

04 SQL优化

The Nature of SQL

南京大学软件学院



关系代数

- E.F.Codd 关系理论之父，关系代数究竟有什么用？

什么是代数？-表达式的等价变换

$$(2+4) \times 3 = 6 \times 3 = 18$$

$$= (2+2+2) \times 3$$

$$= 2 \times 3 + 4 \times 3$$

$$= 2 \times 3 + (2+2) \times 3$$

$$= \dots$$

关系代数也是一样

2、3、4 这些数字对应的就是 关系（表）

+ - × / 这些运算符对应的就是 关系操作

SQL语句

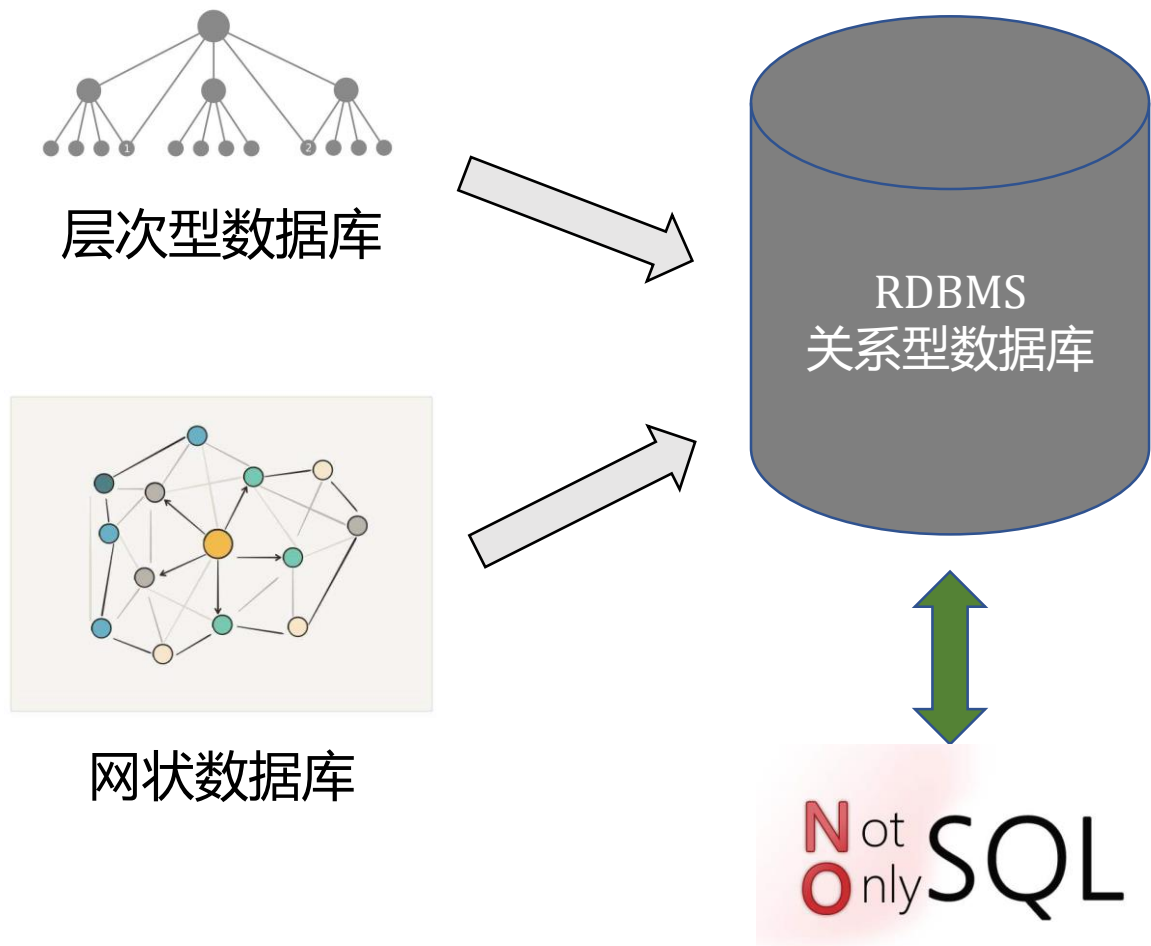
关系表达式

