Vela: PGBench Results

All tests are run with PostgreSQL 17.5.

Test Setup

Storage Setup:

- 1 physical storage host with 2 NUMA nodes
- 2x Intel Xeon 8175 (24 cores each) at 3.9GHz
- 4 storage nodes per physical host (two per NUMA node)
- 2 NVMe devices per storage node (8 per storage host)

PostgreSQL Setup:

- 4 Vela Logical Volumes striped per LVM
- Shared Buffers configuration: 16MB
- Maximum Workers: 16

It is important to note that this PostgreSQL configuration was explicitly limited to a bare minimum of cache and RAM usage to increase disk IO (especially in read use cases) as much as possible. These benchmarks are storage benchmarks. They do not represent the theoretically achievable maximum TPS numbers using PostgreSQL and Vela.

The client running pgbench was on the same machine as the PostgreSQL server instance to reduce the latency introduced between the two. Pgbench used TCP connections, not Unix Domain Sockets.

Each benchmark iteration ran for 30 minutes, with 3 iterations for each benchmark.

Transactions Per Second (TPS)

This graph shows the number of successful transactions completed per second under the TPC-B workload. Higher TPS indicates better throughput and responsiveness under concurrent transactional load.

Transaction Latency

This section presents the average latency per transaction during the TPC-B test. Lower latency means faster response time per transaction, which is crucial for applications that require quick interactions.

Executive Summary

This report presents the results of a comprehensive performance evaluation of Vela's storage software using the pgbench benchmark tool with PostgreSQL 17.5. The benchmarks assess transaction throughput and latency across three core OLTP scenarios: TPC-B (general banking-like workloads), Simple Update (write-heavy transactions), and Select (read-heavy operations).

Key Results:

1. TPC-B (Mixed Workload)
    - Achieved up to about 18,999 TPS at 100 concurrent clients.
    - Latency ranged from 3ms (50 clients) to 12ms (200 clients).
    - Performance remained stable across client loads with moderate variability.

2. Simple Update (Write-Intensive)
    - Peak throughput of over 23,000 TPS at 200 clients.
    - Maintained strong linear scaling with increasing clients.
    - Latency ranged from 2.8ms to 9ms, reflecting robust write handling.

3. Select (Read-Only)
    - Exceptional throughput: up to 287,553 TPS at 200 clients.
    - Disk IOPS utilization capped around 400K; read operations did not saturate disk bandwidth.
    - Latency stabilized at ~500ms, limited by client think time and system clock, not I/O.

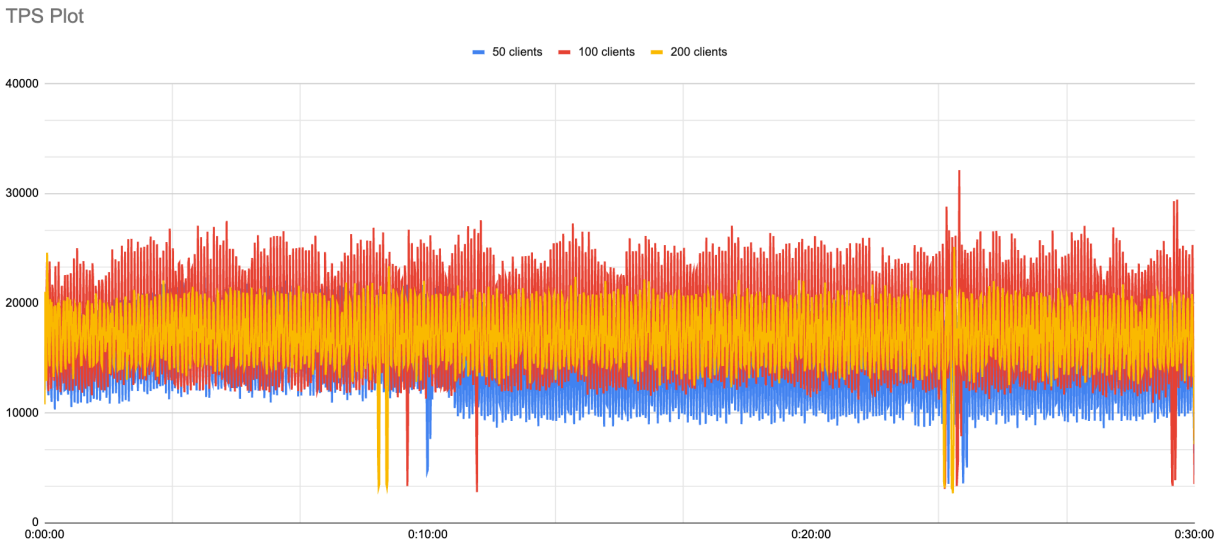
Vela demonstrates high-performance capabilities for PostgreSQL workloads, especially under read-intensive and write-heavy conditions. The system exhibits excellent scalability and low latency, making it well-suited for modern OLTP applications. Notably, even with minimal shared buffers (16MB), Vela maintains consistent high throughput and responsiveness, underlining its suitability for high-demand, cloud-native deployments.

