

Protomaps: Web mapping at any scale

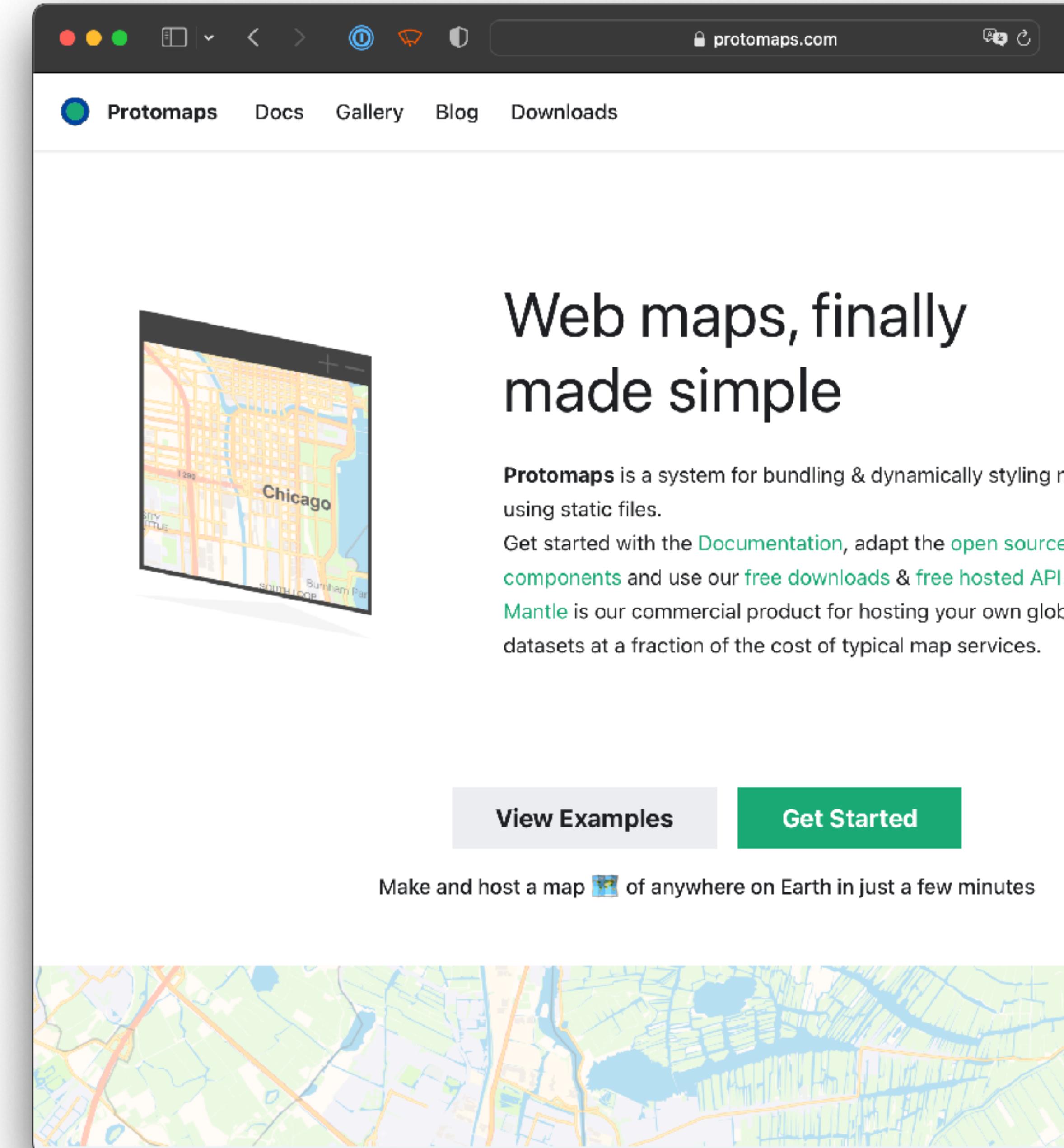
History + future of an independent cartographic stack

Brandon Liu / Protomaps LLC / FOSS4G 2022 Firenze



Protomaps

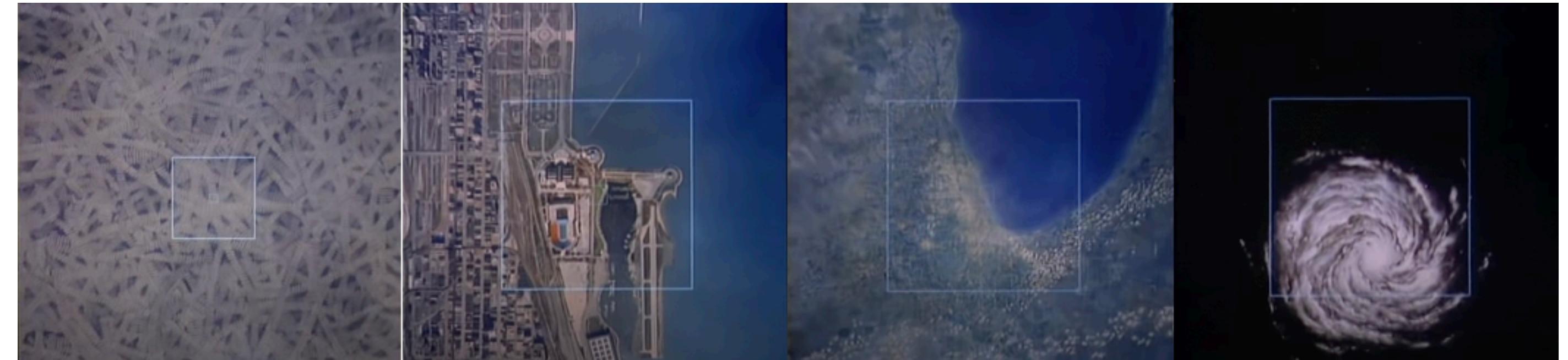
- An **end-to-end** system for bundling and styling vector web maps using **static files**.
- Based on **nearly a decade of experience** building and deploying cartographic web projects.
- **Open-source buffet:** use what you want, interoperable with existing tools and front-ends.



some core values

(the technical ones)

- Works the same at **any scale**



EAMES OFFICE, POWERS OF TEN

- Favor simple technology: **static files** wherever possible
- Favor modularity via **established standards**
- But don't be afraid to build new technology
 - Take a *systems programming approach*

