Databases!
**ADMINISTRIVIA**

**Mid-Term Grade**
- Release today
- Attend instructor’s office hour to check your exam

**Lin’s office hour** changed to Tue @ 2:30pm-4:00pm, GHC 9019

**Project #2** is due on Sun Thu 21\(^{th}\) @ 11:59pm

**Project #3** will be released on Wed Oct 20\(^{th}\)
QUERY OPTIMIZATION

Remember that SQL is declarative.
→ User tells the DBMS what answer they want, not how to get the answer.

There can be a big difference in performance based on plan is used
IBM SYSTEM R

First implementation of a query optimizer from the 1970s.
→ People argued that the DBMS could never choose a query plan better than what a human could write.

Many concepts and design decisions from the System R optimizer are still used today.
Heuristics / Rules

→ Rewrite the query to remove stupid / inefficient things.
→ These techniques may need to examine catalog, but they do not need to examine data.

Cost-based Search

→ Use a model to estimate the cost of executing a plan.
→ Evaluate multiple equivalent plans for a query and pick the one with the lowest cost.
ARCHITECTURE OVERVIEW
ARCHITECTURE OVERVIEW

Application

1 SQL Query

SQL Rewriter
(Optional / Rare)
ARCHITECTURE OVERVIEW

1. SQL Query

2. SQL Query

Parser

SQL Rewriter
(Optional / Rare)

Application
ARCHITECTURE OVERVIEW

1. **SQL Query**
   - SQL Rewriter
     - (Optional / Rare)

2. **SQL Query**
   - Binder

3. **Abstract Syntax Tree**

Application
ARCHITECTURE OVERVIEW

1. SQL Query
   SQL Rewriter
      (Optional / Rare)

2. SQL Query
   Binder

3. Abstract Syntax Tree
   Name → Internal ID

Application
System Catalog
Parser
ARCHITECTURE OVERVIEW

1. SQL Query
   - SQL Rewriter
     - (Optional / Rare)
2. SQL Query
   - Binder
3. Abstract Syntax Tree
   - Parser
4. Logical Plan
   - System Catalog
   - Tree Rewriter
     - (Optional / Common)

Application Name→Internal ID
**ARCHITECTURE OVERVIEW**

1. **Application**
   - SQL Query

2. **SQL Query**
   - **SQL Rewriter** (Optional / Rare)

3. **Abstract Syntax Tree**
   - Name → Internal ID
   - Schema Info

4. **Logical Plan**
   - **Tree Rewriter** (Optional / Common)

**Components:**
- **Parser**
- **Binder**
- **System Catalog**
- **Tree Rewriter**
- **SQL Rewriter**

**Process:**
1. Application submits SQL Query
2. SQL Query goes through SQL Rewriter
3. SQL Query is parsed into Abstract Syntax Tree
4. Binder creates Logical Plan using Schema Info and Dynamic System Catalog

**Note:**
- SQL Rewriter is optional and rare.
- Tree Rewriter is optional and common.
ARCHITECTURE OVERVIEW

1. SQL Query
   - SQL Rewriter (Optional / Rare)

2. SQL Query
   - Binder

3. Abstract Syntax Tree

4. Logical Plan
   - Tree Rewriter (Optional / Common)

5. Logical Plan
   - Optimizer

Application

System Catalog

Schema Info

Name→Internal ID
ARCHITECTURE OVERVIEW

1. SQL Query
   (Optional / Rare)

2. SQL Query

3. Abstract Syntax Tree

4. Logical Plan

5. Logical Plan
   (Optional / Common)

Optimizer

System Catalog

Parser

Binder

Application

Schema Info

Tree Rewriter

Name→Internal ID

SQL Rewriter

Logical Plan
ARCHITECTURE OVERVIEW

1. Application → SQL Query

   SQL Query

2. SQL Query → Parser

   Parse

3. Abstract Syntax Tree

4. Logical Plan

   Binder

5. Logical Plan

   Tree Rewriter

   (Optional / Common)

   (Optional / Rare)

