下面这个SQL要跑5个多小时

SELECT

B.AREA\_ID,

A.PARTY\_ID,

B.AREA\_NAME,

C.NAME CHANNEL\_NAME,

B.NAME PARTY\_NAME,

B.ACCESS\_NUMBER,

B.PROD\_SPEC,

B.START\_DT,

A.BO\_ACTION\_NAME,

A.SO\_STAFF\_ID,

A.ATOM\_ACTION\_ID,

A.PROD\_ID

FROM DW\_CHANNEL C,

DW\_CRM\_DAY\_USER B,

DW\_BO\_ORDER A

WHERE A.PROD\_ID = B.PROD\_ID AND

A.CHANNEL\_ID = C.CHANNEL\_ID AND

A.SO\_STAFF\_ID LIKE '36%' AND

A.BO\_ACTION\_NAME IN ('新装','移机','资费变更') AND

B.PROD\_SPEC IN ('普通电话', 'ADSL','LAN', '手机',

'E8 - 2S','E6移动版', 'E9版1M(老版)',

'普通E9','普通新版E8',

'全省\_紧密融合型E9套餐产品规格',

'(新) 全省\_紧密融合型E9套餐产品规格',

'新春欢乐送之E8套餐',

'新春欢乐送之E6套餐') AND

NOT EXISTS (SELECT \*

FROM DW\_BO\_ORDER D

WHERE D.STAFF\_ID LIKE '36%' AND

A.PARTY\_ID = D.PARTY\_ID AND

A.BO\_ID != D.BO\_ID AND

A.PROD\_ID != D.PROD\_ID AND

A.BO\_ACTION\_NAME IN

('新装', '移机','资费变更') AND

A.COMPLETE\_DT - INTERVAL '7' DAY < D.COMPLETE\_DT);

执行计划和相关信息如下：

Plan hash value: 2142862569

----------------------------------------------------------------------------------------------------------

| Id | Operation | Name | Rows | Bytes | Cost (%CPU)| Time | Pstart| Pstop |

----------------------------------------------------------------------------------------------------------

| 0 | SELECT STATEMENT | | 905 | 121K| 4152K (2)| 13:50:32 | | |

|\* 1 | FILTER | | | | | | | |

|\* 2 | HASH JOIN | | 905 | 121K| 12616 (2)| 00:02:32 | | |

|\* 3 | HASH JOIN | | 905 | 99550 | 12448 (2)| 00:02:30 | | |

| 4 | PARTITION RANGE ALL| | 1979 | 108K| 9168 (2)| 00:01:51 | 1 | 5 |

|\* 5 | TABLE ACCESS FULL | DW\_BO\_ORDER | 1979 | 108K| 9168 (2)| 00:01:51 | 1 | 5 |

|\* 6 | TABLE ACCESS FULL | DW\_CRM\_DAY\_USER | 309K| 15M| 3277 (2)| 00:00:40 | | |

| 7 | TABLE ACCESS FULL | DW\_CHANNEL | 48425 | 1276K| 168 (1)| 00:00:03 | | |

|\* 8 | FILTER | | | | | | | |

| 9 | PARTITION RANGE ALL | | 1 | 29 | 9147 (2)| 00:01:50 | 1 | 5 |

|\* 10 | TABLE ACCESS FULL | DW\_BO\_ORDER | 1 | 29 | 9147 (2)| 00:01:50 | 1 | 5 |

----------------------------------------------------------------------------------------------------------

Predicate Information (identified by operation id):

---------------------------------------------------

1 - filter( NOT EXISTS (SELECT /\*+ \*/ 0 FROM "DW\_BO\_ORDER" "D" WHERE (:B1='新装' OR :B2='移机' OR

:B3='资费变更') AND "D"."PARTY\_ID"=:B4 AND TO\_CHAR("D"."STAFF\_ID") LIKE '36%' AND

"D"."COMPLETE\_DT">:B5-INTERVAL'+07 00:00:00' DAY(2) TO SECOND(0) AND "D"."PROD\_ID"<>:B6 AND

"D"."BO\_ID"<>:B7))

2 - access("A"."CHANNEL\_ID"="C"."CHANNEL\_ID")

3 - access("A"."PROD\_ID"="B"."PROD\_ID")

5 - filter("A"."PROD\_ID" IS NOT NULL AND ("A"."BO\_ACTION\_NAME"='新装' OR

"A"."BO\_ACTION\_NAME"='移机' OR "A"."BO\_ACTION\_NAME"='资费变更') AND TO\_CHAR("A"."SO\_STAFF\_ID") LIKE

'36%')