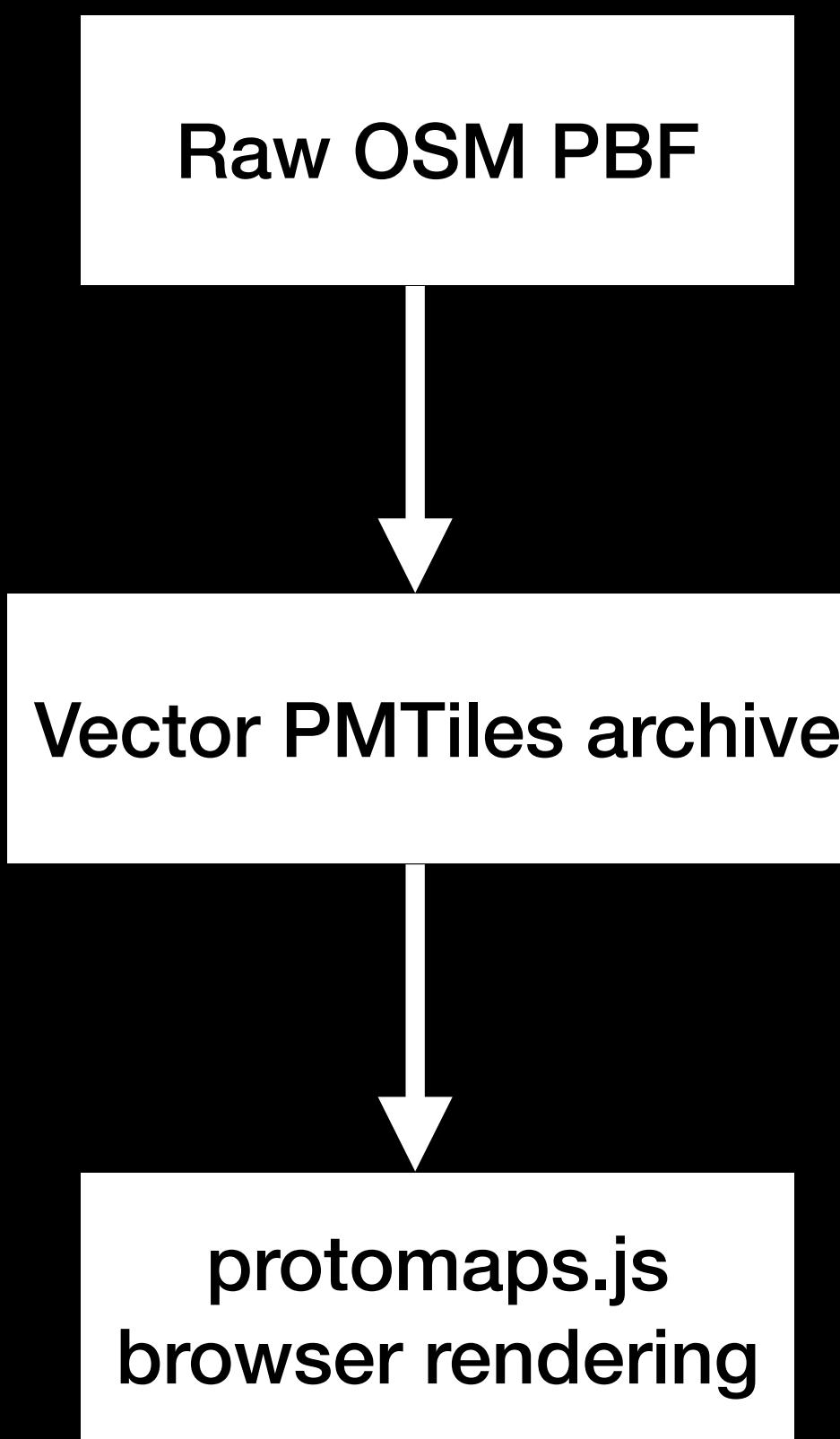
some core values

(the social and organizational ones)

- **hypothesis driven:**
 - determine if approaches viable by developing an MVP
 - *faster horse problem*
- **open source** is key to:
 - establishing confidence in an indie project
 - building an audience inclusive of hobbyist and civic-minded use cases

talk overview

- 3 open source projects:
 - **protomaps.js**
 - **OSM Express**
 - **PMTiles**
- **Hypotheses for why they should exist**
- **Successes, failures and future plans**
- All at GitHub.com/protomaps



Hypothesis

Developers love Leaflet, so we need a lightweight vector renderer designed for it.



leafletjs.com

protomaps.js

Finally a vector tile renderer for Leaflet

- Non-realtime browser renderer built on Canvas2D
 - Leaflet-style zoom UX
 - Labeling + i18n + web fonts
 - Enough for basemaps
 - Full JS styling + extensible API

