Introduction to Query Engines

Scope

Just the **Data Query Language** aspects - that's more or less the bit that handles `SELECT` statements.

Will cover generic aspects of implementation, but will include detail relating to Opteryx.

Key Steps

Query Language Interpretation

Query Planning and Optimization

Execution Engine

Files / Storage

Key Steps

SQL -> **Query Language Interpretation**

Abstract Syntax Tree -> **Query Planning and Optimization**

Query Plan -> **Execution Engine**

Resource Access -> **Files / Storage**

Result Creation

Key Components

Parser / Lexer Interprets SQL into a semantic representation (AST)
Abstract Syntax Tree (AST) First machine processable representation of the query (we can rewrite the query here)
Query Plan Describes the steps to take to fulfil the request
Optimizer Reworks the Query Plan to improve performance
Executor Runs the Query Plan

Fixed Query Plan

Based on Relational Algrebra.

This is the order items are processed before optimizations.

Has implications, e.g. can't `GROUP BY` aliases defined in the `SELECT` clause.

Naive Plan Order

SELECT (5) [project]
DISTINCT (6) [distinct]
FROM (1)
WHERE (2) [select]
GROUP BY (3) [aggregate]
HAVING (4) [select]
ORDER BY (7) [sort]
OFFSET (8)
LIMIT (9)

Plan Optimization

Optimized plan has to create the same result as naive plan.

Get rid of data (rows and columns) as quickly as possible

- Selection (`WHERE`) and Projection (`SELECT`) push-downs
- `LIMIT` push-downs

Algorithm Decisions

- Choose `JOIN` order and algorithm (HASH or SORT MERGE)

Execution Models

- Row Processing (Volcano) **Mabel**
- Block/Column Processing (Vectorized) **Opteryx**

Volcano Model

1. The step at the end of our plan tries to return a record
2. It asks the previous step, which asks the previous step
3. Until we get to the files, which we read line-by-line

All calculations are done on each line, one at a time

Block/Column Processing

1. The step at the end of our plan tries to return a block
2. It asks the previous step, which asks the previous step
3. Util we get to the files, which we read an entire file/block

All calculations are done block at a time.