

Cloud Optimized Geotiff

Open-source GIS Israel
By Guy Doulberg





About Guy Doulberg

- In the last 3 years working for Satellogic
- Leads the team that is responsible for delivering satellogic product



About Satellogic

- Argentinian startup
- Founded in 2011
- The mission is important to the rest of the talk

Satelloptic Accessible Satellite data

Accessible

We provide the most affordable and up-to-date geo-information analytics on planet Earth, enabling access to geospatial-driven insights at the right price point for our customers.

Frequency + Resolution

With strategic orbits and frequent revisits, we provide industry-best frequency at high resolution. Our 1-meter resolution is the sweet spot for monitoring economic activity and high-frequency change.



Satellogic Challenge

- How to store in our data
- How to retrieve our data
- Without being bankrupt
- In a safe place



Satelloptic - implementation

- Store rasters on Azure Blob Storage
- Store rasters as Cloud Optimized GeoTiff (COG)



COG words explained





Why cloud?





Cloud Storage is elastic, cheap and safe

- You pay only for what you are actually using
- When you pay it is cheap compared to other solution
- The cloud providers grant you some data-availability SLA - you don't need to manage it



Cloud storage price list

	PREMIUM	HOT	COOL	ARCHIVE
First 50 terabyte (TB) / month	\$0.195 per GB	\$0.0196 per GB	\$0.01 per GB	\$0.0023 per GB
Next 450 TB / Month	\$0.195 per GB	\$0.0189 per GB	\$0.01 per GB	\$0.0023 per GB
Over 500 TB / Month	\$0.195 per GB	\$0.0181 per GB	\$0.01 per GB	\$0.0023 per GB



**What need to be
Optimized?**





Data retrieval - performance

- COG is intended to be used on big areas of high resolution.
- The client would like to retrieve only the data that is required and no more than that.



Data retrieval - cost

Operations and data transfer prices

	PREMIUM	HOT	COOL	ARCHIVE
Write Operations (per 10,000) ¹	\$0.0228	\$0.054	\$0.10	\$0.12
List and Create Container Operations (per 10,000) ²	\$0.065	\$0.054	\$0.054	\$0.054
Read Operations (per 10,000) ³	\$0.0019	\$0.0043	\$0.01	\$6
All other Operations (per 10,000), except Delete, which is free	\$0.0019	\$0.0043	\$0.0043	\$0.0043
Data Retrieval (per GB) ⁴	Free	Free	\$0.01	\$0.024
Data Write (per GB) ⁴	Free	Free	\$0.0025	Free



Why GeoTiff?





GeoTiff - legacy support

- The community wanted existing systems will work without a change



COG description





COG description

A file format that uses:

- Tiling
- Overviews
- HTTP range
- IFD at the top of a file

Using the above techniques will optimize your data for the cloud

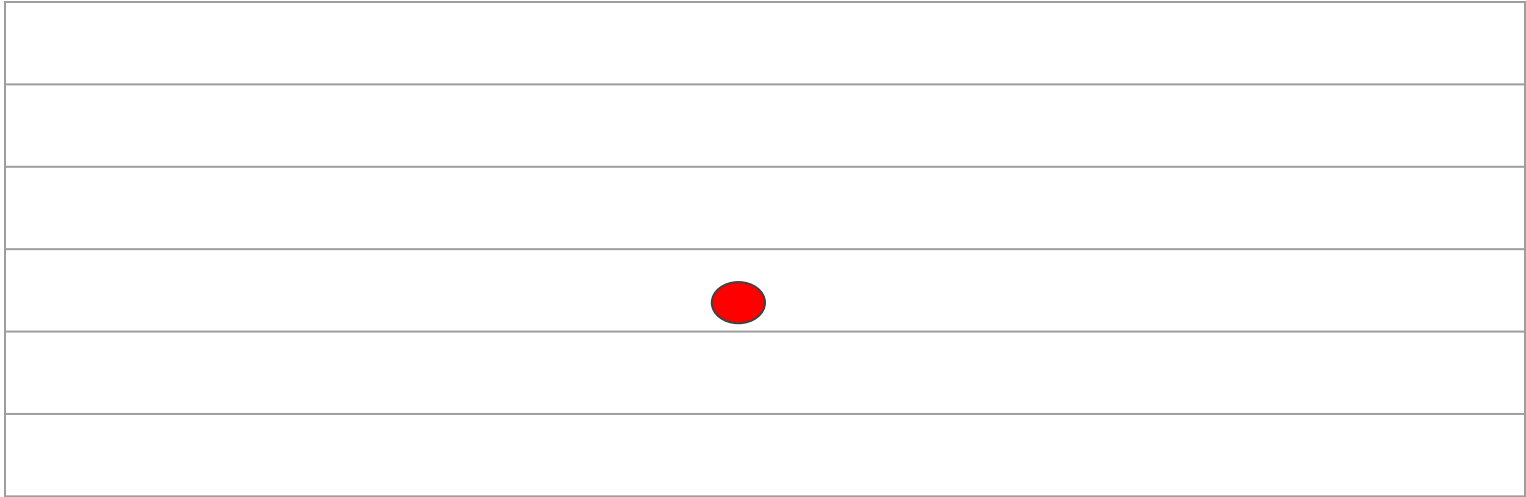


GeoTiff: Tiled

- A Tiled GeoTiff is a Tiff that uses tiles to store its data.
- Tiles instead of Stripes
- A Tile has width and height
- Common tiles sizes: 256X256, 512X512




