GDAL 2.3

20 year already and heading to the cloud!

Even Rouault

SPATIALYS

July 19th 2018

GDAL/OGR : Introduction

- GDAL? Geospatial Data Abstraction Library. The swiss army knife for geospatial.
- Read and write Raster (GDAL) and Vector (OGR) datasets
- 240 (mainly) geospatial formats and protocols.
- Widely used

GRASS GIS
PostGIS
MapServer
open source web mapping
mapnik
ArcGIS
FME
gvSIG
SAGA
Google Earth

(> 100 http://trac.osgeo.org/gdal/wiki/SoftwareUsingGdal)

- MIT/X Open Source license (permissive)

GDAL/OGR : Introduction

- GDAL? Geospatial Data Abstraction Library. The swiss army knife for geospatial.
- Read and write Raster (GDAL) and Vector (OGR) datasets
- 240 (mainly) geospatial formats and protocols.
- Widely used

GRASS GIS
PostGIS
MapServer
open source web mapping
ArcGIS
mapnik
FME
gvSIG
SAGA
Google Earth

(> 100 http://trac.osgeo.org/gdal/wiki/SoftwareUsingGdal)

- MIT/X Open Source license (permissive)

GDAL/OGR 2.3

- v2.3.0: May 4th 2018.   V2.3.1: June 22nd 2018
- 3801 “commits” (total since 1998: 37988)
- 3 RFCs implemented during 2.3 cycle
  - Use (and build requirement) for C++11
  - Migration from Trac+Subversion to GitHub
    - Past tickets remain in Trac database
    - New tickets to be created on GitHub
  - Autodetection of output format from file extension
      - gdal_translate my.tif my.png
      - ogr2ogr my.gpkg my.shp

GDAL/OGR 2.3

- 2 new raster drivers
  - PDS4: Planetary Data System v4 (NASA format)
    - XML header for metadata + raw binary file for imagery
    - Read/Write
  - RDA: DigitalGlobe Raster Data Access. Read acces to images through REST API

- 2 new vector drivers:
  - MVT: MapBox Vector Tiles
  - WFS3

- ESRIJSON et TOPOJSON drivers split off from existing GeoJSON driver (unchanged functional scope)

- Gdal2tiles can be multi-threaded

MVT driver (Mapbox Vector Tiles)

- Read/write
- Variants: tileset of files or MBTILES container

```
/metadata.json        ←        List layers and attributes
/0/
   0/
     0.pbf
/1/
   0/
     0.pbf
     1.pbf
   1/
     0.pbf
     1.pbf
```

MVT driver

- Read:
  - On-disk or HTTP accessible tiles

    ```
    ogrinfo MVT:https://free.tilehosting.com/data/v3/1 -oo tile_extension="pbf.pict?key=${YOUR_KEY}" \
            -oo metadata_file="https://free.tilehosting.com/data/v3.json?key=${YOUR_KEY}"
    ```

  - Geometry clipping or not
- Write settings:
  - Zoom level
  - Resolution and buffer size
  - Geometry simplification
  - Layer assigning per zoom level

WFS3 driver

- *EXPERIMENTAL*! Specification is moving a lot (driver updated in GDAL 2.3.1)
- WFS v3:
  - Nice break w.r.t previous WFS versions
  - REST philosophy
    - GET /collections
    - GET /collections/{name}/items?bbox=160.6,-55.95,-170,-25.89
    - GET /collections/{name}/items/{id}
  - End of all-XML. Content negotiation
  - OpenAPI 3.0
  - Minimal core, easy to implement

https://github.com/opengeospatial/WFS_FES

Cloud Optimized GeoTIFF (COG)

- “Profile” of TIFF format, optimized for reading through HTTP
- Goal: efficient access to subparts of the file without specialized server
- ~ WCS with only a file server
- Client-side: use of HTTP GET requests with Range header
- Increasing adoption by software and data providers

planet.

GDAL

DigitalGlobe

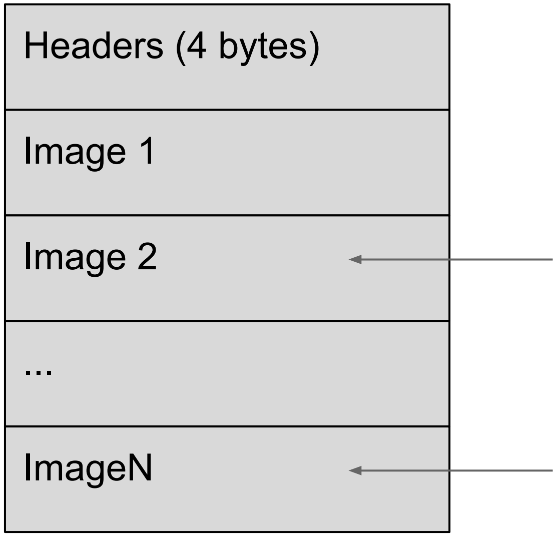
GeoServer

Radiant.Earth
Earth Imagery for Impact

farmshots
growing with you

OpenAerialMap

TIFF format structure: high level



Headers (4 bytes)