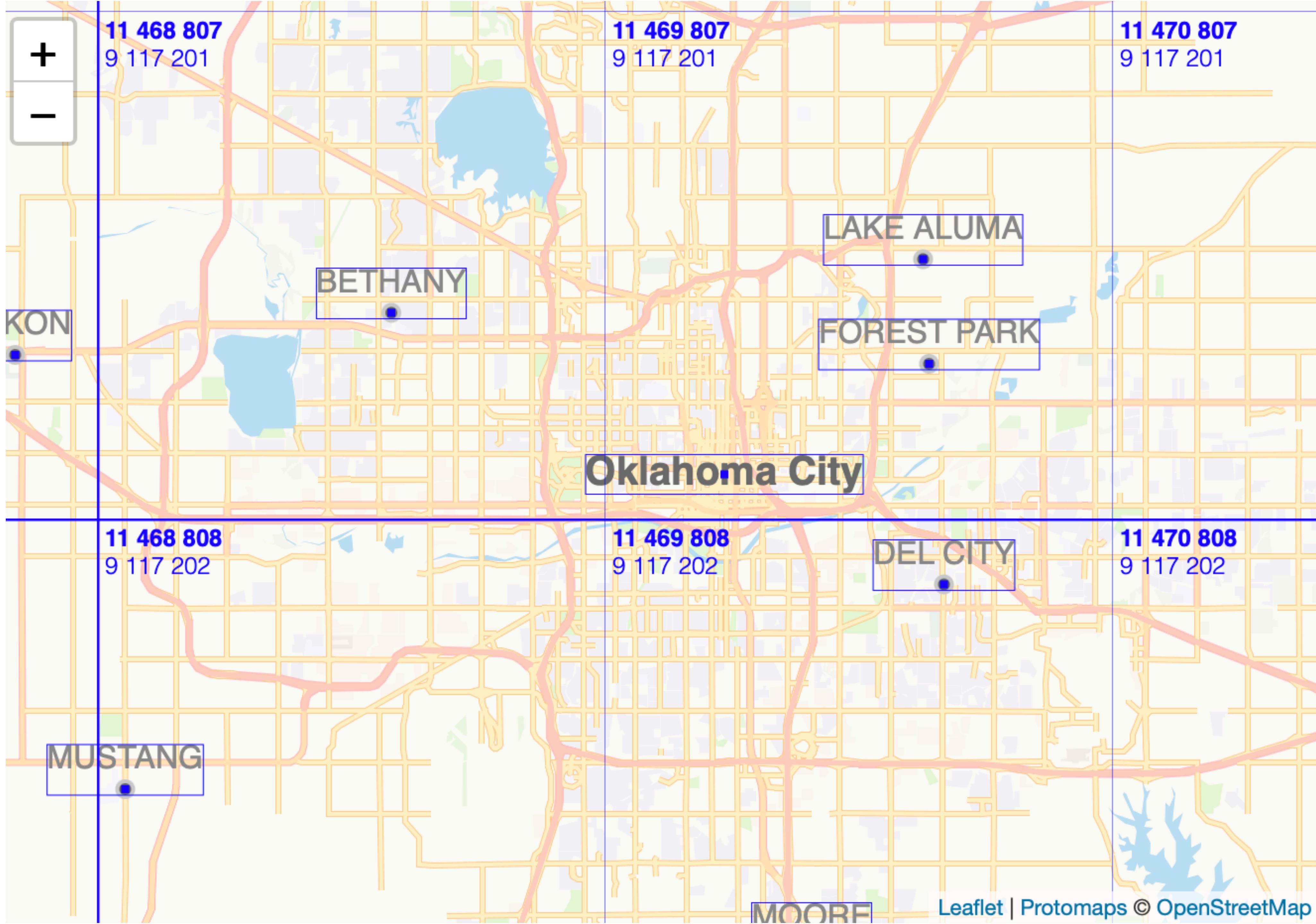


+

-

Lat





protomaps.js: a lightweight vector tile renderer

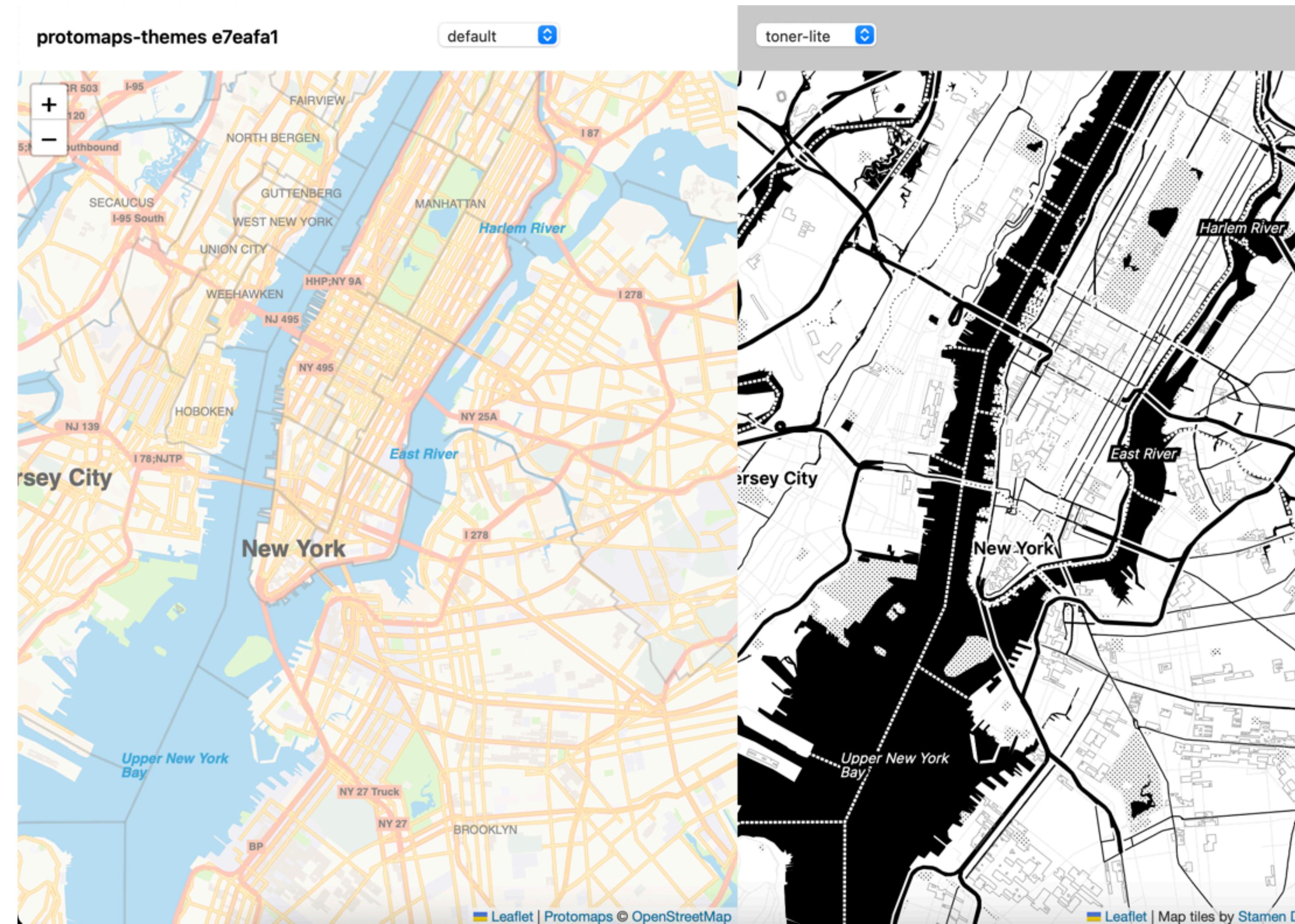
Failures

- “*The best way to never make a map is to write your own map engine*”
- Modern expectations for real time rendering and fractional zoom UX
- No map rotation

protomaps.js: a lightweight vector tile renderer

Future development

- Focus on differentiating use cases vs OL, GL renderers
 - Custom symbology
 - i18n/l10n support
 - Legacy migration path
 - Additional open-source themes



Hypothesis

OpenStreetMap is hard, and users want access to fresh datasets of on-demand areas of interest.

Create OpenStreetMap Extract

Download areas of raw OpenStreetMap data.

• Data updated 1 minute ago

Draw Rectangle

Draw Polygon

Upload Shape

Choose File **no file selected**

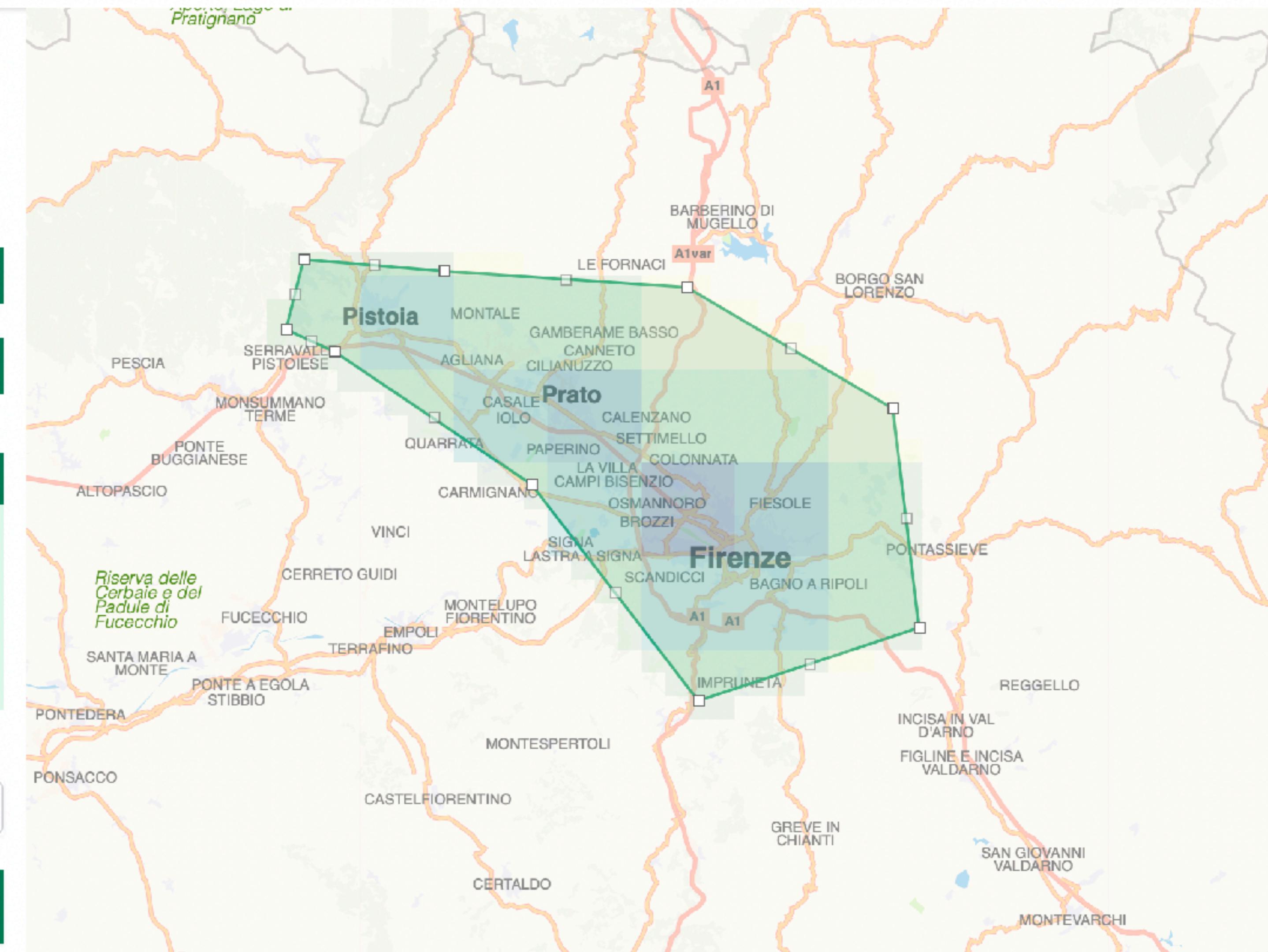
Selected area

This area contains about 1,575,496 nodes, which will take approximately 0.7 minutes.

