# Optimization of Common Table Expressions in MPP Database Systems

Amr El-Helw, Venkatesh Raghavan, Mohamed A. Soliman, George Caragea, Zhongxian Gu, Michalis Petropoulos

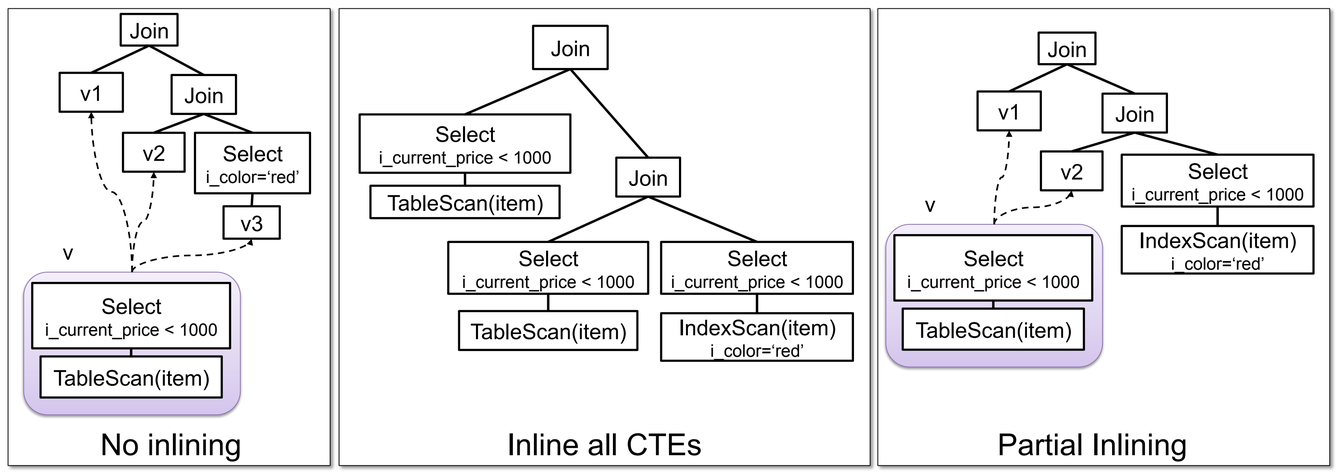
Pivotal

## I. Motivation

**Example Query:**

```
WITH v as (SELECT i_brand, i_color FROM item WHERE i_current_price < 1000)
SELECT v1.* FROM v v1, v v2, v v3
WHERE v1.i_brand = v2.i_brand AND v2.i_brand = v3.i_brand AND v3.i_color = 'red';
```

Index on `i_color`




**Challenges:**

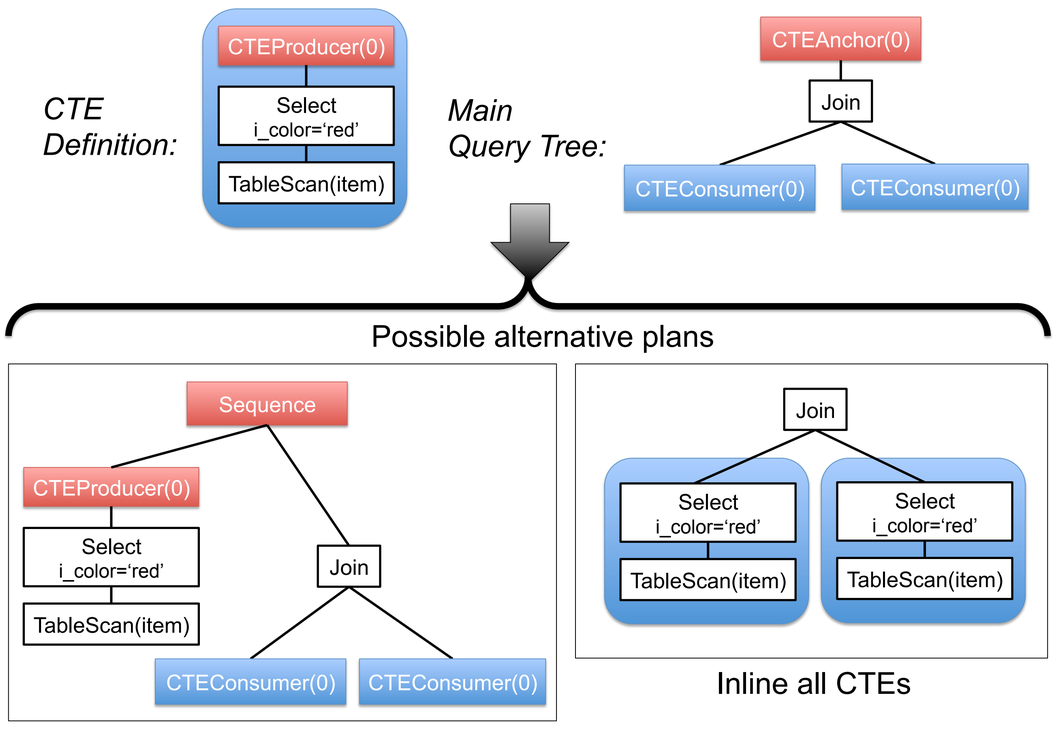
- Deadlock hazard → Hanging queries
- Inlining heuristics → suboptimal plans } Slow execution
- CTEs optimized in isolation from context }

