# Boxplot & Frequency Results

TracerBench on HeadlessChrome/120.0.6099.109
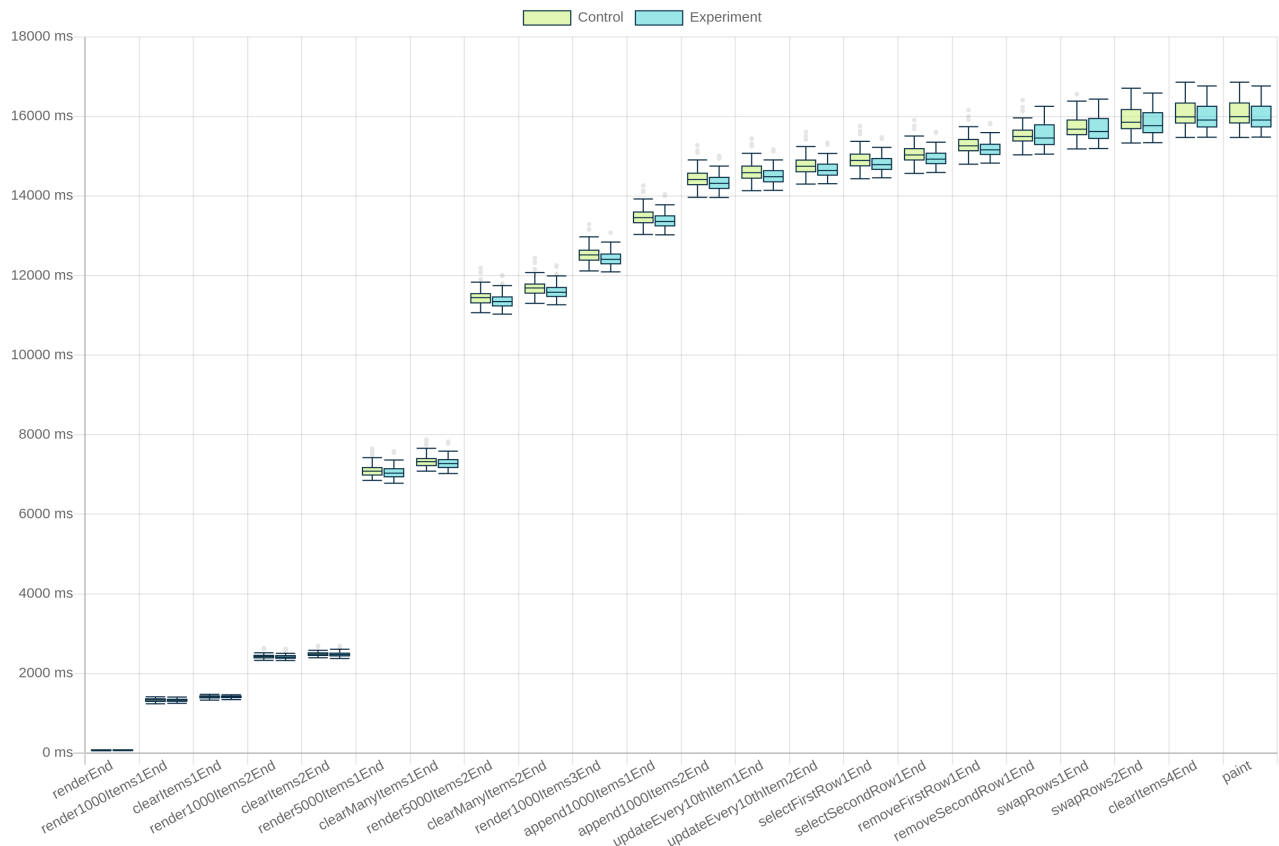
duration (90 ms faster)

Based on the P-value of this benchmark the evidence for a metric shift is **very strong.** TracerBench has determined the results are **significant** meaning they are worth looking at. A statistics estimator *(Hodges–Lehmann estimator)* was used to determine "Experiment" is **faster** by **90 ms**. TracerBench is 95% confident "Experiment" is **faster** between **7 ms to 173 ms** based on 100 samples using a *(confidence interval)*.

duration (lower is better)
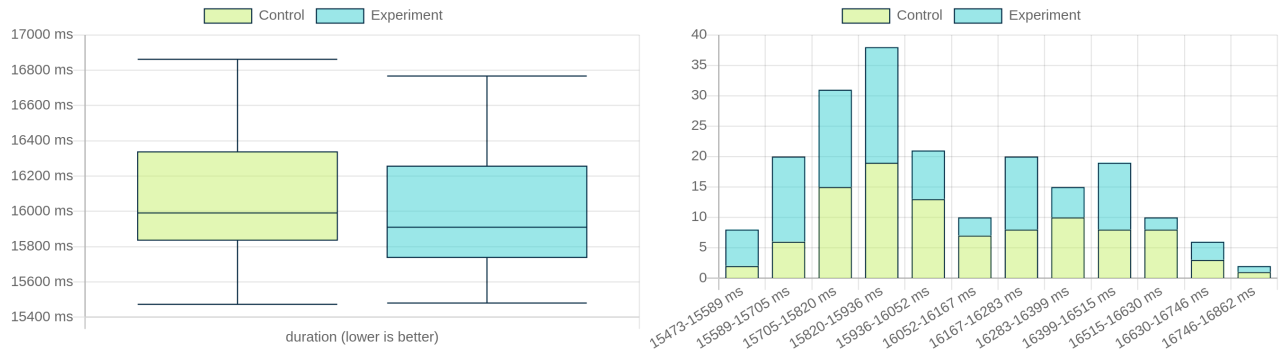
Cumulative sub-phases of duration

