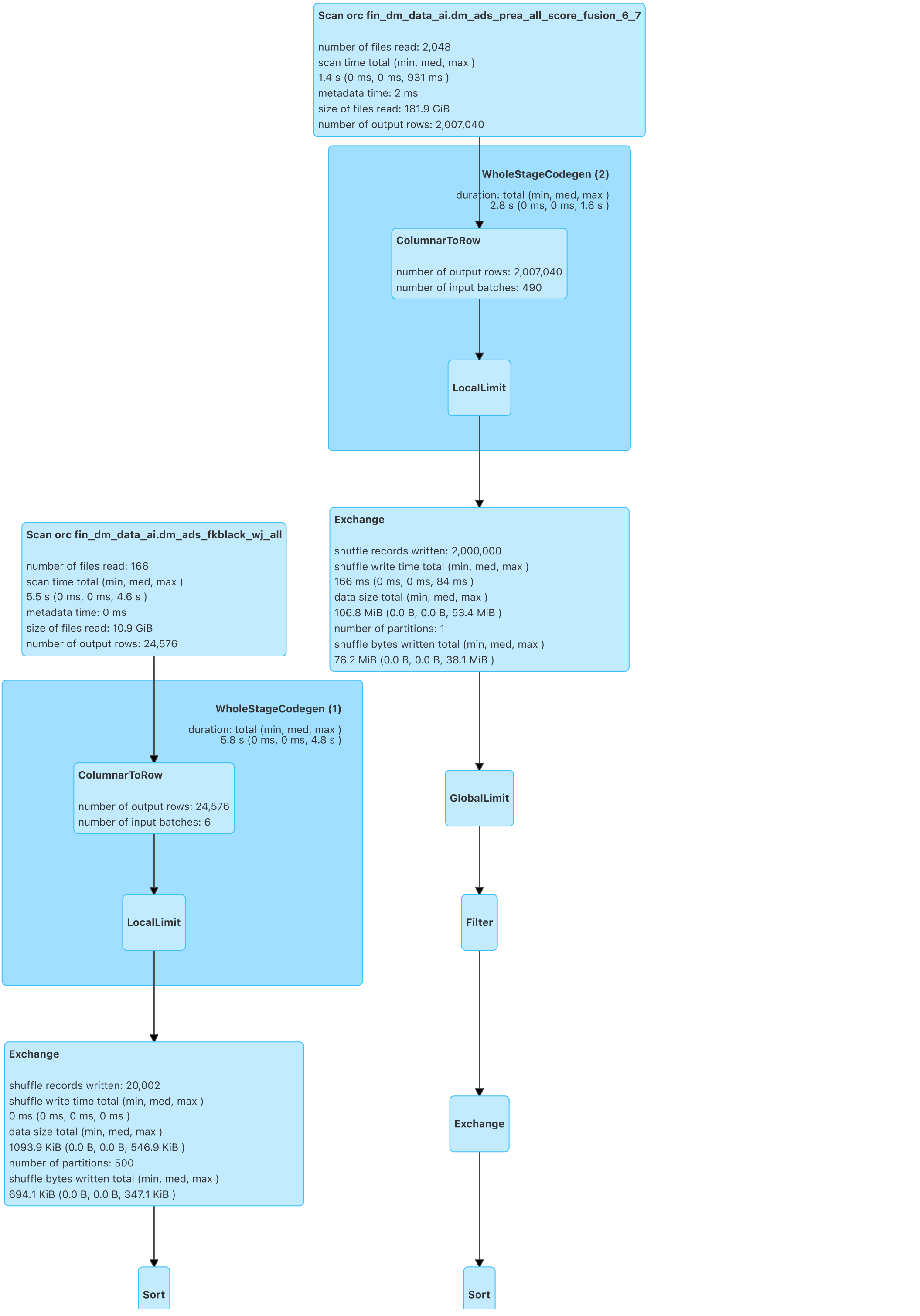
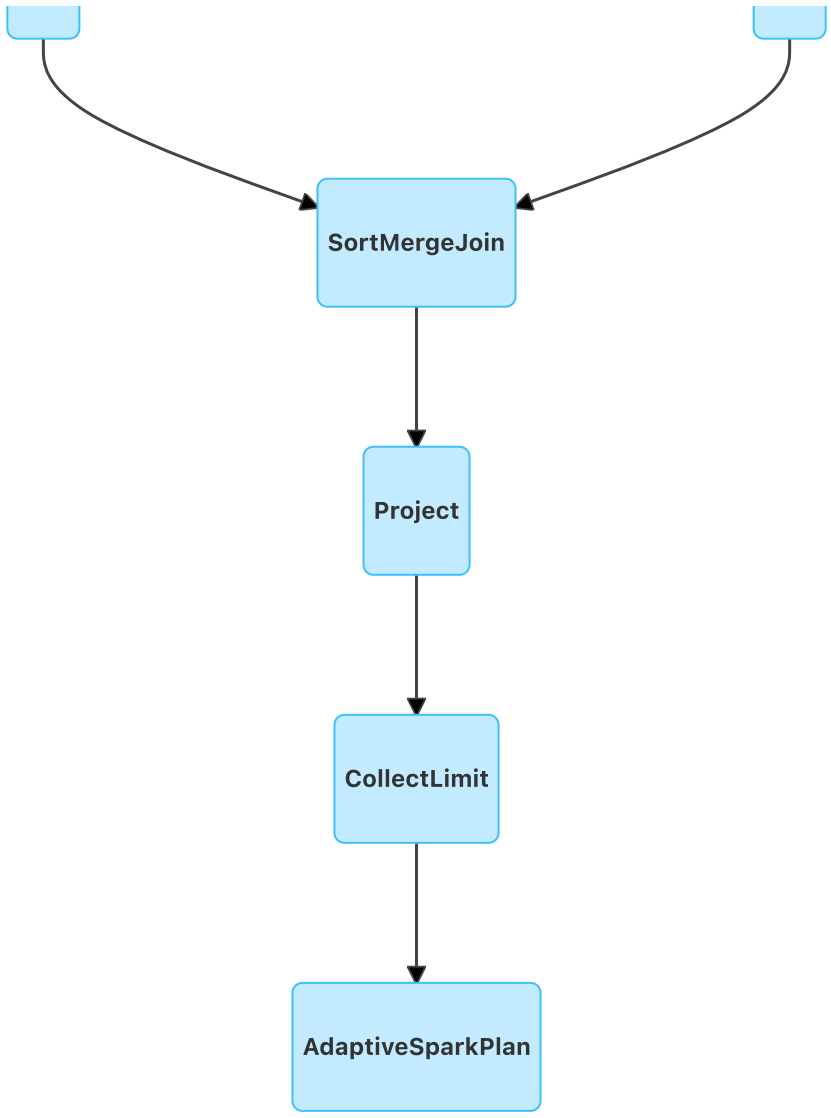
Details for Query 9