**Goals:**

- Cost-based inlining approach
- Contextualized optimization
- Pushing filters, ordering into CTEs
- Deadlock-free execution

## II. CTE Representation in Orca

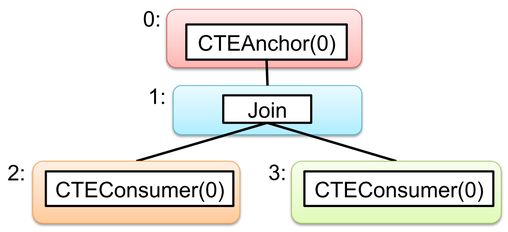
```
WITH v AS (SELECT i_brand FROM item WHERE i_color = 'red')
SELECT * FROM v as v1, v as v2 WHERE v1.i_brand = v2.i_brand;
```



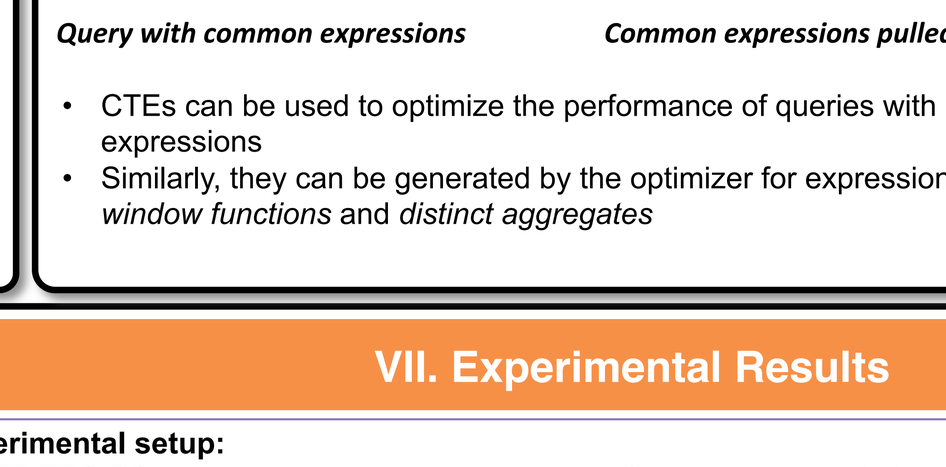

## III. Plan Enumeration

1. Main query tree inserted into compressed MEMO structure

   Each operator initializes a new *group* in the MEMO

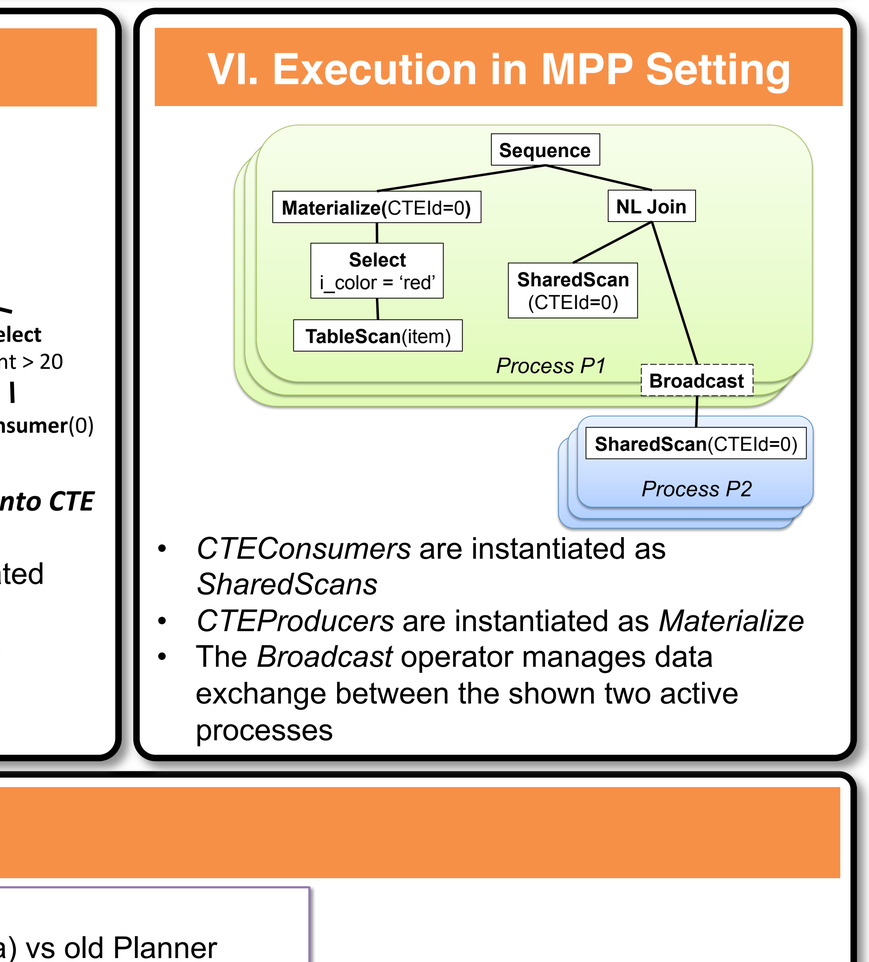



2. Transformation rules are applied to generate equivalent alternatives, which are also added to the MEMO




3. Different combinations of operators can be used together to form different plans

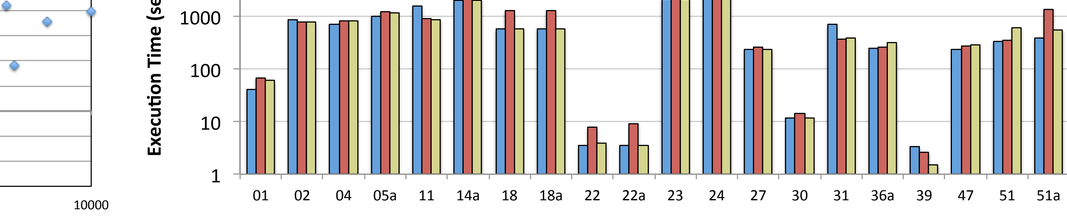
   Properties are used to ensure that only the valid plans are generated




## IV. Predicate Push-down



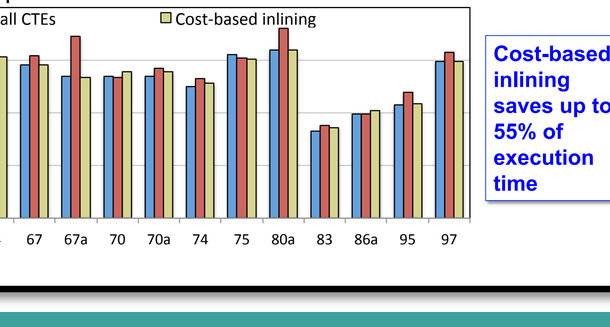

## V. Common Subexpression Elimination




***Query with common expressions*** ***Common expressions pulled out into CTE***

- CTEs can be used to optimize the performance of queries with repeated expressions
- Similarly, they can be generated by the optimizer for expressions like *window functions* and *distinct aggregates*

## VI. Execution in MPP Setting




- *CTEConsumers* are instantiated as *SharedScans*
- *CTEProducers* are instantiated as *Materialize*
- The *Broadcast* operator manages data exchange between the shown two active processes

## VII. Experimental Results

**Experimental setup:**

- 5TB TPC-DS benchmark
- 8-node cluster, CPU: 3.33GHz, RAM: 48GB
- Compare new optimizer (Orca) vs old Planner
- Compare Orca using different CTE inlining settings

Performance improvement of Orca vs. Planner




Performance comparison for different CTE inlining approaches




**Cost-based inlining saves up to 55% of execution time**