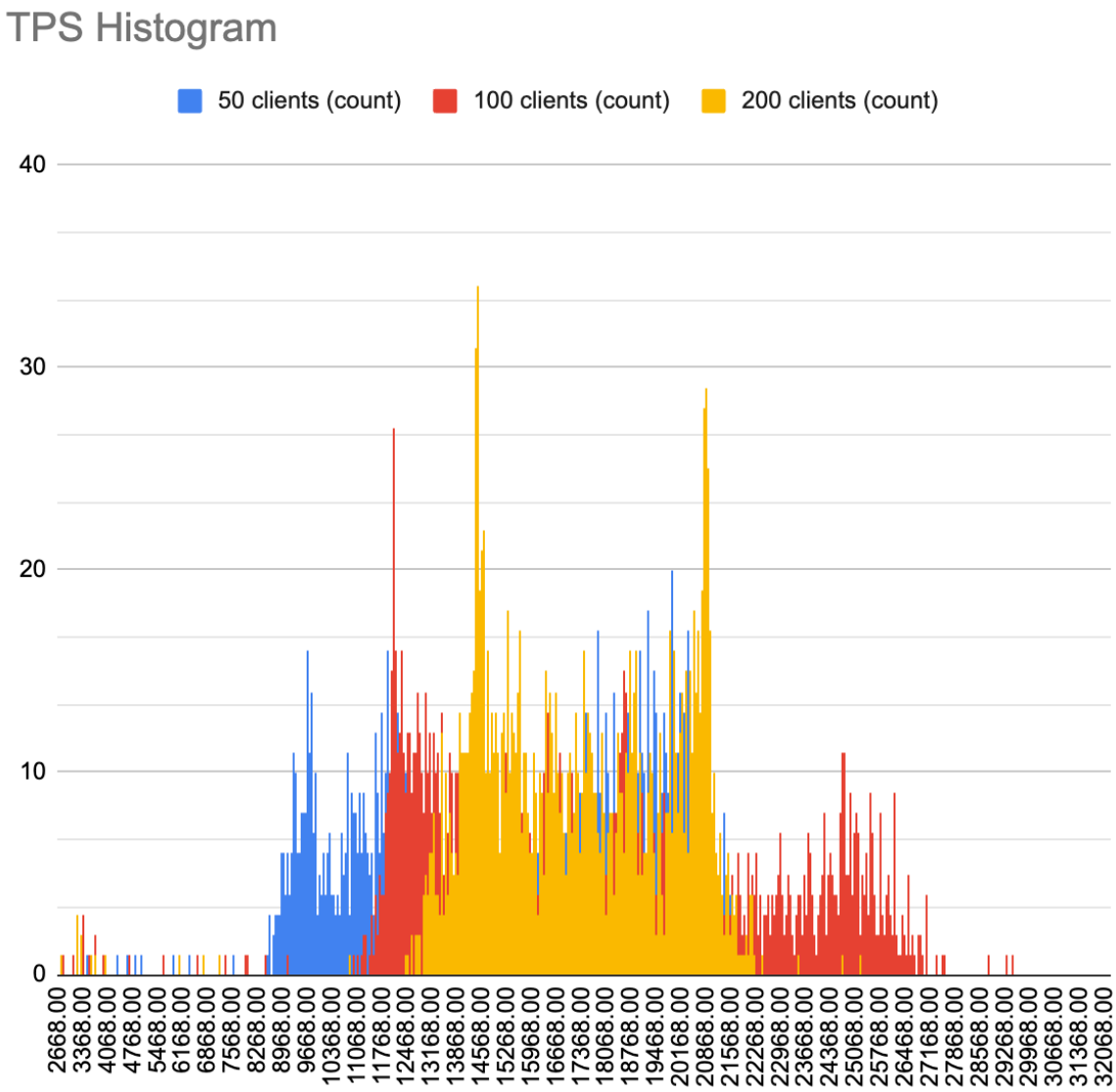
TPC-B Benchmark

The TPC-B benchmark simulates a banking workload involving multiple small transactions that include reading and writing data across several tables. It is a classic standard for evaluating OLTP (Online Transaction Processing) performance.

Transactions Per Second

Clients	Average	Median	StdDev
50	15504	15948	3902
100	17673	17171	4650
200	17138	17089	2779