6 - filter("B"."PROD\_SPEC"='(新) 全省\_紧密融合型E9套餐产品规格' OR "B"."PROD\_SPEC"='ADSL' OR

"B"."PROD\_SPEC"='E6移动版' OR "B"."PROD\_SPEC"='E8 - 2S' OR "B"."PROD\_SPEC"='E9版1M(老版)' OR

"B"."PROD\_SPEC"='LAN' OR "B"."PROD\_SPEC"='普通E9' OR "B"."PROD\_SPEC"='普通电话' OR

"B"."PROD\_SPEC"='普通新版E8' OR "B"."PROD\_SPEC"='全省\_紧密融合型E9套餐产品规格' OR "B"."PROD\_SPEC"='手机' OR

"B"."PROD\_SPEC"='新春欢乐送之E6套餐' OR "B"."PROD\_SPEC"='新春欢乐送之E8套餐')

8 - filter(:B1='新装' OR :B2='移机' OR :B3='资费变更')

10 - filter("D"."PARTY\_ID"=:B1 AND TO\_CHAR("D"."STAFF\_ID") LIKE '36%' AND

"D"."COMPLETE\_DT">:B2-INTERVAL'+07 00:00:00' DAY(2) TO SECOND(0) AND "D"."PROD\_ID"<>:B3 AND

"D"."BO\_ID"<>:B4)

下面是执行计划以及表信息

SQL> select count(\*) from dw\_bo\_order; ----200W数据

COUNT(\*)

----------

2282548

SQL> select count(\*) from dw\_crm\_day\_user; ----40W数据

COUNT(\*)

----------

420918

SQL> select count(\*) from dw\_channel; ---4W数据

COUNT(\*)

----------

48031

谁来说说这个SQL慢在何处？

请注意看执行计划里面的 ID=1 这里，它是不是FILTER？

那么大家回忆一下，FILTER 类似于什么操作？ 是不是嵌套循环操作？

那么大家再请看一下FILTER要过滤的表是不是DW\_BO\_ORDER，它是不是最大的表？

我们假设200W条记录的这个表大概有2GB这么大。

我们也假设ID=2这步，就真的返回905条记录，那么大家算一下，

FILTER是不是要扫描DW\_BO\_ORDER表905次(接近905次)，

那么一共是扫描905\*2GB=1.76TB ，同志们，你们说扫描1.76TB的数据快吗？

那么这就是这个SQL跑得慢的根本原因。

找到了原因之后如何进行优化呢？谁来优化一下？

光找到原因没用啊，要能解决问题才有用。

根据以前讲课知识，我们是不是希望让有 IN /EXISTS 的子查询走JOIN啊？

但是我们这里的查询是

EXISTS (SELECT \*

FROM DW\_BO\_ORDER D

WHERE D.STAFF\_ID LIKE '36%' AND

A.PARTY\_ID = D.PARTY\_ID AND

A.BO\_ID != D.BO\_ID AND

A.PROD\_ID != D.PROD\_ID AND

A.BO\_ACTION\_NAME IN

('新装', '移机','资费变更') AND

A.COMPLETE\_DT - INTERVAL '7' DAY < D.COMPLETE\_DT)

坑爹了，这里有<> ，没法走HASH JOIN啊。

走NESTED LOOPS 肯定不对，NESTED LOOPS 类似FILTER了

走SORT MERGE 也不对啊，要去排序，200多万的数据排序，坑爹。

那咋办呢？

我的解决方法就是改写SQL:

with D as (select /\*+ materialize \*/ PARTY\_ID,BO\_ID,PROD\_ID from DW\_BO\_ORDER where STAFF\_ID LIKE '36%')