   System Catalog

   Schema Info

   Name → Internal ID

   Cost Model

   Estimates

   Optimizer

   SQL Rewriter

   Schema Info
**ARCHITECTURE OVERVIEW**

1. **SQL Query**
   - Application
   - SQL Rewriter
     - Optional / Rare

2. **SQL Query**
   - Parser
     - Abstract Syntax Tree

3. Binder
   - Name→Internal ID
   - Schema Info

4. **Logical Plan**
   - Tree Rewriter
     - Optional / Common

5. Logical Plan
   - Optimizer
     - Cost Model

6. **Physical Plan**
The optimizer generates a mapping of a logical algebra expression to the optimal equivalent physical algebra expression.

Physical operators define a specific execution strategy using an access path.
→ They can depend on the physical format of the data that they process (i.e., sorting, compression).
→ Not always a 1:1 mapping from logical to physical.
QUERY OPTIMIZATION IS NP-HARD

This is the hardest part of building a DBMS.
If you are good at this, you will get paid $$$.

People are starting to look at employing ML to improve the accuracy and efficacy of optimizers.
→ IBM DB2 tried this with LEO in the early 2000s...
TODAY'S AGENDA

Relational Algebra Equivalences
Logical Query Optimization
Nested Queries
Expression Rewriting

Cost Model
RELATIONAL ALGEBRA EQUIVALENCES

Two relational algebra expressions are equivalent if they generate the same set of tuples.

The DBMS can identify better query plans without a cost model.

This is often called query rewriting.
SELECT s.name, e.cid
FROM student AS s, enrolled AS e
WHERE s.sid = e.sid
AND e.grade = 'A'

π\text{name, cid}(\sigma\text{grade}='A'(\text{student} \bowtie \text{enrolled})))
SELECT s.name, e.cid
FROM student AS s, enrolled AS e
WHERE s.sid = e.sid
AND e.grade = 'A'
**SELECT** s.name, e.cid  
**FROM** student AS s, enrolled AS e  
**WHERE** s.sid = e.sid  
**AND** e.grade = 'A'

**PREDICATE PUSHDOWN**

π_{s.name,e.cid} σ_{s.sid=e.sid} student × enrolled →  
π_{s.name,e.cid} σ_{grade='A'} student × enrolled
RELATIONAL ALGEBRA EQUIVALENCES

\[
\pi_{\text{name, cid}}(\sigma_{\text{grade='A'}(\text{student} \bowtie \text{enrolled}))}
\]

\[
= \\
\pi_{\text{name, cid}}(\text{student} \bowtie (\sigma_{\text{grade='A'}(\text{enrolled})))))
\]
Selections:

→ Perform filters as early as possible.
→ Break a complex predicate, and push down

$\sigma_{p_1 \land p_2 \land \ldots \land p_n}(R) = \sigma_{p_1}(\sigma_{p_2}(\ldots \sigma_{p_n}(R))))$

Simplify a complex predicate

→ $(X=Y \text{ AND } Y=3) \rightarrow X=3 \text{ AND } Y=3$
RELATIONAL ALGEBRA EQUIVALENCES

Joins:
→ Commutative, associative
\[ R \bowtie S = S \bowtie R \]
\[ (R \bowtie S) \bowtie T = R \bowtie (S \bowtie T) \]

The number of different join orderings for an \( n \)-way join is a **Catalan Number** \( (\approx 4^n) \)
→ Exhaustive enumeration will be too slow.
RELATIONAL ALGEBRA EQUIVALENCES

Projections:

→ Perform them early to create smaller tuples and reduce intermediate results (if duplicates are eliminated)
→ Project out all attributes except the ones requested or required (e.g., joining keys)

This is not important for a column store...
SELECT s.name, e.cid
FROM student AS s, enrolled AS e
WHERE s.sid = e.sid
AND e.grade = 'A'
PROJECTION PUSHDOWN

\[
\begin{align*}
\text{SELECT } & \text{s.name, e.cid} \\
\text{FROM } & \text{student AS s, enrolled AS e} \\
\text{WHERE } & \text{s.sid = e.sid} \\
\text{AND } & \text{e.grade = 'A'}
\end{align*}
\]
LOGICAL QUERY OPTIMIZATION

Transform a logical plan into an equivalent logical plan using pattern matching rules.

