METRIC	DEFINITION	COMMENTS
(Common metrics)		
QueryId	Unique Identifier for the executed query	
Operator	operator responsible for performing specific tasks within the query execution. E.g: SinkOperator, LogicalValueOperator, PartInMemoryAggregationOperator,	
Operator_id	Identifier for each operator within the execution plan.	
planType	type of planner used (E.g: volcano)	
Thread_duration_max	maximum duration taken amongst all threads to perform a specific operation operation.	
Cost_percent_str	cost percentage represented as string.	
inputRowsPerThread_0_50_75_90_100	Array of 5 values, which shows us how many number of rows are input per thread at different percentiles(wrt how much data is read)	
row_count_in/out	number of rows entering/exiting the operator.	
Partitioned	boolean for whether partitions happened for data or not	
Parallelism	the degree of parallelism employed in executing a specific operator - i.e; we perform certain operation parallelly on different threads at a time generally on different files of data to increase the speed/decrease the time of execution of particular operation.	
Max_memory	maximum memory used while performing certain operation	
num_chunks_in/out:	number of chunks taken in or generated out for an operator	
thread_Duration_0_50_75_90_100	provides information about thread execution duration at different percentiles(wrt how much data is read)	
Cost_percent	percentage of total time spent on performing certain operation	
(Operator specific metrics)		
Total_query_time	total time taken for query execution.	
Was_distributed	Boolean indicating if the query execution/data in database was distributed across multiple systems.	
table_name	name of table on which operation is done	
Stats time	time recorded for gathering all stats.	
totalClientQueryTime		
TotalParquetReadingTime	total time taken to read the parquet files in db (if parallelism is there it is collective time of parquet reading of all threads)	
PartitionPruningDurationMs	Duration (in milliseconds) for partition pruning (partition definition is down)(pruning - decision of removal/skipping unwanted partitions and selecting wanted partitions ready to read)	
executionQueueingTime	Time taken to queue the tasks in order for execution	
ReadIOTime	Duration spent on IO operations to read data(IO - to disk)	
totalBytes	total size of data of output	
fileListingDurationInMs	duration for listing of what files to read	
Query_max_memory	maximum of all max_memorys used for query for all operators	
Queue_blocked_time	time for which queue is blocked with tasks for execution(in % and ms)	
TotalTablescanFilteringTime	total time taken for table scan filtering (if parallelism is there it is collective time of all threads)	
Parsing_time	total time taken for parsing, planning and giving out the plan	
SubmitTasksDurationMs	total time taken for coordinating executor to submit/assign tasks to other executors	
totalOpenDuration		
skipped_pages/skipped_row_groups:	count of pages/rows_groups that are skipped while reading(parquet)	
File_name_max_read_time	name of file which took maximum time for reading	
Parquet_task_cost_percent	cost percent associated specifically with parquet reading	
Seek_io_time/count	time spent on IO seeks and number of IO seeks (seek is moving the header/pointer(points to top of chunk/column of rowgroup in parquet) to other chunck's top to skip the reading of unnecessary chunk)	
read_io_bytes/count	number of IO bytes read and read_IO count	
taskInitializationDuration	max time taken to start the task on all threads(parallelism)	
Tasks	number of tasks to be executed	
Partitions	number of partitions used (partitons are made on large datasets so that to ignore unnecessary reading of some files)	
Files	number of files processed	
cacheHits	we store data in cache in order to ignore unnecessary loading of same data from s3. if we run the query which used same data, it checks in cache for reading, if it is matched it is considered as cacheHit.	
Stream_close_time	close time of stream(pipeline) from s3 to engine	
Page_filter_creation_time_max	maximum time for filtering pages	
Open_time_percent	percentage of time spent in open state	
readColumnChunkStreamsDuration	time taken to read column chunks	
fileReaderOpenTime	total duration for which file reading is happen	
task_rowsInCount_0_50_75_90_100	count of rows that are processed per task across various percentiles	
Total_row_groups	total number of row_groups are invovled	
pageReadFromChunkDuration	duration of reading of pages from chunks	
Read_io_time_percent	percentage of time spent in reading from IO(disk) from total_query_time	
totalRowGroupReadTimeMillis	time taken to read all involved row_groups	
totalRowGroupFilteringTime	time taken to filter all involved row_groups	
filtering_cost_percent	percentage of time spent on filtering	
InMemoryAggregationOperator_max	(in view of parallelism) max time taken among all threads to do aggregation operation individually	

SinkOperator_max	(in view of parallelism) max time taken among all threads to do sink operation individually								
TableScanOperator_max	(in view of parallelism) max time taken among all threads to do table scan operation individually								
PartInMemoryAggregationOperator_max	(in view of parallelism) max time taken among all threads to do aggregation operation individually								