

Environment

▼ Runtime Information

Name	Value
Java Home	/usr/lib/jvm/temurin-17-jdk-amd64
Java Version	17.0.14 (Eclipse Adoptium)
Scala Version	version 2.13.12

▼ Spark Properties

Name	Value
spark.app.id	app-20250219140302-0000
spark.app.initial.jar.urls	spark://gdpic-srvls-batch-f4489043-46c1-4edd-9692-753fea951ac8-m.c.gcp-daci-activity-prd.internal:34011/jars/dataproc-empty-jar-1739973769003.jar,spark://gdpic-srvls-batch-f4489043-46c1-4edd-9692-753fea951ac8-m.c.gcp-daci-activity-prd.internal:34011/jars/digital-clickstream-parser.jar,spark://gdpic-srvls-batch-f4489043-46c1-4edd-9692-753fea951ac8-m.c.gcp-daci-activity-prd.internal:34011/jars/spark-3.5-bigquery-0.36.4.jar,spark://gdpic-srvls-batch-f4489043-46c1-4edd-9692-753fea951ac8-m.c.gcp-daci-activity-prd.internal:34011/jars/digital-clickstream-parser-shadow.jar
spark.app.name	digital-clickstream-parser
spark.app.startTime	1739973780070
spark.app.submitTime	1739973778732
spark.authenticate	true
spark.authenticate.secret	***** (redacted)
spark.checkpoint.compress	true
spark.dataproc.advanced.infer.filter.enabled	true
spark.dataproc.appContext.enabled	true
spark.dataproc.appContext.endpoint	dataproc-us-east4.googleapis.com
spark.dataproc.age.remove.redundant.groupby.keys.enabled	true
spark.dataproc.enhanced.execution.enabled	true
spark.dataproc.enhanced.optimizer.enabled	true
spark.dataproc.listeners	com.google.cloud.spark.performance.DataprocMetricsListener
spark.dataproc.metrics.listener.max.stages	5000
spark.dataproc.sql.optimizer.join.fusion.enabled	true
spark.dataproc.sql.optimizer.scalar.subquery.fusion.enabled	true
spark.datasource.bigquery.intermediateFormat	orc
spark.decommission.enabled	true
spark.default.parallelism	1000
spark.driver.cores	16
spark.driver.extraJavaOptions	-Djava.net.preferIPv6Addresses=false -XX: +IgnoreUnrecognizedVMOptions --add-opens=java.base/java.lang=ALL-UNNAMED --add-opens=java.base/java.lang.invoke=ALL-UNNAMED --add-opens=java.base/java.lang.reflect=ALL-UNNAMED --add-opens=java.base/java.io=ALL-UNNAMED --add-opens=java.base/java.net=ALL-UNNAMED --add-opens=java.base/java.nio=ALL-UNNAMED --add-opens=java.base/java.util=ALL-UNNAMED --add-opens=java.base/java.util.concurrent=ALL-UNNAMED --add-opens=java.base/java.util.concurrent.atomic=ALL-UNNAMED --add-opens=java.base/jdk.internal.ref=ALL-UNNAMED --add-opens=java.base/sun.nio.ch=ALL-UNNAMED --add-opens=java.base/sun.nio.cs=ALL-UNNAMED --add-opens=java.base/sun.security.action=ALL-UNNAMED --add-opens=java.base/sun.util.calendar=ALL-UNNAMED --add-opens=java.security.jgss/sun.security.krb5=ALL-UNNAMED -Djdk.reflect.useDirectMethodHandle=false -Dprofile=prod

Name	Value
spark.driver.host	gdpic-srvls-batch-f4489043-46c1-4edd-9692-753fea951ac8-m.c.gcp-daci-activity-prd.internal
spark.driver.maxResultSize	26214m
spark.driver.memory	32g
spark.driver.memoryOverhead	4g
spark.driver.port	34011
spark.dynamicAllocation.enabled	true
spark.dynamicAllocation.executorAllocationRatio	1
spark.dynamicAllocation.executorIdleTimeout	120s
spark.dynamicAllocation.initialExecutors	80
spark.dynamicAllocation.maxExecutors	560
spark.dynamicAllocation.minExecutors	80
spark.dynamicAllocation.shuffleTracking.enabled	true
spark.eventLog.compress	true
spark.eventLog.dir	gs://daci-activity-phs-eventlog-prd/dataproc-serverless/f4489043-46c1-4edd-9692-753fea951ac8/spark-job-history
spark.eventLog.enabled	true
spark.executor.cores	8
spark.executor.defaultJavaOptions	-XX:MaxHeapFreeRatio=30 -XX:MinHeapFreeRatio=10
spark.executor.extraJavaOptions	-Djava.net.preferIPv6Addresses=false -XX:MaxHeapFreeRatio=30 -XX:MinHeapFreeRatio=10 -XX:+IgnoreUnrecognizedVMOptions --add-opens=java.base/java.lang=ALL-UNNAMED --add-opens=java.base/java.lang.invoke=ALL-UNNAMED --add-opens=java.base/java.lang.reflect=ALL-UNNAMED --add-opens=java.base/java.io=ALL-UNNAMED --add-opens=java.base/java.net=ALL-UNNAMED --add-opens=java.base/java.nio=ALL-UNNAMED --add-opens=java.base/java.util=ALL-UNNAMED --add-opens=java.base/java.util.concurrent=ALL-UNNAMED --add-opens=java.base/java.util.concurrent.atomic=ALL-UNNAMED --add-opens=java.base/jdk.internal.ref=ALL-UNNAMED --add-opens=java.base/sun.nio.ch=ALL-UNNAMED --add-opens=java.base/sun.nio.cs=ALL-UNNAMED --add-opens=java.base/sun.security.action=ALL-UNNAMED --add-opens=java.base/sun.util.calendar=ALL-UNNAMED --add-opens=java.security.jgss/sun.security.krb5=ALL-UNNAMED -Djdk.reflect.useDirectMethodHandle=false -XX:MaxHeapFreeRatio=30 -XX:MinHeapFreeRatio=10
spark.executor.id	driver
spark.executor.instances	80
spark.executor.memory	24g
spark.executor.memoryOverhead	4g
spark.files.maxPartitionBytes	256MB
spark.hadoop.fs.AbstractFileSystem.gs.impl	com.google.cloud.hadoop.fs.gcs.GoogleHadoopFS
spark.hadoop.fs.defaultFS	gs://dataproc-temp-us-east4-252876682928-on949chf
spark.hadoop.fs.gs.application.name.suffix	-dataproc zone/us-east4-a (GPN:dataproc)
spark.hadoop.fs.gs.impl	com.google.cloud.hadoop.fs.gcs.GoogleHadoopFileSystem
spark.hadoop.fs.gs.project.id	gcp-daci-activity-prd
spark.hadoop.fs.gs.reported.permissions	777
spark.hadoop.hive.exec.local.scratchdir	/var/dataproc/tmp/spark/tmp/hive
spark.hadoop.hive.exec.scratchdir	/var/dataproc/tmp/spark/tmp/hive
spark.hadoop.mapreduce.fileoutputcommitter.algorithm.version	2
spark.history.fs.logDirectory	gs://dataproc-temp-us-east4-252876682928-on949chf/f4489043-46c1-4edd-9692-753fea951ac8/spark-job-history
spark.io.compression.lz4.blockSize	256k