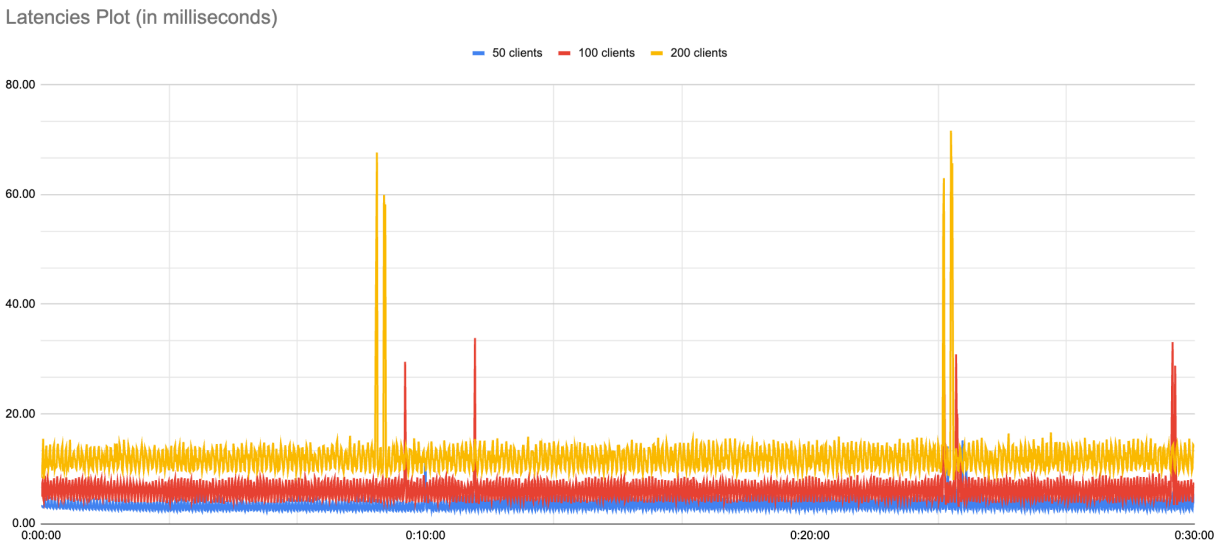


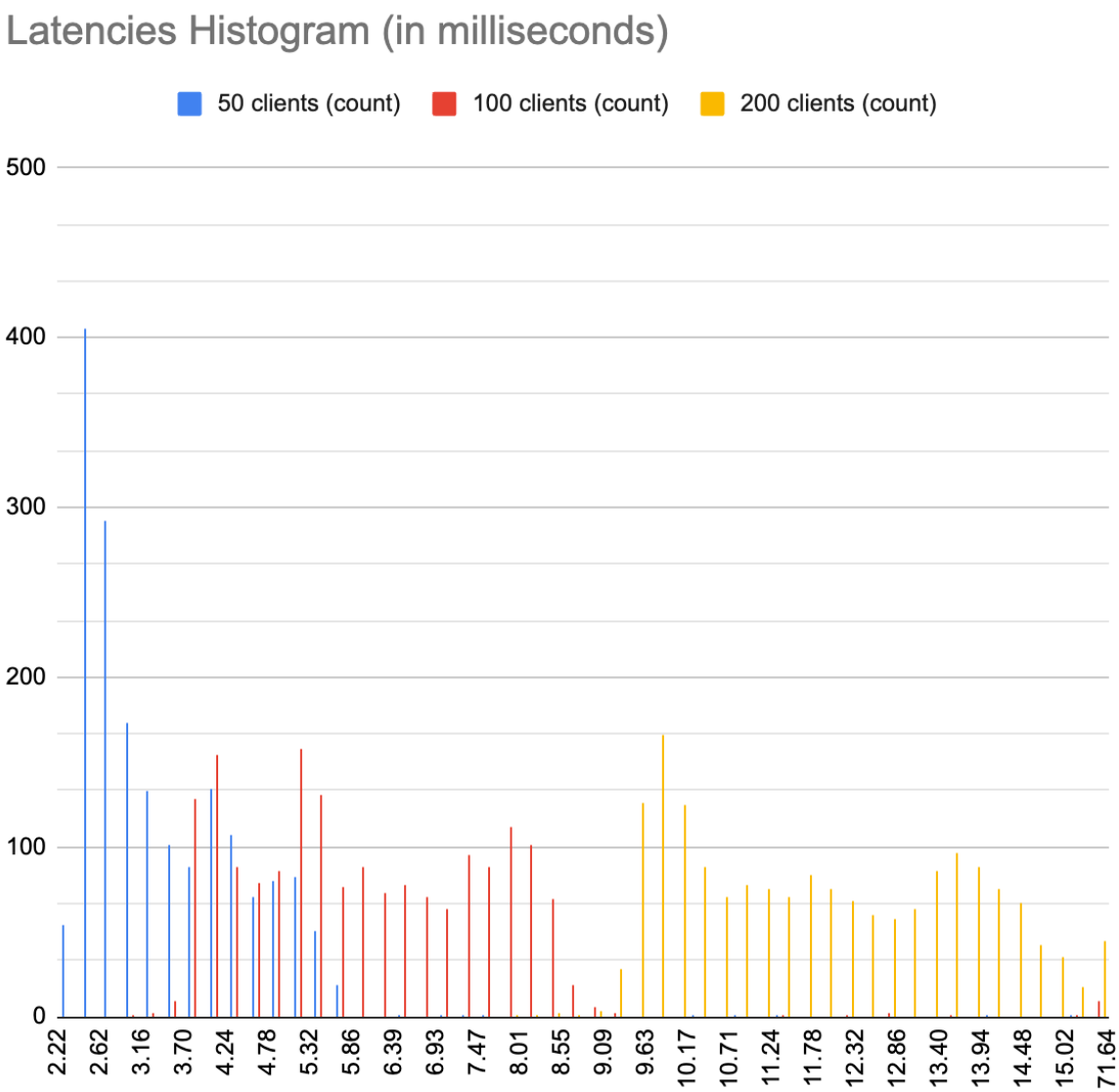

TPS Histogram
50 clients (count)
100 clients (count)
200 clients (count)
40
30
20
10
0
2668.00
3368.00
4068.00
4768.00
5468.00
6168.00
6868.00
7568.00
8268.00
8968.00
9668.00
10368.00
11068.00
11768.00
12468.00
13168.00
13868.00
14568.00
15268.00
15968.00
16668.00
17368.00
18068.00
18768.00
19468.00
20168.00
20868.00
21568.00
22268.00
22968.00
23668.00
24368.00
25068.00
25768.00
26468.00
27168.00
27868.00
28568.00
29268.00
29968.00
30668.00
31368.00
32068.00

