

Environment

▼ Runtime Information

Name	Value
Java Home	/usr/lib/jvm/temurin-11-jdk-amd64
Java Version	11.0.20.1 (Eclipse Adoptium)
Scala Version	version 2.12.18

▼ Spark Properties

Name	Value
spark.app.id	application_1747991848066_0005
spark.app.name	Many Small Files Demo – full
spark.app.startTime	1748547274466
spark.app.submitTime	1748547274324
spark.checkpoint.compress	true
spark.dataproc.advanced.infer.filter.enabled	true
spark.dataproc.aqe.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.metrics.collector.hostname	spark-cluster-m
spark.dataproc.sql.joinConditionReorder.enabled	true
spark.dataproc.sql.local.rank.pushdown.enabled	true
spark.dataproc.sql.optimizer.join.fusion.enabled	true
spark.dataproc.sql.optimizer.leftsemijoin.conversion.enabled	true
spark.dataproc.sql.optimizer.scalar.subquery.fusion.enabled	true
spark.dataproc.sql.parquet.enableFooterCache	true
spark.decommission.enabled	true
spark.driver.appUIAddress	http://spark-cluster-m.europe-west4-c.c.genai-security.internal:38195
spark.driver.extraJavaOptions	-Djava.net.preferIPv6Addresses=false -XX:+IgnoreUnrecognizedVMOptions -Djava.lang=ALL-UNNAMED --add-opens=java.base/java.lang=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.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.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
spark.driver.host	spark-cluster-m.europe-west4-c.c.genai-security.internal
spark.driver.maxResultSize	2048m
spark.driver.memory	4096m
spark.driver.port	44769
spark.dynamicAllocation.enabled	true
spark.dynamicAllocation.maxExecutors	10000
spark.dynamicAllocation.minExecutors	1
spark.eventLog.dir	gs://genai-security-dataproc-utility-temp/spark-cluster/spark-job-history
spark.eventLog.enabled	true
spark.executor.cores	2
spark.executor.extraJavaOptions	-Djava.net.preferIPv6Addresses=false -XX:+IgnoreUnrecognizedVMOptions -Djava.lang=ALL-UNNAMED --add-opens=java.base/java.lang=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/javz UNNAMED --add-opens=java.base/java.util.concurrent=ALL-UNNAMED --ad opens=java.base/java.util.concurrent.atomic=ALL-UNNAMED --add- opens=java.base/jdk.internal.ref=ALL-UNNAMED --add-opens=java.base/sur 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
spark.executor.id	driver
spark.executor.instances	2
spark.executor.memory	6157m
spark.executorEnv.OPENBLAS_NUM_THREADS	1
spark.executorEnv.PYTHONPATH	{{PWD}}/pyspark.zip<CPS>{{PWD}}/py4j-0.10.9.7-src.zip
spark.hadoop.hive.execution.engine	mr
spark.hadoop.mapreduce.fileoutputcommitter.algorithm.version	2
spark.history.fs.gs.outputstream.type	FLUSHABLE_COMPOSITE
spark.history.fs.logDirectory	gs://genai-security-dataproc-utility-temp/spark-cluster/spark-job-history
spark.master	yarn
spark.metrics.namespace	app_name##\${spark.app.name}.app_id##\${spark.app.id}
spark.org.apache.hadoop.yarn.server.webproxy.amfilter.AmIpFilter.param.PROXY_HOSTS	spark-cluster-m.europe-west4-c.c.genai-security.internal.
spark.org.apache.hadoop.yarn.server.webproxy.amfilter.AmIpFilter.param.PROXY_URI_BAS ES	http://ezxe6h27ljavrmfp3rdpdgpxfe-dot-europe- west4.dataproc.googleusercontent.com:80/gateway/default/yarn/proxy/applic 1848066_0005
spark.plugins.defaultList	com.google.cloud.dataproc.DataprocSparkPlugin
spark.rdd.compress	True
spark.rpc.message.maxSize	512
spark.scheduler.minRegisteredResourcesRatio	0.0
spark.scheduler.mode	FAIR
spark.serializer.objectStreamReset	100
spark.shuffle.service.enabled	true
spark.sql.adaptive.enabled	true
spark.sql.autoBroadcastJoinThreshold	46m
spark.sql.catalogImplementation	hive
spark.sql.cbo.enabled	true
spark.sql.cbo.joinReorder.enabled	true
spark.sql.optimizer.runtime.bloomFilter.join.pattern.enabled	true
spark.sql.shuffle.partitions	1000
spark.stage.maxConsecutiveAttempts	10
spark.storage.decommission.enabled	true
spark.storage.decommission.rddBlocks.enabled	true
spark.storage.decommission.shuffleBlocks.enabled	true
spark.submit.deployMode	client
spark.submit.pyFiles	
spark.task.maxFailures	10
spark.ui.filters	org.apache.hadoop.yarn.server.webproxy.amfilter.AmIpFilter
spark.ui.port	0
spark.ui.showConsoleProgress	true
spark.yarn.am.attemptFailuresValidityInterval	1h
spark.yarn.am.memory	640m
spark.yarn.executor.failuresValidityInterval	1h
spark.yarn.historyServer.address	spark-cluster-m:18080
spark.yarn.isPython	true
spark.yarn.jars	local:/usr/lib/spark/jars/*

spark.yarn.unmanagedAM.enabled

true

▼ Resource Profiles

Resource Profile Id	Resource Profile Contents
0	<div>Executor Reqs:<div>cores: [amount: 2]memory: [amount: 6157]offHeap: [amount: 0]</div>Task Reqs:<div>cpus: [amount: 1.0]</div></div>

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