

Details for Query 6

Submitted Time: 2023/11/07 21:57:37

Duration: 53 s

Succeeded Jobs: [67](#)

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)
31.1 m (333 ms, 634 ms, 21.9 s)
metadata time: 2 ms
size of files read: 181.9 GiB
number of output rows: 209,715,200

WholeStageCodegen (1)

duration: total (min, med, max)
33.3 m (380 ms, 692 ms, 22.0 s)

ColumnarToRow

number of output rows: 209,715,200
number of input batches: 51,200

LocalLimit

Exchange

shuffle records written: 204,800,000
shuffle write time total (min, med, max)
26.9 s (6 ms, 9 ms, 918 ms)
records read: 100,000
local bytes read: 49.6 MiB
fetch wait time: 0 ms
remote bytes read: 38.1 MiB
local blocks read: 13
remote blocks read: 10
data size total (min, med, max)
10.7 GiB (5.3 MiB, 5.3 MiB, 5.3 MiB)
number of partitions: 1
remote bytes read to disk: 0.0 B
shuffle bytes written total (min, med, max)
7.6 GiB (3.8 MiB, 3.8 MiB, 3.8 MiB)

WholeStageCodegen (2)

duration: 611 ms

GlobalLimit

AdaptiveSparkPlan

Execute InsertIntoHadoopFsRelationCommand

task commit time: 23 ms
number of written files: 1
job commit time: 14 ms
number of output rows: 100,000
number of dynamic part: 0
written output: 4.2 MiB

```

== Physical Plan ==
Execute InsertIntoHadoopFsRelationCommand (11)
+- AdaptiveSparkPlan (10)
  +- == Final Plan ==
    * GlobalLimit (6)
    +- ShuffleQueryStage (5), Statistics(sizeInBytes=10.7 GiB, rowCount=2.05E+8)
      +- Exchange (4)
        +- * LocalLimit (3)
          +- * ColumnarToRow (2)
            +- Scan orc fin_dm_data_ai.dm_ads_prea_all_score_fusion_6_7 (1)
  +- == Initial Plan ==
    GlobalLimit (9)
    +- Exchange (8)
      +- LocalLimit (7)
        +- Scan orc fin_dm_data_ai.dm_ads_prea_all_score_fusion_6_7 (1)

(1) 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>

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

(3) LocalLimit [codegen id : 1]
Input [2]: [deviceid#0, score#1]
Arguments: 100000

(4) Exchange
Input [2]: [deviceid#0, score#1]
Arguments: SinglePartition, ENSURE_REQUIREMENTS, [plan_id=109]

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

(6) GlobalLimit [codegen id : 2]
Input [2]: [deviceid#0, score#1]
Arguments: 100000

(7) LocalLimit
Input [2]: [deviceid#0, score#1]
Arguments: 100000

(8) Exchange
Input [2]: [deviceid#0, score#1]
Arguments: SinglePartition, ENSURE_REQUIREMENTS, [plan_id=96]

(9) GlobalLimit
Input [2]: [deviceid#0, score#1]
Arguments: 100000

(10) AdaptiveSparkPlan
Output [2]: [deviceid#0, score#1]
Arguments: isFinalPlan=true

(11) Execute InsertIntoHadoopFsRelationCommand
Input [2]: [deviceid#0, score#1]
Arguments: hdfs://360jinronglycc/user/tianzhipeng-jk/temp1107, false, CSV, [path=/user/tianzhipeng-jk/temp1107], ErrorIfExists, [deviceid, score]

```

► [SQL / DataFrame Properties](#)