The goal is to increase the likelihood of enumerating the optimal plan in the search.

Cannot compare plans because there is no cost model but can "direct" a transformation to a preferred side.
LOGICAL QUERY OPTIMIZATION

Split Conjunctive Predicates
Predicate Pushdown
Replace Cartesian Products with Joins
Projection Pushdown
SELECT ARTIST.NAME
FROM ARTIST, APPEARS, ALBUM
WHERE ARTIST.ID=APPEARS.ARTIST_ID
    AND APPEARS.ALBUM_ID=ALBUM.ID
    AND ALBUM.NAME="Andy's OG Remix"

Decompose predicates into their simplest forms to make it easier for the optimizer to move them around.
SELECT ARTIST.NAME  
FROM ARTIST, APPEARS, ALBUM  
WHERE ARTIST.ID=APPEARS.ARTIST_ID  
AND APPEARS.ALBUM_ID=ALBUM.ID  
AND ALBUM.NAME="Andy's OG Remix"

Decompose predicates into their simplest forms to make it easier for the optimizer to move them around.
SELECT ARTIST.NAME
FROM ARTIST, APPEARS, ALBUM
WHERE ARTIST.ID=APPEARS.ARTIST_ID
AND APPEARS.ALBUM_ID=ALBUM.ID
AND ALBUM.NAME="Andy's OG Remix"

Move the predicate to the lowest applicable point in the plan.
Move the predicate to the lowest applicable point in the plan.

SELECT ARTIST.NAME
FROM ARTIST, APPEARS, ALBUM
WHERE ARTIST.ID=APPEARS.ARTIST_ID
AND APPEARS.ALBUM_ID=ALBUM.ID
AND ALBUM.NAME="Andy's OG Remix"
SELECT ARTIST.NAME
FROM ARTIST, APPEARS, ALBUM
WHERE ARTIST.ID=APPEARS.ARTIST_ID
AND APPEARS.ALBUM_ID=ALBUM.ID
AND ALBUM.NAME="Andy's OG Remix"

Replace all Cartesian Products with inner joins using the join predicates.
SELECT ARTIST.NAME
FROM ARTIST, APPEARS, ALBUM
WHERE ARTIST.ID=APPEARS.ARTIST_ID
AND APPEARS.ALBUM_ID=ALBUM.ID
AND ALBUM.NAME="Andy's OG Remix"

Replace all Cartesian Products with inner joins using the join predicates.
SELECT ARTIST.NAME
FROM ARTIST, APPEARS, ALBUM
WHERE ARTIST.ID=APPEARS.ARTIST_ID
    AND APPEARS.ALBUM_ID=ALBUM.ID
    AND ALBUM.NAME="Andy's OG Remix"

Eliminate redundant attributes before pipeline breakers to reduce materialization cost.
SELECT ARTIST.NAME
FROM ARTIST, APPEARS, ALBUM
WHERE ARTIST.ID=APPEARS.ARTIST_ID
AND APPEARS.ALBUM_ID=ALBUM.ID
AND ALBUM.NAME="Andy's OG Remix"

Eliminate redundant attributes before pipeline breakers to reduce materialization cost.
The DBMS treats nested sub-queries in the where clause as functions that take parameters and return a single value or set of values.

