

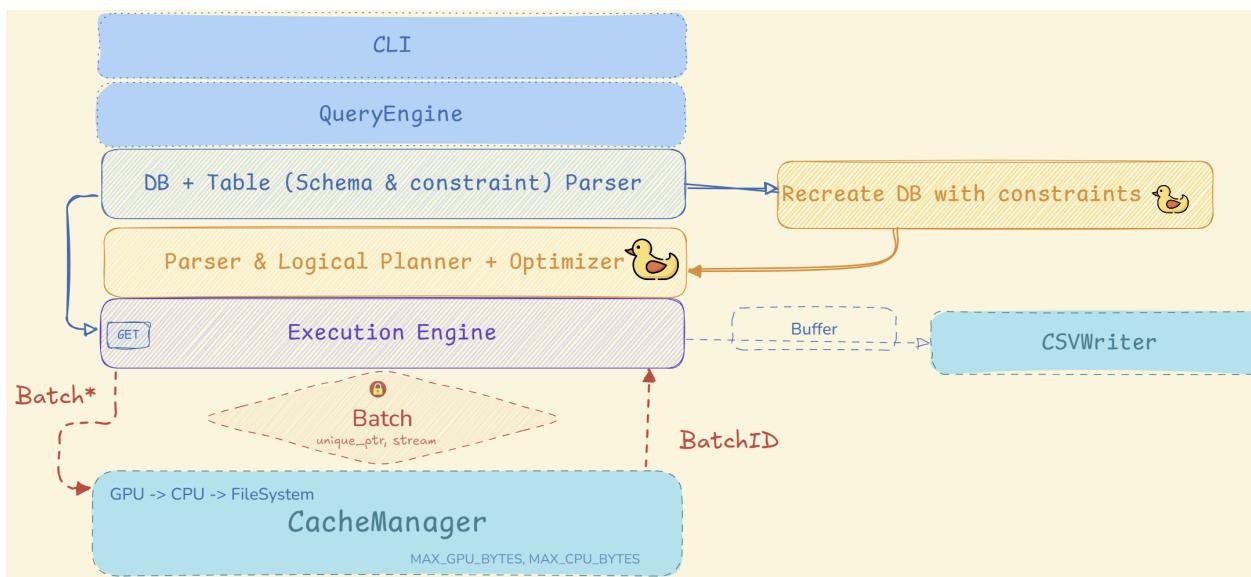


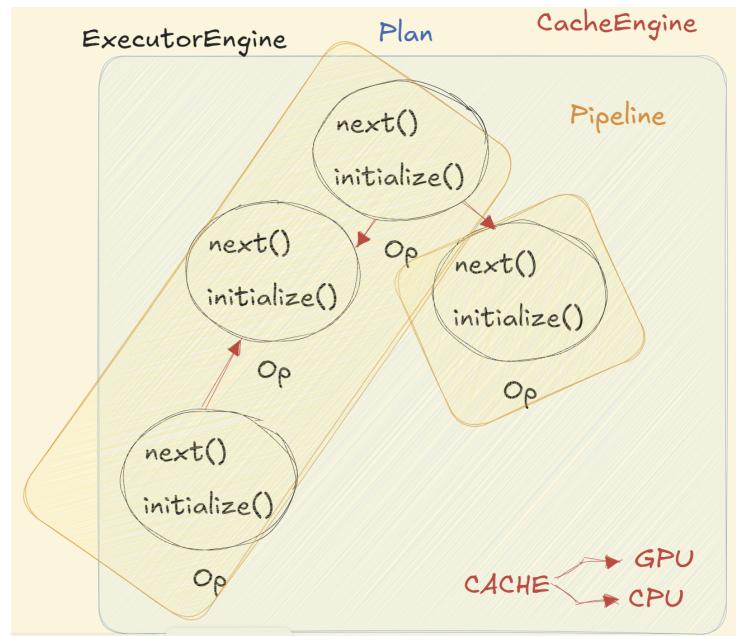
# PC Project Report

## Team 10: YallaSQL GPU QueryProcessor

Names	Bn	Sec
Aliaa Abdelazize Ali Abduallah	27	1
Ahmed Osama Helmy	5	1
Omar Mahmoud	29	1
Abdallah Ahmed	25	1

## Architecture





## Kernels Implementation:

- Binary, Cast operators, Comparison operators → utilizing **Coarsening**, **Streaming**
  - Filters → getting Mask → prefix sum to get index → move rows
  - Aggregates
    - → reduction on each wrap using optimized algorithm used in lecture
    - → reduction of all streams, blocks connected to one place
    - → float casted as double for precision in sum then return back to sum
  - Sorting →
    - local sort on each batch using **Radix Sort** → cache it
    - out-of-core merge sort
      - each RUN contain cached batches
      - run merge sort on key columns & store indexs
    - same algorithms in lecture slides
  - Join →
    - with  $2 * \text{MAX\_ROWS\_SELECTIVITY}$  store idxs to select and pass it
  - Cross Product →
    - $\text{LeftBatchSize} \times \text{RightBatchSize}$

