Workload Evaluation Report

genotyper-xray-log

Report Date:06/29/2018

Overview

|  |  |
| --- | --- |
| Results File(s): | genotyper-xray-log |
| Spark Version: | 2.1.1 |
| Cluster Size (number of servers): | local[\*] |
| Cluster Location: | AWS EMR |
| Node or Instance Type: |  |
| Workload Description(s): |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Number of Steps | Number of Jobs | Number of Stages | UDFs (Y/N) | Data Format (csv, avro, parquet) | Data Frame/ RDD | Total Run Time (s) |
| 1 | 0 | 7 | No |  | RDD | 42.914 |

**Note:**On Task Summary charts and Task Detail charts there are outlier tasks marked with red color. Task is considered to be an outlier when its run time differs by more than two standard deviations from the mean and at the same time more than 25% from the mean.

Overall Performance Summary

|  |  |  |
| --- | --- | --- |
| Workload | Accelerated? | Time (s) |
| Sampling Stages | No | 42.914 |
| Total App |  | 42.914 |

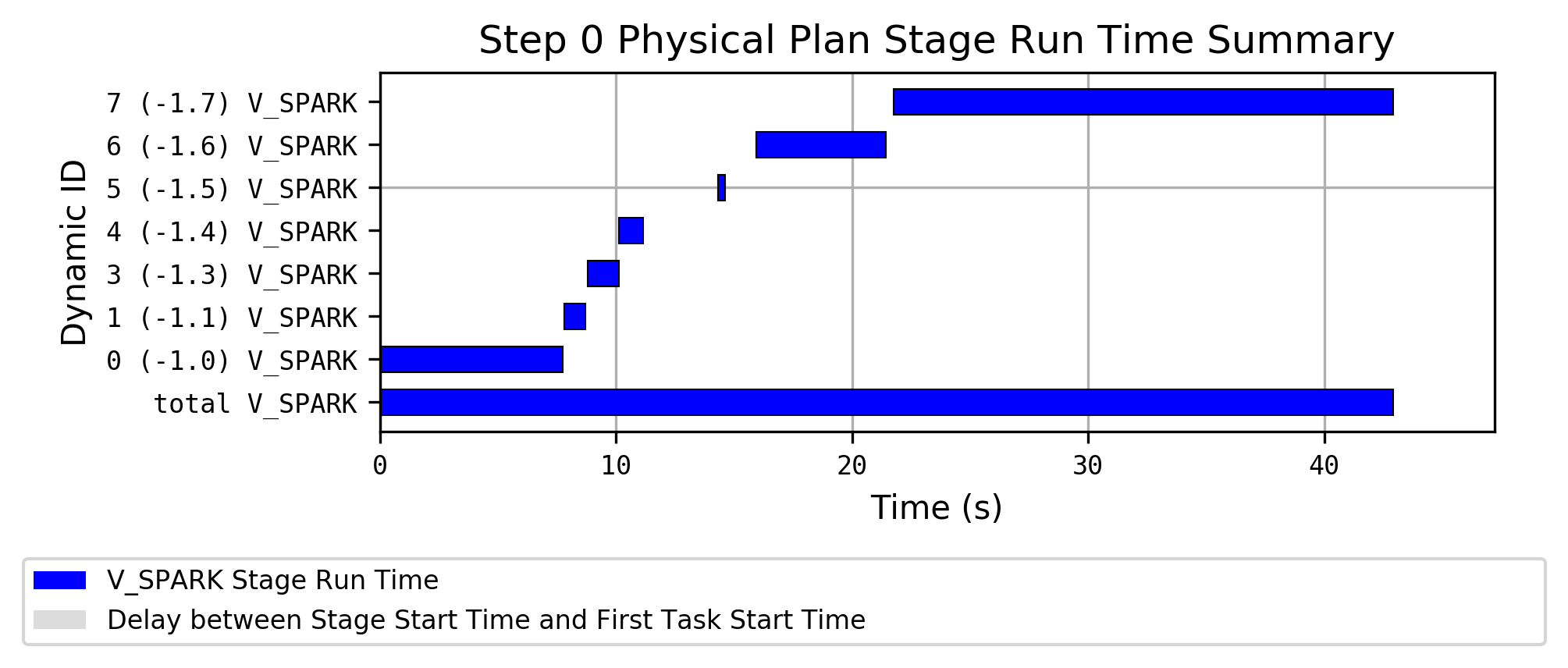
Workload details on the next pages.

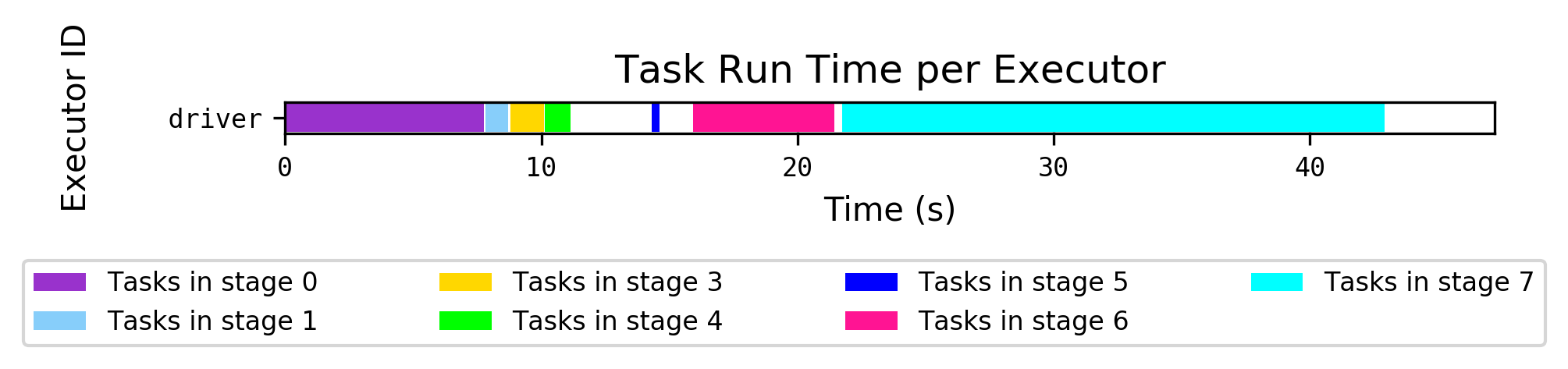
Workload details:

Step0: genotyper-xray-log

Execution Time:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Number of Jobs | Number of Stages | UDFs (Y/N) | Data Format (csv, avro, parquet) | Data Frame/ RDD |
| 0 | 7 | No |  | RDD |





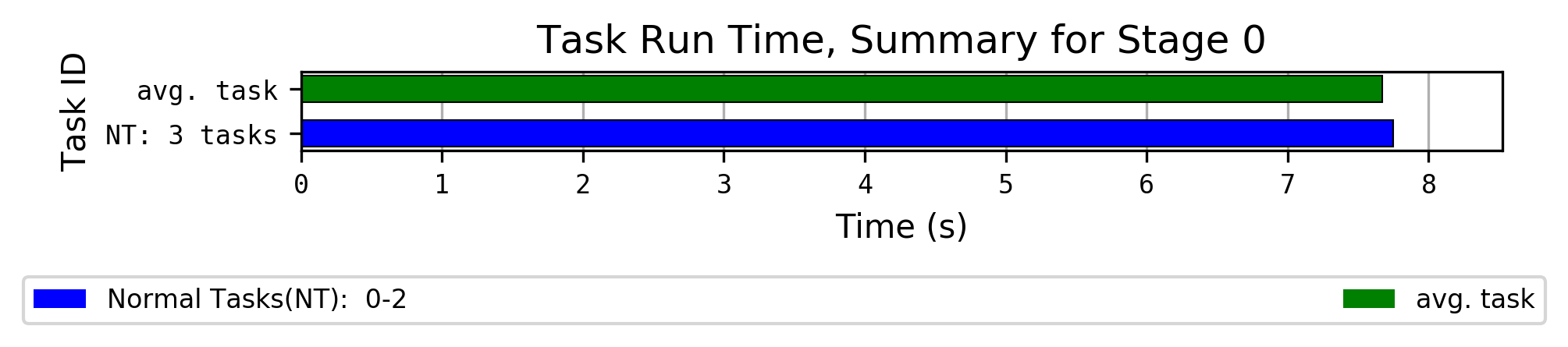
Hostname to Executor ID Mapping:

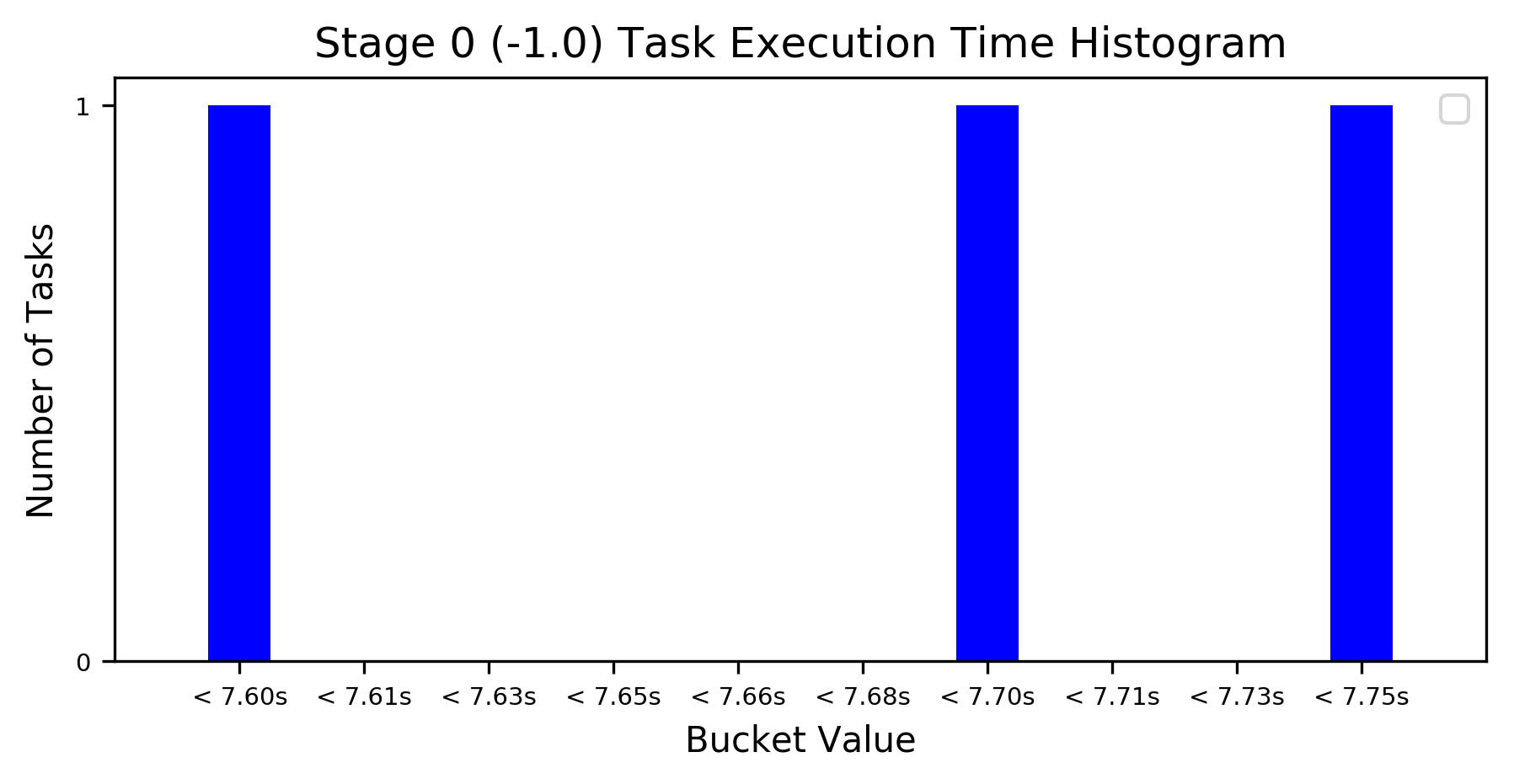
|  |  |  |
| --- | --- | --- |
| Hostname | IP | Executor ID List |
| localhost |  | driver |