Two Approaches:
→ Rewrite to de-correlate and/or flatten them
→ Decompose nested query and store result to temporary table
SELECT name FROM sailors AS S
WHERE EXISTS (  
    SELECT * FROM reserves AS R
    WHERE S.sid = R.sid
    AND R.day = '2018-10-15'
  )
SELECT name FROM sailors AS S
WHERE EXISTS (
    SELECT * FROM reserves AS R
    WHERE S.sid = R.sid
    AND R.day = '2018-10-15'
)
**NESTED SUB-QUERIES: REWRITE**

```
SELECT name FROM sailors AS S
WHERE EXISTS (  
    SELECT * FROM reserves AS R  
    WHERE S.sid = R.sid  
    AND R.day = '2018-10-15'
)

SELECT name
FROM sailors AS S, reserves AS R
WHERE S.sid = R.sid
AND R.day = '2018-10-15'
```
For each sailor with the highest rating (over all sailors) and at least two reservations for red boats, find the sailor id and the earliest date on which the sailor has a reservation for a red boat.

```
SELECT S.sid, MIN(R.day)
FROM sailors S, reserves R, boats B
WHERE S.sid = R.sid
    AND R.bid = B.bid
    AND B.color = 'red'
    AND S.rating = (SELECT MAX(S2.rating)
                    FROM sailors S2)
GROUP BY S.sid
HAVING COUNT(*) > 1
```
For each sailor with the highest rating (over all sailors) and at least two reservations for red boats, find the sailor id and the earliest date on which the sailor has a reservation for a red boat.

```
SELECT S.sid, MIN(R.day)
FROM sailors S, reserves R, boats B
WHERE S.sid = R.sid
    AND R.bid = B.bid
    AND B.color = 'red'
    AND S.rating = (SELECT MAX(S2.rating)
                     FROM sailors S2)
GROUP BY S.sid
HAVING COUNT(*) > 1
```
DECOMPOSING QUERIES

For harder queries, the optimizer breaks up queries into blocks and then concentrates on one block at a time.

Sub-queries are written to a temporary table that are discarded after the query finishes.
SELECT S.sid, MIN(R.day) 
FROM sailors S, reserves R, boats B 
WHERE S.sid = R.sid 
  AND R.bid = B.bid 
  AND B.color = 'red' 
  AND S.rating = (SELECT MAX(S2.rating) 
                  FROM sailors S2) 
GROUP BY S.sid 
HAVING COUNT(*) > 1
SELECT S.sid, MIN(R.day) 
FROM sailors S, reserves R, boats B 
WHERE S.sid = R.sid 
    AND R.bid = B.bid 
    AND B.color = 'red' 
    AND S.rating = (SELECT MAX(S2.rating) 
                    FROM sailors S2) 
GROUP BY S.sid 
HAVING COUNT(*) > 1
**DECOMPOSING QUERIES**

```
SELECT MAX(rating) FROM sailors
```

```
SELECT S.sid, MIN(R.day)
FROM sailors S, reserves R, boats B
WHERE S.sid = R.sid
AND R.bid = B.bid
AND B.color = 'red'
AND S.rating = (SELECT MAX(S2.rating)
                FROM sailors S2)
GROUP BY S.sid
HAVING COUNT(*) > 1
```
DECOMPOSING QUERIES

```
SELECT MAX(rating) FROM sailors

SELECT S.sid, MIN(R.day)
FROM sailors S, reserves R, boats B
WHERE S.sid = R.sid
   AND R.bid = B.bid
   AND B.color = 'red'
   AND S.rating = (SELECT MAX(S2.rating) FROM sailors S2)
GROUP BY S.sid
HAVING COUNT(*) > 1
```

**Nested Block**

```
DECOMPOSING QUERIES

SELECT MAX(rating) FROM sailors

SELECT S.sid, MIN(R.day)
FROM sailors S, reserves R, boats B
WHERE S.sid = R.sid
  AND R.bid = B.bid
  AND B.color = 'red'
  AND S.rating = ###
GROUP BY S.sid
HAVING COUNT(*) > 1
```sql
SELECT MAX(rating) FROM sailors

SELECT S.sid, MIN(R.day)
FROM sailors S, reserves R, boats B
WHERE S.sid = R.sid
AND R.bid = B.bid
AND B.color = 'red'
AND S.rating = ###
GROUP BY S.sid
HAVING COUNT(*) > 1
```

**Outer Block**
An optimizer transforms a query's expressions (e.g., WHERE clause predicates) into the optimal/minimal set of expressions.

Implemented using if/then/else clauses or a pattern-matching rule engine.
→ Search for expressions that match a pattern.
→ When a match is found, rewrite the expression.
→ Halt if there are no more rules that match.
Impossible / Unnecessary Predicates