Image 1

Image 2

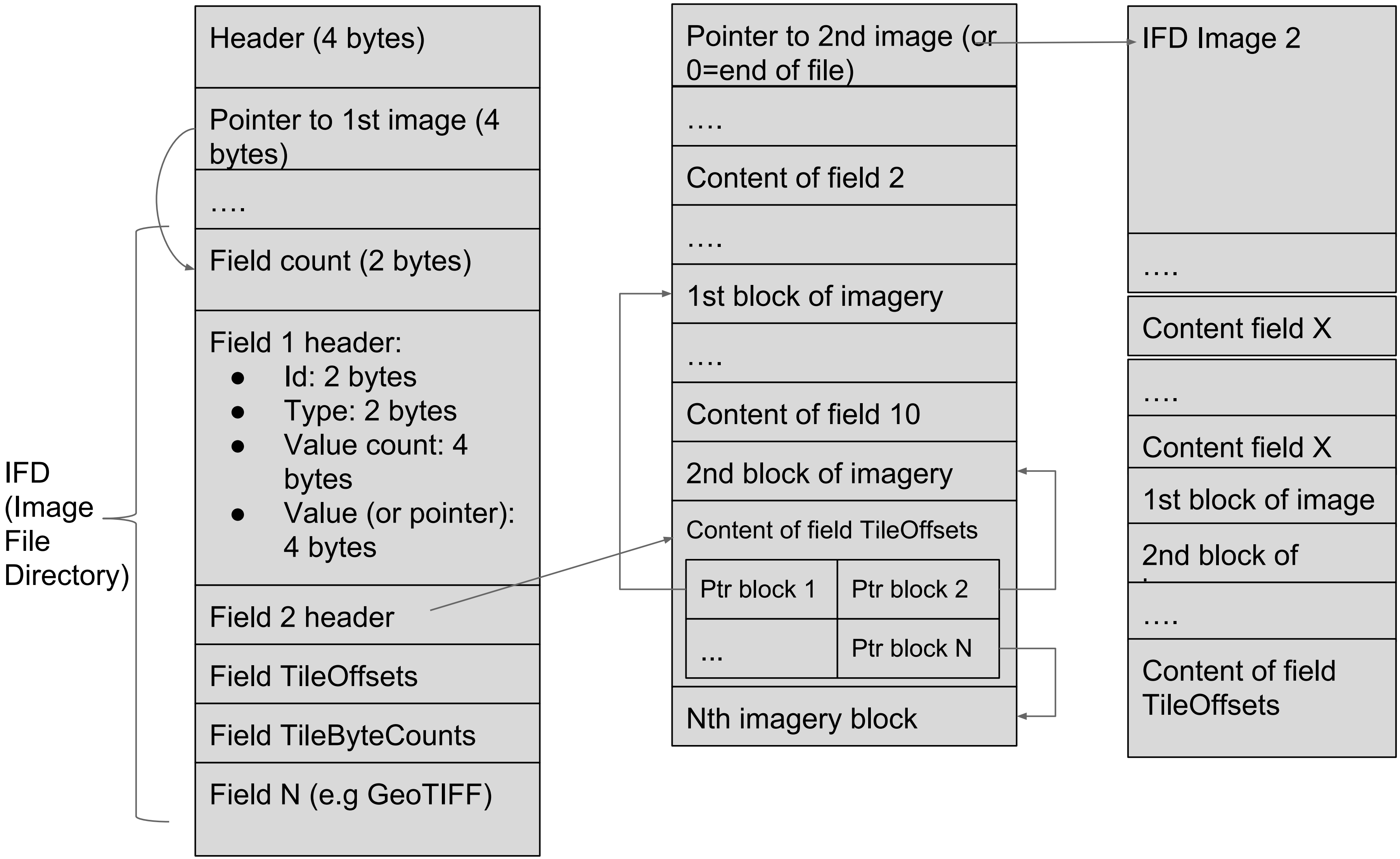
...

ImageN

Typically: overview factor 2. Or mask

Typically: overview factor 2^(N-1)

TIFF format structure: low level



Header (4 bytes)

Pointer to 1st image (4 bytes)

….

Field count (2 bytes)

Field 1 header:

- Id: 2 bytes
- Type: 2 bytes
- Value count: 4 bytes
- Value (or pointer): 4 bytes

Field 2 header

Field TileOffsets

Field TileByteCounts

Field N (e.g GeoTIFF)

IFD (Image File Directory)

Pointer to 2nd image (or 0=end of file)

….

Content of field 2

….

1st block of imagery

….

Content of field 10

2nd block of imagery

Content of field TileOffsets

Ptr block 1	Ptr block 2
...	Ptr block N

Nth imagery block

IFD Image 2

….