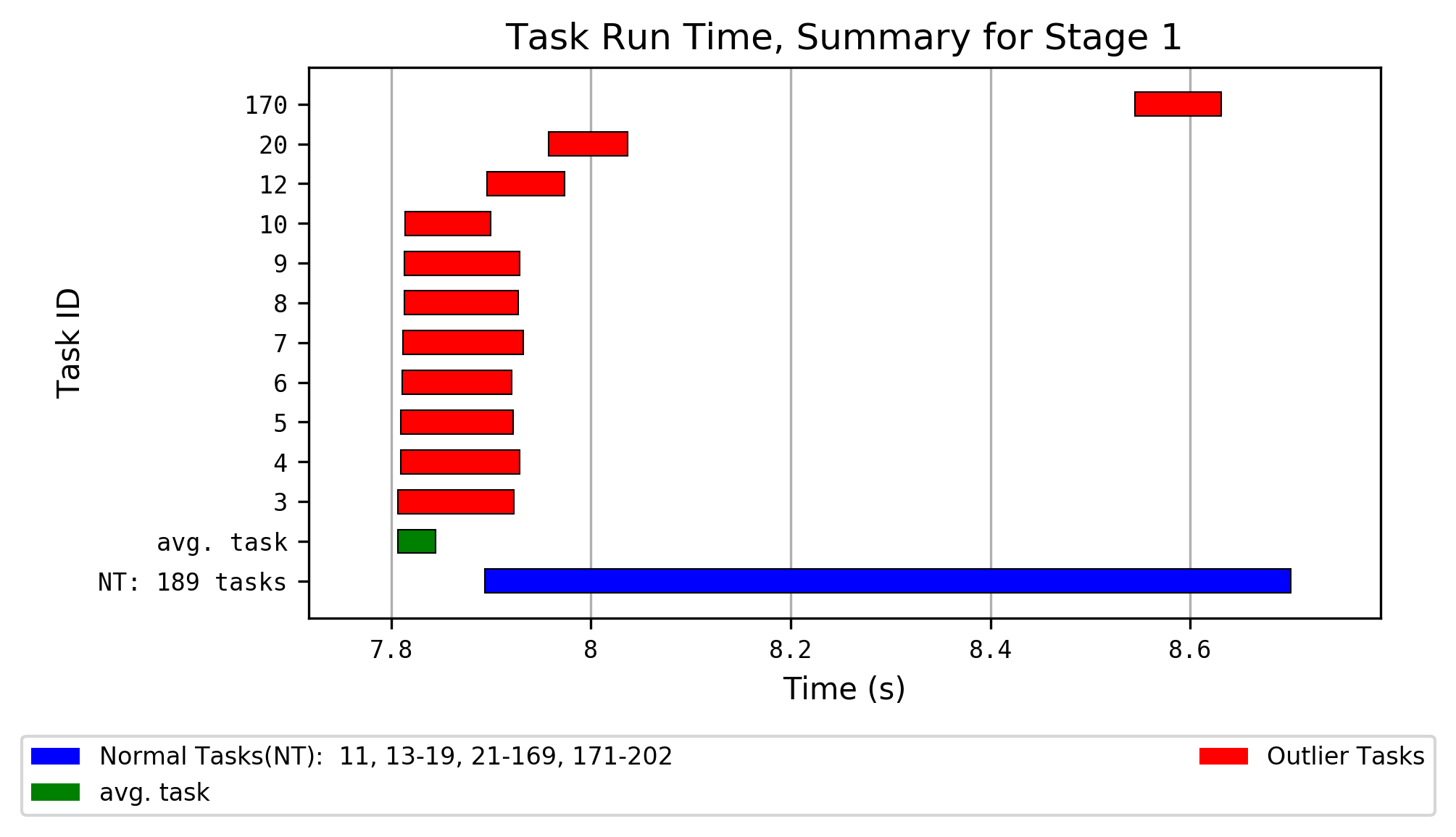
Sampling Stages Details

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Stage ID | Operators | Number of Tasks | Number of UDFs | Total Runtime (Task Runtime) (s) |
| 0 | (0) | 3 | 0 | 7.755 (7.746) |
| 1 | (0) | 200 | 0 | 0.895 (0.894) |
| 3 | (0) | 200 | 0 | 1.315 (1.315) |
| 4 | (0) | 200 | 0 | 0.996 (0.996) |
| 5 | (0) | 8 | 0 | 0.324 (0.323) |
| 6 | (0) | 3 | 0 | 5.494 (5.493) |
| 7 | (0) | 200 | 0 | 21.154 (21.154) |

|  |  |  |  |
| --- | --- | --- | --- |
| Stage | Time (s) | Average Task Time (s) | Task Standard Deviation (ms) |
| 0 | 7.755 | 7.67 | 69 |
| 1 | 0.895 | 0.038 | 20 |
| 3 | 1.315 | 0.054 | 20 |
| 4 | 0.996 | 0.04 | 11 |
| 5 | 0.324 | 0.258 | 23 |
| 6 | 5.494 | 5.434 | 43 |
| 7 | 21.154 | 0.842 | 230 |
| **TOTAL** | **42.914** |  |  |