The chart below shows the finish times (a point in the page load duration) of the sub-phases for experiment and control. It gives a high level view on what changed (if any).
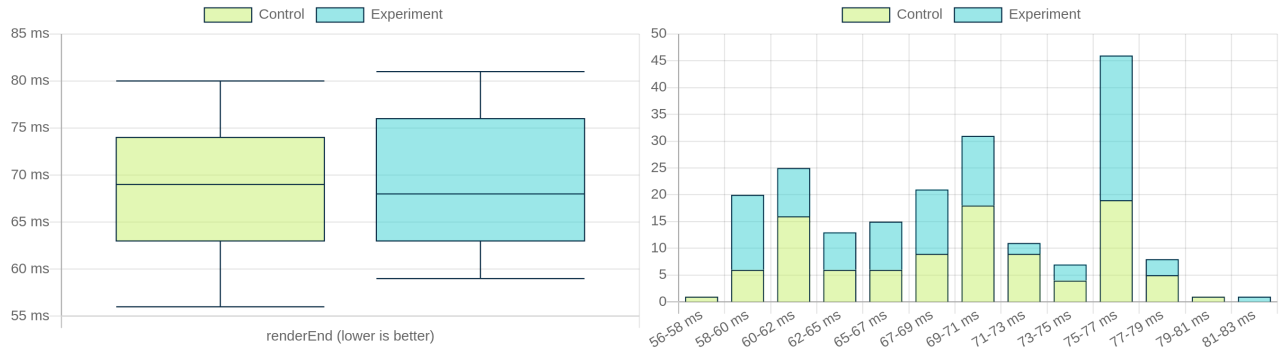
You can view more details about the sub-phases in the section below "Isolated sub-phases of duration".
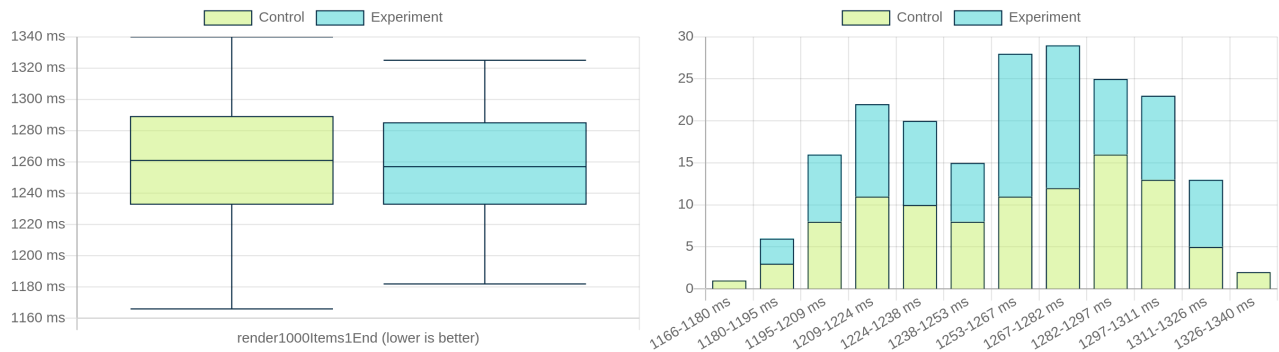
Isolated sub-phases of duration

renderEnd (No/Borderline Difference)

Based on the P-value of this benchmark the evidence for a metric shift is **weak.** TracerBench has determined the results are **not significant**.
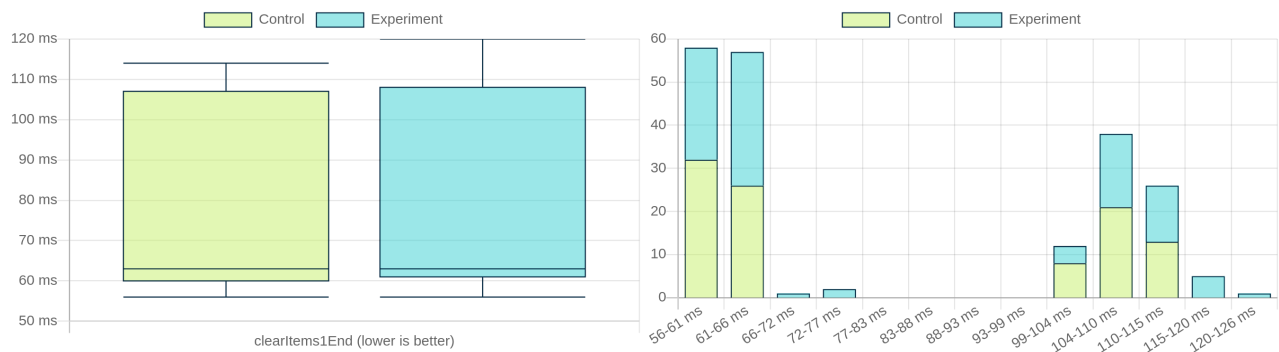

renderEnd (lower is better)

render1000Items1End (No/Borderline Difference)

Based on the P-value of this benchmark the evidence for a metric shift is **weak.** TracerBench has determined the results are **not significant**.