Content field X

….

Content field X

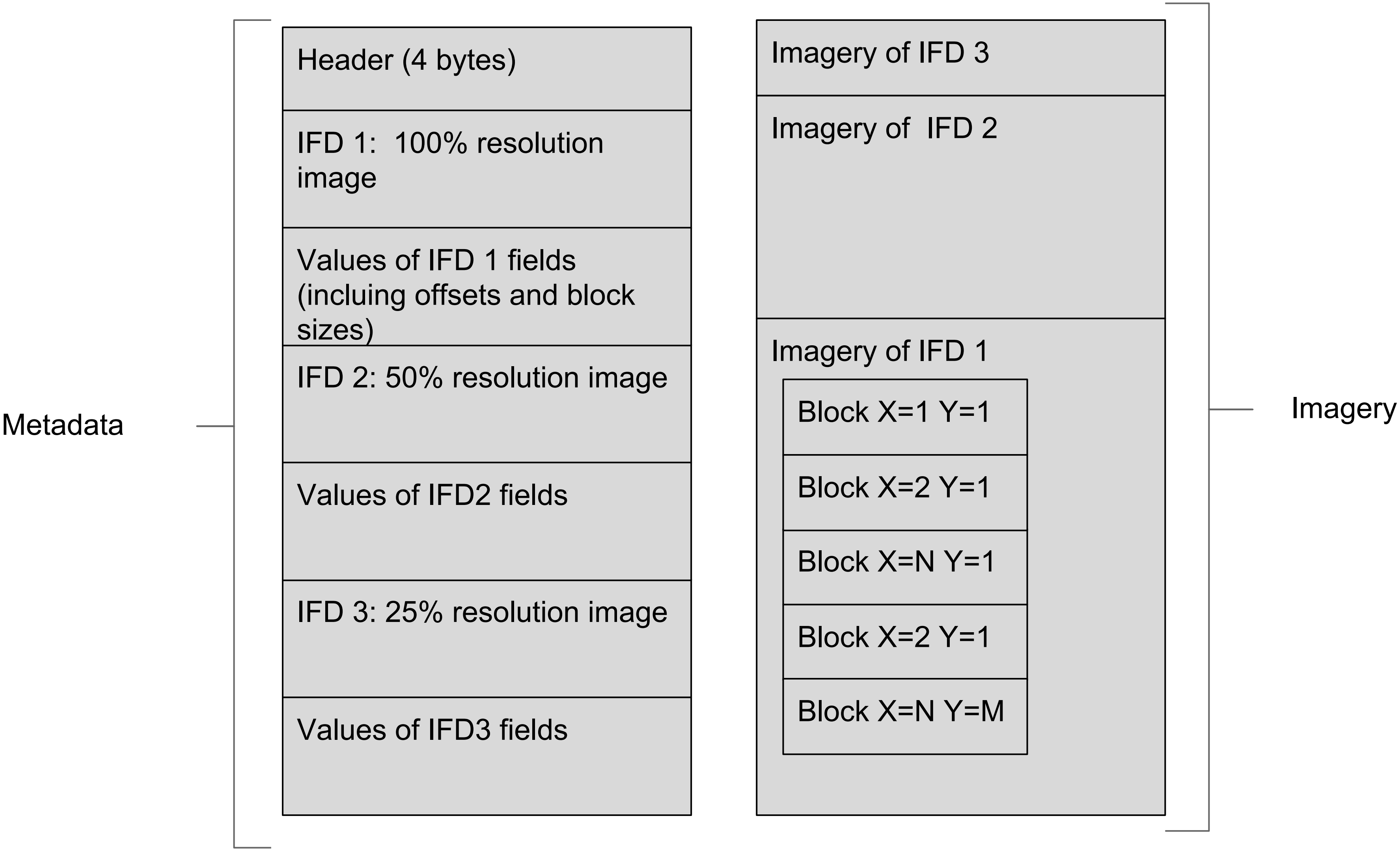
1st block of image

2nd block of

….

Content of field TileOffsets

COG file structure



Metadata

Header (4 bytes)

IFD 1: 100% resolution image

Values of IFD 1 fields (incluing offsets and block sizes)

IFD 2: 50% resolution image

Values of IFD2 fields

IFD 3: 25% resolution image

Values of IFD3 fields

Imagery of IFD 3

Imagery of IFD 2

Imagery of IFD 1

Block X=1 Y=1

Block X=2 Y=1

Block X=N Y=1

Block X=2 Y=1

Block X=N Y=M

Imagery

Creation of COG file

1. Creation of GeoTIFF file:

   -
     ```
     gdal_translate source temp.tif
     ```

2. Creation of its overviews:

   -
     ```
     gdaladdo -r average temp.tif
     ```

