

Obtain timing information for data flow executions

Last updated by | Jackie Huang | Jan 4, 2022 at 12:24 AM PST

Sometimes the Data Flow execution time does not fully show where all the time goes during an execution:



The screenshot shows the 'SinkToLake' Data Flow execution summary. The total processing time is 63m 51s 406ms. Below the summary is a table with three columns: TRANSFORM, ROWS, and TIME. The table lists 12 transformations with their respective row counts and execution times.

TRANSFORM	ROWS	TIME
IndexWeightsForADV	-	-
FixedIncomeADVReferen...	22	3s 533ms
BenchmarkData	32m	
BenchmarkAsOfDateFilter	25m	23m 38s 584ms
GetBenchmarkLatestAsO...	458	893ms
PickDistinctRecords	10	2s 85ms
DerivedAdvGenericInGBP	378k	
JoinWithLatestBenchmark	378k	
SplitOnCurrency@Includ...	12482	
SubIndexBenchmark	13874	9s 682ms
SplitOnCurrency@Other...	1392	14s 647ms
UnionWithLGACRPUSDH	19532	

[How to get number of rows processed, time taken for different stages for Dataflow activity.](#)

This TSG will show you more detailed timing information that should help determine the timeline of specific transformations in your Data Flow execution.

Few points to check out in a case like this:

- What is the compute acquisition time.
- Also check the support files, and check if customer is using wildcard path in source, we are aware of the issue where file discovery time is not reported. We are aware of the issue, but if customer is looking for making dataflow faster, he should look for more targeted wild card path.

Also get the activity output to know answers better, like if customer has post SQL statement, thats might not be showing in UX, but might be reported in activity output.

[Ava conversation related to this](#) 