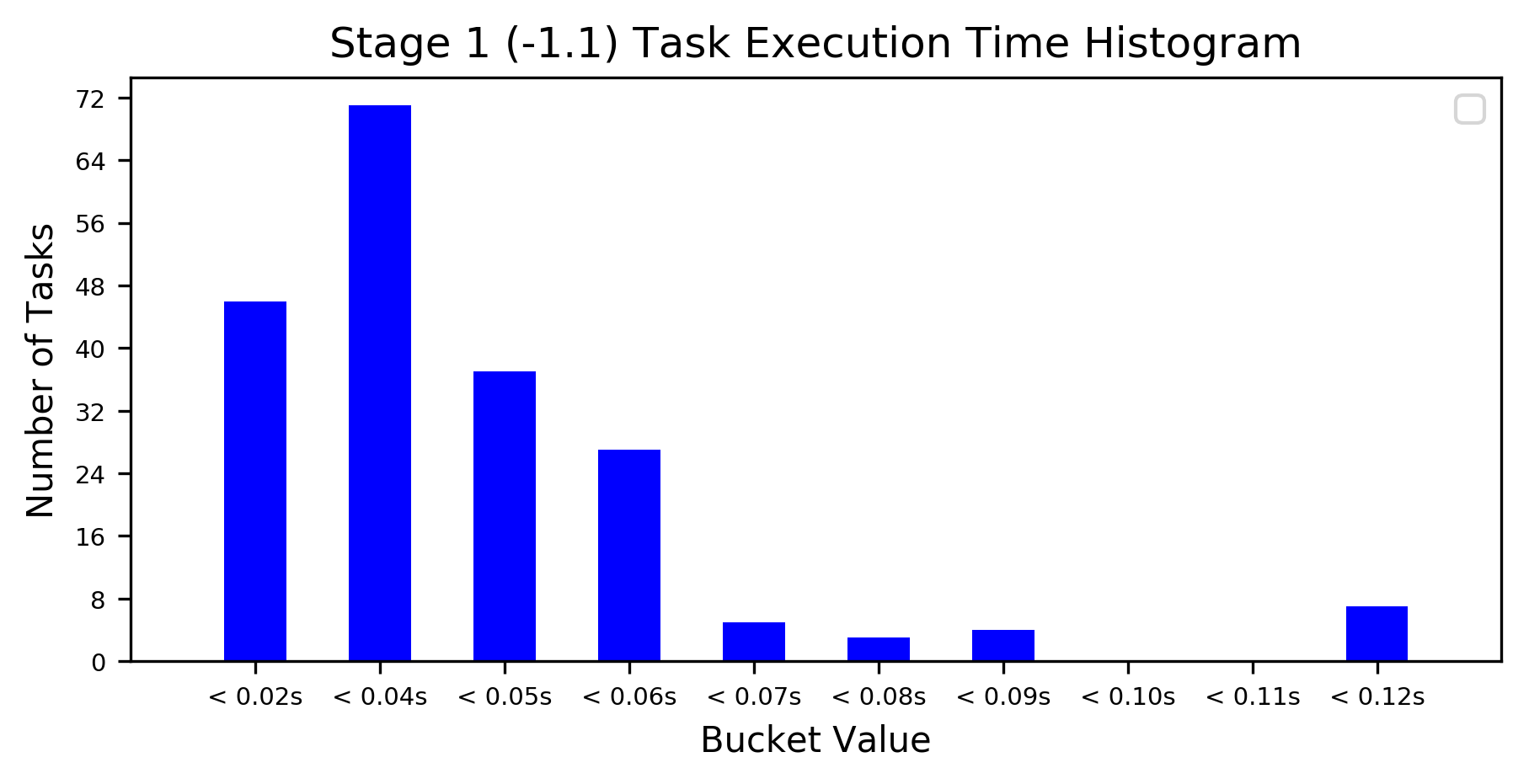


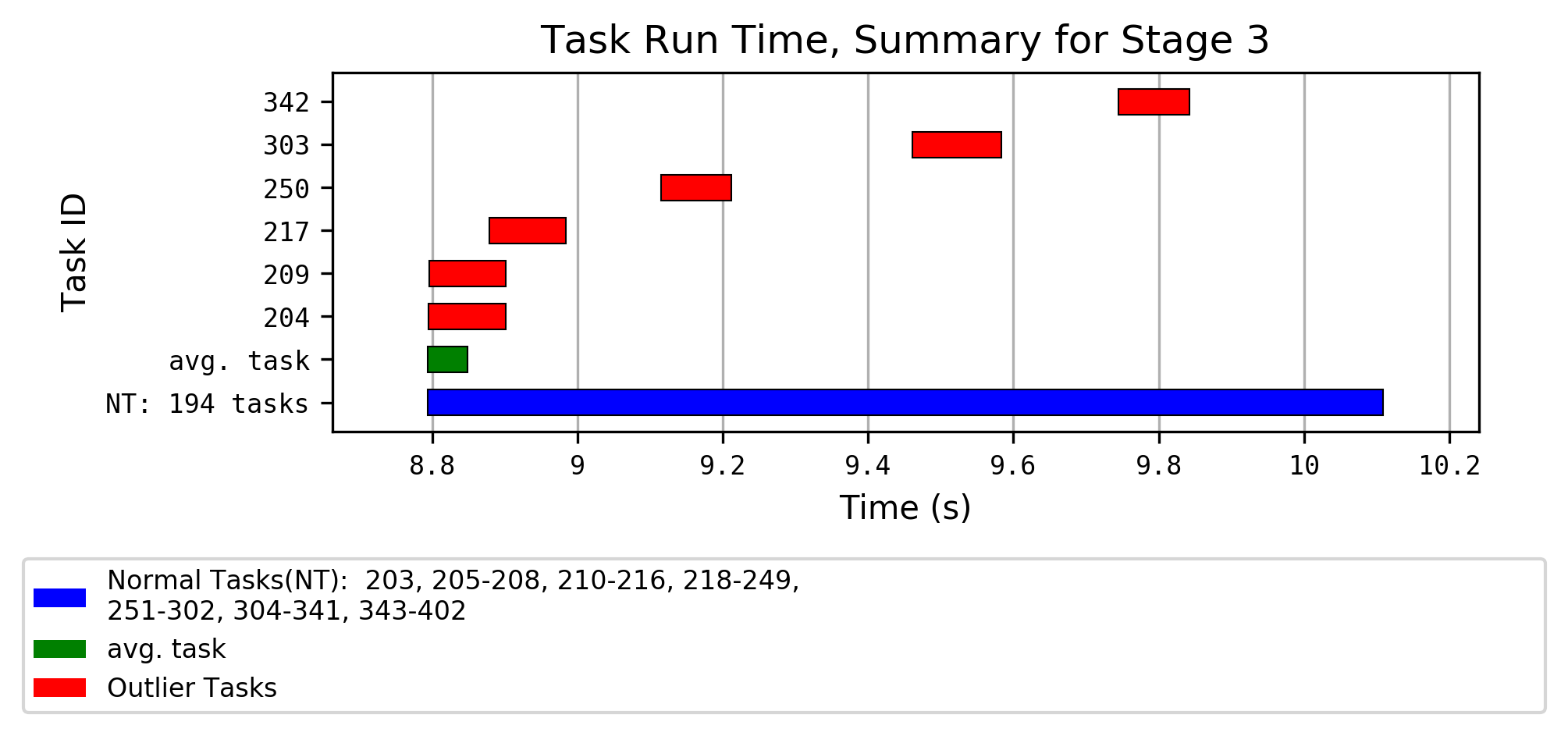
Task Outliers for Stage 1, part 1

| Task ID | Executor ID | Executor Run Time | Total Records Read | Local Bytes Read | Executor Deserialize CPU Time | Result Serialization Time | Result Size |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Stage Avg. |  | 11 ms | 102.50 | 2.83 KiB | 4 ms | 0 ms | 2.73 KiB |
| Slow Outliers Avg. |  | 53 ms | 101.27 | 2.78 KiB | 6 ms | 0 ms | 2.75 KiB |
| 3 | driver | 68 ms | 114 | 3.07 KiB | 6 ms | 0 ms | 2.72 KiB |
| 4 | driver | 74 ms | 106 | 2.77 KiB | 9 ms | 0 ms | 2.74 KiB |
| 5 | driver | 72 ms | 95 | 2.66 KiB | 9 ms | 0 ms | 2.74 KiB |
| 6 | driver | 68 ms | 97 | 2.61 KiB | 6 ms | 0 ms | 2.72 KiB |
| 7 | driver | 71 ms | 90 | 2.48 KiB | 6 ms | 0 ms | 2.74 KiB |
| 8 | driver | 56 ms | 107 | 3 KiB | 6 ms | 1 ms | 2.79 KiB |
| 9 | driver | 63 ms | 107 | 2.96 KiB | 5 ms | 1 ms | 2.84 KiB |
| 10 | driver | 52 ms | 94 | 2.71 KiB | 4 ms | 0 ms | 2.7 KiB |
| 12 | driver | 19 ms | 120 | 3.1 KiB | 5 ms | 0 ms | 2.72 KiB |
| 20 | driver | 6 ms | 86 | 2.49 KiB | 5 ms | 0 ms | 2.75 KiB |
| 170 | driver | 38 ms | 98 | 2.69 KiB | 3 ms | 1 ms | 2.82 KiB |

Task Outliers for Stage 1, part 2

| Task ID | Executor ID | Executor CPU Time | Executor Deserialize Time | Task Run Time | Aggregate Time Total | Duration Total | Number Of Output Rows |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Stage Avg. |  | 5 ms | 7 ms | 38 ms | 2 ms | 5 ms | 31.59 |
| Slow Outliers Avg. |  | 16 ms | 11 ms | 0.1 s | 10 ms | 32 ms | 19.45 |
| 3 | driver | 20 ms | 13 ms | 0.1 s | 16 ms | 43 ms | 1 |
| 4 | driver | 19 ms | 10 ms | 0.1 s | 14 ms | 47 ms | 2 |
| 5 | driver | 25 ms | 10 ms | 0.1 s | 16 ms | 45 ms | 2 |
| 6 | driver | 30 ms | 10 ms | 0.1 s | 14 ms | 43 ms | 1 |
| 7 | driver | 17 ms | 9 ms | 0.1 s | 12 ms | 45 ms | 2 |
| 8 | driver | 18 ms | 10 ms | 0.1 s | 10 ms | 32 ms | 103 |
| 9 | driver | 19 ms | 6 ms | 0.1 s | 16 ms | 44 ms | 3 |
| 10 | driver | 15 ms | 4 ms | 86 ms | 14 ms | 36 ms | 94 |
| 12 | driver | 7 ms | 17 ms | 78 ms | 1 ms | 14 ms | 1 |
| 20 | driver | 6 ms | 32 ms | 79 ms | 1 ms | 3 ms | 3 |
| 170 | driver | 5 ms | 3 ms | 86 ms | 1 ms | 3 ms | 2 |



