PFLOCK Report

Andres Calderon

University of California, Riverside

May 22, 2020

Solving issues in local quadtree creation

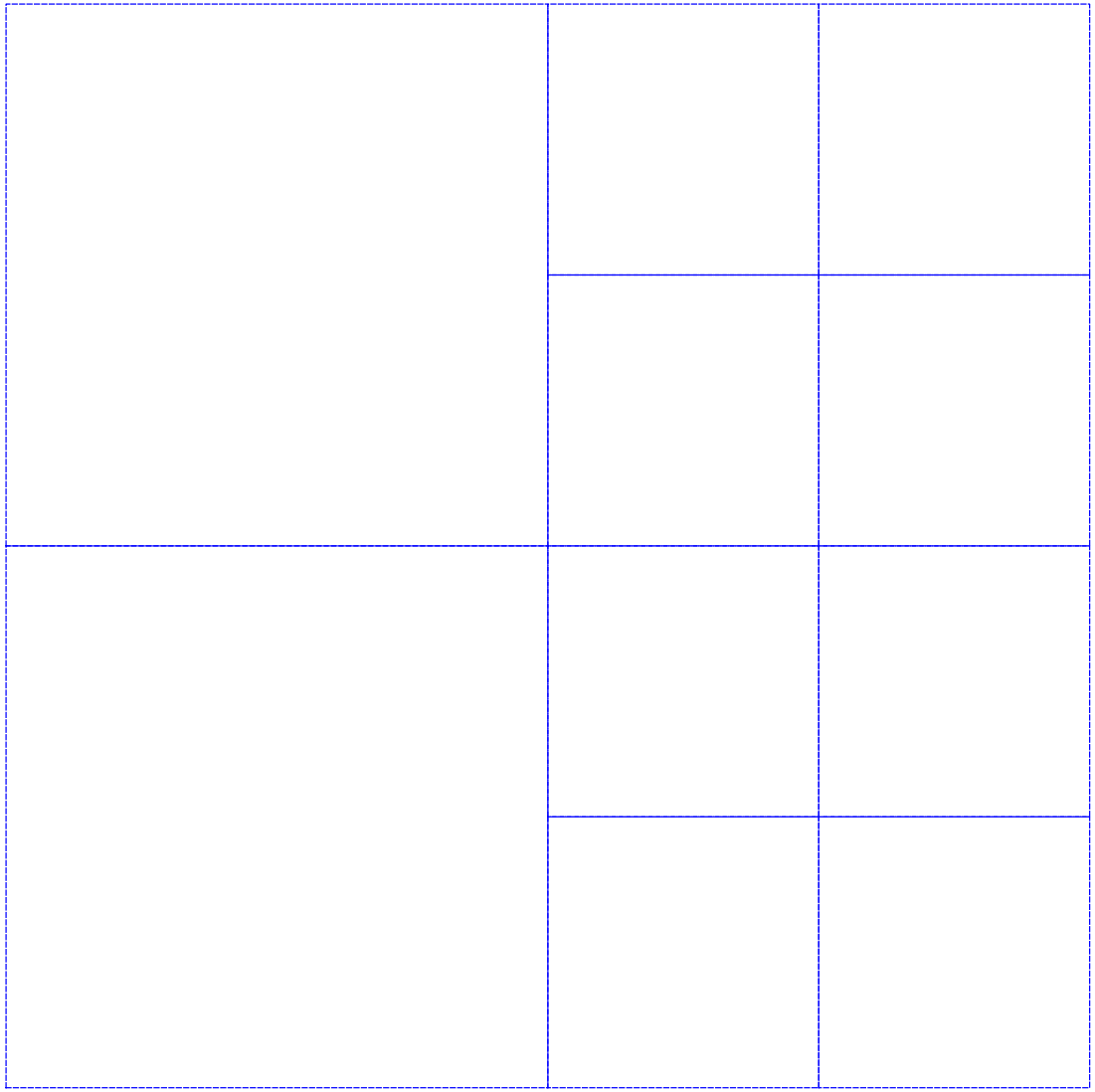
- The GeoSpark Quadtree library receives three parameters: maximum items per node (capacity), a fraction of the total size of the input to extract a sample (fraction), and a number of levels to limit the depth of the tree (levels).
- I fixed a bug which set incorrectly the value of levels. It keeps a default value of 8 which allows up to $4^8 = 64563$ possible leaves.
- As we expect relatively small partitions, I have set this value to 5 (no more than 1024 leaves per partition).