Submitted Time: 2023/11/07 22:06:36
Duration: 10 s
Running Jobs: 10 11

☐ Show the Stage ID and Task ID that corresponds to the max metric


Scan orc fin_dm_data_ai.dm_ads_prea_all_score_fusion_6_7
number of files read: 2,048
scan time total (min, med, max)
1.4 s (0 ms, 0 ms, 931 ms)
metadata time: 2 ms
size of files read: 181.9 GiB
number of output rows: 2,007,040
WholeStageCodegen (2)
duration: total (min, med, max)
2.8 s (0 ms, 0 ms, 1.6 s)
ColumnarToRow
number of output rows: 2,007,040
number of input batches: 490
LocalLimit
Exchange
shuffle records written: 2,000,000
shuffle write time total (min, med, max)
166 ms (0 ms, 0 ms, 84 ms)
data size total (min, med, max)
106.8 MiB (0.0 B, 0.0 B, 53.4 MiB)
number of partitions: 1
shuffle bytes written total (min, med, max)
76.2 MiB (0.0 B, 0.0 B, 38.1 MiB)
Scan orc fin_dm_data_ai.dm_ads_fkblack_wj_all
number of files read: 166
scan time total (min, med, max)
5.5 s (0 ms, 0 ms, 4.6 s)
metadata time: 0 ms
size of files read: 10.9 GiB
number of output rows: 24,576
WholeStageCodegen (1)
duration: total (min, med, max)
5.8 s (0 ms, 0 ms, 4.8 s)
ColumnarToRow
number of output rows: 24,576
number of input batches: 6
LocalLimit
GlobalLimit
Filter
Exchange
shuffle records written: 20,002
shuffle write time total (min, med, max)
0 ms (0 ms, 0 ms, 0 ms)
data size total (min, med, max)
1093.9 KiB (0.0 B, 0.0 B, 546.9 KiB)
number of partitions: 500
shuffle bytes written total (min, med, max)
694.1 KiB (0.0 B, 0.0 B, 347.1 KiB)
Exchange
Sort
Sort