Name this area

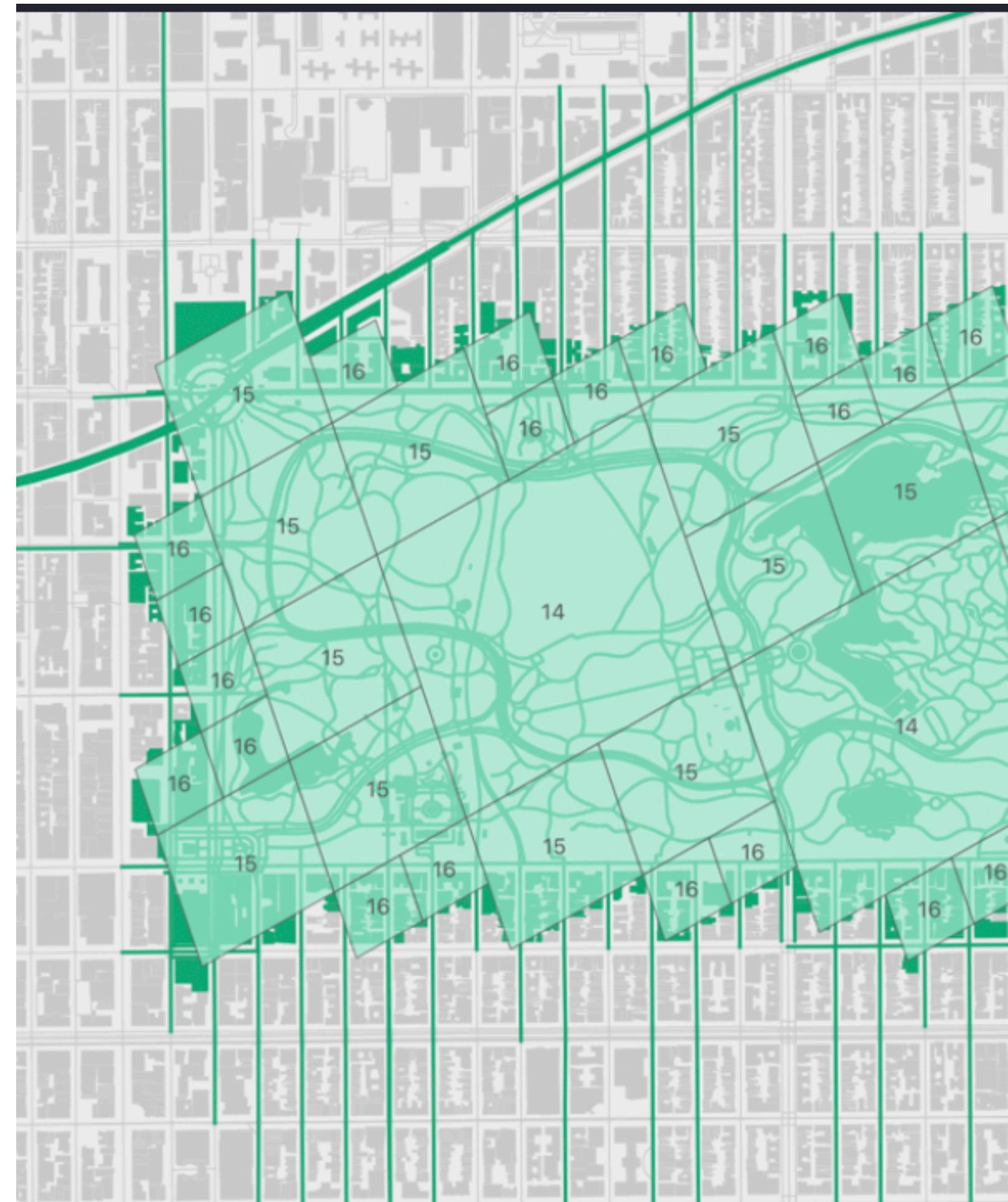
Null Island, Earth

Create Extract



OSM Express: a fast database for OpenStreetMap

- Transactional embedded database specific to OSM
- Consumes minute-level updates
- S2Cell indexing for efficient spatial queries
- Hosted on-demand UI at protomaps.com/downloads
 - Over 15,000 created so far
- Used in production at other companies as low-level OSM infrastructure



OSM Express: a fast database for OpenStreetMap

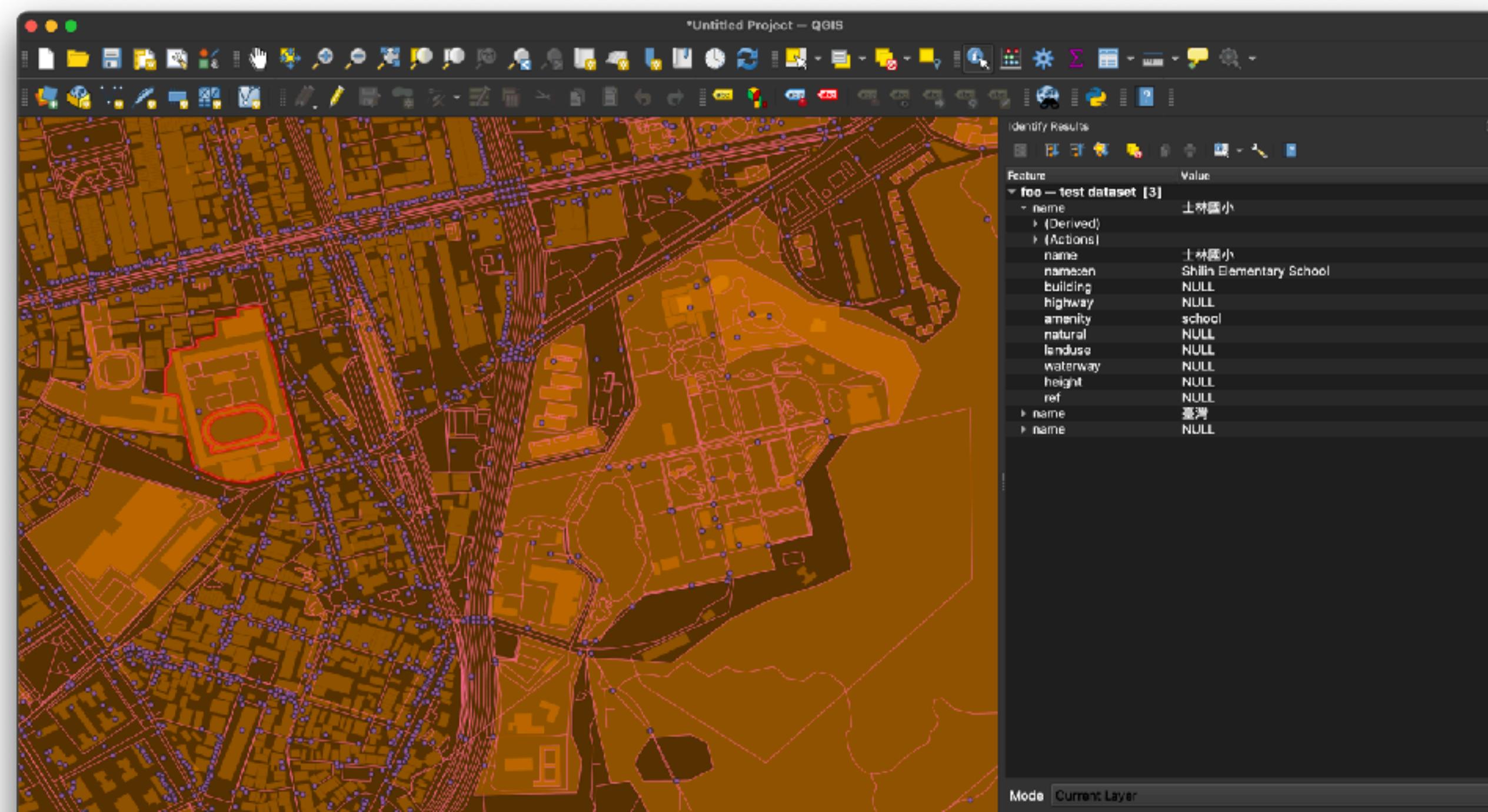
Failures

- “Raw” OSM is exotic and only for an audience of OSM ecosystem devs
- Minute-level updates to data is a niche requirement
- Difficult to build and deploy

OSM Express: a fast database for OpenStreetMap

Future development

- Tabular **GIS output** (FlatGeobuf in GDAL-friendly format)
- Programmatic API for automated extracts via hosted server



