

# SQL优化的其它问题

南京大学软件学院



# Where子句的比较运算符

①  $a = 1000$

②  $a = 500 + 500$

CBO只能将 ② 自动转化成 ①，无法改变 ③，这是f().....

---

③  $a - 500 = 500$

在全表扫描的时候，② ① 是一个比较操作，③ 是一个减法操作



# Where子句的比较运算符

- 请看下面这个例子，分析3条SQL语句的差别

```
create table t as select × from dba_objects;  
insert into t select *from t;
```

创建一个40万+的表

统计3条SQL语句的差别，都是统计2019年4月20日这一天每个用户下的对象的个数



1



set time on

```
select count(*) from t
```

```
where cteated>=to_date('2019-04-20 00:00:00','yyyy-mm-dd hh24:mi:ss')
```

```
and created<to_date('2019-04-21 00:00:00','yyyy-mm-dd hh24:mi:ss')
```

```
group by owner;
```

---

2



```
select count(*) from t
```

```
where to_char(created,'yyyy-mm-dd')='2019-04-20'
```

```
group by owner;
```

---

3



```
select count(*) from t
```

```
where to_char(created,'yyyy-mm-dd hh24:mi:ss')>='2019-04-20 00:00:00'
```

```
and to_char(created,,'yyyy-mm-dd hh24:mi:ss')<= '2019-04-20 23:59:59'
```

```
group by owner
```



# 比较运算符的转化

sales_qty>any (:in_qty1,:in_qty2)	sales_qty>in_qty1 or sales_qty>:in_qty2
where 10000>any (select sal from emp where job='clerk')	where exists ( select sal from emp where job='clerk' and 10000>sal)
sales_qty>all (:in_qty1,:in_qty2)	sales_qty>in_qty1 and sales_qty>:in_qty2
sales_qty between 100and 200	sales_qty >= 100 and sales_qty <= 200
not (sal<30000 or comm is null) not sal<30000 and comm is not null	sal >= 30000 and comm is not null
not deptno = (select deptno from emp where empno =9527)	deptno<> (select deptno from emp where empno = 9527)



# 大数据量查询

- 越快剔除不需要的数据，查询的后续阶段必须处理的数据量就越少，查询效率就越高

```
Select ...  
From A, B, C, D, E1  
Where (condition on E1)  
and (join and other conditions)  
Union  
Select ...  
From A,B,C,D,E2  
Where (condition on E2)  
and (join and other conditions)
```



```
Select...  
From A,B,C,D,  
  (select ...  
   from E1  
   where (condition on E1)  
   union  
   select...  
   from E2  
   where (condition on E2)  
  )E  
Where (joins and other conditions)
```



# 将子查询转换为JOIN

- 不包含聚合函数，不出现多种条件选择可以不需要子查询

Jobs (employee, title) Ranks (title, rank) Salary (rank, payment)

```
Select payment from salary where rank=  
  (select rank from ranks where title=  
    (select title from jobs where employee = '...'))
```



```
Select payment from salary, ranks, jobs  
Where salary.rank = ranks.rank  
  And ranks.title = jobs.title  
  And jobs.employee = '...'
```



# 查询不存在的内容

- 在salary表中查询是否存在某个等级当前没有分配职位，显示等级和薪水

Jobs (employee, title) Ranks (title, rank) Salary (rank, payment)

```
Select salary.ranks salary.payment from salary  
Where rank NOT IN (select rank from ranks)
```



```
Select salary.ranks salary.payment  
From salary  
LEFT OUTER JOIN ON(salary.rank = ranks.rank)  
Where ranks.rank IS NULL
```





# 将聚合子查询转换为JOIN或内嵌视图

- 在订单完成前有不同状态，记录在orderstatus (ordid,status,statusdate) 中
- 需求是：列出所有尚未标记为完成状态的订单的下列字段：订单号，客户名，订单的最后状态，以及设置状态的时间。



# 再回头看订单和客户的例子

```
select c.custname, o.ordid, os.status, os.statusdate
from customers c,
     orders o,
     orderstatus os
```

```
where o.ordid = os.ordid
```

```
and not exists (select null
                from orderstatus os2
                where os2.status = 'COMPLETE'
                and os2.ordid = o.ordid)
```

```
and os.statusdate = (select max(statusdate)
                    from orderstatus os3
                    where os3.ordid = o.ordid)
```

```
and o.custid = c.custid
```

需求是：列出所有尚未标记为完成状态的订单的下列字段：订单号，客户名，订单的最后状态，以及设置状态的时间

```
and (o.ordid, os.statusdate)
    = (select ordid, max(statusdate)
      from orderstatus
      group by ordid)
```



# 非关联子查询变成内嵌视图

```
select c.custname, o.ordid, os.status, os.statusdate
  from customers c,
       orders o,
       orderstatus os,
       (select ordid, max(statusdate) laststatusdate
        from orderstatus
        group by ordid) x
 where o.ordid = os.ordid
        and not exists (select null
                        from orderstatus os2
                        where os2.status = 'COMPLETE'
                        and os2.ordid = o.ordid)
        and os.statusdate = x.laststatusdate
        and os.ordid = x.ordid
        and o.custid = c.custid
```

需求是：列出所有尚未标记为完成状态的订单的下列字段：订单号，客户名，订单的最后状态，以及设置状态的时间



# 非关联子查询变成内嵌视图

```
select c.custname, o.ordid, os.status, os.statusdate
from customers c,
     orders o,
     orderstatus os,
     (select ordid, max(statusdate) laststatusdate
      from orderstatus
      group by ordid) x
where o.ordid = os.ordid
     and os.statusdate = x.laststatusdate
     and os.ordid = x.ordid
     and os.status != 'COMPLETE'
     and o.custid = c.custid
```



# 思考题

- Orders (custid, ordered, totalitems)
- 需要显示每一个客户购物件数最多的日期，如何用连接改写这个SQL的子查询

```
Select custid, ordered, totalitems  
From orders o1  
Where o1.ordered = (  
    select max(ordered)  
    from orders o2  
    where o1.custid = o2.custid )
```



# End

下一部分，进入SQL的一些主题

