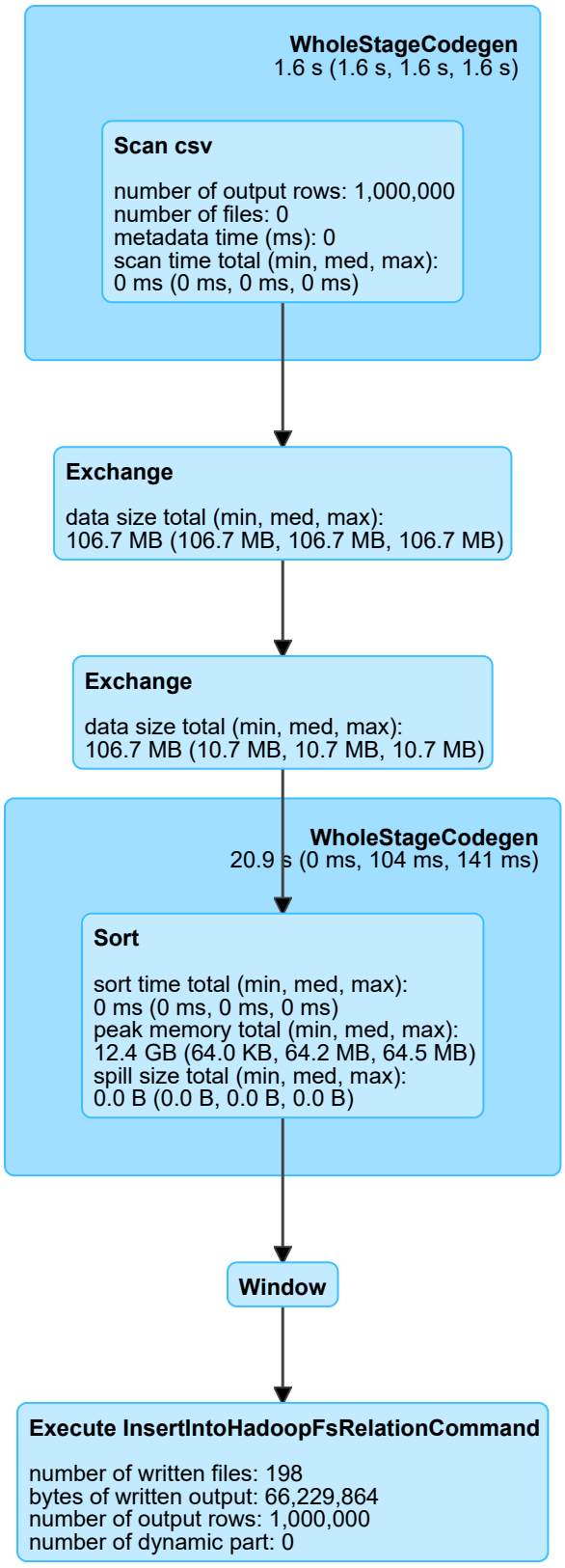
Details for Query 3

Submitted Time: 2023/11/18 20:22:22
Duration: 37 s
Succeeded Jobs: 4 (/jobs/job?id=4)


WholeStageCodegen
1.6 s (1.6 s, 1.6 s, 1.6 s)
Scan csv
number of output rows: 1,000,000
number of files: 0
metadata time (ms): 0
scan time total (min, med, max):
0 ms (0 ms, 0 ms, 0 ms)
Exchange
data size total (min, med, max):
106.7 MB (106.7 MB, 106.7 MB, 106.7 MB)
Exchange
data size total (min, med, max):
106.7 MB (10.7 MB, 10.7 MB, 10.7 MB)
WholeStageCodegen
20.9 s (0 ms, 104 ms, 141 ms)
Sort
sort time total (min, med, max):
0 ms (0 ms, 0 ms, 0 ms)
peak memory total (min, med, max):
12.4 GB (64.0 KB, 64.2 MB, 64.5 MB)
spill size total (min, med, max):
0.0 B (0.0 B, 0.0 B, 0.0 B)
Window
Execute InsertIntoHadoopFsRelationCommand
number of written files: 198
bytes of written output: 66,229,864
number of output rows: 1,000,000
number of dynamic part: 0

▼ Details

```
== Parsed Logical Plan ==
InsertIntoHadoopFsRelationCommand file:/c:/temp/window.csv, false, CSV, Map(header -> true, path ->
c:/temp/window.csv), Overwrite, [id#10, parentid#11, data#12, date#13, url#14, count#98L]
+- AnalysisBarrier
      +- Project [id#10, parentid#11, data#12, date#13, url#14, count#98L]
         +- Project [id#10, parentid#11, data#12, date#13, url#14, count#98L, count#98L]
            +- Window [count(1) windowspecdefinition(parentid#11, specifiedwindowframe(RowFrame,
unboundedpreceding$(), unboundedfollowing$())) AS count#98L], [parentid#11]
               +- Project [id#10, parentid#11, data#12, date#13, url#14]
                  +- Repartition 10, true
                     +- Relation[id#10,parentid#11,data#12,date#13,url#14] csv

== Analyzed Logical Plan ==
InsertIntoHadoopFsRelationCommand file:/c:/temp/window.csv, false, CSV, Map(header -> true, path ->
c:/temp/window.csv), Overwrite, [id#10, parentid#11, data#12, date#13, url#14, count#98L]
+- Project [id#10, parentid#11, data#12, date#13, url#14, count#98L]
   +- Project [id#10, parentid#11, data#12, date#13, url#14, count#98L, count#98L]
      +- Window [count(1) windowspecdefinition(parentid#11, specifiedwindowframe(RowFrame,
unboundedpreceding$(), unboundedfollowing$())) AS count#98L], [parentid#11]
         +- Project [id#10, parentid#11, data#12, date#13, url#14]
            +- Repartition 10, true
               +- Relation[id#10,parentid#11,data#12,date#13,url#14] csv

== Optimized Logical Plan ==
InsertIntoHadoopFsRelationCommand file:/c:/temp/window.csv, false, CSV, Map(header -> true, path ->
c:/temp/window.csv), Overwrite, [id#10, parentid#11, data#12, date#13, url#14, count#98L]
+- Window [count(1) windowspecdefinition(parentid#11, specifiedwindowframe(RowFrame,
unboundedpreceding$(), unboundedfollowing$())) AS count#98L], [parentid#11]
   +- Repartition 10, true
      +- Relation[id#10,parentid#11,data#12,date#13,url#14] csv

== Physical Plan ==
Execute InsertIntoHadoopFsRelationCommand InsertIntoHadoopFsRelationCommand file:/c:/temp/window.csv,
false, CSV, Map(header -> true, path -> c:/temp/window.csv), Overwrite, [id#10, parentid#11, data#12,
date#13, url#14, count#98L]
+- Window [count(1) windowspecdefinition(parentid#11, specifiedwindowframe(RowFrame,
unboundedpreceding$(), unboundedfollowing$())) AS count#98L], [parentid#11]
   +- *(2) Sort [parentid#11 ASC NULLS FIRST], false, 0
      +- Exchange hashpartitioning(parentid#11, 200)
         +- Exchange RoundRobinPartitioning(10)
            +- *(1) FileScan csv [id#10,parentid#11,data#12,date#13,url#14] Batched: false, Format: CSV,
Location: InMemoryFileIndex[file:/C:/temp/data.csv], PartitionFilters: [], PushedFilters: [], ReadSchema:
struct<id:int,parentid:int,data:string,date:string,url:string>
```