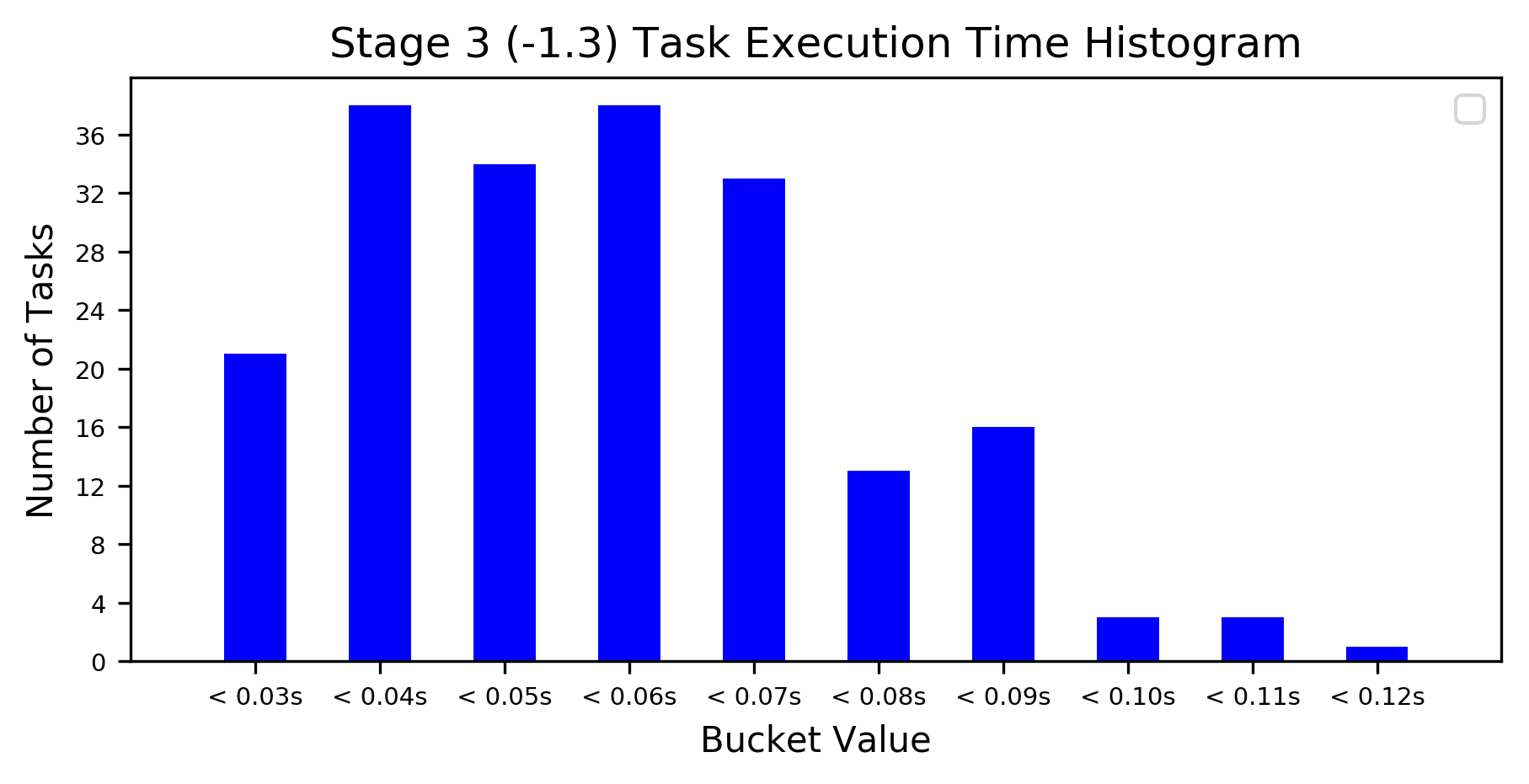


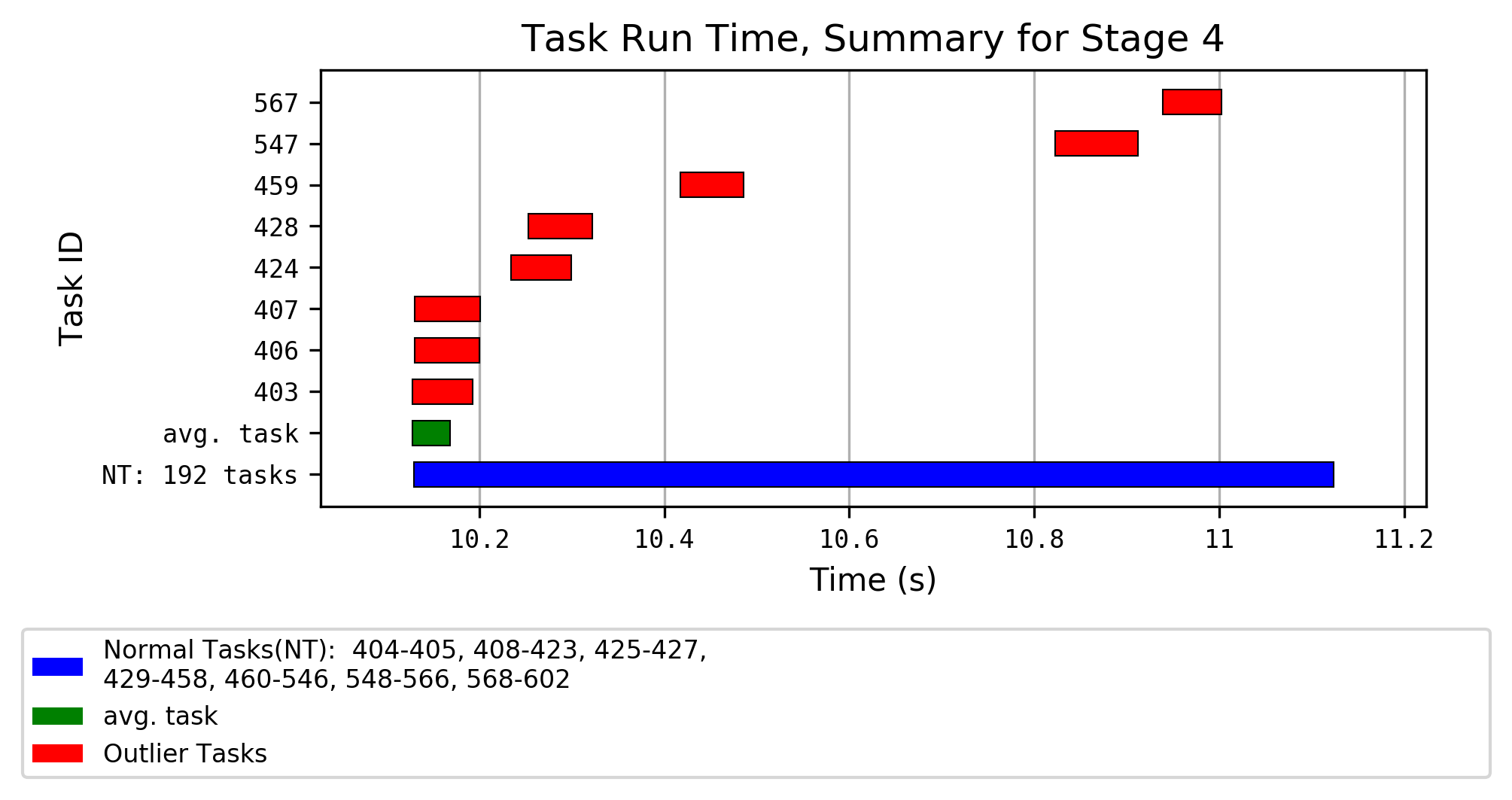
Task Outliers for Stage 3, part 1

| Task ID | Executor ID | Executor Run Time | Total Records Read | Local Bytes Read | JVM GC Time | Executor Deserialize CPU Time | Shuffle Bytes Written | Result Size | Shuffle Write Time |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Stage Avg. |  | 21 ms | 102.50 | 2.83 KiB | 1 ms | 10 ms | 112.58 bytes | 3.11 KiB | 4 ms |
| Slow Outliers Avg. |  | 52 ms | 106.33 | 2.85 KiB | 3 ms | 12 ms | 257.5 bytes | 3.19 KiB | 6 ms |
| 204 | driver | 48 ms | 106 | 2.77 KiB | 0 ms | 12 ms | 197 bytes | 3.18 KiB | 4 ms |
| 209 | driver | 58 ms | 107 | 2.96 KiB | 0 ms | 15 ms | 288 bytes | 3.18 KiB | 3 ms |
| 217 | driver | 67 ms | 98 | 2.64 KiB | 0 ms | 12 ms | 391 bytes | 3.18 KiB | 3 ms |
| 250 | driver | 45 ms | 124 | 3.37 KiB | 0 ms | 10 ms | 382 bytes | 3.18 KiB | 4 ms |
| 303 | driver | 75 ms | 105 | 2.72 KiB | 17 ms | 11 ms | 193 bytes | 3.25 KiB | 16 ms |
| 342 | driver | 18 ms | 98 | 2.63 KiB | 0 ms | 9 ms | 94 bytes | 3.18 KiB | 3 ms |

Task Outliers for Stage 3, part 2

| Task ID | Executor ID | Executor CPU Time | Shuffle Records Written | Executor Deserialize Time | Task Run Time | Aggregate Time Total | Duration Total | Number Of Output Rows |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Stage Avg. |  | 12 ms | 1.18 | 18 ms | 54 ms | 1 ms | 16 ms | 31.59 |
| Slow Outliers Avg. |  | 23 ms | 2.67 | 32 ms | 0.1 s | 4 ms | 46 ms | 2.67 |
| 204 | driver | 23 ms | 2 | 30 ms | 0.1 s | 0 ms | 45 ms | 2 |
| 209 | driver | 29 ms | 3 | 22 ms | 0.1 s | 14 ms | 55 ms | 3 |
| 217 | driver | 32 ms | 4 | 26 ms | 0.1 s | 0 ms | 64 ms | 4 |
| 250 | driver | 29 ms | 4 | 36 ms | 97 ms | 8 ms | 40 ms | 4 |
| 303 | driver | 16 ms | 2 | 39 ms | 0.1 s | 0 ms | 57 ms | 2 |
| 342 | driver | 11 ms | 1 | 36 ms | 97 ms | 0 ms | 15 ms | 1 |