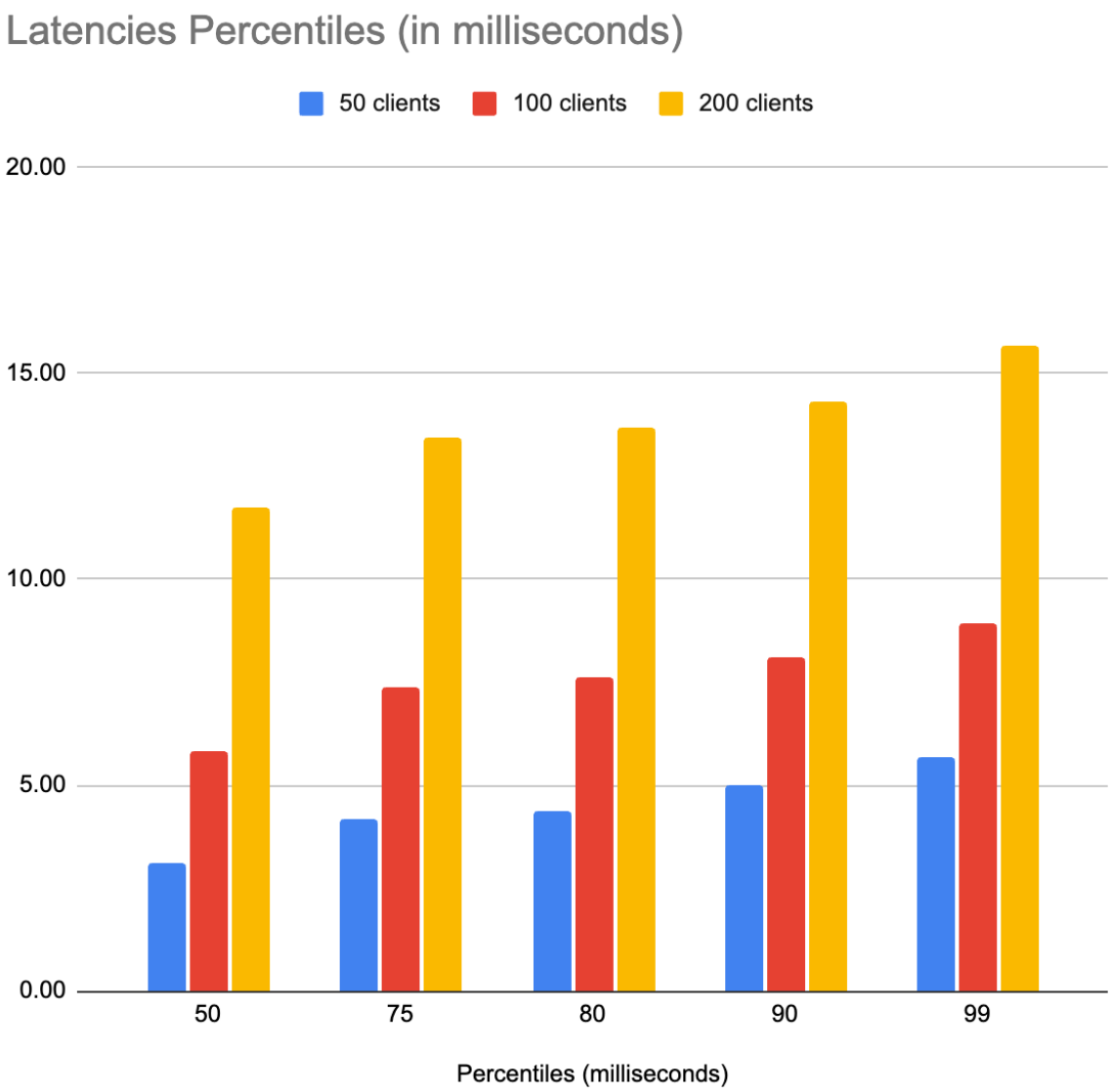
Transaction Latency

Clients	Average	Median	StdDev
50	3.468	3.124	1.072
100	6.123	5.814	2.253
200	12.15	11.726	4.137

Clients	50th	75th	80th	90th	99th
50	3.12	4.18	4.38	5.01	5.67
100	5.81	7.39	7.63	8.11	8.94
200	11.73	13.43	13.69	14.29	15.66




Latencies Histogram (in milliseconds)
50 clients (count)
100 clients (count)
200 clients (count)
500
400
300
200
100
0
2.22
2.62
3.16
3.70
4.24
4.78
5.32
5.86
6.39
6.93
7.47
8.01
8.55
9.09
9.63
10.17
10.71
11.24
11.78
12.32
12.86
13.40
13.94
14.48
15.02
71.64