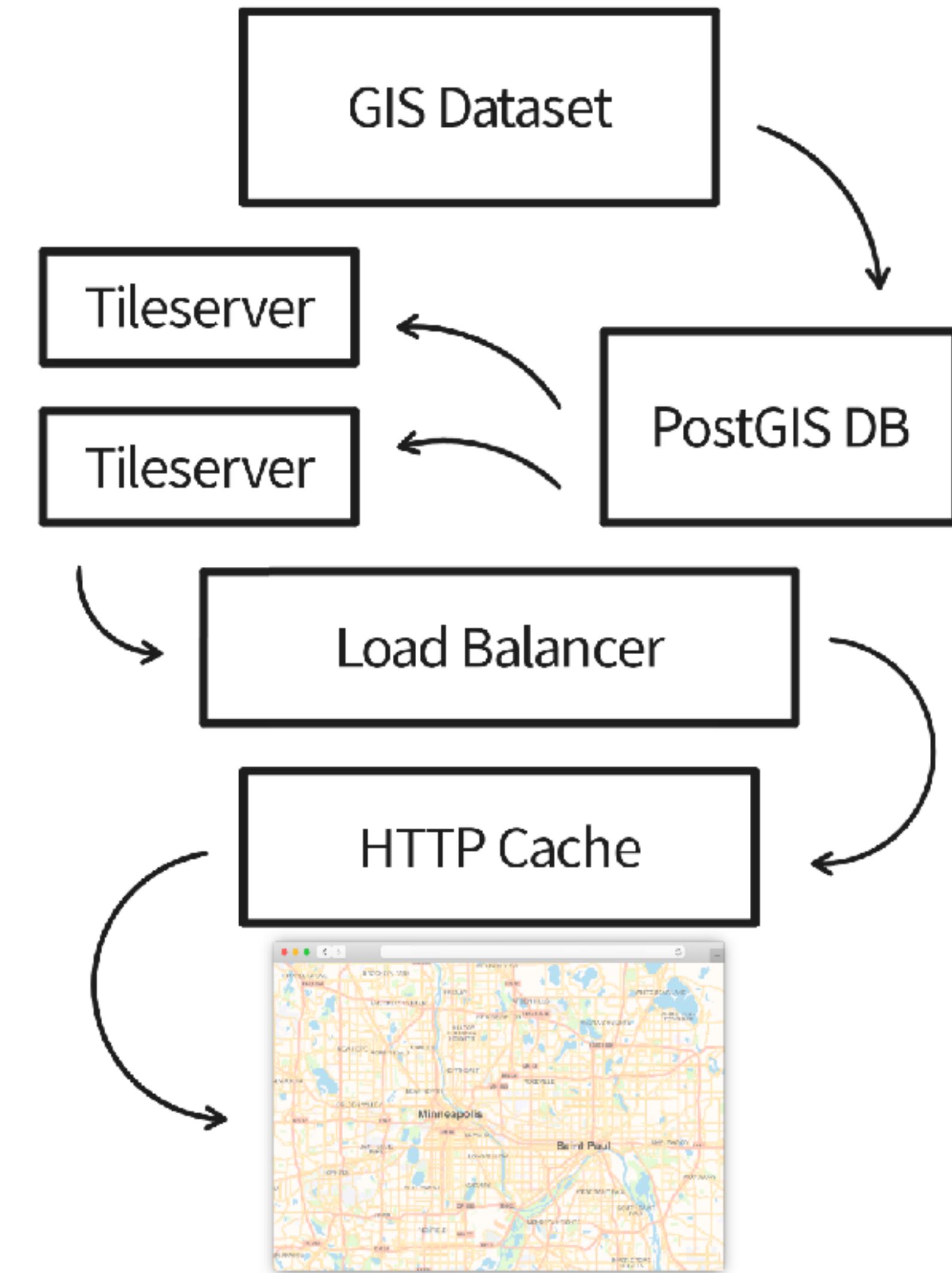
Hypothesis

Storing and serving read-only web map tiles can be made much simpler with a cloud-native archive format.

PMTiles motivation

MBTiles status quo

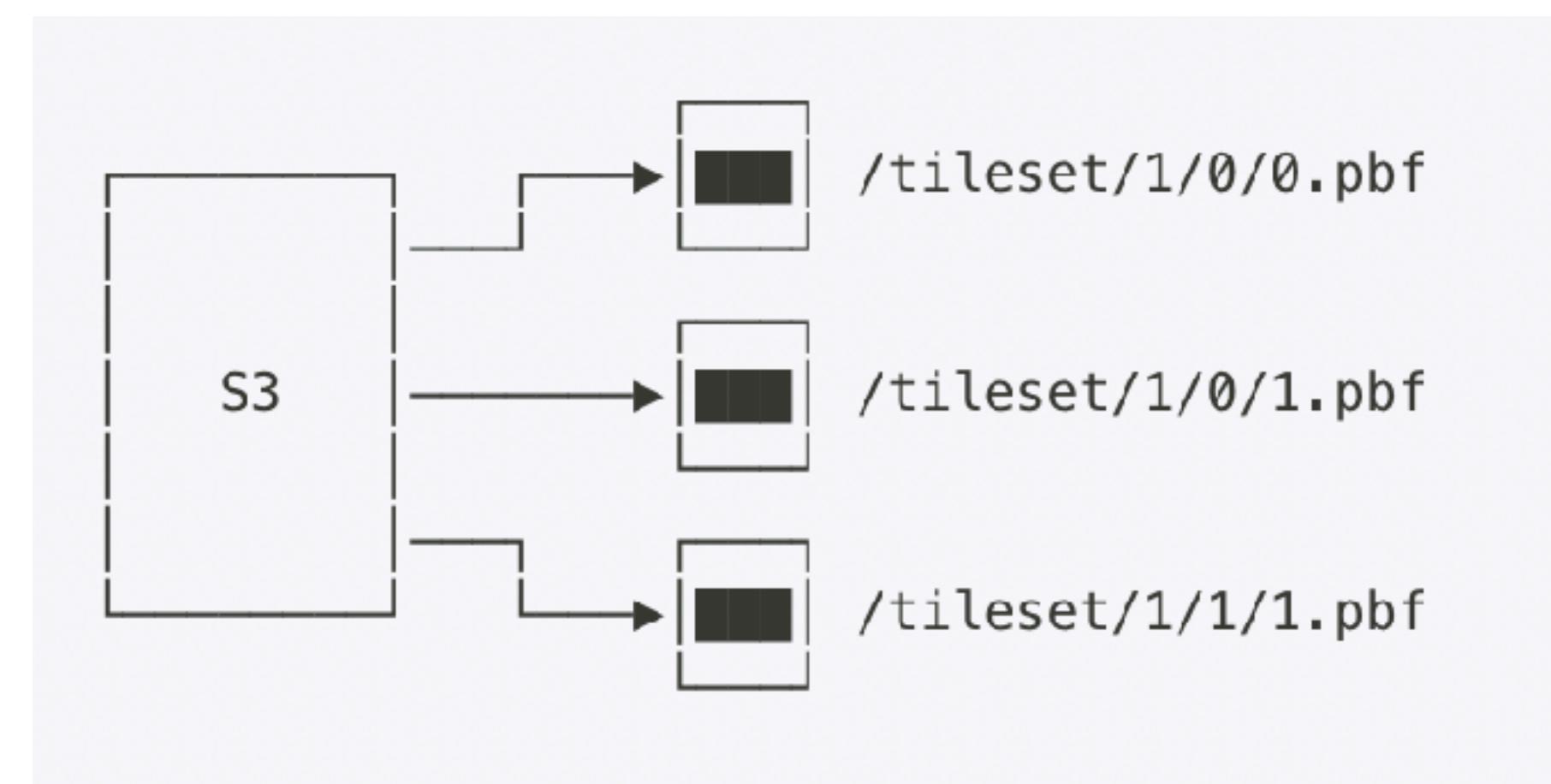
- MBTiles is great solution!
 - **Simple spec** with easy implementation based on SQLite
 - Few use database features; majority use case **read-only**
 - Spec does not cover efficiently encoding planet-scale datasets
 - All readers need **SQLite C** dependency
 - Goal: **replace MBTiles** and make switching compelling
 - Make it a cloud-optimized format



PMTiles motivation

Static tiles on S3-like storage?

