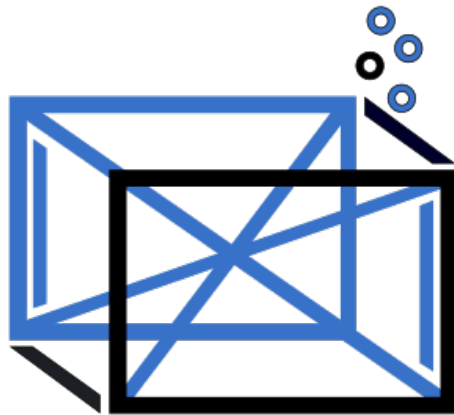


TOAST, Config, and Other Minutia

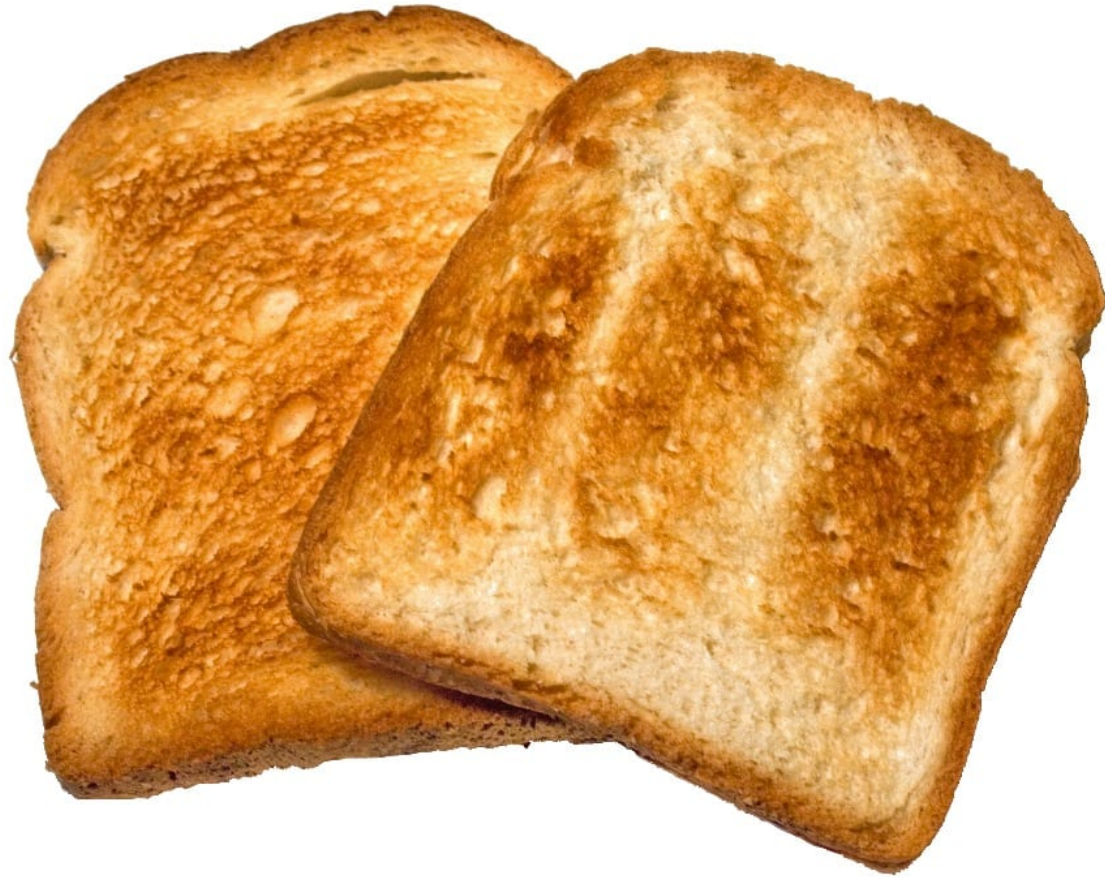


**RustProof Labs**

bringing you data



TOAST



# TOAST

- **T**he
- **O**versized
- **A**tttribute
- **S**torage
- **T**echnique

<https://www.postgresql.org/docs/current/storage-toast.html>

## TOAST and PostGIS

- Know it's there
- Minimize (simplify) geometry data when possible

# Remember the place polygon sizes

- Washington State: 21 kB
- Olympic National Park: 238 kB

# Remember the place polygon sizes

- Washington State: 21 kB
- Olympic National Park: 238 kB
  
- Split into ~ 2 kB chunks
- Compressed

Doesn't compression slow down queries?

Probably not

- Always test your scenario if performance matters
- Partial decompression helps (Postgres 12+)

<https://blog.rustprooflabs.com/2020/07/postgres-storing-large-text>

<https://www.postgresql.org/docs/release/12.0/>





## Memory in Postgres

- Postgres is conservative
- Delegates a lot to OS
- aka: use \*nix

# Memory across Queries

- OS Cache
- `shared_buffers`

<https://www.postgresql.org/docs/15/runtime-config-resource.html>

## Memory after Reboot

- `pg_prewarm`

<https://www.postgresql.org/docs/current/pgprewarm.html>

## Memory within Queries

- `work_mem`

*Used by* `ORDER BY`

## Memory behind the scenes

- `maintenance_work_mem`
- `autovacuum_work_mem`

# shared\_buffers

Postgres default: 128MB

Postgres recommendation: 25%

# work\_mem

Postgres default: 4 MB

Ryan's suggestion: Start at 10 MB

Monitor performance, track temp files & slow queries



`work_mem = 10MB`

- Use `EXPLAIN` to determine how to set `work_mem` safely and smartly!
- `pg_stat_statements` -> identify slow queries and temp file usage

work\_mem

- Set per user type (analyst, webapp, power user)

PostGIS uses Functions  
and Functions Ain't Free

# PostGIS uses Functions

- Poor row estimates ( `COST ROWS 1000` )
- RBAR calculations
- Window functions with PostGIS are cool, but each is a sort

# Index Nuances

- Bounding boxes are great
- Not a silver bullet (see row estimates)

# Index Nuances

- Bounding boxes are great
- Not a silver bullet (see row estimates)
  
- Postgres does not have Clustered Indexes...
- ... and Order still matters

<https://blog.rustprooflabs.com/2020/08/postgres-integer-index-performance>

# Performance Nuances

- Some things just don't run in Parallel
- Spatial Joins often result in Bad Row Estimates

# Performance Nuances

- Filter early, filter often
- Geometries get Complicated
- Simplify when possible
- Consider sub-dividing





# Recap

- Spatial Joins
- Spatial Indexes
- Performance
- Opens door to Amazing Analysis

# Temperature Check

How Familiar with PostGIS? Postgres?



