



ClickHouse



Real Estate
Analytics



Real Estate Analytics

Yi Sam Lee

Data Lead @ Real Estate Analytics (REA)



TABLE OF CONTENT

0

About REA

1

Slicing & Dicing

2

Extraction & Parsing



REAL ESTATE ANALYTICS (REA)

The Leading Real Estate AI Provider in Asia

At Real Estate Analytics (REA), we revolutionise the real estate industry with cutting-edge AI technology. Leveraging advanced data science and machine learning, we offer tailored data solutions for real estate professionals and enthusiasts.



Founded in Singapore in 2019 by Dr. Jean-Michel Paul, Patrick Yan, Jiajia Yao. Chairman Pr. R Edelstein



HQ in Singapore, active in Australia, Hong Kong, and Malaysia



Real Estate
Analytics



REAL Agent

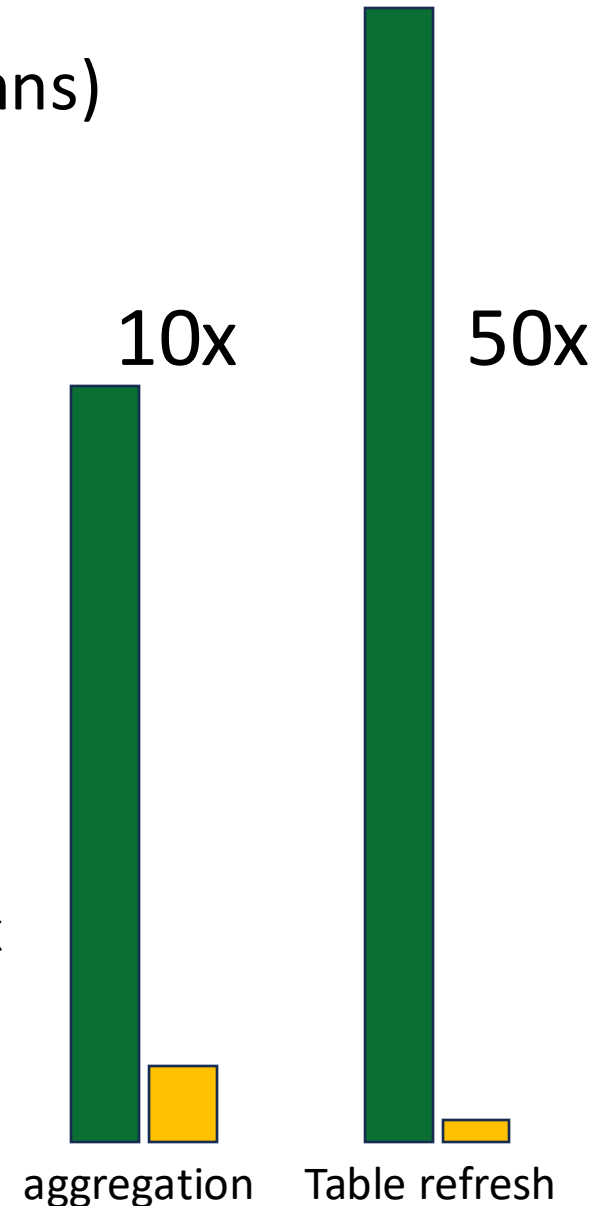


REAL  insight

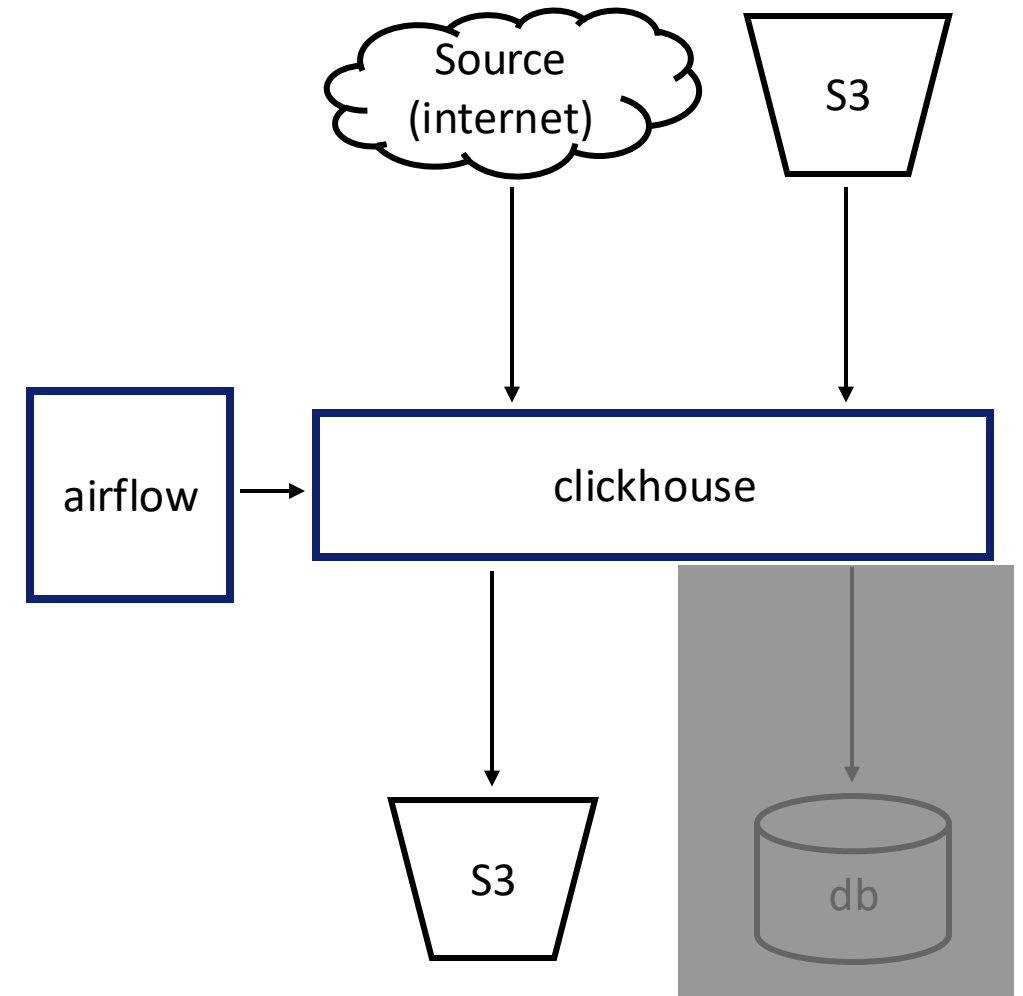
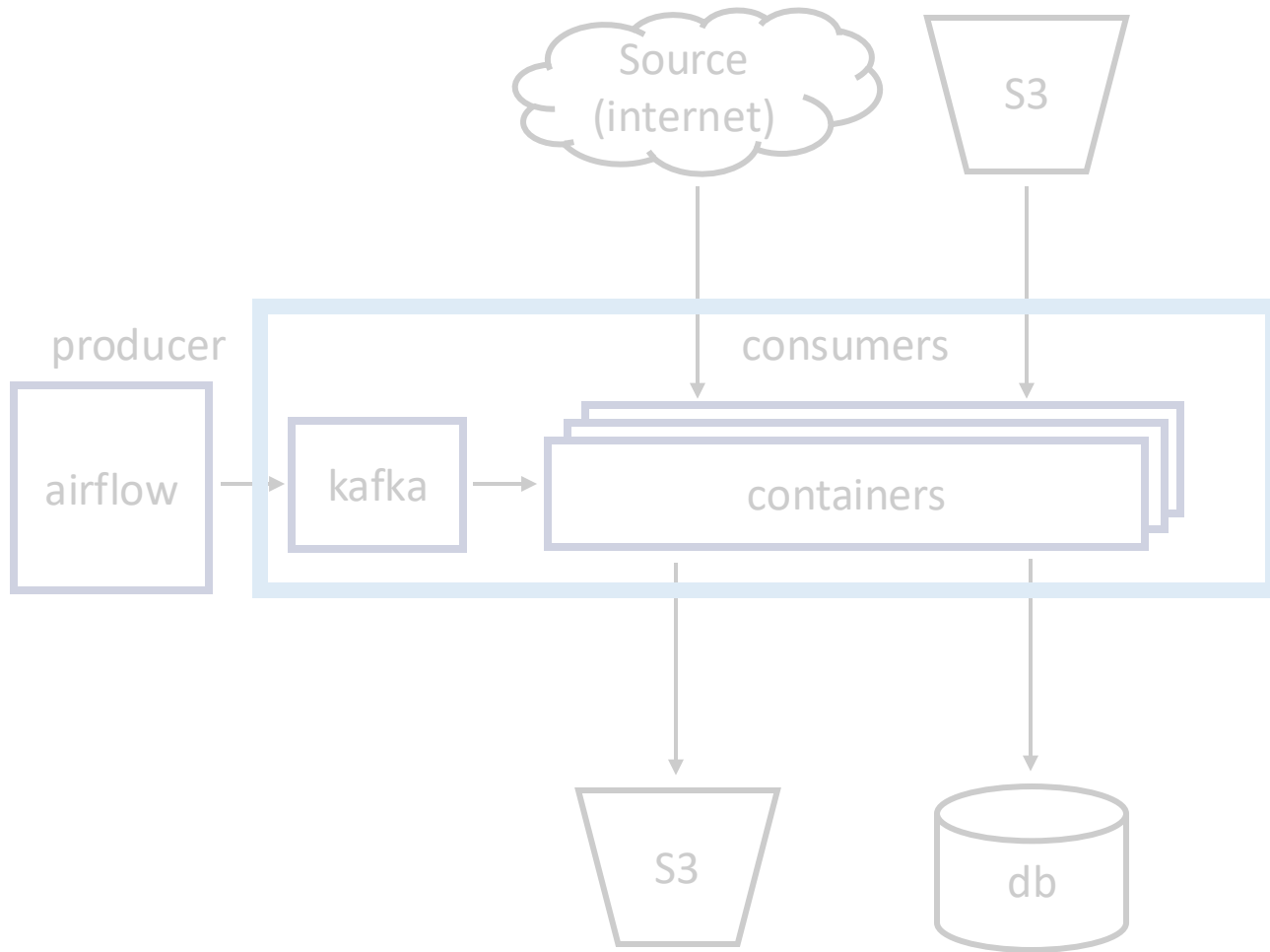
Slicing and Dicing - Aggregation & point search



- For data size in the range of 3 - 20 mil rows (> 200 columns)
- Mongodb vs Clickhouse
 - Aggregation
 - about 10x faster in clickhouse
 - Point search
 - Comparable speed for comparable resource – with index
 - Speed for table refresh is significantly faster in clickhouse
 - <30s
 - 25-30min (1500-1800s)
 - Mongodb -> 10% additional time for each additional index



Extraction and Parsing



Why not use clickhouse



- Globs / path functions
 - 1..10
 - * / **
 - zip file reading
- Insert into function s3
 - Save file to s3
 - partition by
- Parsing functions
 - extract / extractAll
 - JSON_QUERY...



- Demo
 - data.gov.sg - [link](#)
 - Property website - [link](#)
 - Medical centre website - [link](#)

Some issues we currently face



- Date starts from 1970 (date) or 1900 (datetime64)
- No wkt format for geometry for display in dbeaver
- No auto generated sequence