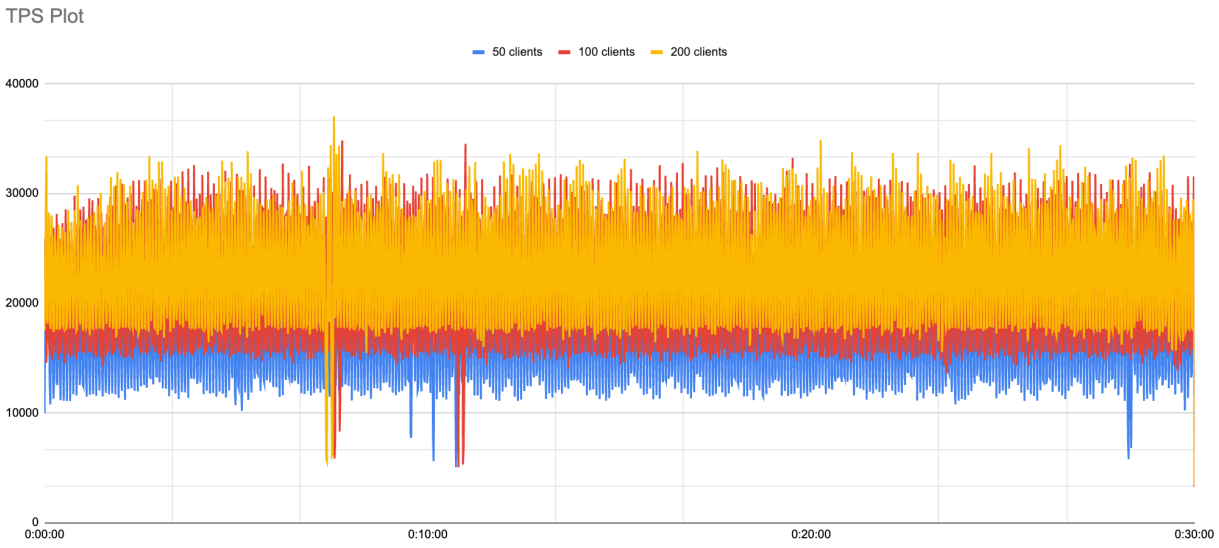

Latencies Percentiles (in milliseconds)
50 clients
100 clients
200 clients
20.00
15.00
10.00
5.00
0.00
50
75
80
90
99
Percentiles (milliseconds)