## Benchmark SSB-Benchmark

all this statistics done using 5 warmups and 10 runs & all end to end from moment giving query till the writing of whole csv

6M rows - 270mb

```
Database switched successfully
14:29 yallaSQL λ SELECT s.supkey, s.name, l.supplier_supkey, l.supplycost FROM lineorder l, supplier s WHERE l.supplier_supkey
= s.supkey;
< ⏱ execute time>
Mean: 5197.3ms, Median: 5143 ms, Min: 4929 ms, Max: 5912 ms, StdDev: 280.499 ms
Query executed successfully
14:30 yallaSQL λ duckdb SELECT s.supkey, s.name, l.supplier_supkey, l.supplycost FROM lineorder l, supplier s WHERE l.supplier_supkey
= s.supkey;
< ⏱ execute time>
Mean: 5951.5ms, Median: 6057 ms, Min: 5362 ms, Max: 6550 ms, StdDev: 442.109 ms
Query executed successfully
```

```
11:52 yallaSQL λ SELECT count(*) from lineorder;
< ⏱ execute time>
Mean: 2852.2ms, Median: 2853 ms, Min: 2710 ms, Max: 3008 ms, StdDev: 99.1835 ms
Query executed successfully
11:53 yallaSQL λ SELECT AVG(discount) from lineorder;
< ⏱ execute time>
Mean: 2627.6ms, Median: 2607 ms, Min: 2525 ms, Max: 2891 ms, StdDev: 104.06 ms
Query executed successfully
```

```
select min(revenue) as min_revenue from lineorder;
//Mean: 2898.4ms, Median: 2886 ms, Min: 2537 ms, Max: 3566 ms, StdDev: 303.469 ms
```

```
SELECT revenue, quantity FROM lineorder where quantity < 20;
// more than 600mb
```

```

23:13  yallaSQL λ SELECT
    lo.orderkey,
    lo.linenumber,
    c.name,
    s.name,
    p.name,
    d.date
  FROM
    lineorder lo,
    customer c,
    supplier s,
    part p,
    date d
 WHERE
    lo.customer_custkey = c.custkey
    AND lo.supplier_suppkey = s.suppkey
    AND lo.part_partkey = p.partkey
    AND lo.date_datekey_orderdate = d.datekey
    AND c.region = 'AMERICA'
    AND s.region = 'AMERICA'
    AND p.category = 'MFGR#12'
    AND d.year = 1997;
<⌚ execute time>
Mean: 10467.5ms, Median: 10108 ms, Min: 8197 ms, Max: 17276 ms, StdDev: 2409.32 ms

```

## Experiments with different parameters

- using MAX\_BYTES\_PER\_BATCH = 8mb, BLOCK\_DIM=256, COARSENING\_FACTOR=5
  1. Mean: 2858.6ms, Median: 2855 ms, Min: 2628 ms, Max: 3239 ms, StdDev: 188.652 ms
- using MAX\_BYTES\_PER\_BATCH = 16mb, BLOCK\_DIM=256, COARSENING\_FACTOR=5
  1. Mean: 3011.8ms, Median: 2930 ms, Min: 2603 ms, Max: 3716 ms, StdDev: 358.025 ms
- using MAX\_BYTES\_PER\_BATCH = 16mb, BLOCK\_DIM=256, COARSENING\_FACTOR=10
  1. Mean: 2993.7ms, Median: 3017 ms, Min: 2799 ms, Max: 3172 ms, StdDev: 123.509 ms
- using MAX\_BYTES\_PER\_BATCH = 16mb, BLOCK\_DIM=512, COARSENING\_FACTOR=10
  1. Mean: 3238.7ms, Median: 3257 ms, Min: 3098 ms, Max: 3326 ms, StdDev: 57.5449 ms
- using MAX\_BYTES\_PER\_BATCH = 16mb, BLOCK\_DIM=512, COARSENING\_FACTOR=5
  1. Mean: 3276.7ms, Median: 3273 ms, Min: 3025 ms, Max: 3780 ms, StdDev: 195.156 ms
- using MAX\_BYTES\_PER\_BATCH = 4mb, BLOCK\_DIM=256, COARSENING\_FACTOR=5
  1. Mean: 3022.4ms, Median: 3046 ms, Min: 2872 ms, Max: 3204 ms, StdDev: 109.107 ms
  2. Mean: 2788.5ms, Median: 2805 ms, Min: 2683 ms, Max: 2864 ms, StdDev: 49.261 ms
- using MAX\_BYTES\_PER\_BATCH = 4mb, BLOCK\_DIM=256, COARSENING\_FACTOR=2
  1. Mean: 2750.5ms, Median: 2749 ms, Min: 2705 ms, Max: 2822 ms, StdDev: 34.3344 ms

2. Mean: 2683.4ms, Median: 2673 ms, Min: 2542 ms, Max: 2865 ms, StdDev: 88.98 ms