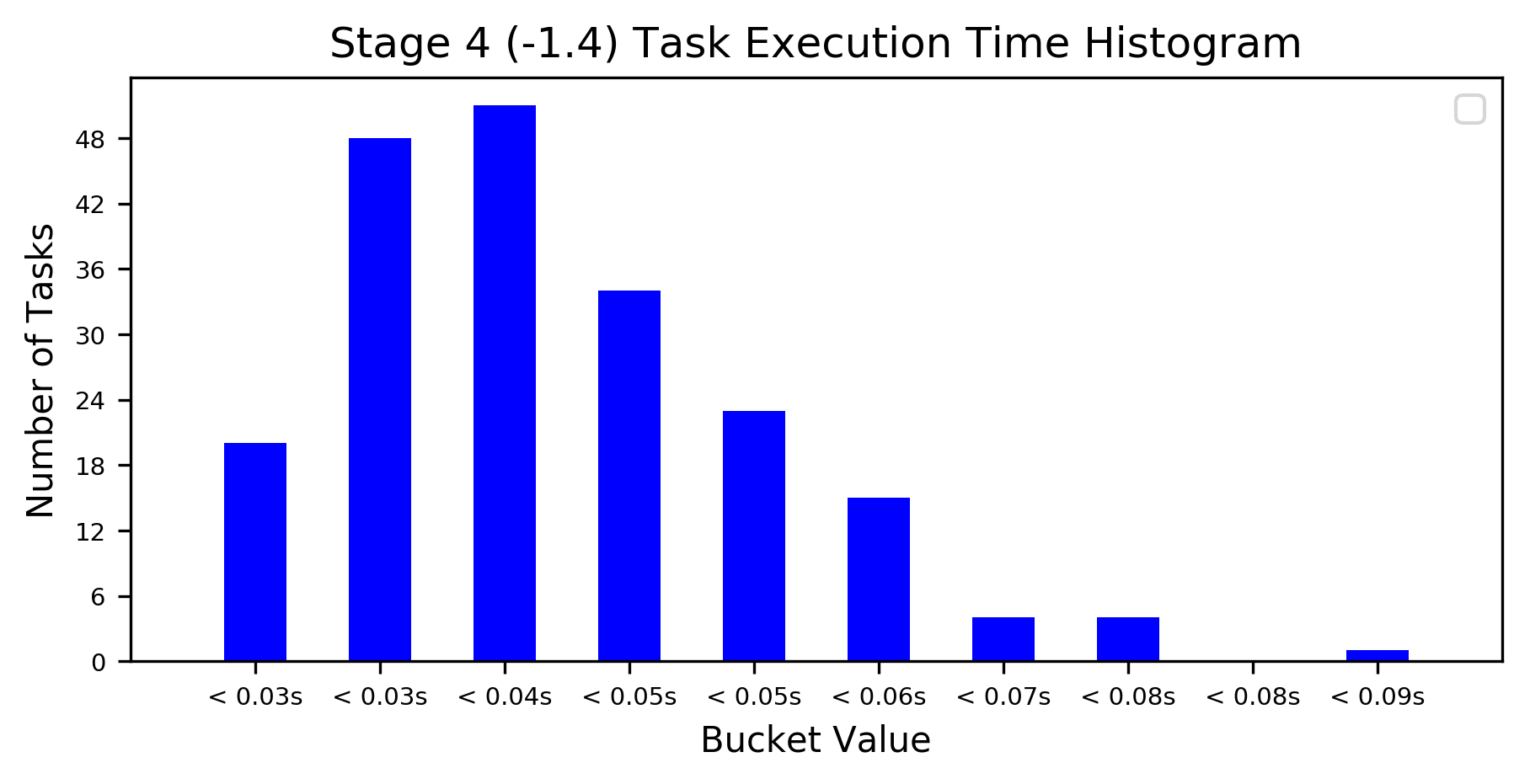


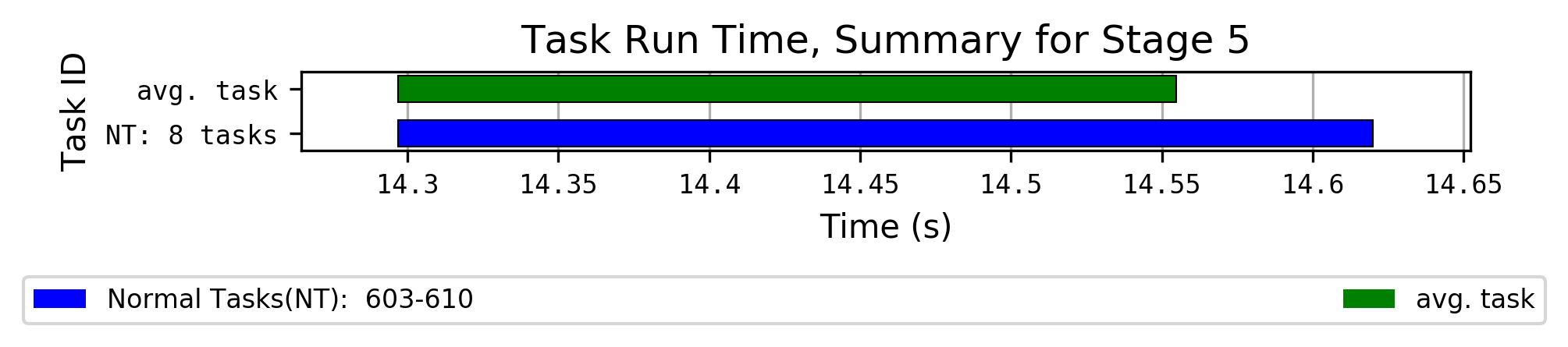
Task Outliers for Stage 4, part 1

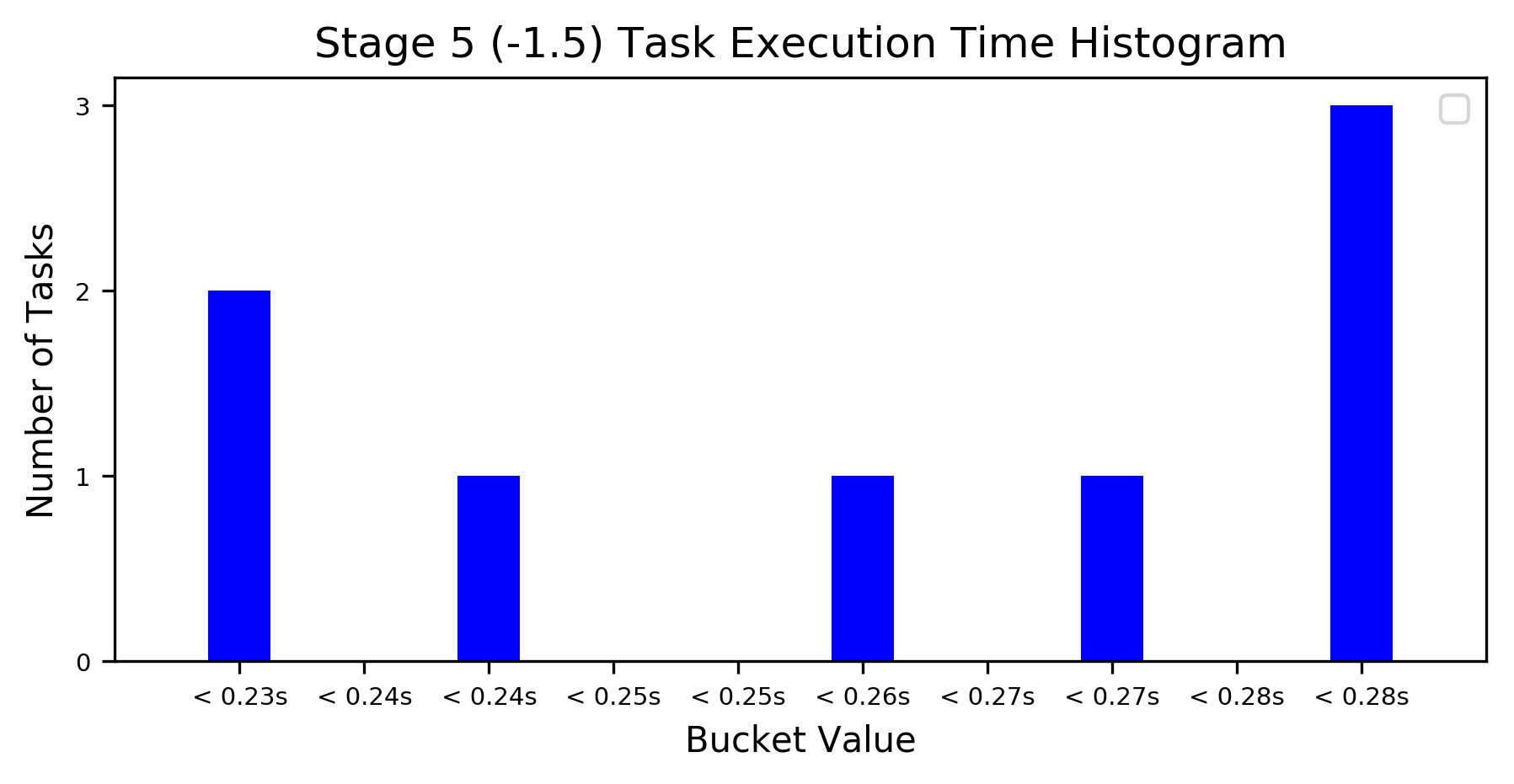
| Task ID | Executor ID | Executor Run Time | Total Records Read | Local Bytes Read | JVM GC Time | Executor Deserialize CPU Time | Result Size |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Stage Avg. |  | 18 ms | 1.18 | 112.58 bytes | 0 ms | 6 ms | 1.72 KiB |
| Slow Outliers Avg. |  | 35 ms | 1.50 | 145.62 bytes | 1 ms | 7 ms | 1.76 KiB |
| 403 | driver | 37 ms | 5 | 497 bytes | 0 ms | 6 ms | 2.06 KiB |
| 406 | driver | 43 ms | 1 | 93 bytes | 0 ms | 7 ms | 1.7 KiB |
| 407 | driver | 42 ms | 1 | 93 bytes | 0 ms | 7 ms | 1.7 KiB |
| 424 | driver | 24 ms | 1 | 100 bytes | 0 ms | 7 ms | 1.73 KiB |
| 428 | driver | 27 ms | 1 | 94 bytes | 0 ms | 7 ms | 1.7 KiB |
| 459 | driver | 39 ms | 1 | 100 bytes | 0 ms | 7 ms | 1.73 KiB |
| 547 | driver | 37 ms | 1 | 94 bytes | 11 ms | 7 ms | 1.77 KiB |
