Query Optimization Exercise Session 12

Andrey Gubichev

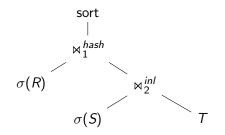
January 19, 2015

SQL Subqueries

Background: execution model

Iterator model:

- open()
- next()
- close()



Background: aggregation

select max(age)
from students

Background: aggregation

select max(age)
from students

F_{max(age)} ∣ students

Example (1)

How to execute it?

Example (2)

select *
from R
where R.a in (select S.b
 from S
 where S.c = R.d)

How to execute it?

Example (2)

```
select *
from R
where R.a in (select S.b
      from S
      where S.c = R.d)
```

How to execute it? *Correlated subquery*. Subquery unnesting – getting rid of subqueries.

Type A subqueries

Evaluate inner block only once – build side of a join. Outer query becomes the probe side of a join

Type N subqueries

Same technique. Subquery becomes a build side of a semi-join

Type JA subqueries

Type JA subqueries

Group by s2.dept_id, compute max(s2.age). Then join wiht the s1

Type JA subqueries: COUNT bug

Type JA subqueries: COUNT bug

 $COUNT(\emptyset) = 0$ Use outer-join for unnesting COUNT queries

Dealing with quantifiers

- ▶ ... < ANY (select ...) \Rightarrow ...< (select max() ...)
- ... < ALL (select ...)</p>

Dealing with quantifiers

- ▶ ... < ANY (select ...) \Rightarrow ...< (select max() ...)
- ▶ < ALL (select ...) \Rightarrow ...< (select min() ...)
- ► for >: flip the rules
- EXISTS (correlated): semi-join

TPC-H

TPC-H - "22 most well studied SQL queries in history"

part (p_)	partsupp (ps_)	lineitem (l_)	orders (o_)
SF*200k	SF*800k	SF*6000k	SF*1500k
partkey	partkey 🚽	orderkey	orderkey
integer	integer	integer	integer
name	suppkey 🚽	partkey	custkey
char 55	integer	- integer	integer
mfgr	availqty	suppkey	orderstatus
char 25	integer	integer	char 1
brand	supplycost	linenumber	totalprice
char 10	decimal	integer	decimal
type	comment	quantity	orderdate
varchar 25	varchar 199	decimal	date
size		extendedprice	orderpriority
integer	customer (c)	decimal	char 15
container	SF*150k	discount	clerk
char 10	custkey	decimal	char 15
retailprice	integer	tax	shippriority
decimal	name	decimal	integer
comment	char 25	returnflag	comment
varchar 23	address	char 1	varchar 79
	varchar 40	linestatus	lar and the
supplier (s)		char 1	
	nationkey	shipdate	
SF*10k	integer	date	
suppkey 🚽	phone	commitdate	
integer	char 15	date	
name	acctbal	receiptdate	
char 25	decimal		
address	mktsegment	date	
varchar 40	char 10	shipinstruct	
nationkey	comment	char 25	
integer	varchar 117	shipmode	
phone		char 10	
char 15	nation (n_)	comment	
acctbal	25	varchar 44	
decimal	nationkey]
comment	integer	region (r)	
varchar 101	name	5	
	char 25	regionkey	
	regionkey	integer	
	integer	name	

12/16

TPC-H Query 4

```
select
        o_orderpriority,
        count(*) as order_count
from
        orders
where
        o orderdate >= date '1993-07-01'
        and o orderdate < date '1993-10-01'
        and exists (
                select *
                from lineitem
                where l_orderkey = o_orderkey
                        and l_commitdate < l_receiptdate
        )
group by o_orderpriority
order by o_orderpriority
```

TPC-H Query 16

```
select p_brand, p_type, p_size,
        count(distinct ps_suppkey) as supplier_cnt
from partsupp, part
where p_partkey = ps_partkey
        and p_brand <> 'Brand#45'
        and p_type not like 'MEDIUM POLISHED%'
        and p_size in (49, 14, 23, 45, 19, 3, 36, 9)
        and ps_suppkey not in (
                select s_suppkey
                from supplier
                where s_comment like '%Customer%Complaints%
group by p_brand, p_type, p_size
order by supplier_cnt desc, p_brand, p_type, p_size
```

TPC-H Query 17

select sum(l_extendedprice) / 7.0 as avg_yearly
from

```
lineitem,
        part
where
        p_partkey = l_partkey
        and p_brand = 'Brand#23'
        and p_container = 'MED BOX'
        and l_quantity < (
                 select
                         0.2 * avg(l_quantity)
                from
                         lineitem
                where
                         l_partkey = p_partkey
        )
```

Info

Exercises due January 26, 2015.