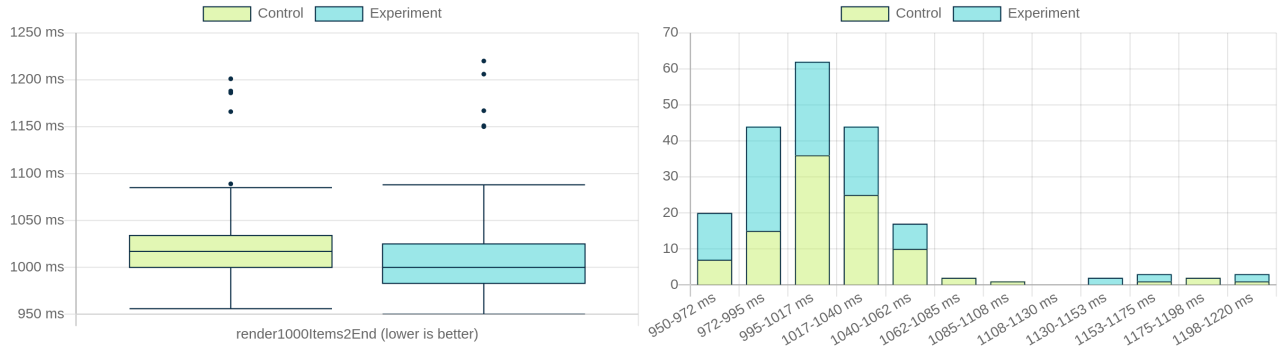render1000Items1End (lower is better)

clearItems1End (No/Borderline Difference)

Based on the P-value of this benchmark the evidence for a metric shift is **weak.** TracerBench has determined the results are **not significant**.
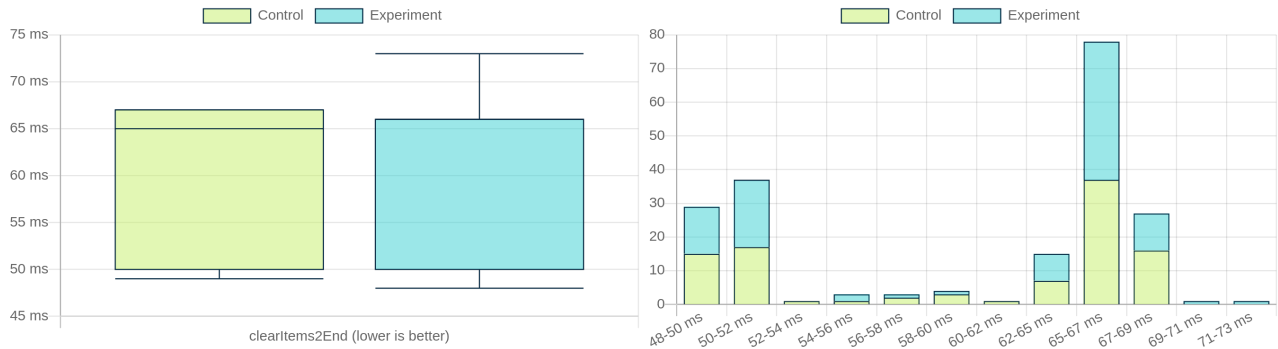

clearItems1End (lower is better)

render1000Items2End (12 ms faster)

Based on the P-value of this benchmark the evidence for a metric shift is **very strong.** TracerBench has determined the results are **significant** meaning they are worth looking at. A statistics estimator *(Hodges–Lehmann estimator)* was used to determine "Experiment" is **faster** by **12 ms**. TracerBench is 95% confident "Experiment" is **faster** between **1 ms to 18 ms** based on 100 samples using a *(confidence interval)*.
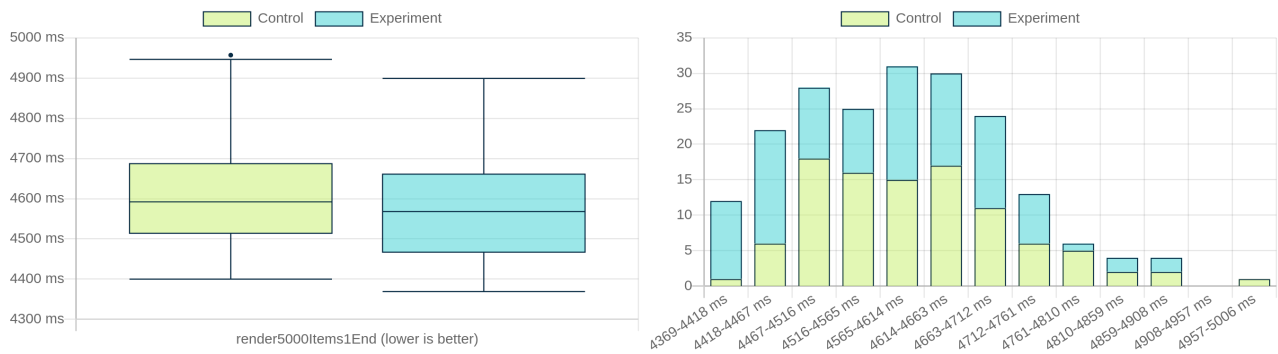

render1000Items2End (lower is better)

clearItems2End (No/Borderline Difference)

Based on the P-value of this benchmark the evidence for a metric shift is **weak.** TracerBench has determined the results are **not significant**.

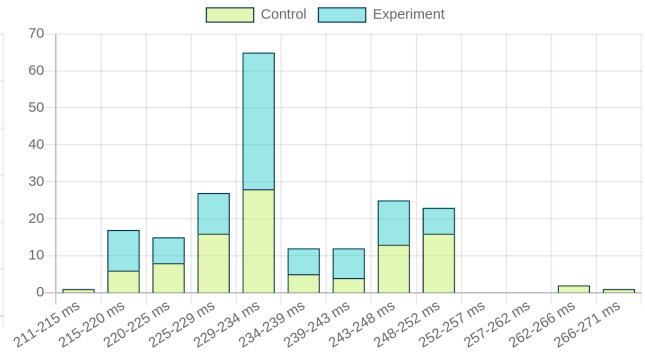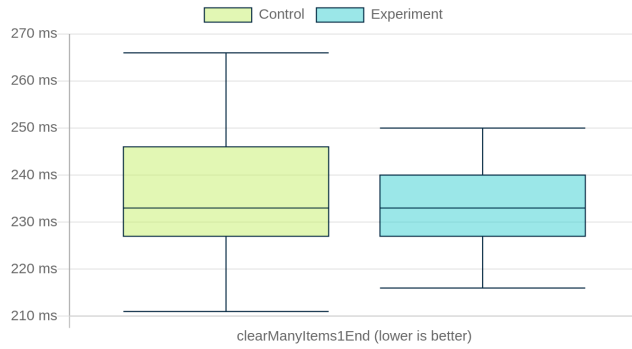
clearItems2End (lower is better)

render5000Items1End (No/Borderline Difference)

Based on the P-value of this benchmark the evidence for a metric shift is **strong.** TracerBench has determined the results are **not significant**.

