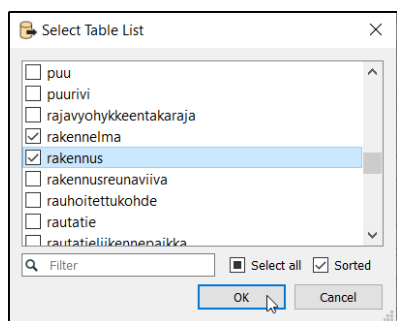
Mitatut ajat kuvaavat vain tietokantahaun nopeutta. Sovelluksissa tehdään tietokannasta haetuille tiedoille lisäksi prosessointia, jotta kohteet voidaan näyttää kartalla, joten käyttäjän kokema nopeus ole yhtä hyvä kuin puhdas tietokannan nopeus.

GeoPackage-tietokannan pienentäminen

Maastotietokannan GeoPackage-tietokantatiedoston kokoa voidaan pienentää poistamalla tietokannasta käyttäjälle tarpeettomia tauluja tai valitsemalla tietokannasta vain osa kohteista aluerajauksen perusteella. Levytilan säästämisen lisäksi ohjelmat, jotka eivät käytä GeoPackage-tietokantaa optimaalisella tavalla, voivat myös toimia nopeammin.

Taulujen valinta ja alueleikkaus FME Quick Translator -ohjelmalla

FME Quick Translator -ohjelmalla voidaan valita, mitkä GeoPackage-tietokantaan kuuluvat taulut otetaan mukaan uuteen GeoPackage-tietokantaan, sekä määrittää aluerajaus antamalla X- ja Y-koordinaattien minimi- ja maksimiarvot. Esimerkkinä rakennusten ja rakennelmien valinta Saarijärven alueelta.



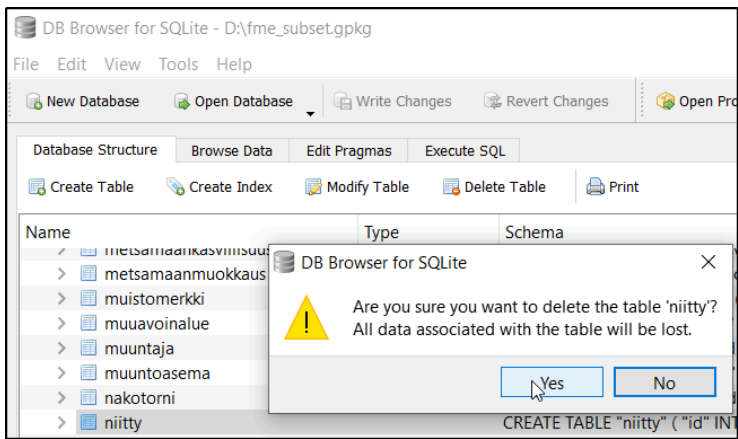
Taulujen valinta ja alueleikkaus GDAL ogr2ogr-ohjelmalla

Uuden GeoPackage-tietokannan muodostaminen Saarijärven rakennuksista ja rakennelmista voidaan tehdä komentoriviltä avoimen lähdekoodin ogr2ogr-ohjelmalla.

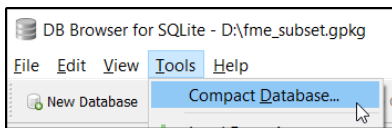
```
ogr2ogr -f gpkg -spat 372400 6933200 433500 6976300  
gdal_rakennus_leikattu.gpkg mtk.gpkg rakennus rakennelma
```

Taulujen poistaminen ja tyhjän tilan vapauttaminen ogrinfo-ohjelmalla

GeoPackage-tietokannan kokoa voidaan pienentää poistamalla siitä käyttäjälle tarpeettomia tauluja. Tämä tehdään kuten muissakin SQL:ää käytävissä tietokannoissa komennolla "DROP TABLE taulun_nimi;", ja komennon antamiseen voidaan käyttää mitä tahansa sovellusta tai menetelmää, jolla voidaan lähettää SQL-komentoja SQLite-tietokantaan. Esimerkiksi avoimen lähdekoodin ohjelma "DB Browser for SQLite" tarjoaa tätä varten graafisen käyttöliittymän.



Taulun poistaminen ei vielä vapauta levytilaa, vaan se tapahtuu vasta kun on suoritettu SQL komento "VACUUM". DB Browser for SQLite-ohjelmassa tätä varten on oma toiminto



Taulun poistaminen ja tietokannan tiivistäminen voidaan tehdä myös ogrinfo-ohjelmalla komentoriviltä. Kannattaa huomioida, että VACUUM-toiminto johtaa koko tietokantatiedoston kirjoittamiseen uudestaan ja siitä syystä kestää oman aikansa.

```
ogrinfo -sql "DROP TABLE niitty" fme_subset.gpkg
```

```
INFO: Open of `fme_subset.gpkg'  
using driver `GPKG' successful.
```

```
ogrinfo -sql "VACUUM" fme_subset.gpkg
```

```
INFO: Open of `fme_subset.gpkg'  
using driver `GPKG' successful.
```