3. Creation of COG file:

   -
     ```
     gdal_translate temp.tif cog.tif \
       -co TILED=YES \
       -co COMPRESS=DEFLATE \  (or JPEG -co PHOTOMETRIC=YCBCR)
       -co COPY_SRC_OVERVIEWS=YES
     ```

Network virtual file systems

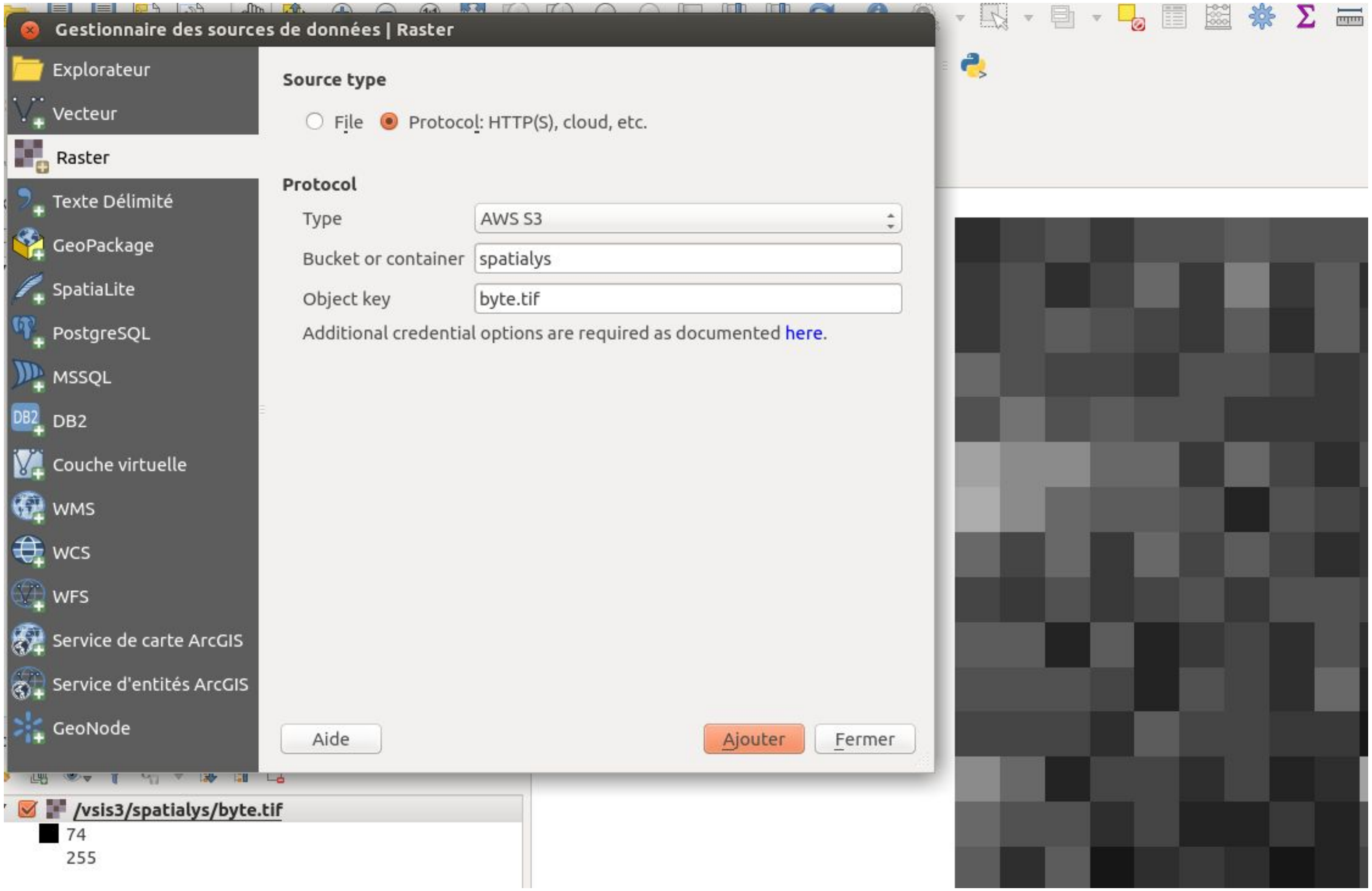
- /vsicurl/
  - gdalinfo /vsicurl/http://landsat-pds.s3.amazonaws.com/L8/139/045/LC81390452014295LGN00/LC81390452014295LGN00_B1.TIF
- /vsis3/ : Amazon AWS S3
- /vsigs/ : Google Cloud Storage
- /vsiaz/ : Microsoft Azure Blob storage
- /vsiswift/: OpenStack SWIFT
- /vsioss/: Alibaba Object Storage Service
- All available in (random) reading / (sequential) writing
- Streamed version for reading: /vsicurl_streaming/, /vsis3_streaming/ etc,...
- VSIGetSignedURL() function to get a signed URL

Network virtual file systems

- Several authentication methods
  - /vsis3:
    - AWS_SECRET_ACCESS_KEY + AWS_ACCESS_KEY_ID
    - ~/.aws/credentials (AWS_PROFILE)
    - ~/.aws/config
    - Automatic identification when running on AWS EC2
  - /vsigs:
    - GS_SECRET_ACCESS_KEY + GS_ACCESS_KEY_ID
    - GDAL_HTTP_HEADER_FILE
    - GS_OAUTH2_REFRESH_TOKEN
    - GS_OAUTH2_PRIVATE_KEY + GS_OAUTH2_CLIENT_EMAIL
    - ~/.boto
    - Automatic identification when running on Google Compute Engine