Name	Value
spark.jars	file:///var/dataproc/tmp/srvls-batch-f4489043-46c1-4edd-9692-753fea951ac8/spark-3.5-bigquery-0.36.4.jar,file:///var/dataproc/tmp/srvls-batch-f4489043-46c1-4edd-9692-753fea951ac8/digital-clickstream-parser-shadow.jar,file:///var/dataproc/tmp/srvls-batch-f4489043-46c1-4edd-9692-753fea951ac8/digital-clickstream-parser.jar,file:/var/dataproc/tmp/srvls-batch-f4489043-46c1-4edd-9692-753fea951ac8/dataproc-empty-jar-1739973769003.jar
spark.jars.ivySettings	/etc/spark/conf/ivysettings.xml
spark.kryo.unsafe	true
spark.kryoserializer.buffer	256k
spark.master	spark:///gdpic-srvls-batch-f4489043-46c1-4edd-9692-753fea951ac8-m:7077
spark.metrics.namespace	app_name:\${spark.app.name}.app_id:\${spark.app.id}
spark.network.crypto.enabled	true
spark.network.crypto.keyLength	256
spark.network.crypto.saslFallback	false
spark.network.timeout	600s
spark.plugins.defaultList	org.apache.spark.dataproc.buffer.metrics.DataprocBufferSparkPlugin
spark.rdd.compress	true
spark.reducer.fetchMigratedShuffle.enabled	true
spark.repl.local.jars	file:///var/dataproc/tmp/srvls-batch-f4489043-46c1-4edd-9692-753fea951ac8/spark-3.5-bigquery-0.36.4.jar,file:///var/dataproc/tmp/srvls-batch-f4489043-46c1-4edd-9692-753fea951ac8/digital-clickstream-parser-shadow.jar,file:///var/dataproc/tmp/srvls-batch-f4489043-46c1-4edd-9692-753fea951ac8/digital-clickstream-parser.jar
spark.rpc.message.maxSize	512
spark.scheduler.excludeShuffleSkewExecutors	true
spark.scheduler.mode	FIFO
spark.scheduler.shuffleSkew.maxExecutorsRatio	1.0
spark.serializer	org.apache.spark.serializer.KryoSerializer
spark.shuffle.file.buffer	256k
spark.shuffle.io.maxRetries	8
spark.shuffle.io.retryWait	60s
spark.shuffle.mapOutput.minSizeForBroadcast	128m
spark.shuffle.server.chunkFetchHandlerThreadsPercent	100
spark.sql.adaptive.enabled	true
spark.sql.autoBroadcastJoinThreshold	200m
spark.sql.broadcastTimeout	900s
spark.sql.cbo.joinReorder.enabled	true
spark.sql.debug.maxToStringFields	5000
spark.sql.files.maxPartitionBytes	256MB
spark.sql.maxMetadataStringLength	5000
spark.sql.shuffle.partitions	1000
spark.sql.sources.partitionOverwriteMode	dynamic
spark.storage.decommission.enabled	true
spark.storage.decommission.rddBlocks.enabled	true
spark.storage.decommission.shuffleBlocks.enabled	true
spark.submit.deployMode	client
spark.submit.pyFiles	

Name	Value
spark.task.maxFailures	8
spark.ui.port	4040
viewsEnabled	true

▼ Resource Profiles

Resource Profile Id	Resource Profile Contents
0	<div>Executor Reqs: memoryOverhead: [amount: 4096] cores: [amount: 8] memory: [amount: 24576] offHeap: [amount: 0] Task Reqs: cpus: [amount: 1.0]</div>

- ▶ Hadoop Properties
- ▶ System Properties
- ▶ Metrics Properties
- ▶ Classpath Entries