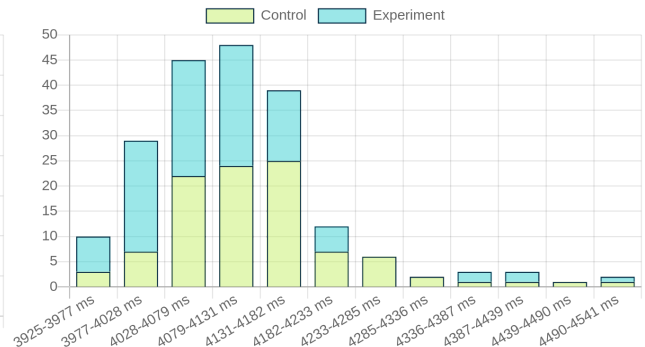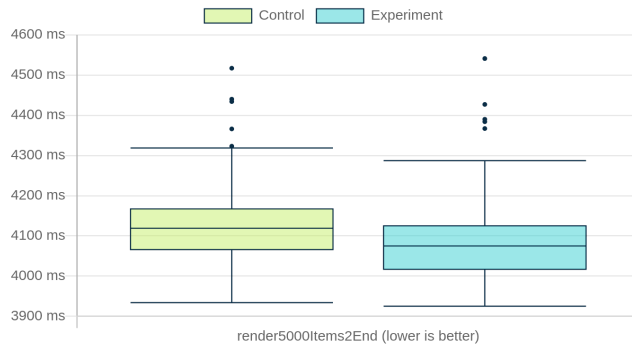render5000Items1End (lower is better)

clearManyItems1End (No/Borderline Difference)

Based on the P-value of this benchmark the evidence for a metric shift is **weak.** TracerBench has determined the results are **not significant**.
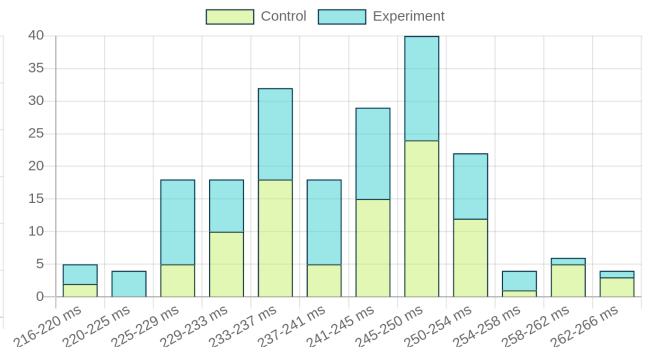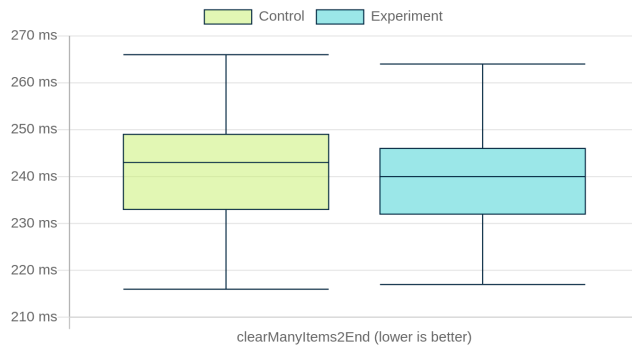

render5000Items2End (44 ms faster)

Based on the P-value of this benchmark the evidence for a metric shift is **very strong.** TracerBench has determined the results are **significant** meaning they are worth looking at. A statistics estimator *(Hodges–Lehmann estimator)* was used to determine "Experiment" is **faster** by **44 ms**. TracerBench is 95% confident "Experiment" is **faster** between **22 ms to 66 ms** based on 100 samples using a *(confidence interval)*.
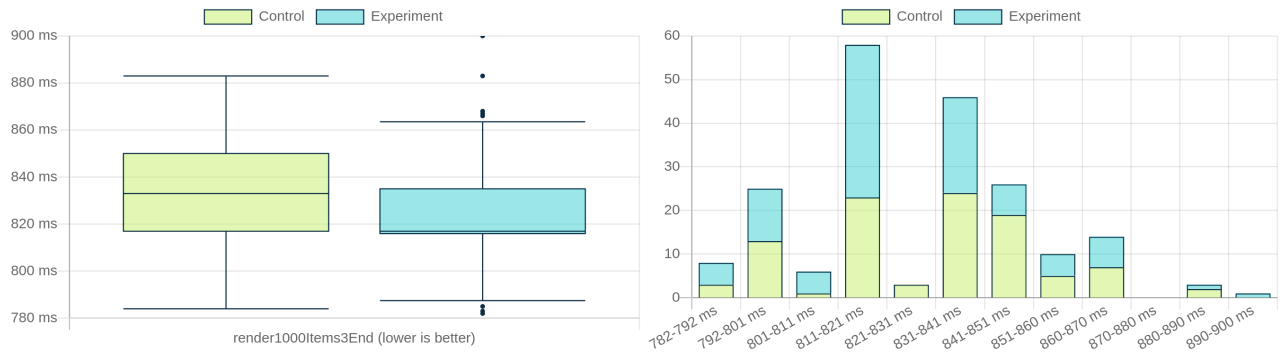

clearManyItems2End (3 ms faster)