- Utilities:
  - https://github.com/OSGeo/gdal/tree/master/gdal/swig/python/samples
    - gdal_ls.py
    - gdal_cp.py
    - gdal_rm.py

Network virtual file systems

- Easy reading in QGIS 3.2 (*Mathieu Pellerin, iMHere Asia*)



Gestionnaire des sources de données | Raster

Source type

○ File ● Protocol: HTTP(S), cloud, etc.

Protocol

Type: AWS S3

Bucket or container: spatialys

Object key: byte.tif

Additional credential options are required as documented here.

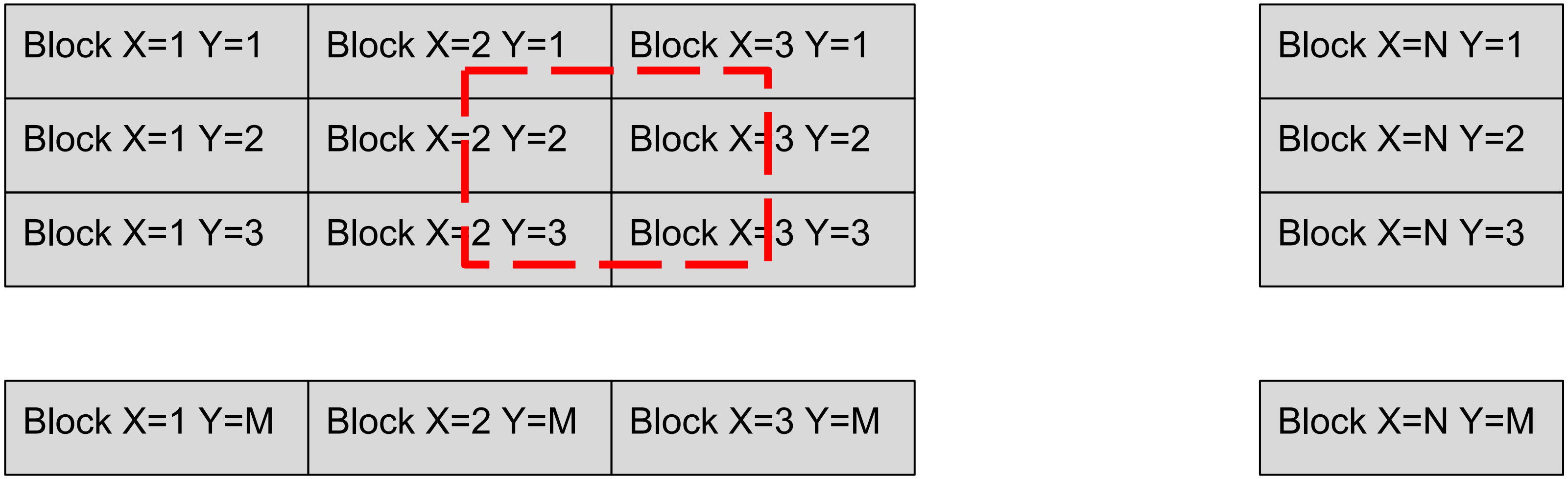
☑ /vsis3/spatialys/byte.tif

■ 74

255

Network virtual file systems

GeoTIFF driver improvements for more efficient HTTP requests



Block X=1 Y=1	Block X=2 Y=1	Block X=3 Y=1	Block X=N Y=1
Block X=1 Y=2	Block X=2 Y=2	Block X=3 Y=2	Block X=N Y=2
Block X=1 Y=3	Block X=2 Y=3	Block X=3 Y=3	Block X=N Y=3
Block X=1 Y=M	Block X=2 Y=M	Block X=3 Y=M	Block X=N Y=M

⇒ 3 parallel GET Range requests

- ○ HTTP/1.1 parallel connections
- ○ HTTP/2 multiplexing

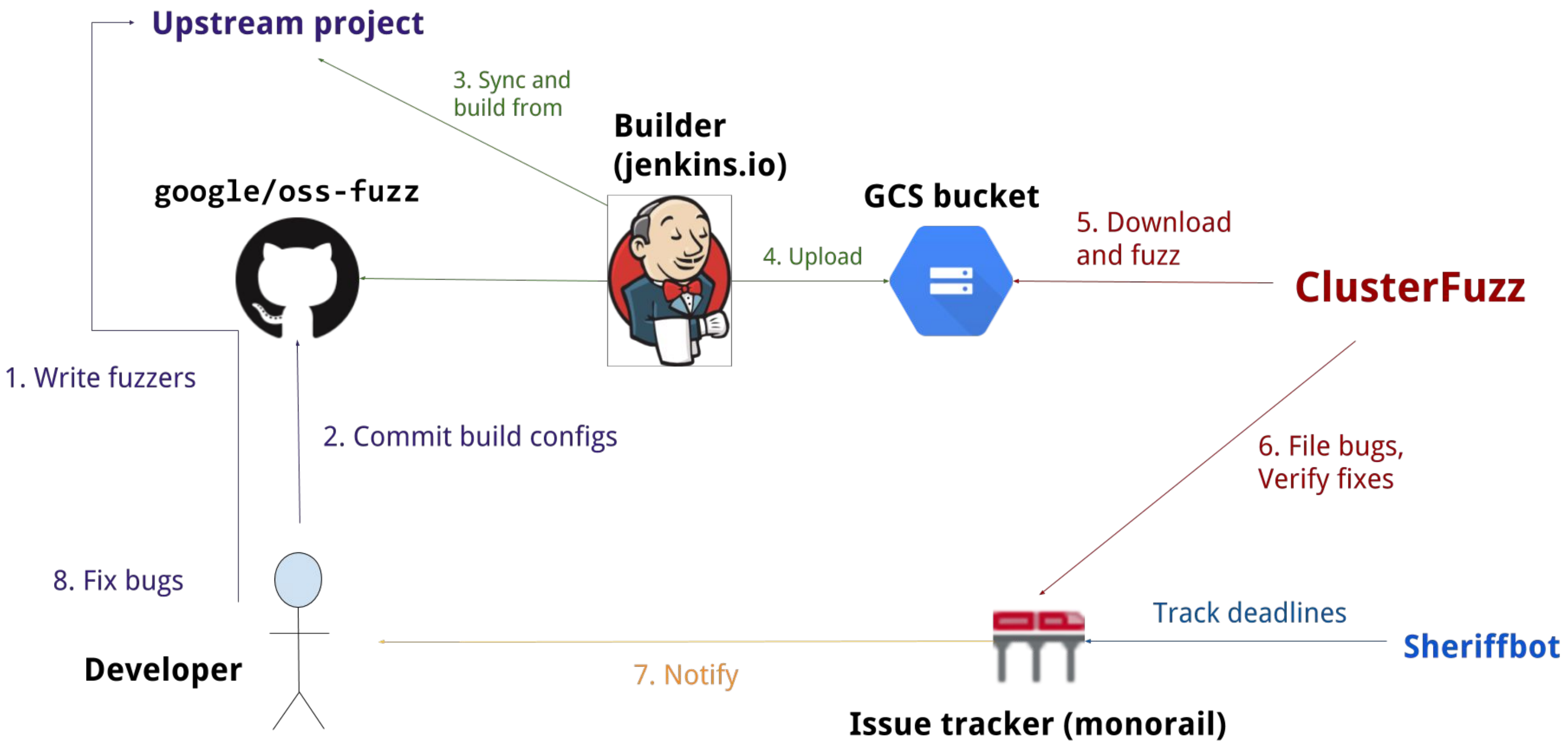
COG and virtual file systems

- General documentation
  - http://www.cogeo.org
  - https://github.com/cogeotiff/cog-spec/blob/master/spec.md
- Javascript clients
  - https://geotiffjs.github.io/cog-explorer (EOX)