GeoTiff: retrieving a pixel in a stripe





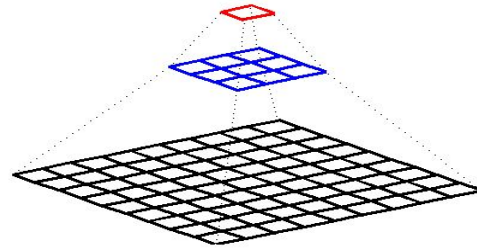
GeoTiff: retrieving a pixel in a tile

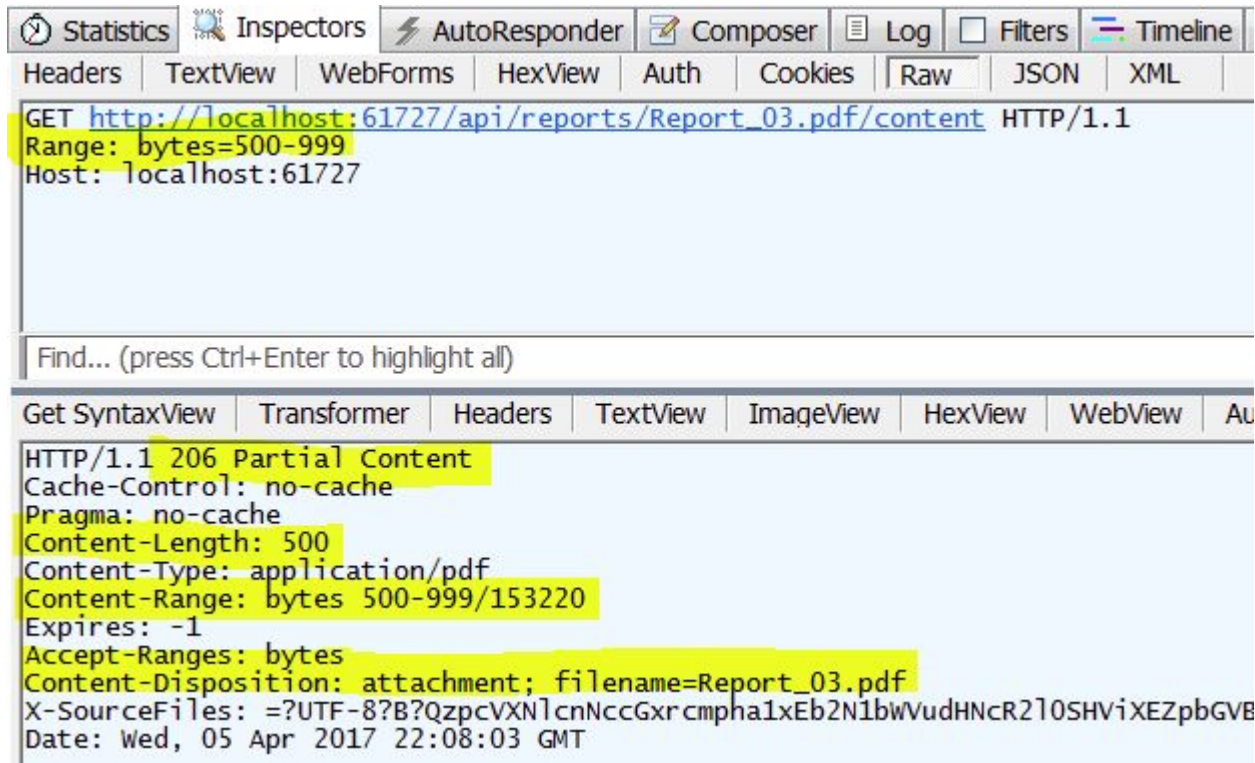


Geotiff: Overviews

- A Geotiff with overviews is a Geotiff constructed out of several Geotiffs.
- Each Geotiff is for a different resolution (pyramid)
- Overviews are used to cache resampled lower resolution pixels
-



HTTP Ranges request



The screenshot shows the 'Headers' tab of a web browser's developer tools. The request is a GET for `http://localhost:61727/api/reports/Report_03.pdf/content` with a `Range: bytes=500-999` header. The response is an `HTTP/1.1 206 Partial Content` with various headers including `Cache-Control: no-cache`, `Pragma: no-cache`, `Content-Length: 500`, `Content-Type: application/pdf`, `Content-Range: bytes 500-999/153220`, `Expires: -1`, `Accept-Ranges: bytes`, and `Content-Disposition: attachment; filename=Report_03.pdf`. The `X-SourceFiles` header contains a long alphanumeric string.

Statistics Inspectors AutoResponder Composer Log Filters Timeline

Headers TextView WebForms HexView Auth Cookies Raw JSON XML

GET http://localhost:61727/api/reports/Report_03.pdf/content HTTP/1.1
Range: bytes=500-999
Host: localhost:61727

Find... (press Ctrl+Enter to highlight all)

Get SyntaxView Transformer Headers TextView ImageView HexView WebView Au

HTTP/1.1 206 Partial Content
Cache-Control: no-cache
Pragma: no-cache
Content-Length: 500
Content-Type: application/pdf
Content-Range: bytes 500-999/153220
Expires: -1
Accept-Ranges: bytes
Content-Disposition: attachment; filename=Report_03.pdf
X-SourceFiles: =?UTF-8?B?QzpcVXNlcnNccGxr cmpha1xEb2N1bWVudHNcR2l0SHVixEZpbGV
Date: Wed, 05 Apr 2017 22:08:03 GMT



COG - IFD Locations

- An IFD is a pointer to the location of the first tile
- A GeoTiff with overviews has several IFD one for each overview
- A COG must have its IFDs located in the first data block



Putting it all together



Data Retrieval example

- Most cases user needs are different from the way the raster was stored.
- A user defines an Area of Interest (AOI) and resolution
- For example: WMS, XYZ requests



COG - Classic data retrieval

- Fetch the header of a cog file (<16KB)
- Pick the IFD by resolution
- Calculate the tiles needed according to the tiles offsets in the IFD

Demos





Using COGs





Landsat file

- https://landsat-pds.s3.amazonaws.com/c1/L8/189/027/LC08_L1TP_189027_20170403_20170414_01_T1/LC08_L1TP_189027_20170403_20170414_01_T1_B4.TIF



Qgis

- Adding a cog as a layer



Rasterio

- [notebook](#)



Creating COGs





Gdal_translate

- Follow the steps here:
<https://trac.osgeo.org/gdal/wiki/CloudOptimizedGeoTIFF#Preparation>



Telluric

- <https://github.com/satellopic/telluric>
- A library developed by satellopic (including me)



Rasterio

- <https://github.com/cogeotiff/rio-cogeo>



Nice tools





Cog Explorer

1. Use javascript library to visualize cogs
2. <https://geotiffjs.github.io/cog-explorer/#long=16.370&lat=48.210&zoom=5&scene=&bands=&pipeline=>



marblecutter

- Flask or Lambda to server xyz of cogs
- <https://hi.stamen.com/stamen-aws-lambda-tiler-blog-post-76fc1138a145>

Thank you