- Upload each tile individually to S3
- Works fine for small tilesets
- Each write operation incurs transaction cost and time
 - Planet-scale tileset can have 300+ million tiles

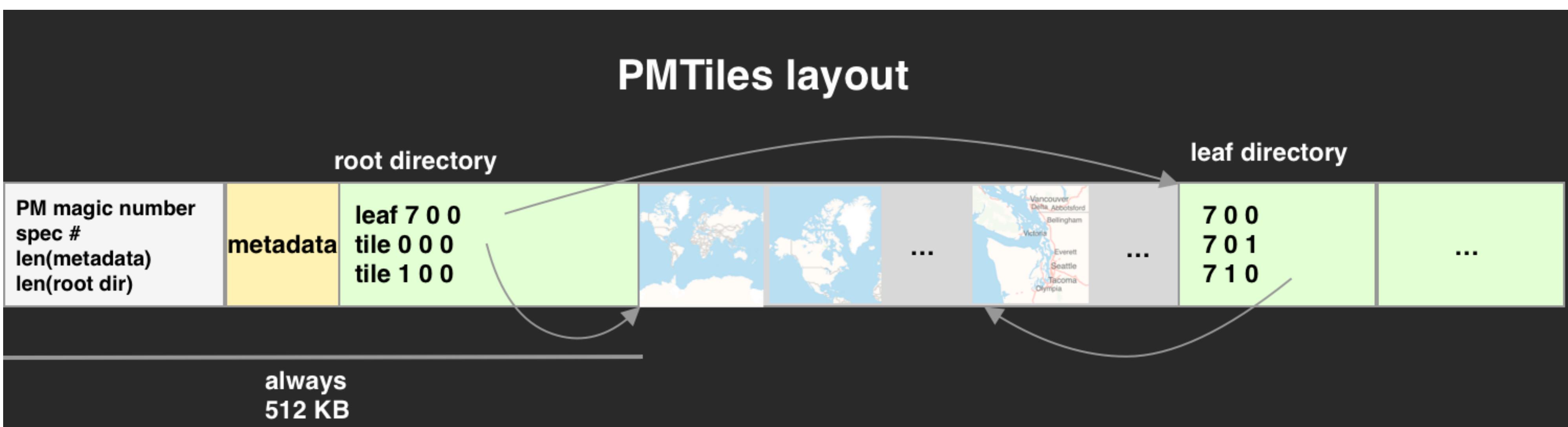


PMTiles design

A cloud-optimized tile archive

- Single archive, agnostic to tile content: raster, vector, DEM...
 - Address single tiles via **HTTP Range Requests** (COG technique)
 - constrained to Z/X/Y pyramid, scale to **huge tilesets**
- Deduplication of identical tiles: same offset, length
- **Direct integration** with Leaflet and MapLibre GL (v1.15+)

PMTiles layout



my_neighborhood.pmtiles 10 MB



my_neighborhood.pmtiles 10 MB

my_city.pmtiles 500 MB



my_neighborhood.pmtiles 10 MB

my_city.pmtiles 500 MB

my_planet.pmtiles 80 GB





TimSalabim

@TimSalabim3 · Aug 20

...

Thanks to the great help by [@bdon](#) we now have a working prototype for rendering [#pmtiles](#) using [@protomaps](#) in [#rstats](#) [#rspatial](#)  {leafem}. Here's rendering all ~3.3 million building footprints of New Zealand by [@LandInfoNZ](#) straight from an S3 bucket.



PMTiles

Outcomes

- Cloud-native format enables **new use cases**
 - Shareable data science notebook visualizations
 - Offline or air-gapped networks
 - Affordable small-scale civic tech projects
- Already in production for many companies and products

PMTiles spec v3!

PMTiles v3

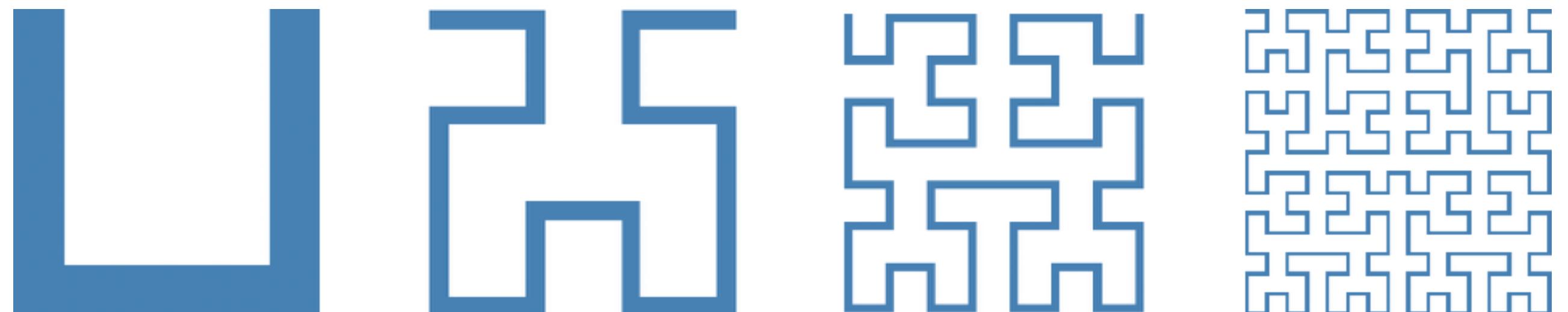
Motivation

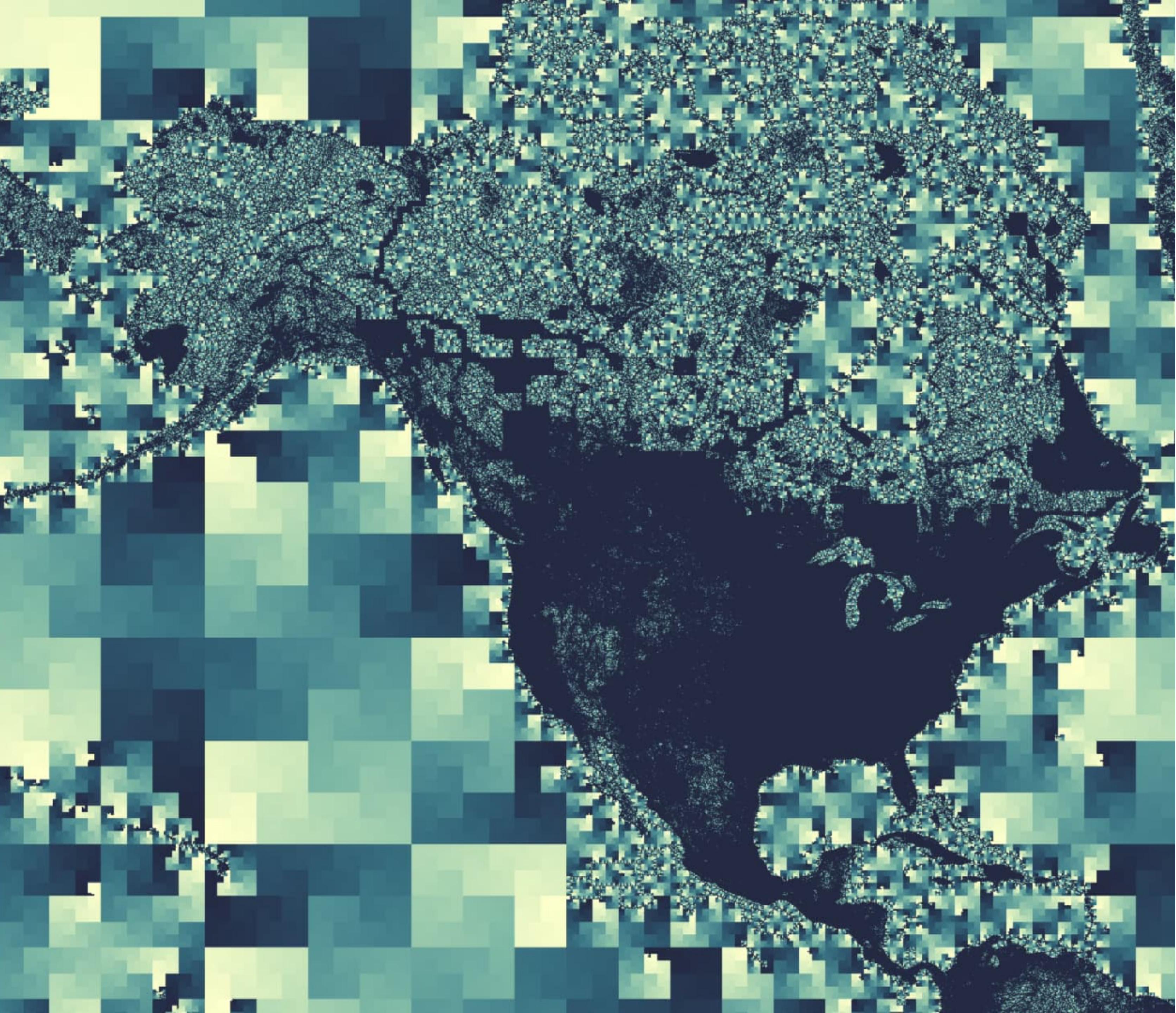
- V2: optimization *not* a design goal; just get it working on many environments
- V3 introduces compression:
 - By default for tile data
 - Compressed index structures
 - Overhead generally goes from ~500 kB to < 50 kB
- Also introduces optional archive **clustering on space-filling curve**