  - http://www.cogeo.org/map/ (Radiant Earth) + tiles.rdnt.io (server)
- Validation:
  - http://cog-validate.radiant.earth/html
  - https://github.com/rouault/cog_validator
- GDAL GeoTIFF driver documentation:
  - http://gdal.org/frmt_gtiff.html
- GDAL virtual file system documentation:
  - http://gdal.org/gdal_virtual_file_systems.html

GDAL/OGR 2.3: other changes

- PROJ v5 support
- Upgrade SRS database to EPSG v9.2
- Better identification of EPSG codes for shapefiles
- Improvements in following drivers:
  - GeoJSON: arbitrary large file reading
  - GRIB2 write support
  - WCS:
    - WCS 2.0 protocol (*Ari Jolma*)
    - Metadata disk cache
    - Various options to handle server non-conformities
  - DXF: many improvements (*Alan Thomas / ThinkSpatial thinkspatial.com.au*): LEADER, DIMENSION, BLOCK, ...

Integration to oss-fuzz initiative



(source: https://github.com/google/oss-fuzz)

⇒ > 1000 bugs fixed

Community activity

- 19 developers with direct GitHub push rights
- 95 contributors during 2.3.0 cycle
- 2295 subscribers to gdal-dev.

  1924 messages may 2017→ may 2018
- ~470 tickets opened / 12 last months (total: 7310). ~480 opened

What’s next ?

gdalbarn.com

“GDAL SRS barn”: gdalbarn.com