Based on the P-value of this benchmark the evidence for a metric shift is **very strong.** TracerBench has determined the results are **significant** meaning they are worth looking at. A statistics estimator *(Hodges–Lehmann estimator)* was used to determine "Experiment" is **faster** by **3 ms**. TracerBench is 95% confident "Experiment" is **faster** between **0 ms to 6 ms** based on 100 samples using a *(confidence interval)*.

render1000Items3End (No/Borderline Difference)

Based on the P-value of this benchmark the evidence for a metric shift is **strong.** TracerBench has determined the results are **not significant**.

render1000Items3End (lower is better)

append1000Items1End (No/Borderline Difference)

Based on the P-value of this benchmark the evidence for a metric shift is **weak.** TracerBench has determined the results are **not significant**.
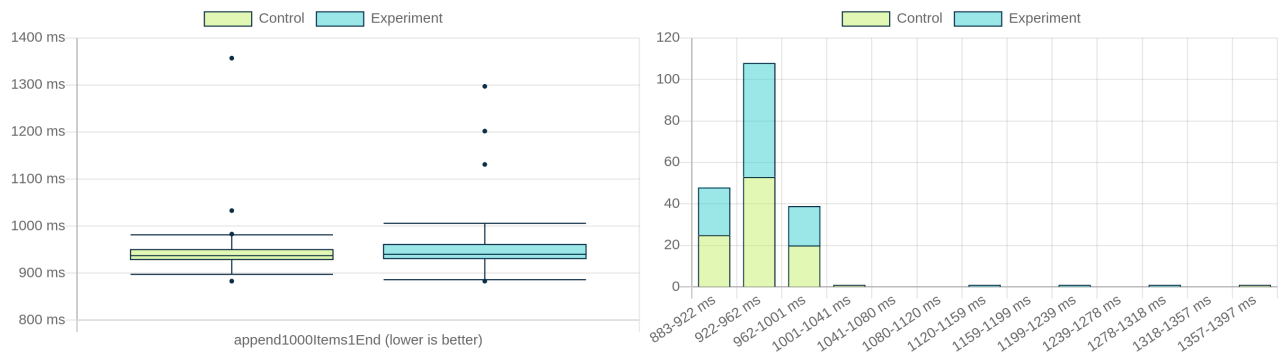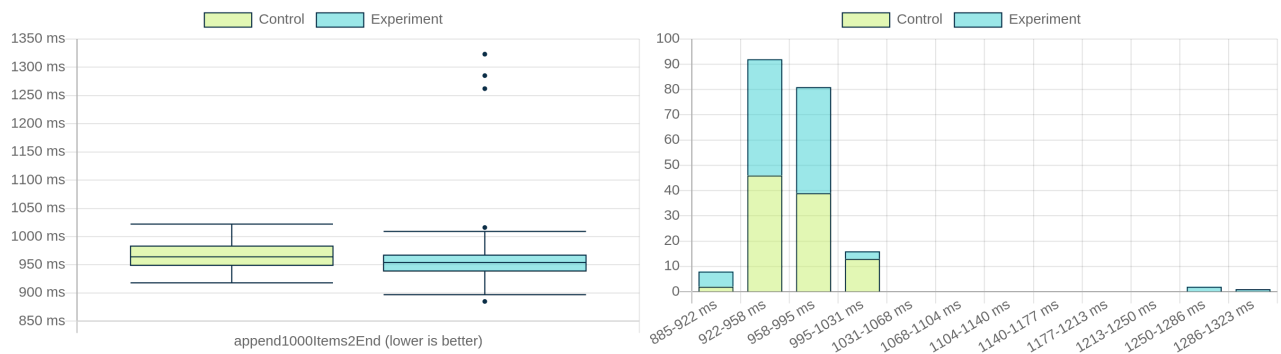

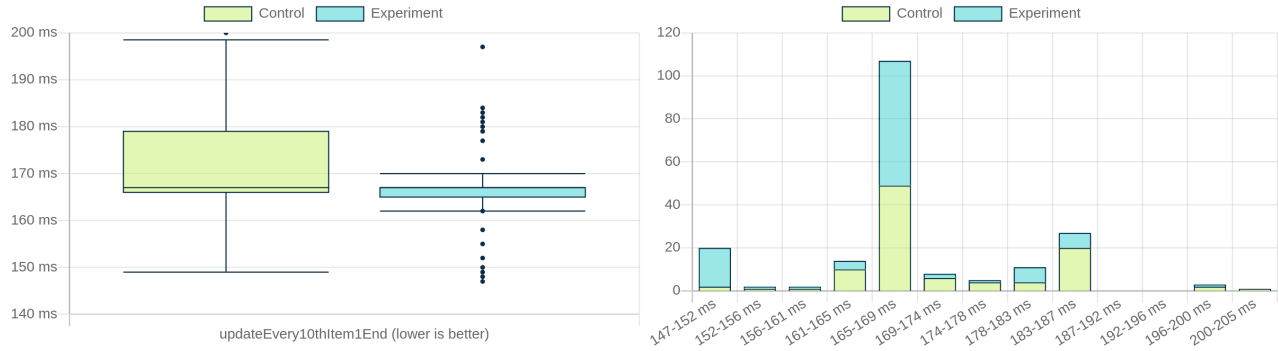append1000Items1End (lower is better)

append1000Items2End (No/Borderline Difference)

Based on the P-value of this benchmark the evidence for a metric shift is **strong.** TracerBench has determined the results are **not significant**.

