TEST 2 : Hive view test for job submission. Test with single user only as could not create users on Humboldt server yet.

Number of Threads: 3

Ramp-up: 1 sec

Loop Count: 10

* No errors were encountered.
* There were no errors in the ambari server logs.
* All the jobs succeeded from ambari side.
* All the jobs succeeded in yarn cluster. (checked from yarn ui)
* For more details see the attached logs.

Description of Test:

Three threads were spawned. Each thread corresponds to a browser session. Right now same user was used for login into all three sessions. In each session below “procedure” was executed 10 times. So total jobs submitted were 30. In each session next job is submitted once “result” of current job is received.

“procedure”:

1. Submitting the job: the POST REST call /jobs is used for submitting the job.

Query submitted is a join query below.

select \* from table1 as t1 join table2 as t2 on t1.col1 = t2.col1;

After submission the job id is retrieved. Hive view creates query file and logs file in HDFS corresponding to this query. This query file is fetched (REST GET) once using the /resources/file/queryFile

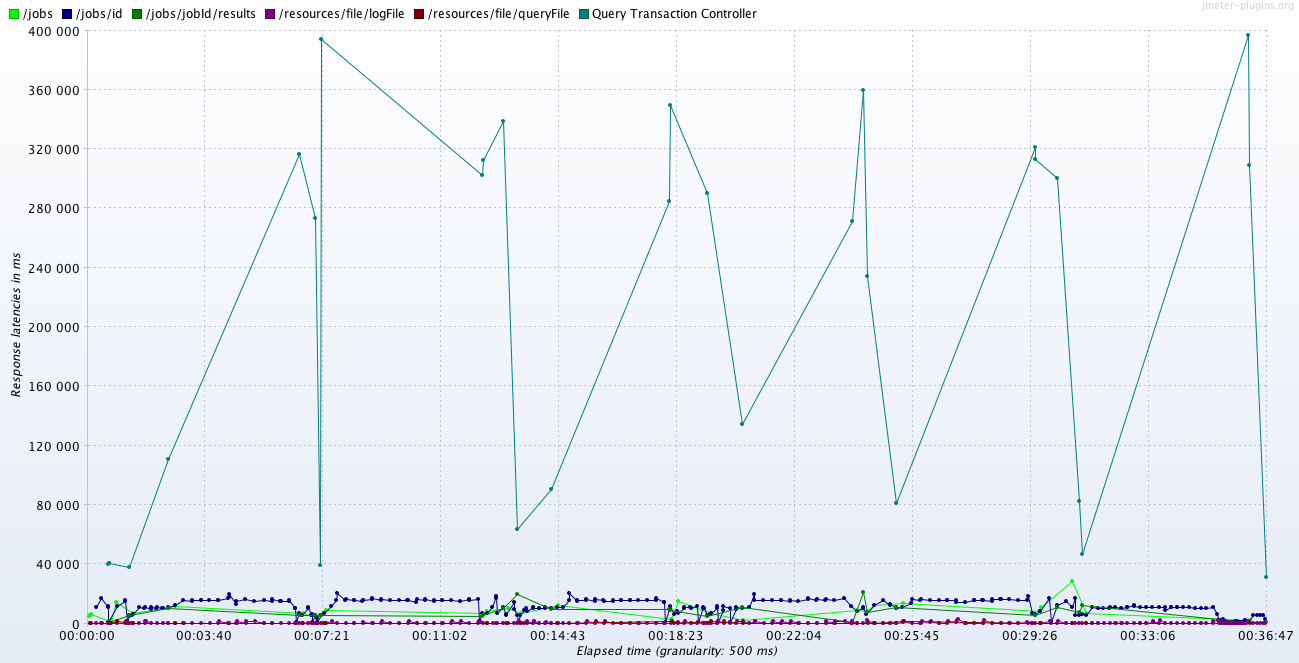
1. Polling for job status: /jobs/id REST GET call is used to fetch the status of current job. This is done continuously till the job success or error status is received. In each of this loop /resources/file/logFile REST GET request is also sent to get the current logs for the job.

1. Fetching the results: if job executed successfully and there were no exceptions encountered then the job result is fetched once using the REST GET /jobs/jobId/results call.