```
SELECT * FROM A WHERE 1 = 0;
```
Impossible / Unnecessary Predicates

```sql
SELECT * FROM A WHERE 1 = 0;
```
Impossible / Unnecessary Predicates

SELECT * FROM A WHERE 1 = 0; 

Source: Lukas Eder
Impossible / Unnecessary Predicates

SELECT * FROM A WHERE 1 = 0;  

SELECT * FROM A WHERE 1 = 1;
Impossible / Unnecessary Predicates

\[
\text{SELECT } * \text{ FROM } A \text{ WHERE } 1 = 0; \times
\]

\[
\text{SELECT } * \text{ FROM } A \text{ WHERE } 1 = 1;
\]
Impossible / Unnecessary Predicates

```
SELECT * FROM A WHERE 1 = 0;  // X
SELECT * FROM A;
```

CREATE TABLE A (
  id INT PRIMARY KEY,
  val INT NOT NULL);

MORE EXAMPLES

Source: Lukas Eder
Impossible / Unnecessary Predicates

- SELECT * FROM A WHERE 1 = 0;
- SELECT * FROM A;

Join Elimination

- SELECT A1.*
  FROM A AS A1 JOIN A AS A2
  ON A1.id = A2.id;
Impossible / Unnecessary Predicates

- SELECT * FROM A WHERE 1 = 0;
- SELECT * FROM A;

Join Elimination

- SELECT A1.*
  FROM A AS A1 JOIN A AS A2
  ON A1.id = A2.id;
Impossible / Unnecessary Predicates

\[
\text{SELECT } * \text{ FROM } A \text{ WHERE } 1 = 0; \times
\]

\[
\text{SELECT } * \text{ FROM } A;
\]

Join Elimination

\[
\text{SELECT } * \text{ FROM } A;
\]
MORE EXAMPLES

Join Elimination with Sub-Query

```
SELECT * FROM A AS A1
WHERE EXISTS(SELECT val FROM A AS A2
WHERE A1.id = A2.id);
```

CREATE TABLE A (
    id INT PRIMARY KEY,
    val INT NOT NULL
);
CREATE TABLE A ( id INT PRIMARY KEY, val INT NOT NULL );

MORE EXAMPLES

Join Elimination with Sub-Query

```
SELECT * FROM A AS A1
WHERE EXISTS(SELECT val FROM A AS A2
             WHERE A1.id = A2.id);
```
CREATE TABLE A (
  id INT PRIMARY KEY,
  val INT NOT NULL
);

MORE EXAMPLES

Join Elimination with Sub-Query

SELECT * FROM A;
MORE EXAMPLES

Join Elimination with Sub-Query

```sql
SELECT * FROM A;
```

Merging Predicates

```sql
SELECT * FROM A
WHERE val BETWEEN 1 AND 100
OR val BETWEEN 50 AND 150;
```
CREATE TABLE A (id INT PRIMARY KEY, val INT NOT NULL);

MORE EXAMPLES

Join Elimination with Sub-Query

```
SELECT * FROM A;
```

Merging Predicates

```
SELECT * FROM A
WHERE val BETWEEN 1 AND 100
  OR val BETWEEN 50 AND 150;
```
MORE EXAMPLES

Join Elimination with Sub-Query