关系表达式等价变换

选择最优路径执行



关系代数使数据库变成了科学而不是艺术



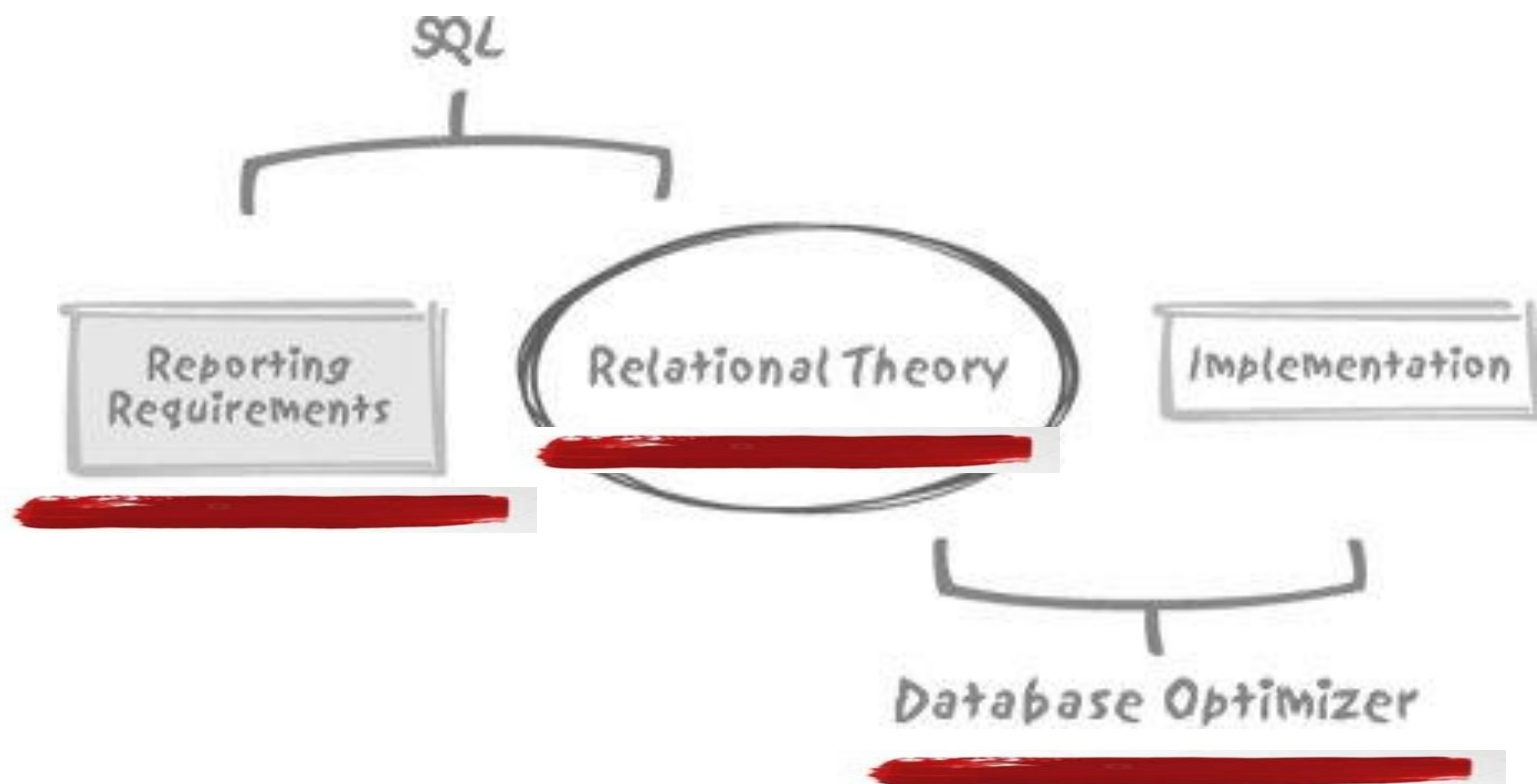
SQL操作：不需要去考虑操作实现细节

关系设计：拥有理论上的普遍规则

关系数据库变成了科学，也变成了黑盒

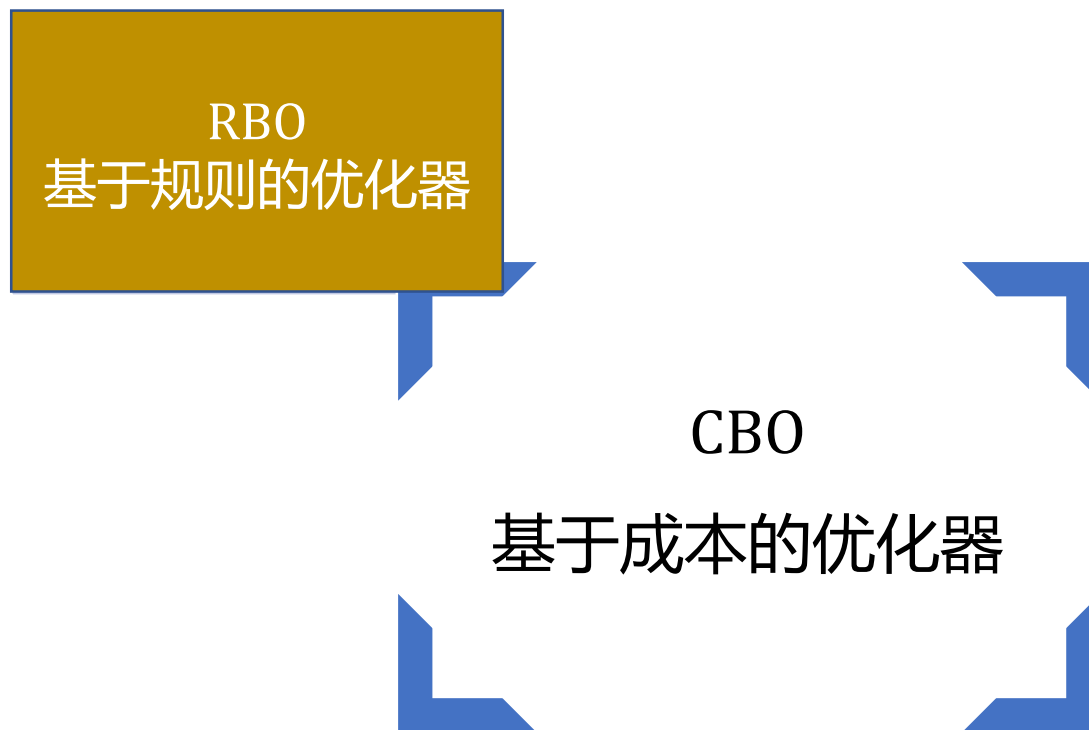


SQL与查询优化器

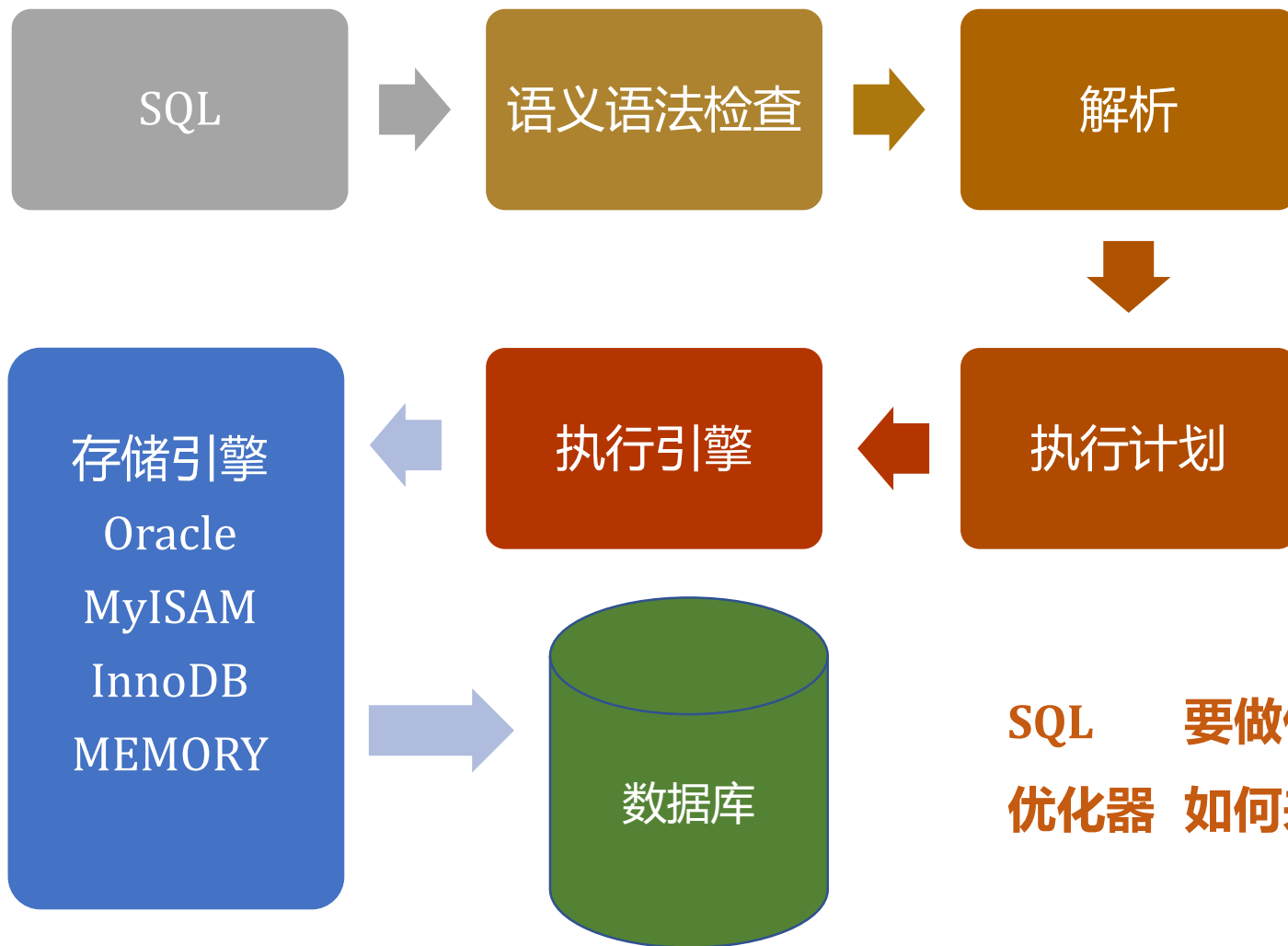


SQL和查询优化器

- 优化器借助关系理论提供的语义无误的原始查询进行**有效的等价变换**
- 优化器根据数据库的实际实现情况对理论上等价的不同优化方案做出**权衡**,
- 产生可能的最优查询执行方案



SQL的执行顺序



SQL 要做什么? (声明性语言)
优化器 如何做!



优化器只能对关系领域进行优化

- 忽略这点很可能出现错误

例子：查询不是经理的员工当中，哪五个人收入最高？

```
select empname, salary
  from employees
 where status != 'EXECUTIVE'
    and rownum <= 5
 order by salary desc
```



```
select *
  from (select empname, salary
        from employees
       where status != 'EXECUTIVE'
      order by salary desc)
 where rownum <= 5
```



优化器的有效范围

- 优化器需要借助数据库中找到的信息
- 能够进行数学意义上的等价变换
- 优化器考虑整体响应时间
- 优化器改善的是独立的查询



思考题

- Oracle的rownum是一个非常讨厌的SQL方言，但它是Oracle数据库中唯一的限定返回行数的函数，其它数据库也有类似的方言
 - DB2使用FETCH FIRST子句
 - MySQL和PostgreSQL使用LIMIT子句
 - SQL Server使用TOP关键字
- 请你用你手上常用的数据库试一下本课程那个限定返回行数查询的例子，看看有没有Oracle出现的问题
- 如果你是用Oracle，你试一下，你能通过rownum=5，来返回第5行记录嘛？



End

下一讲，我们将开始进入SQL优化的部分