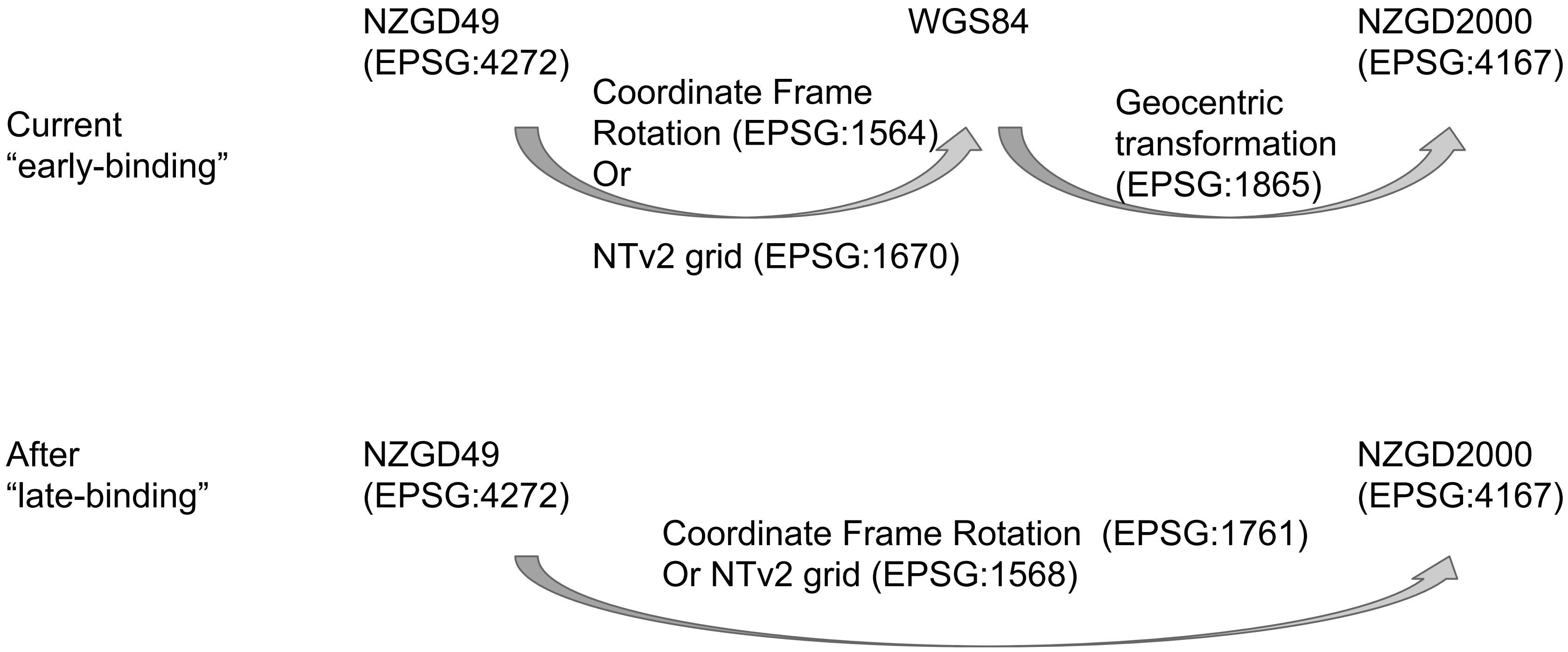
- Modernization and rework of spatial reference system management in GDAL, libgeotiff and PROJ
- Opportunity linked to PROJ 5.0 release
- Adoption of OGC WKTv2 standard (12-063r5) / ISO 19162
  - Solves interoperability issues
  - Temporal dimension management
  - Better handling of vertical component

“GDAL SRS barn”: gdalbarn.com

- SQLite database to store SRS definition
  - No longer CSV files duplicated among software
  - Better query capabilities
- Use of area of uses
- Temporal component handling
  - Helmert transformation with derived terms
  - Transformations with deformation models

“GDAL SRS barn”: gdalbarn.com

- WGS84 no longer as pivot system for datum transformation



Current “early-binding”

NZGD49 (EPSG:4272)

Coordinate Frame Rotation (EPSG:1564) Or

NTv2 grid (EPSG:1670)

WGS84

Geocentric transformation (EPSG:1865)

NZGD2000 (EPSG:4167)

After “late-binding”

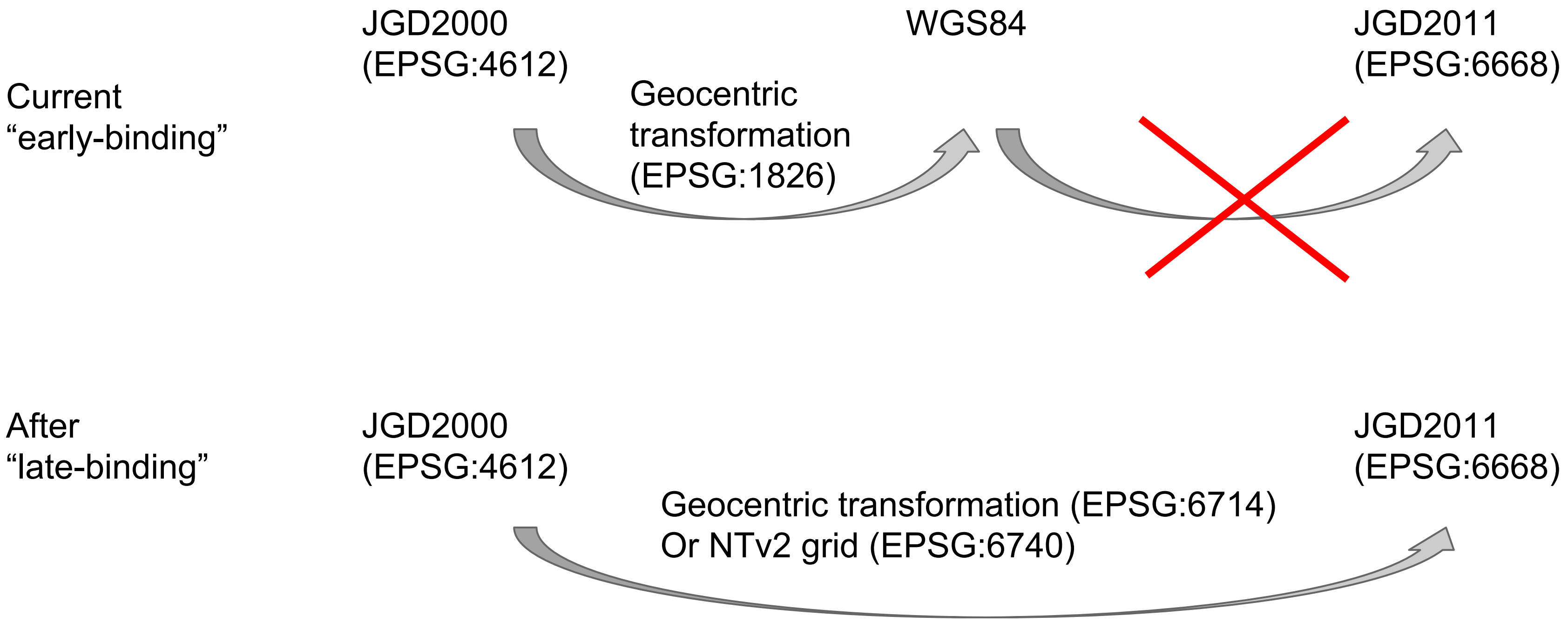
NZGD49 (EPSG:4272)

Coordinate Frame Rotation (EPSG:1761) Or NTv2 grid (EPSG:1568)

NZGD2000 (EPSG:4167)

“GDAL SRS barn”: gdalbarn.com

- WGS84 no longer as pivot system for datum transformation



Current “early-binding”

JGD2000 (EPSG:4612)