```sql
SELECT * FROM A;
```

Merging Predicates

```sql
SELECT * FROM A
WHERE val BETWEEN 1 AND 150;
```
Heuristics / Rules

→ Rewrite the query to remove stupid / inefficient things.
→ These techniques may need to examine catalog, but they do not need to examine data.

Cost-based Search

→ Use a model to estimate the cost of executing a plan.
→ Enumerate multiple equivalent plans for a query and pick the one with the lowest cost.
**QUERY OPTIMIZATION**

**Heuristics / Rules**
- Rewrite the query to remove stupid / inefficient things.
- These techniques may need to examine catalog, but they do not need to examine data.

**Cost-based Search**
- Use a model to estimate the cost of executing a plan.
- Enumerate multiple equivalent plans for a query and pick the one with the lowest cost.
Generate an estimate of the cost of executing a particular query plan for the current state of the database.

→ Estimates are only meaningful internally.

This is independent of the plan enumeration step that we will talk about next class.
COST MODEL COMPONENTS

Choice #1: Physical Costs
→ Predict CPU cycles, I/O, cache misses, RAM consumption, pre-fetching, etc…
→ Depends heavily on hardware.

Choice #2: Logical Costs
→ Estimate result sizes per operator.
→ Independent of the operator algorithm.
→ Need estimations for operator result sizes.

Choice #3: Algorithmic Costs
→ Complexity of the operator algorithm implementation.
DISK-BASED DBMS COST MODEL

The number of disk accesses will always dominate the execution time of a query.
→ CPU costs are negligible.
→ Must consider sequential vs. random I/O.

This is easier to model if the DBMS has full control over buffer management.
→ We will know the replacement strategy, pinning, and assume exclusive access to disk.
POSTGRES COST MODEL

Uses a combination of CPU and I/O costs that are weighted by “magic” constant factors.

Default settings are obviously for a disk-resident database without a lot of memory:
→ Processing a tuple in memory is \(400\times\) faster than reading a tuple from disk.
→ Sequential I/O is \(4\times\) faster than random I/O.
19.7.2. Planner Cost Constants

The cost variables described in this section are measured on an arbitrary scale. Only their relative values matter, hence scaling them all up or down by the same factor will result in no change in the planner’s choices. By default, these cost variables are based on the cost of sequential page fetches; that is, seq_page_cost is conventionally set to 1.0 and the other cost variables are set with reference to that. But you can use a different scale if you prefer, such as actual execution times in milliseconds on a particular machine.

Note: Unfortunately, there is no well-defined method for determining ideal values for the cost variables. They are best treated as averages over the entire mix of queries that a particular installation will receive. This means that changing them on the basis of just a few experiments is very risky.

seq_page_cost (floating point)

Sets the planner’s estimate of the cost of a disk page fetch that is part of a series of sequential fetches. The default is 1.0. This value can be overridden for tables and indexes in a particular tablespace by setting the tablespace parameter of the same name (see ALTER TABLESPACE).

random_page_cost (floating point)
19.7.2. Planner Cost Constants

The cost variables described in this section are measured on an arbitrary scale. Only their relative values matter, hence scaling them all up or down by the same factor will result in no change in the planner's choices. By default, these cost variables are based on the cost of sequential page fetches; that is, seq_page_cost is conventionally set to 1.0 and the other cost variables are set with reference to that. But you can use a different scale if you prefer, such as actual execution times in milliseconds on a particular machine.

**Note:** Unfortunately, there is no well-defined method for determining ideal values for the cost variables. They are best treated as averages over the entire mix of queries that a particular installation will receive. This means that changing them on the basis of just a few experiments is very risky.

seq_page_cost (floating point)

Sets the planner's estimate of the cost of a disk page fetch that is part of a series of sequential fetches. The default is 1.0. This value can be overridden for tables and indexes in a particular tablespace by setting the tablespace parameter of the same name (see ALTER TABLESPACE).

random_page_cost (floating point)
IBM DB2 Cost Model

Database characteristics in system catalogs
Hardware environment (microbenchmarks)
Storage device characteristics (microbenchmarks)
Communications bandwidth (distributed only)
Memory resources (buffer pools, sort heaps)

Concurrent Environment
→ Average number of users
→ Isolation level / blocking
→ Number of available locks

Source: Guy Lohman
CONCLUSION

We can use static rules and heuristics to optimize a query plan without needing to understand the contents of the database.

We use cost model to help perform more advanced query optimizations.
NEXT CLASS

Statistics and plan enumeration