SELECT

B.AREA\_ID,

A.PARTY\_ID,

B.AREA\_NAME,

C.NAME CHANNEL\_NAME,

B.NAME PARTY\_NAME,

B.ACCESS\_NUMBER,

B.PROD\_SPEC,

B.START\_DT,

A.BO\_ACTION\_NAME,

A.SO\_STAFF\_ID,

A.ATOM\_ACTION\_ID,

A.PROD\_ID

FROM DW\_CHANNEL C,

DW\_CRM\_DAY\_USER B,

DW\_BO\_ORDER A

WHERE A.PROD\_ID = B.PROD\_ID AND

A.CHANNEL\_ID = C.CHANNEL\_ID AND

A.SO\_STAFF\_ID LIKE '36%' AND

A.BO\_ACTION\_NAME IN ('新装','移机','资费变更') AND

B.PROD\_SPEC IN ('普通电话', 'ADSL','LAN', '手机',

'E8 - 2S','E6移动版', 'E9版1M(老版)',

'普通E9','普通新版E8',

'全省\_紧密融合型E9套餐产品规格',

'(新) 全省\_紧密融合型E9套餐产品规格',

'新春欢乐送之E8套餐',

'新春欢乐送之E6套餐') AND

NOT EXISTS (SELECT \*

FROM D

WHERE D.STAFF\_ID LIKE '36%' AND

A.PARTY\_ID = D.PARTY\_ID AND

A.BO\_ID != D.BO\_ID AND

A.PROD\_ID != D.PROD\_ID AND

A.BO\_ACTION\_NAME IN

('新装', '移机','资费变更') AND

A.COMPLETE\_DT - INTERVAL '7' DAY < D.COMPLETE\_DT);

SQL> set timi on

SQL> WITH D AS

2 (SELECT /\*+ materialize \*/

3 PARTY\_ID,

4 BO\_ID,

5 PROD\_ID,

6 COMPLETE\_DT

7 FROM DW\_BO\_ORDER

8 WHERE STAFF\_ID LIKE '36%' AND

9 BO\_ACTION\_NAME IN ('新装',

10 '移机',

11 '资费变更'))

12 SELECT

13 B.AREA\_ID,

14 A.PARTY\_ID,

15 B.AREA\_NAME,

16 C.NAME CHANNEL\_NAME,

17 B.NAME PARTY\_NAME,

18 B.ACCESS\_NUMBER,

19 B.PROD\_SPEC,

20 B.START\_DT,

21 A.BO\_ACTION\_NAME,

22 A.SO\_STAFF\_ID,

23 A.ATOM\_ACTION\_ID,

24 A.PROD\_ID

25 FROM DW\_CHANNEL C,

26 DW\_CRM\_DAY\_USER B,

27 DW\_BO\_ORDER A

28 WHERE A.PROD\_ID = B.PROD\_ID AND

29 A.CHANNEL\_ID = C.CHANNEL\_ID AND

30 A.SO\_STAFF\_ID LIKE '36%' AND

31 A.BO\_ACTION\_NAME IN ('新装','移机','资费变更') AND

32 B.PROD\_SPEC IN ('普通电话', 'ADSL','LAN', '手机',

33 'E8 - 2S','E6移动版', 'E9版1M(老版)',

34 '普通E9','普通新版E8',

35 '全省\_紧密融合型E9套餐产品规格',

36 '(新) 全省\_紧密融合型E9套餐产品规格',

37 '新春欢乐送之E8套餐',

38 '新春欢乐送之E6套餐') AND

39 NOT EXISTS (SELECT \*

40 FROM D

41 WHERE A.PARTY\_ID = D.PARTY\_ID AND

42 A.BO\_ID != D.BO\_ID AND

43 A.PROD\_ID != D.PROD\_ID AND

44 A.COMPLETE\_DT - INTERVAL '7' DAY < D.COMPLETE\_DT);

已选择49245行。

已用时间: 00: 00: 12.37

执行计划

--------------------------------------------------------------------------------------------------------------------------

Plan hash value: 2591883460

--------------------------------------------------------------------------------------------------------------------------

| Id | Operation | Name | Rows | Bytes | Cost (%CPU)| Time | Pstart| Pstop |

--------------------------------------------------------------------------------------------------------------------------

