

Обработка растровых данных на GPU в геоинформационных системах

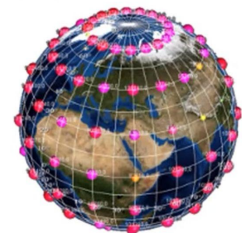
К.И. Сулейманов, Р.А. Родригес Залепинос
Национальный исследовательский университет «Высшая школа экономики»

Введение

Растровая модель данных – это широко распространенный метод хранения географических данных.

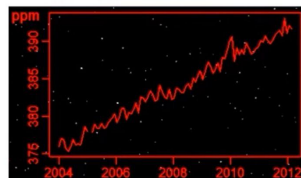
Растры особенно хорошо подходят для хранения непрерывных данных.

Datasets

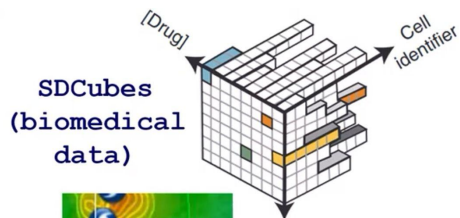


Geospatial grid

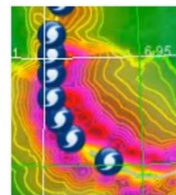
<http://wikience.org>



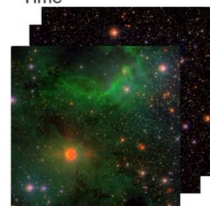
Timeseries



SDCubes
(biomedical data)



Tracks
& swaths

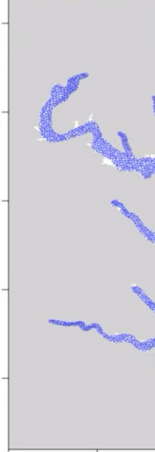


Astronomical
observations

<https://www.sdss.org/>

Mesh

<https://www.ncl.ucla.edu/Application/trimesh.shtml>



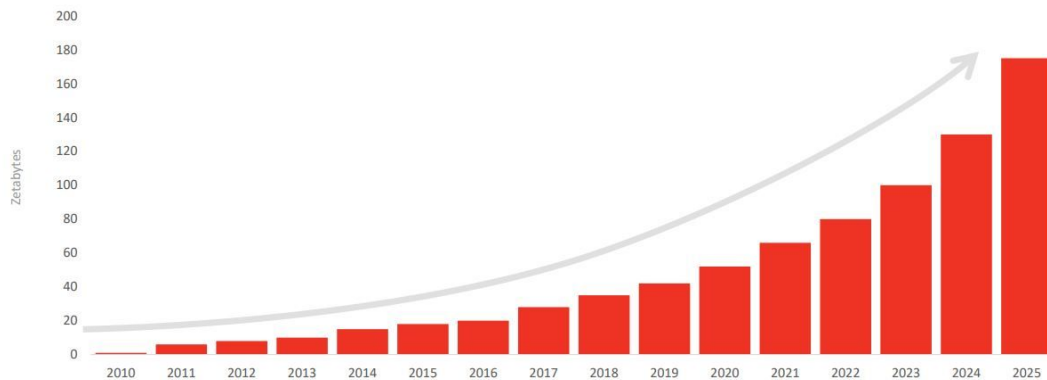
- ArcGIS book. 2021. <https://learn.arcgis.com/en/arcgis-imagery-book/>
- M. Bjorn et al. "Adaptive informatics for multifactorial and high-content biological data." *Nature methods* 8.6 (2011).
- <https://www.hdfgroup.org/portfolio-item/astronomy/>

source: youtu.be/T8n8gQhWHN0

Проблема

Проблема обработки растровых данных заключается в их огромных объёмах. При увеличении разрешения растрового изображения (т.е. уменьшении размера пиксела) объём хранимых данных резко растёт.

Annual Size of Global Digital Data Generated (ZB)



iAB Proprietary Research

Sources: <https://www.forbes.com/sites/tomcoughlin/2018/11/27/175-zettabytes-by-2025/#3314b7575459>; <https://www.networkworld.com/article/3325397/dc-expect-175-zettabytes-of-data-worldwide-by-2025.html>

Цель работы

Разработка универсальных GPU-ускоренных подходов для пакетной обработки геопривязанных растровых данных в ГИС. Ожидается достичь 5–10 кратного сокращения времени анализа спутниковых данных и аэрофотоснимков и гибкая интеграция с популярными ГИС-платформами с целью масштабирования обработки растровых данных.

Существующие GPU-библиотеки и утилиты

Аналогов полностью предполагающих решение для массивно-параллельной обработки данных ГИС нет.

- CUDA C дало ускорение в 7 раз по сравнению с ArcGIS.
- В QGIS с PyCUDA ускорение было в 3 раза.

Спасибо!