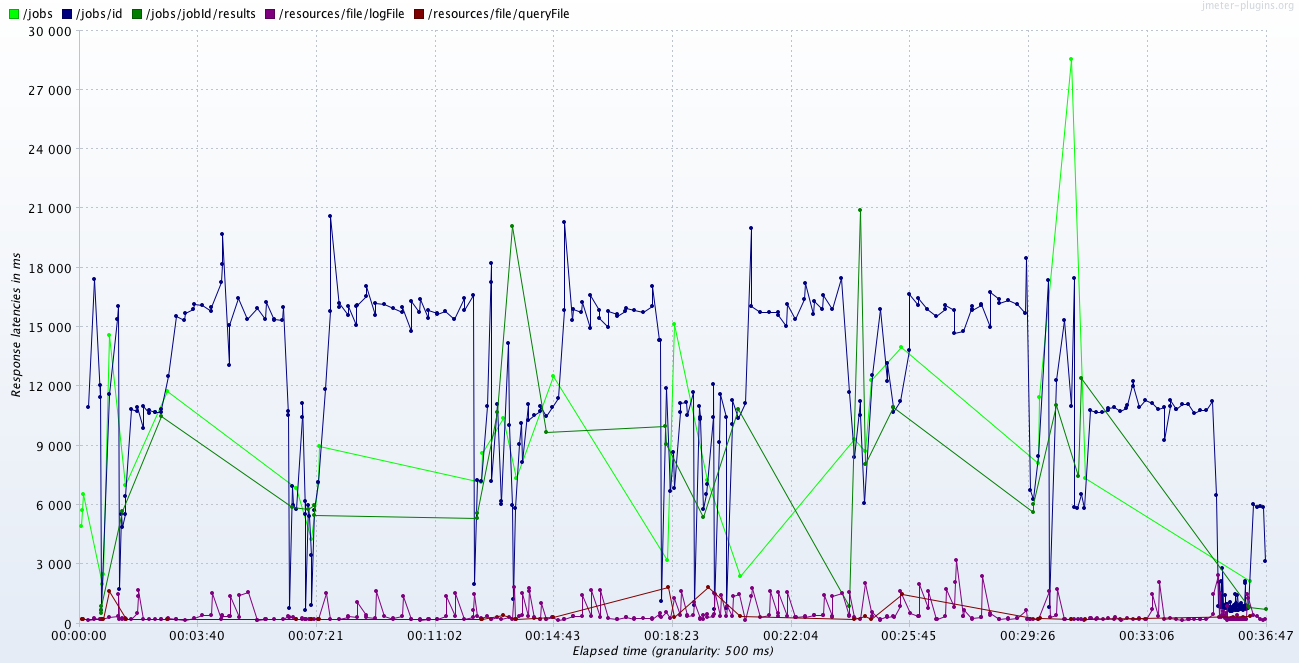
Description of graph:

Graph is plotted with Response Latencies in milliseconds on Y-axis and Elapsed time on X-axis. Special parameter “Query Transaction Controller” for each job is captured in graphs and reports. This represents the total time taken for one execution of “procedure” described above (job submission + poll for status + fetching of result). So query transaction latency also includes the time taken by Hive for the job execution.

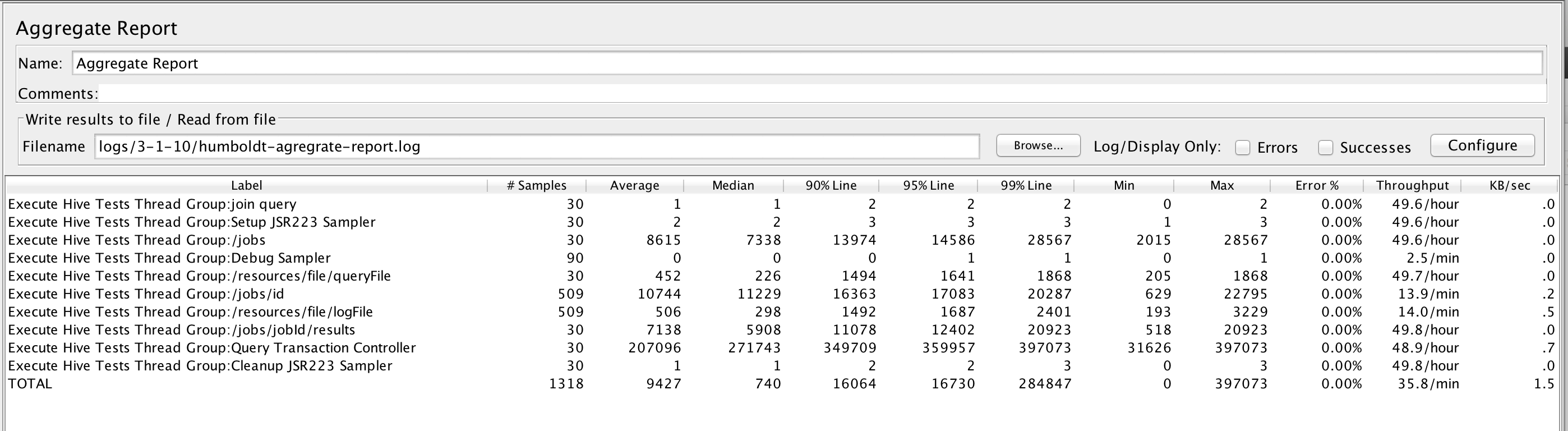
1. Graph with query transaction.



2. Same graph without query transaction (for greater visibility of other parameters)



3. Aggregate Report



4. Summary Report