Solving issues in local quadtree creation

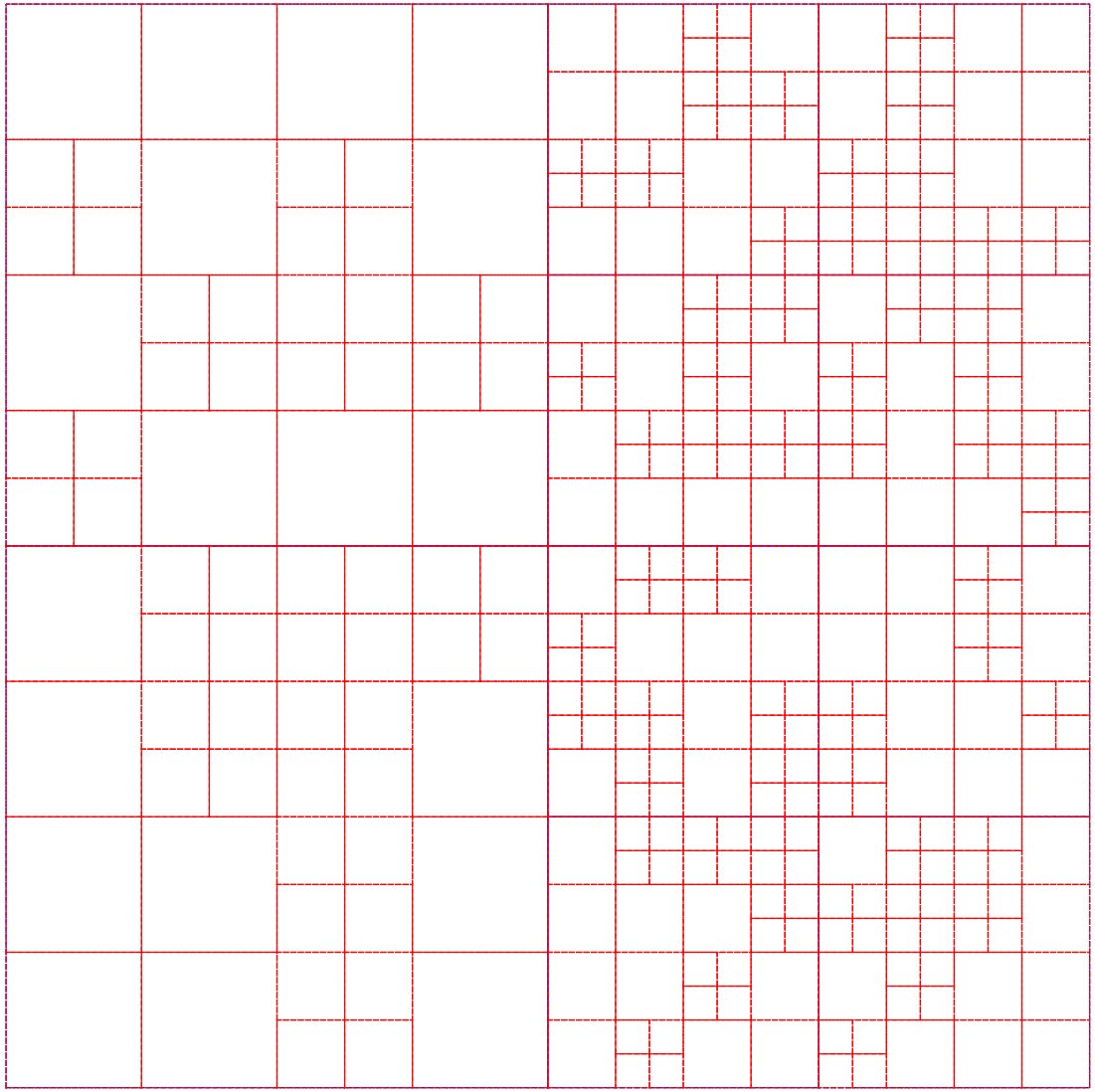
- Follows a similar strategy on GeoSpark repository[1,2] to create global quadtrees.
- Given a number of desired subgrids (x) and a input size (n):
  - capacity = n / x
  - fraction = a scaled value between 1 (for n ¡ 1000) and 0.01 (for n ¿ 10000)
- It works but with a fixed number of subgrids balancing can be a problem...