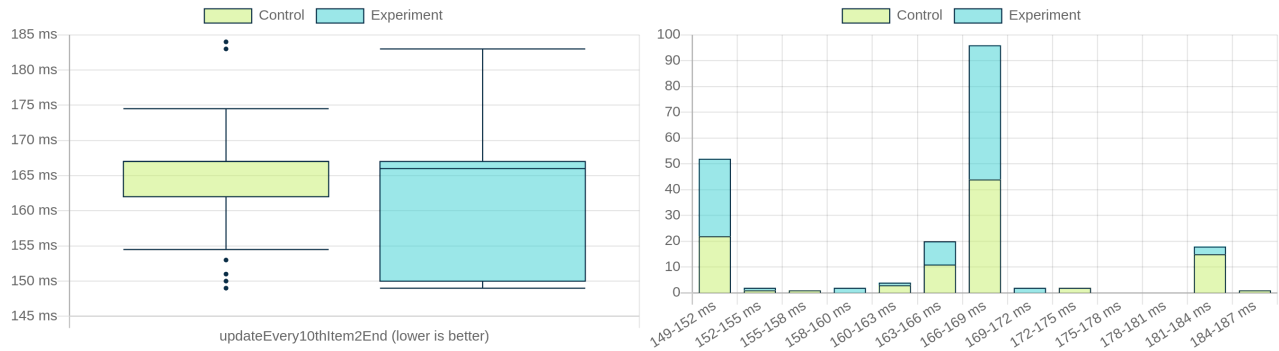append1000Items2End (lower is better)

updateEvery10thItem1End (1 ms faster)

Based on the P-value of this benchmark the evidence for a metric shift is **very strong.** TracerBench has determined the results are **significant** meaning they are worth looking at. A statistics estimator *(Hodges–Lehmann estimator)* was used to determine "Experiment" is **faster** by **1 ms**. TracerBench is 95% confident "Experiment" is **faster** between **0 ms to 5 ms** based on 100 samples using a *(confidence interval)*.
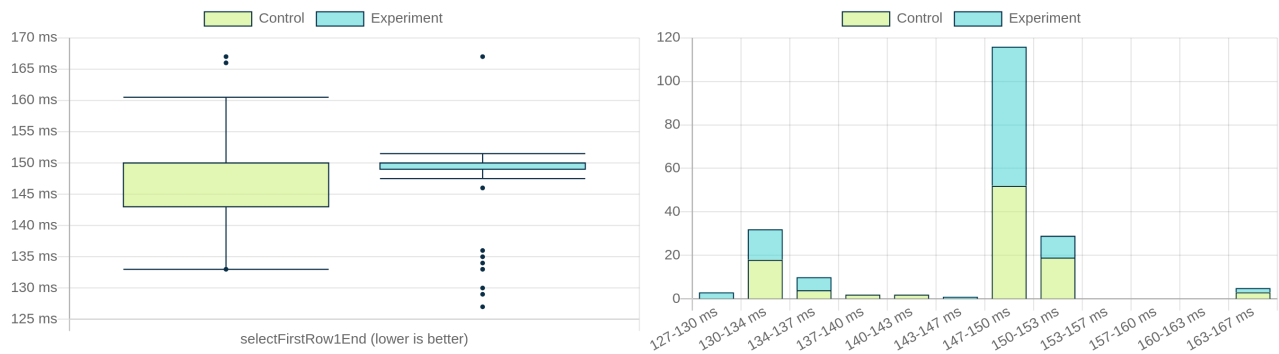

updateEvery10thItem2End (No/Borderline Difference)

Based on the P-value of this benchmark the evidence for a metric shift is **strong.** TracerBench has determined the results are **not significant**.
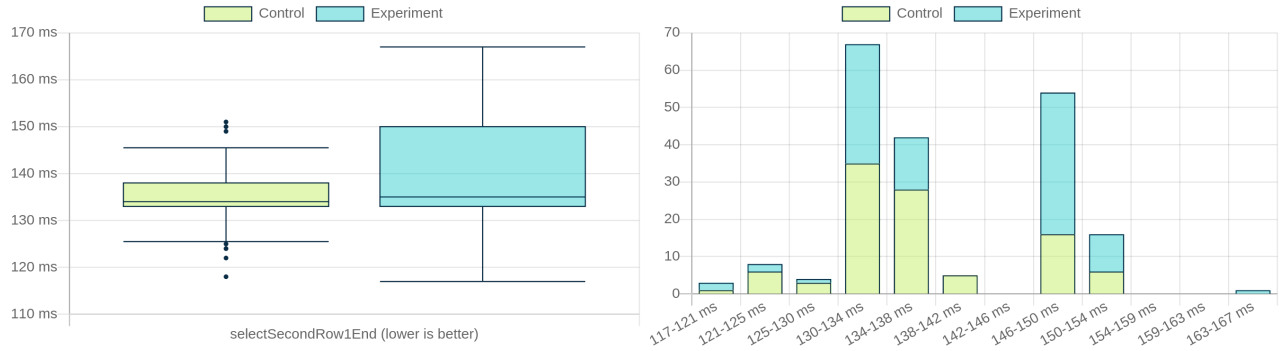

selectFirstRow1End (No/Borderline Difference)

Based on the P-value of this benchmark the evidence for a metric shift is **weak.** TracerBench has determined the results are **not significant**.

selectSecondRow1End (1 ms slower)

Based on the P-value of this benchmark the evidence for a metric shift is **very strong.** TracerBench has determined the results are **significant** meaning they are worth looking at. A statistics estimator *(Hodges–Lehmann estimator)* was used to determine "Experiment" is **slower** by **1 ms**. TracerBench is 95% confident "Experiment" is **slower** between **0 ms to 6 ms** based on 100 samples using a *(confidence interval)*.