PMTiles v3 addressing

Hilbert + Run-length encoding

- Many raster or vector tile datasets have massive **repetition**
- Internal tile addresses no longer plain z , x , y
 - New way: id is Hilbert pyramid position
 - Encode run length in Hilbert space

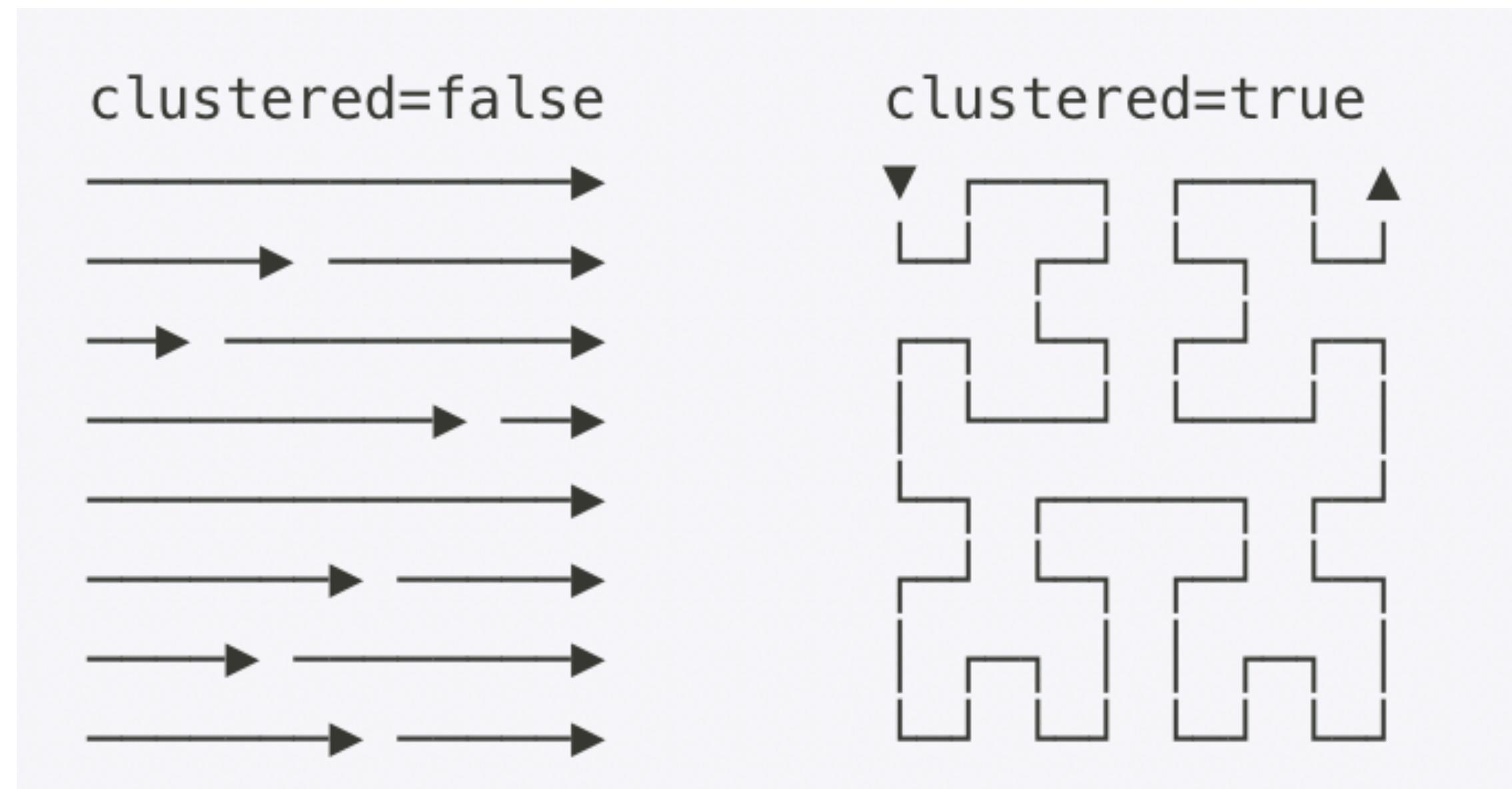




PMTiles v3 addressing

Directory compression

- Test dataset: 357,913,941 tile vector basemap tileset z0-14
- V2 total directory size: **6.0 GB**
- V3 total directory size: **92 MB** with clustered archive



PMTiles v3 addressing

Why not use Z-order curves/Morton codes/Binary quadkeys?

- Hilbert is much more expensive to compute:



vs.



- Hilbert has better locality, all steps are ± 1 in X/Y
- ~3% improvement in RLE, ~5% improvement in total index size
- PMTiles designed for write-once, read-only; I/O is dominant latency factor

PMTiles v3

Ecosystem WIP

- Faster conversion tool from MBTiles, **static binary** with no installation
- Interactive web and local file inspector
- Direct output from popular tile generator tools
 - `tippecanoe input.fgb -o output.pmtiles`
- Readers+writers in more languages

PMTiles Viewer

Specify a remote URL

Load URL

Select a local file

Drag + drop a file here, or click to select

Load an example

https://protomaps-static.sfo3.digitaloceanspaces.com/osm_carto.pmtiles

<https://protomaps-static.sfo3.digitaloceanspaces.com/mantle-trial.pmtiles>

https://protomaps-static.sfo3.digitaloceanspaces.com/cb_2018_us_zcta510_500k_nolimit.pmtiles

https://protomaps-static.sfo3.digitaloceanspaces.com/cb_2018_us_zcta510_500k_gzip.pmtiles

Q Tile Inspector

Leaflet

MapLibre

https://protomaps-static.sfo3.digitaloceanspaces.com/cb_2018_us_zcta510_500k_gzip.pmtiles