[1] `https://tinyurl.com/y8fy4ys7`
[2] `https://tinyurl.com/ycsb29gz`

Example



Example

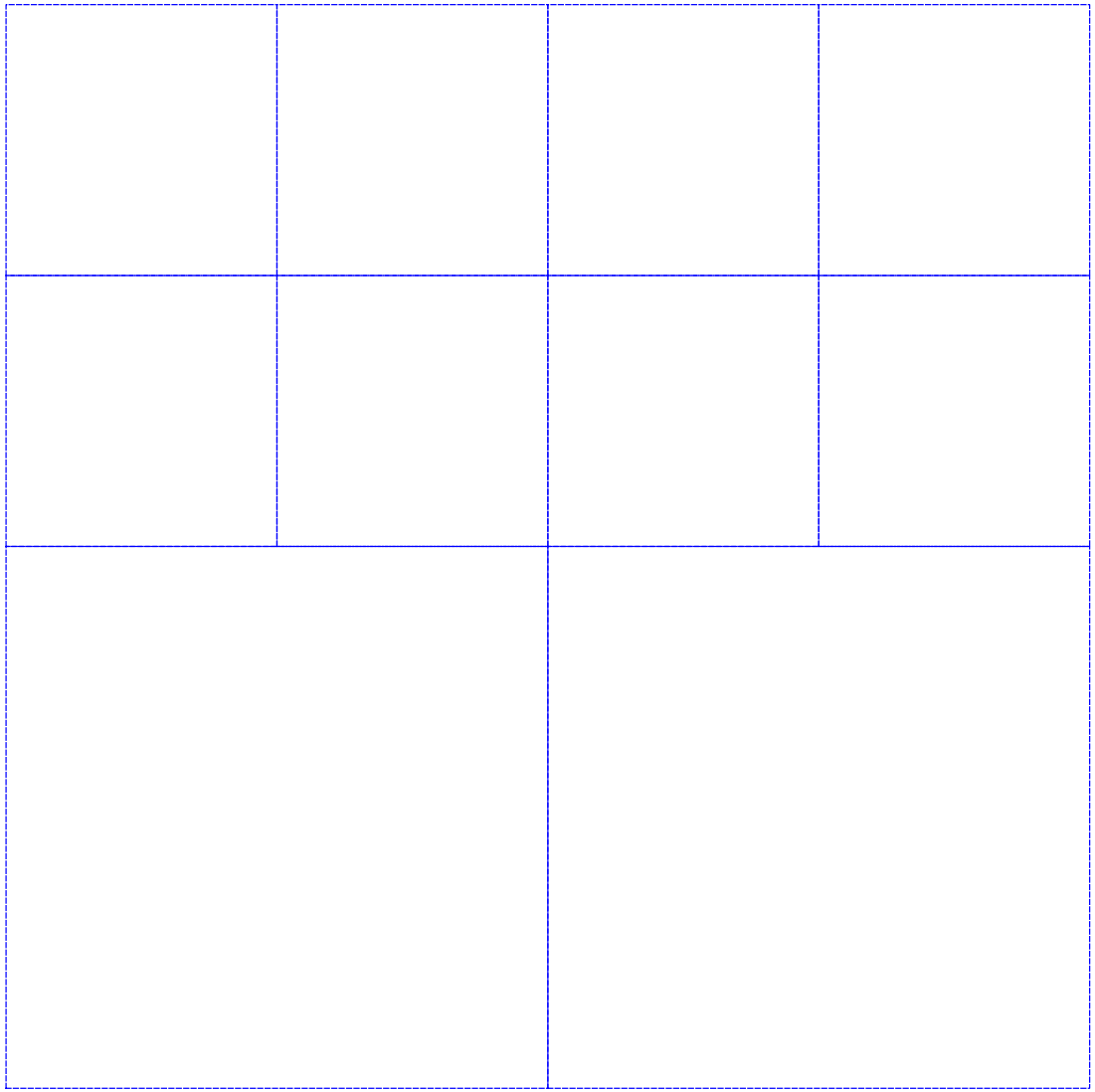


Solving issues in local quadtree creation

- Currently working on optimal values for capacity and fraction.
- Tradeoff between balance and cost.

Example

capacity=50, fraction=1



Example