Details

```
== Physical Plan ==
AdaptiveSparkPlan (31)
+- == Current Plan ==
   CollectLimit (18)
   +- Project (17)
      +- SortMergeJoin LeftOuter (16)
         :- Sort (6)
         :  +- ShuffleQueryStage (5)
         :     +- Exchange (4)
         :        +- * LocalLimit (3)
         :           +- * ColumnarToRow (2)
         :              +- Scan orc fin_dm_data_ai.dm_ads_fkblack_wj_all (1)
         +- Sort (15)
            +- Exchange (14)
               +- Filter (13)
                  +- GlobalLimit (12)
                     +- ShuffleQueryStage (11)
                        +- Exchange (10)
                           +- * LocalLimit (9)
                              +- * ColumnarToRow (8)
                                 +- Scan orc fin_dm_data_ai.dm_ads_prea_all_score_fusion_6_7 (7)
+- == Initial Plan ==
   CollectLimit (30)
   +- Project (29)
      +- SortMergeJoin LeftOuter (28)
         :- Sort (21)
         :  +- Exchange (20)
         :     +- LocalLimit (19)
         :        +- Scan orc fin_dm_data_ai.dm_ads_fkblack_wj_all (1)
         +- Sort (27)
            +- Exchange (26)
               +- Filter (25)
                  +- GlobalLimit (24)
                     +- Exchange (23)
                        +- LocalLimit (22)
                           +- Scan orc fin_dm_data_ai.dm_ads_prea_all_score_fusion_6_7 (7)


(1) Scan orc fin_dm_data_ai.dm_ads_fkblack_wj_all
Output [2]: [deviceid#50, score#51]
Batched: true
Location: InMemoryFileIndex [hdfs://360jinronglycc/user/hive/warehouse/fin_dm_data_ai.db/dm_ads_fkblack_wj_all]
ReadSchema: struct<deviceid:string,score:double>

(2) ColumnarToRow [codegen id : 1]
Input [2]: [deviceid#50, score#51]

(3) LocalLimit [codegen id : 1]
Input [2]: [deviceid#50, score#51]
Arguments: 10001

(4) Exchange
Input [2]: [deviceid#50, score#51]
Arguments: hashpartitioning(deviceid#50, 500), ENSURE_REQUIREMENTS, [plan_id=222]

(5) ShuffleQueryStage
Output [2]: [deviceid#50, score#51]
Arguments: 0

(6) Sort
Input [2]: [deviceid#50, score#51]
Arguments: [deviceid#50 ASC NULLS FIRST], false, 0

(7) Scan orc fin_dm_data_ai.dm_ads_prea_all_score_fusion_6_7
Output [2]: [deviceid#0, score#1]
Batched: true
Location: InMemoryFileIndex [hdfs://360jinronglycc/user/hive/warehouse/fin_dm_data_ai.db/dm_ads_prea_all_score_fusion_6_7]
ReadSchema: struct<deviceid:string,score:float>

(8) ColumnarToRow [codegen id : 2]
Input [2]: [deviceid#0, score#1]

(9) LocalLimit [codegen id : 2]
Input [2]: [deviceid#0, score#1]
Arguments: 1000000

(10) Exchange
Input [2]: [deviceid#0, score#1]
Arguments: SinglePartition, ENSURE_REQUIREMENTS, [plan_id=246]

(11) ShuffleQueryStage
Output [2]: [deviceid#0, score#1]
Arguments: 1
```

```
(12) GlobalLimit
Input [2]: [deviceid#0, score#1]
Arguments: 1000000

(13) Filter
Input [2]: [deviceid#0, score#1]
Condition : isnotnull(deviceid#0)

(14) Exchange
Input [2]: [deviceid#0, score#1]
Arguments: hashpartitioning(deviceid#0, 500), ENSURE_REQUIREMENTS, [plan_id=255]

(15) Sort
Input [2]: [deviceid#0, score#1]
Arguments: [deviceid#0 ASC NULLS FIRST], false, 0

(16) SortMergeJoin
Left keys [1]: [deviceid#50]
Right keys [1]: [deviceid#0]
Join condition: None

(17) Project
Output [4]: [deviceid#50, cast(score#51 as string) AS score#72, deviceid#0, cast(score#1 as string) AS score#74]
Input [4]: [deviceid#50, score#51, deviceid#0, score#1]

(18) CollectLimit
Input [4]: [deviceid#50, score#72, deviceid#0, score#74]
Arguments: 10001

(19) LocalLimit
Input [2]: [deviceid#50, score#51]
Arguments: 10001

(20) Exchange
Input [2]: [deviceid#50, score#51]
Arguments: hashpartitioning(deviceid#50, 500), ENSURE_REQUIREMENTS, [plan_id=204]

(21) Sort
Input [2]: [deviceid#50, score#51]
Arguments: [deviceid#50 ASC NULLS FIRST], false, 0

(22) LocalLimit
Input [2]: [deviceid#0, score#1]
Arguments: 1000000

(23) Exchange
Input [2]: [deviceid#0, score#1]
Arguments: SinglePartition, ENSURE_REQUIREMENTS, [plan_id=199]

(24) GlobalLimit
Input [2]: [deviceid#0, score#1]
Arguments: 1000000

(25) Filter
Input [2]: [deviceid#0, score#1]
Condition : isnotnull(deviceid#0)

(26) Exchange
Input [2]: [deviceid#0, score#1]
Arguments: hashpartitioning(deviceid#0, 500), ENSURE_REQUIREMENTS, [plan_id=205]

(27) Sort
Input [2]: [deviceid#0, score#1]
Arguments: [deviceid#0 ASC NULLS FIRST], false, 0

(28) SortMergeJoin
Left keys [1]: [deviceid#50]
Right keys [1]: [deviceid#0]
Join condition: None

(29) Project
Output [4]: [deviceid#50, cast(score#51 as string) AS score#72, deviceid#0, cast(score#1 as string) AS score#74]
Input [4]: [deviceid#50, score#51, deviceid#0, score#1]

(30) CollectLimit
Input [4]: [deviceid#50, score#72, deviceid#0, score#74]
Arguments: 10001

(31) AdaptiveSparkPlan
Output [4]: [deviceid#50, score#72, deviceid#0, score#74]
Arguments: isFinalPlan=false
```

▸ SQL / DataFrame Properties