| 0 | SELECT STATEMENT | | 905 | 121K| 62428 (2)| 00:12:30 | | |

| 1 | TEMP TABLE TRANSFORMATION | | | | | | | |

| 2 | LOAD AS SELECT | DW\_BO\_ORDER | | | | | | |

| 3 | PARTITION RANGE ALL | | 114K| 3228K| 9127 (2)| 00:01:50 | 1 | 5 |

|\* 4 | TABLE ACCESS FULL | DW\_BO\_ORDER | 114K| 3228K| 9127 (2)| 00:01:50 | 1 | 5 |

|\* 5 | FILTER | | | | | | | |

|\* 6 | HASH JOIN | | 905 | 121K| 12616 (2)| 00:02:32 | | |

|\* 7 | HASH JOIN | | 905 | 99550 | 12448 (2)| 00:02:30 | | |

| 8 | PARTITION RANGE ALL | | 1979 | 108K| 9168 (2)| 00:01:51 | 1 | 5 |

|\* 9 | TABLE ACCESS FULL | DW\_BO\_ORDER | 1979 | 108K| 9168 (2)| 00:01:51 | 1 | 5 |

|\* 10 | TABLE ACCESS FULL | DW\_CRM\_DAY\_USER | 309K| 15M| 3277 (2)| 00:00:40 | | |

| 11 | TABLE ACCESS FULL | DW\_CHANNEL | 48425 | 1276K| 168 (1)| 00:00:03 | | |

|\* 12 | FILTER | | | | | | | |

|\* 13 | VIEW | | 114K| 6791K| 90 (3)| 00:00:02 | | |

| 14 | TABLE ACCESS FULL | SYS\_TEMP\_0FD9D662E\_D625B872 | 114K| 3228K| 90 (3)| 00:00:02 | | |

--------------------------------------------------------------------------------------------------------------------------

Predicate Information (identified by operation id):

---------------------------------------------------

4 - filter(TO\_CHAR("STAFF\_ID") LIKE '36%')

5 - filter( NOT EXISTS (SELECT /\*+ \*/ 0 FROM (SELECT /\*+ CACHE\_TEMP\_TABLE ("T1") \*/ "C0" "STAFF\_ID","C1"

"PARTY\_ID","C2" "BO\_ID","C3" "PROD\_ID","C4" "COMPLETE\_DT" FROM "SYS"."SYS\_TEMP\_0FD9D662E\_D625B872" "T1") "D"

WHERE (:B1='新装' OR :B2='移机' OR :B3='资费变更') AND TO\_CHAR("D"."STAFF\_ID") LIKE '36%' AND "D"."PARTY\_ID"=:B4 AND

"D"."BO\_ID"<>:B5 AND "D"."PROD\_ID"<>:B6 AND "D"."COMPLETE\_DT">:B7-INTERVAL'+07 00:00:00' DAY(2) TO SECOND(0)))

6 - access("A"."CHANNEL\_ID"="C"."CHANNEL\_ID")

7 - access("A"."PROD\_ID"="B"."PROD\_ID")

9 - filter("A"."PROD\_ID" IS NOT NULL AND ("A"."BO\_ACTION\_NAME"='新装' OR "A"."BO\_ACTION\_NAME"='移机' OR

"A"."BO\_ACTION\_NAME"='资费变更') AND TO\_CHAR("A"."SO\_STAFF\_ID") LIKE '36%')

10 - filter("B"."PROD\_SPEC"='(新) 全省\_紧密融合型E9套餐产品规格' OR "B"."PROD\_SPEC"='ADSL' OR "B"."PROD\_SPEC"='E6移动版' OR

"B"."PROD\_SPEC"='E8 - 2S' OR "B"."PROD\_SPEC"='E9版1M(老版)' OR "B"."PROD\_SPEC"='LAN' OR "B"."PROD\_SPEC"='普通E9' OR

"B"."PROD\_SPEC"='普通电话' OR "B"."PROD\_SPEC"='普通新版E8' OR "B"."PROD\_SPEC"='全省\_紧密融合型E9套餐产品规格' OR "B"."PROD\_SPEC"='手机'

OR "B"."PROD\_SPEC"='新春欢乐送之E6套餐' OR "B"."PROD\_SPEC"='新春欢乐送之E8套餐')

12 - filter(:B1='新装' OR :B2='移机' OR :B3='资费变更')

13 - filter(TO\_CHAR("D"."STAFF\_ID") LIKE '36%' AND "D"."PARTY\_ID"=:B1 AND "D"."BO\_ID"<>:B2 AND

"D"."PROD\_ID"<>:B3 AND "D"."COMPLETE\_DT">:B4-INTERVAL'+07 00:00:00' DAY(2) TO SECOND(0))

统计信息

----------------------------------------------------------

2 recursive calls

29 db block gets

110506 consistent gets

22 physical reads

656 redo size

2438096 bytes sent via SQL\*Net to client

449 bytes received via SQL\*Net from client

11 SQL\*Net roundtrips to/from client

0 sorts (memory)

0 sorts (disk)

49245 rows processed

谁来回答一下，我为什么要这么改写SQL?

首先，with as 里面只查询了3个字段，原始表呢我没检查过，不过它肯定不止3个字段是不是？那么3个字段，加上我们的where过滤条件，过滤后我们假设它有10MB的数据，扫描905次，也就是扫描10MB\*905=4.4GB的数据，扫描1个多T 的数据与扫描4.4GB相比，你们自己说哪个快？根本不是一个数量级的东西。

大家说这个案例经典吗？

总结一下，通过这个案例你学到了什么？

是不是要深刻理解FILTER的含义？