| 567 | driver | 28 ms | 1 | 94 bytes | 0 ms | 7 ms | 1.7 KiB |

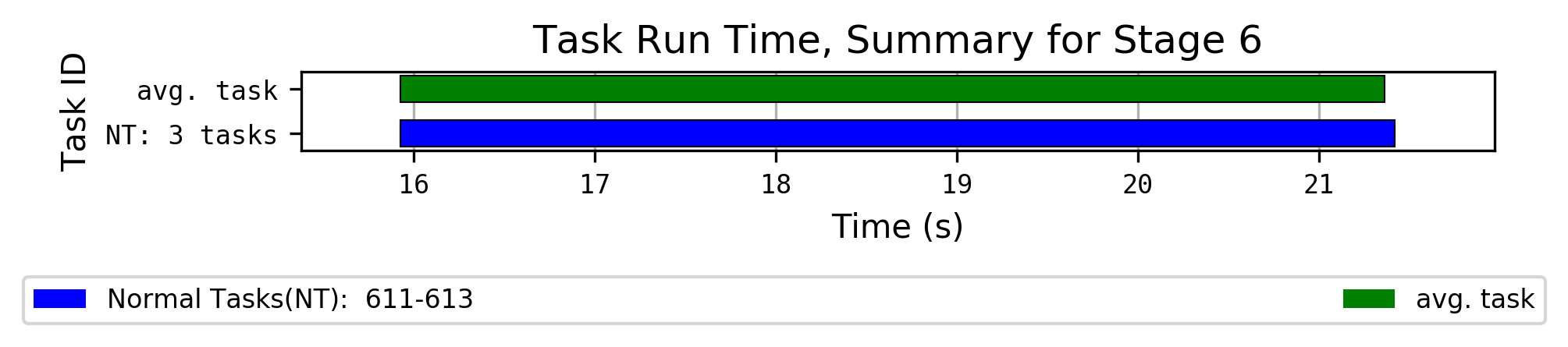
Task Outliers for Stage 4, part 2

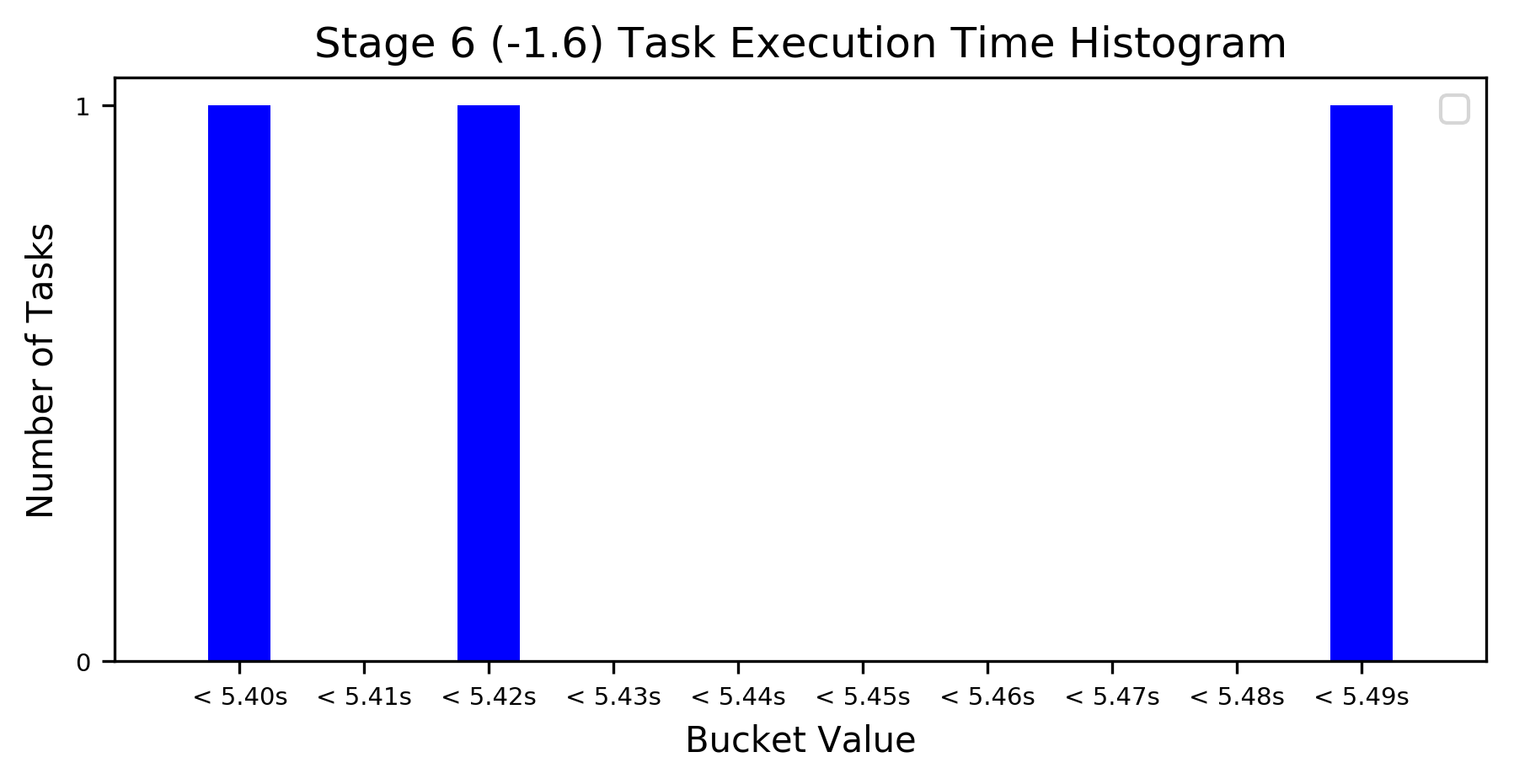
| Task ID | Executor ID | Executor CPU Time | Executor Deserialize Time | Local Blocks Fetched | Task Run Time | Peak Execution Memory |
| --- | --- | --- | --- | --- | --- | --- |
| Stage Avg. |  | 10 ms | 12 ms | 1.18 | 40 ms | 3.32 KiB |
| Slow Outliers Avg. |  | 13 ms | 18 ms | 1.50 | 70 ms | 3.46 KiB |
| 403 | driver | 22 ms | 11 ms | 5 | 65 ms | 5.01 KiB |
| 406 | driver | 15 ms | 7 ms | 1 | 70 ms | 3.23 KiB |
| 407 | driver | 21 ms | 8 ms | 1 | 71 ms | 3.23 KiB |
| 424 | driver | 11 ms | 30 ms | 1 | 65 ms | 3.24 KiB |
| 428 | driver | 10 ms | 33 ms | 1 | 69 ms | 3.23 KiB |
| 459 | driver | 10 ms | 18 ms | 1 | 69 ms | 3.24 KiB |
| 547 | driver | 10 ms | 19 ms | 1 | 89 ms | 3.23 KiB |
| 567 | driver | 9 ms | 16 ms | 1 | 64 ms | 3.23 KiB |