Geocentric transformation (EPSG:1826)

WGS84

JGD2011 (EPSG:6668)

After “late-binding”

JGD2000 (EPSG:4612)

Geocentric transformation (EPSG:6714)
Or NTv2 grid (EPSG:6740)

JGD2011 (EPSG:6668)

Thanks to the sponsors of GDAL barn !

AGI

AppGeo®

APPLIED IMAGERY

azavea

Boundless

CARTO

DYNAMIC GRAPHICS, INC.

GeoCue

Intevation GmbH

LIZARDTECH™

Land Information New Zealand
Toitū te whenua

esri

mapbox

OSLANDIA

radiant
SOLUTIONS

SEPTIMA

RIEGL®
LASER MEASUREMENT SYSTEMS

Spatial Networks

SAFE SOFTWARE™

Mobile Geographics

>sparkgeo

TERRANODO

Timoney Group

Questions?

Links:

- http://www.gdal.org/
- https://trac.osgeo.org/gdal/wiki/Release/2.3.0-News
- https://trac.osgeo.org/gdal/wiki/RfcList

Contact: even.rouault@spatialys.com

GDAL/OGR 2.2 in a nutshell

- V2.2.0: may 2017 → V2.2.4: march 2018
- 4 RFCs during 2.2 cycle including:
  - Sparse dataset improvements for GeoTIFF and VRT
  - New geometry types: polyhedral surfaces, triangulated irregular networks (TIN)
  - Null field vs empty field (JSon and GML)
- 7 new drivers including:
  - CAD: read DWG R2000
  - DGNv8: read/write DGN v8 proprietary format (through proprietary SDK)
  - GMLAS: read/write XML/GML driven by application schemas (“Complex features”)