Metadata

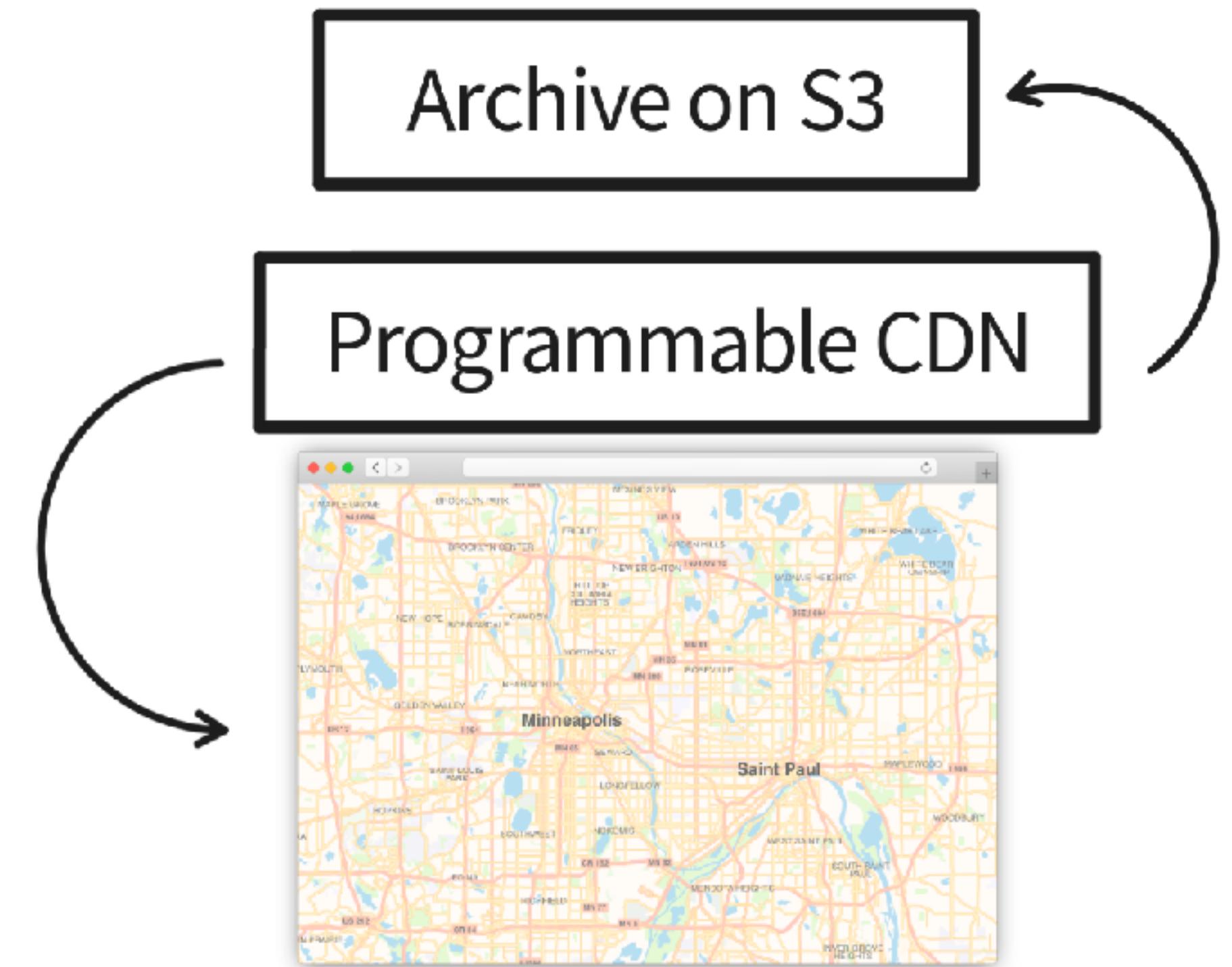
50	8	5443612	231
50	9	5443616	316
50	10	5443209	407
50	16	5443070	139
50	17	5442734	336
51	6	5467805	249
51	7	5464117	3688
51	8	5456708	7409
51	9	5450144	6564
51	10	5449059	1085
51	13	5447575	1484
51	14	5444223	3352
52	6	5520401	611
52	7	5512377	8024
52	8	5492055	20322
52	9	5472626	19429
52	14	5468054	4572
53	7	5538553	3475
53	8	5527730	10823
53	9	5521012	6718
54	9	5570584	7244
54	10	5566191	4393
54	11	5548359	17832
54	12	5542028	6331
55	10	5968907	5168
55	11	5772238	196669
55	12	5578236	194002



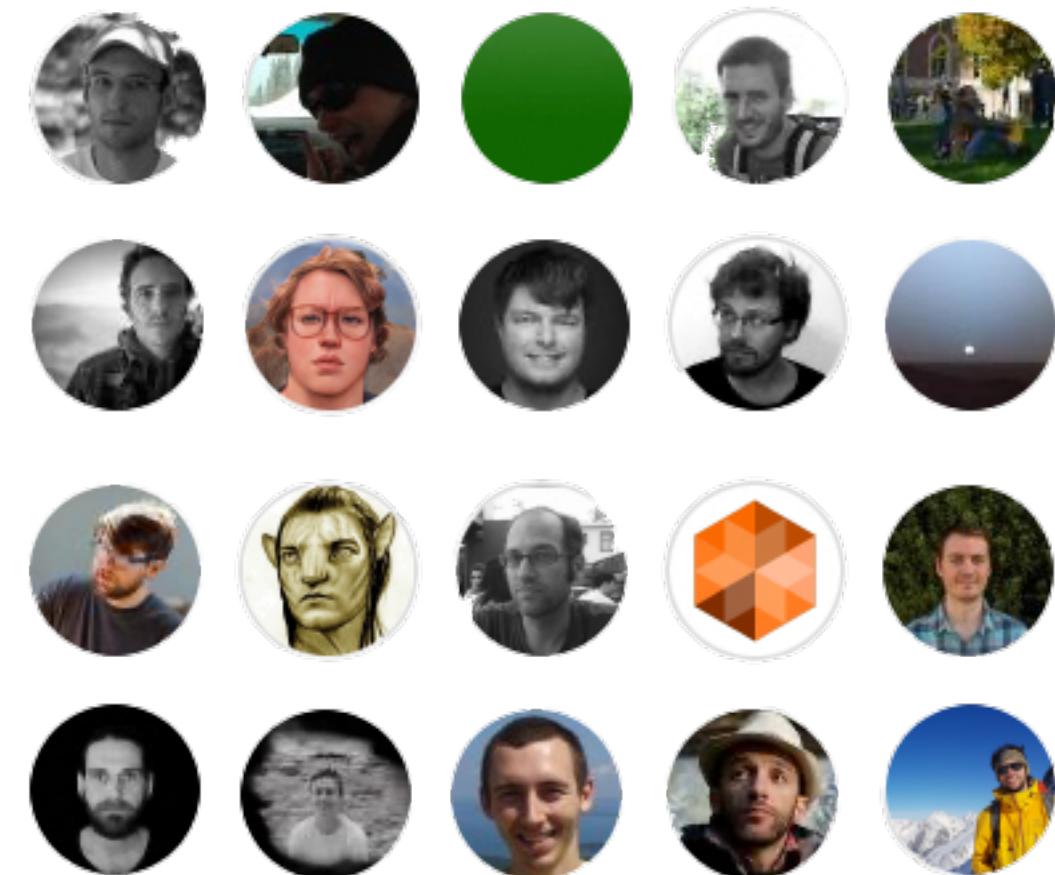
PMTiles v3 ecosystem

Serverless FaaS implementations

- Simulate a traditional Z/X/Y tile endpoint with no server
 - **CDN Edge caching + HTTP2/3 + private buckets**
- Multi-vendor: AWS Lambda + S3, Cloudflare Workers + R2
- ~**2KB** bundled package, copy-and-paste deployment
- `/serverless/` @ GitHub.com/protomaps/PMTiles



***THANK YOU
COLLABORATORS!***



Protomaps

Next steps

- **Commercial support** for organizations:
 - Feature prioritization for open source roadmap
 - Deployment and work with specific datasets
- Contact me!
 - brandon@protomaps.com
 - @bdon / <https://bdon.org>