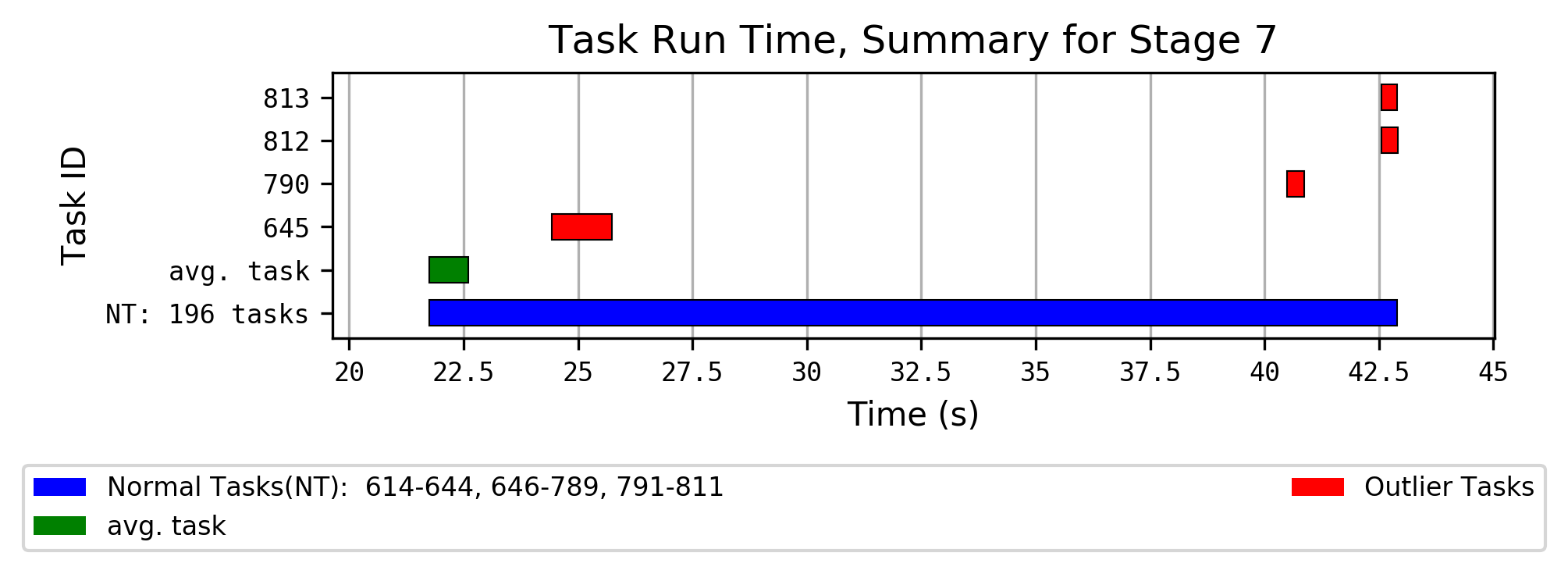












Task Outliers for Stage 7, part 1

| Task ID | Executor ID | Executor Run Time | Total Records Read | Local Bytes Read | JVM GC Time | Executor Deserialize CPU Time | Bytes Written | Result Size | Executor CPU Time | Records Written |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Stage Avg. |  | 0.3 s | 1.68 | 297.01 bytes | 0.5 s | 0.2 s | 26.35 KiB | 3.46 KiB | 33 ms | 0.51 |
| Fast Outliers Avg. |  | 0.2 s | 2.67 | 498.33 bytes | 0.2 s | 0.1 s | 32.68 KiB | 3.5 KiB | 28 ms | 1 |
| 645 | driver | 0.9 s | 0 | 0 bytes | 0.9 s | 0.1 s | 22.48 KiB | 3.42 KiB | 25 ms | 0 |
| 790 | driver | 62 ms | 2 | 383 bytes | 0.1 s | 0.2 s | 32.66 KiB | 3.5 KiB | 38 ms | 1 |
| 812 | driver | 0.2 s | 3 | 656 bytes | 0.2 s | 0.1 s | 32.73 KiB | 3.5 KiB | 24 ms | 1 |
| 813 | driver | 0.2 s | 3 | 456 bytes | 0.2 s | 0.1 s | 32.65 KiB | 3.5 KiB | 23 ms | 1 |

Task Outliers for Stage 7, part 2

| Task ID | Executor ID | Executor Deserialize Time | Local Blocks Fetched | Task Run Time | Aggregate Time Total | Duration Total | Number Of Output Rows | Peak Execution Memory | Peak Memory Total |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Stage Avg. |  | 0.5 s | 1.28 | 0.8 s | 12 ms | 0.2 s | 1.18 | 11.37 MiB | 11.37 MiB |
| Fast Outliers Avg. |  | 0.2 s | 2.33 | 0.4 s | 1 ms | 81 ms | 1.67 | 16.25 MiB | 16.25 MiB |
| 645 | driver | 0.4 s | 0 | 1.3 s | 0 ms | 0.1 s | 0 | 256 KiB | 256 KiB |
| 790 | driver | 0.3 s | 2 | 0.4 s | 1 ms | 41 ms | 2 | 16.25 MiB | 16.25 MiB |
| 812 | driver | 0.2 s | 3 | 0.4 s | 1 ms | 0.2 s | 1 | 16.25 MiB | 16.25 MiB |
| 813 | driver | 0.1 s | 2 | 0.3 s | 0 ms | 13 ms | 2 | 16.25 MiB | 16.25 MiB |