removeFirstRow1End (No/Borderline Difference)

Based on the P-value of this benchmark the evidence for a metric shift is **weak.** TracerBench has determined the results are **not significant**.
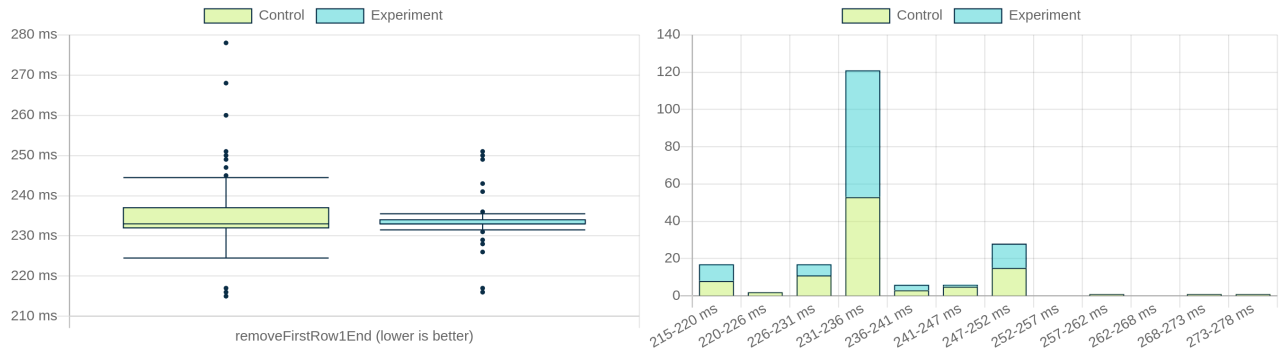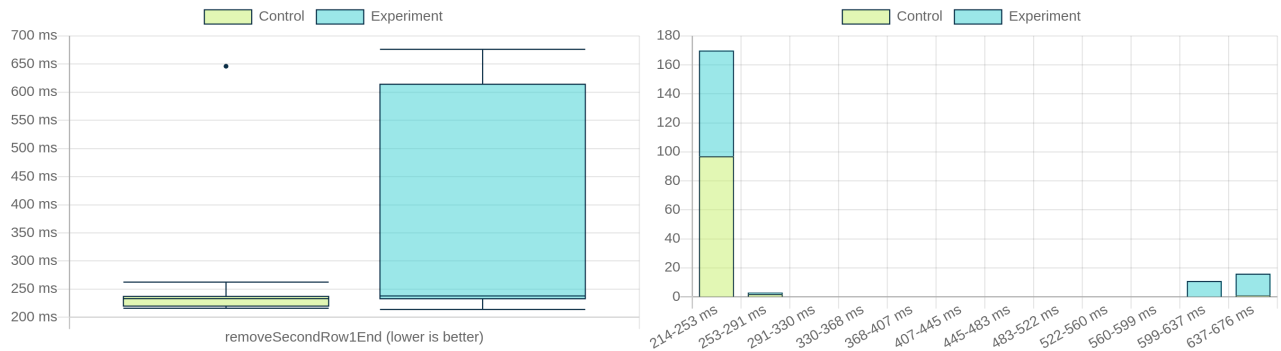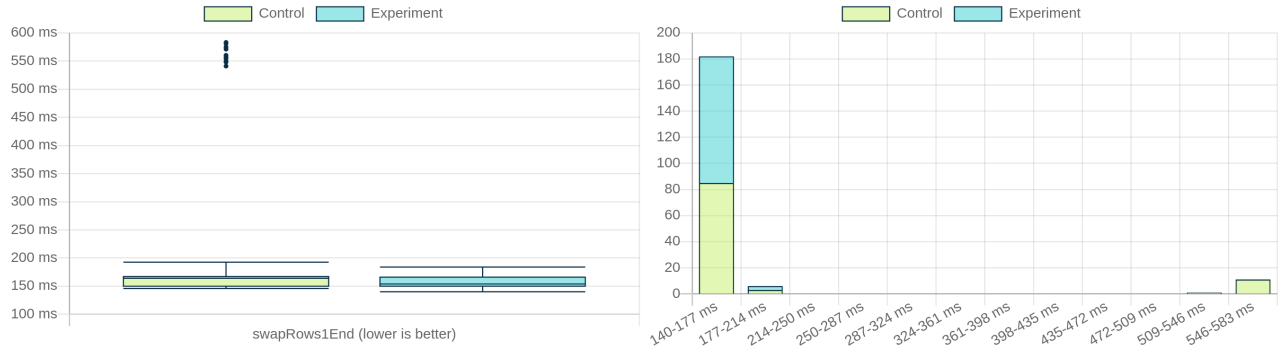

removeSecondRow1End (14 ms slower)

Based on the P-value of this benchmark the evidence for a metric shift is **very strong.** TracerBench has determined the results are **significant** meaning they are worth looking at. A statistics estimator *(Hodges–Lehmann estimator)* was used to determine "Experiment" is **slower** by **14 ms**. TracerBench is 95% confident "Experiment" is **slower** between **7 ms to 16 ms** based on 100 samples using a *(confidence interval)*.
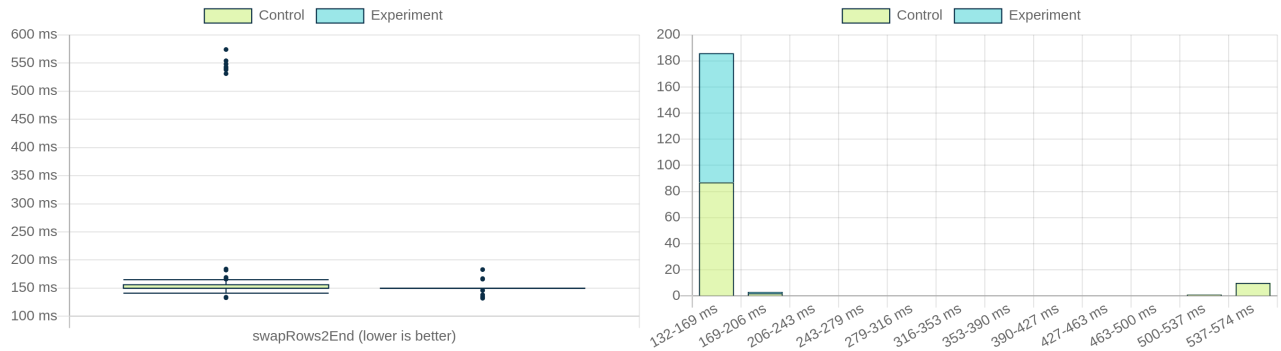
swapRows1End (3 ms faster)

Based on the P-value of this benchmark the evidence for a metric shift is **very strong.** TracerBench has determined the results are **significant** meaning they are worth looking at. A statistics estimator *(Hodges–Lehmann estimator)* was used to determine "Experiment" is **faster** by **3 ms**. TracerBench is 95% confident "Experiment" is **faster** between **0 ms to 7 ms** based on 100 samples using a *(confidence interval)*.
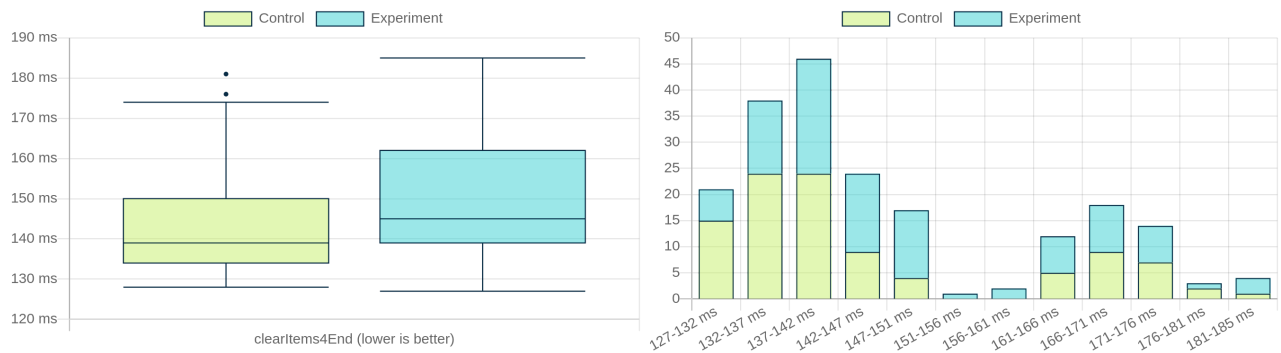

swapRows2End (No/Borderline Difference)

Based on the P-value of this benchmark the evidence for a metric shift is **very strong.** TracerBench has determined the results are **not significant**.


clearItems4End (4 ms slower)

Based on the P-value of this benchmark the evidence for a metric shift is **very strong.** TracerBench has determined the results are **significant** meaning they are worth looking at. A statistics estimator *(Hodges–Lehmann estimator)* was used to determine "Experiment" is **slower** by **4 ms**. TracerBench is 95% confident "Experiment" is **slower** between **1 ms to 7 ms** based on 100 samples using a *(confidence interval)*.

paint (No/Borderline Difference)

Based on the P-value of this benchmark the evidence for a metric shift is **weak.** TracerBench has determined the results are **not significant**.

paint (lower is better)
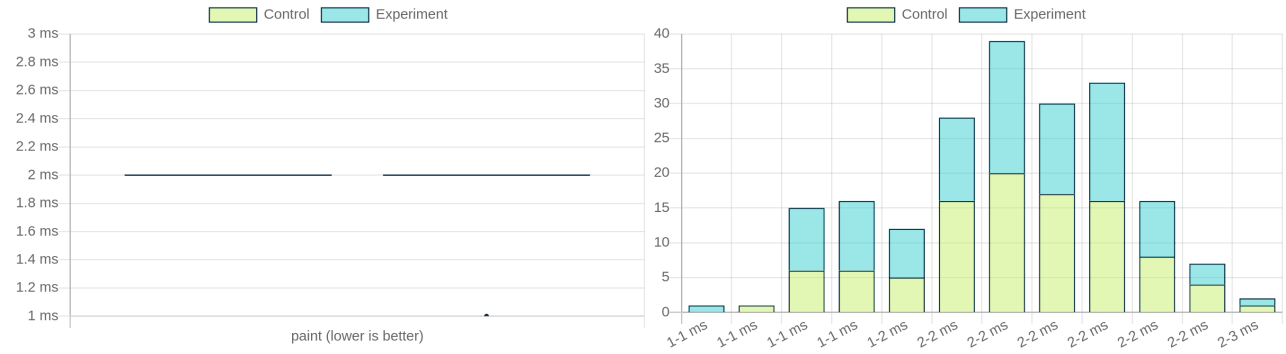
Resources

- [Stats Primer](#)
- [Understanding Boxplots](#)
- [Wilcoxon Rank-Sum Test](#)

Configs Used

```
    {
  "tbResultsFolder": "/home/runner/work/glimmer-vm/glimmer-vm/tracerbench-results",
  "config": "undefined",
  "isCIEnv": false,
  "plotTitle": "TracerBench"
}
```