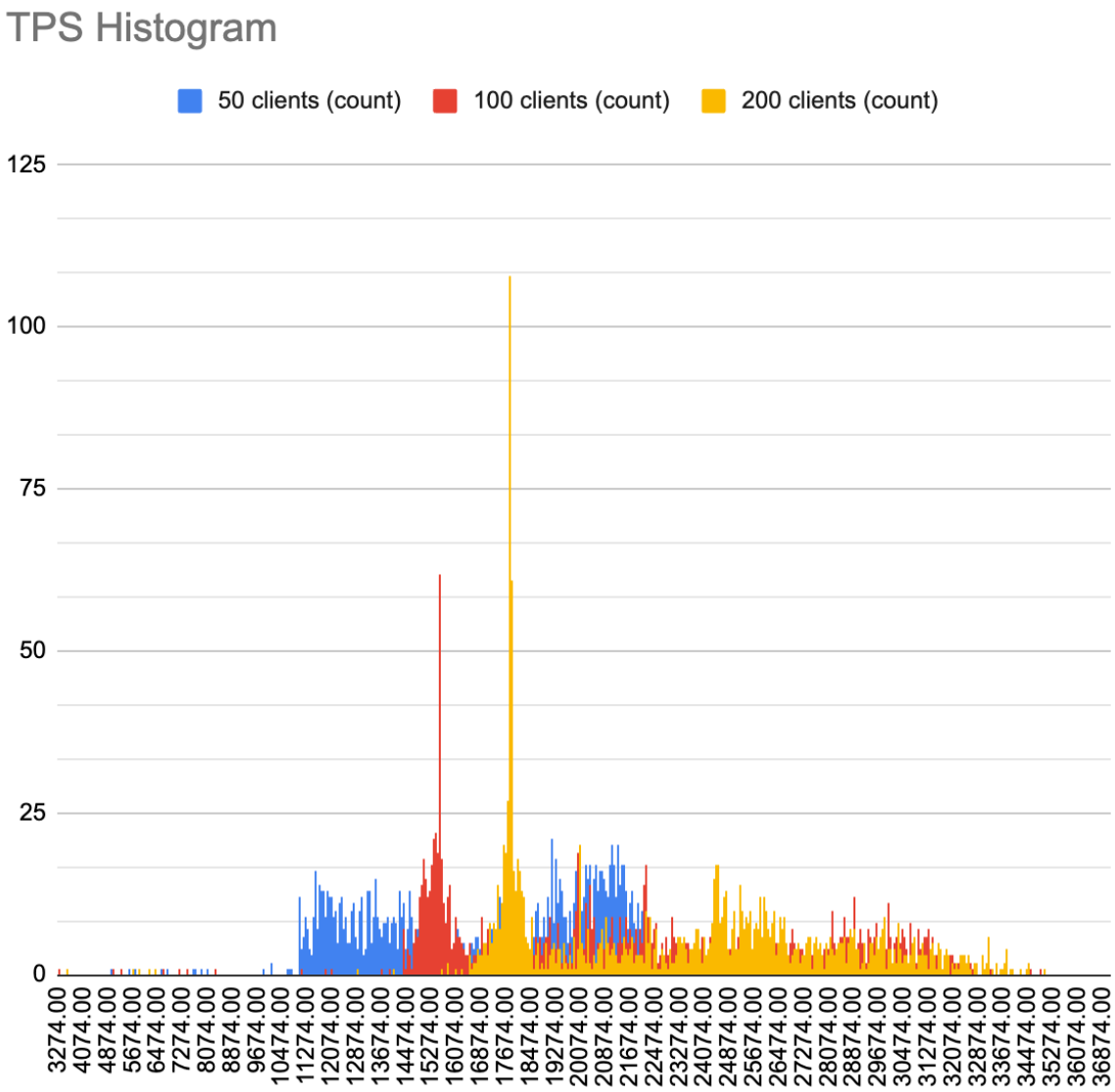
Simple Update Benchmark

The simple update benchmark focuses on a narrower transaction type, only consisting of simple updates to existing rows. It isolates the performance of write operations, particularly under sustained load.

Transactions Per Second

Clients	Average	Median	StdDev
50	17137	17982	3600
100	21720	21014	5630
200	23069	23285	4932

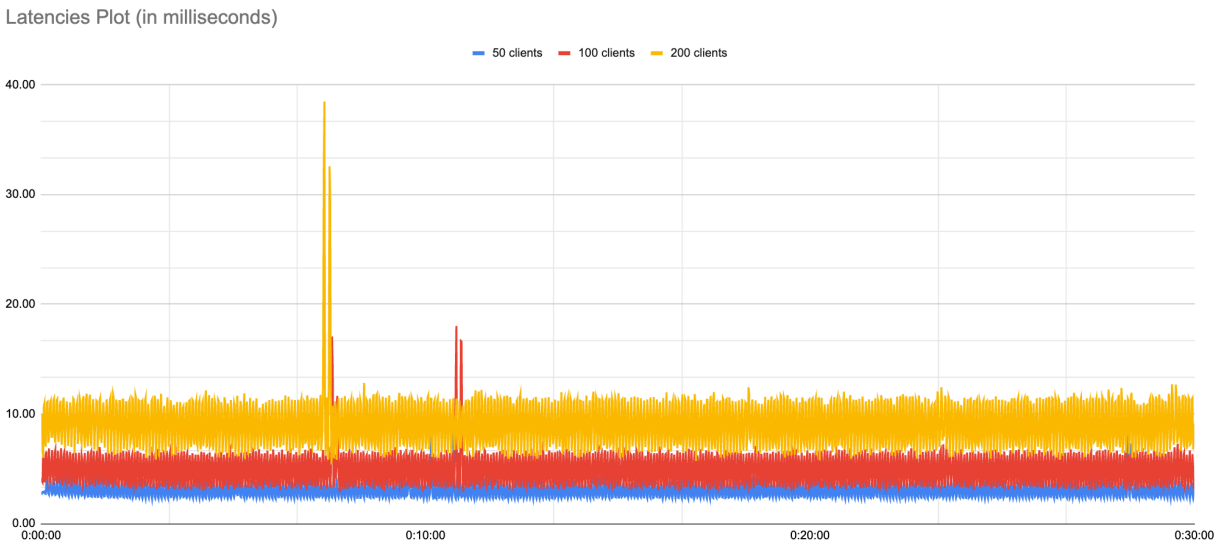


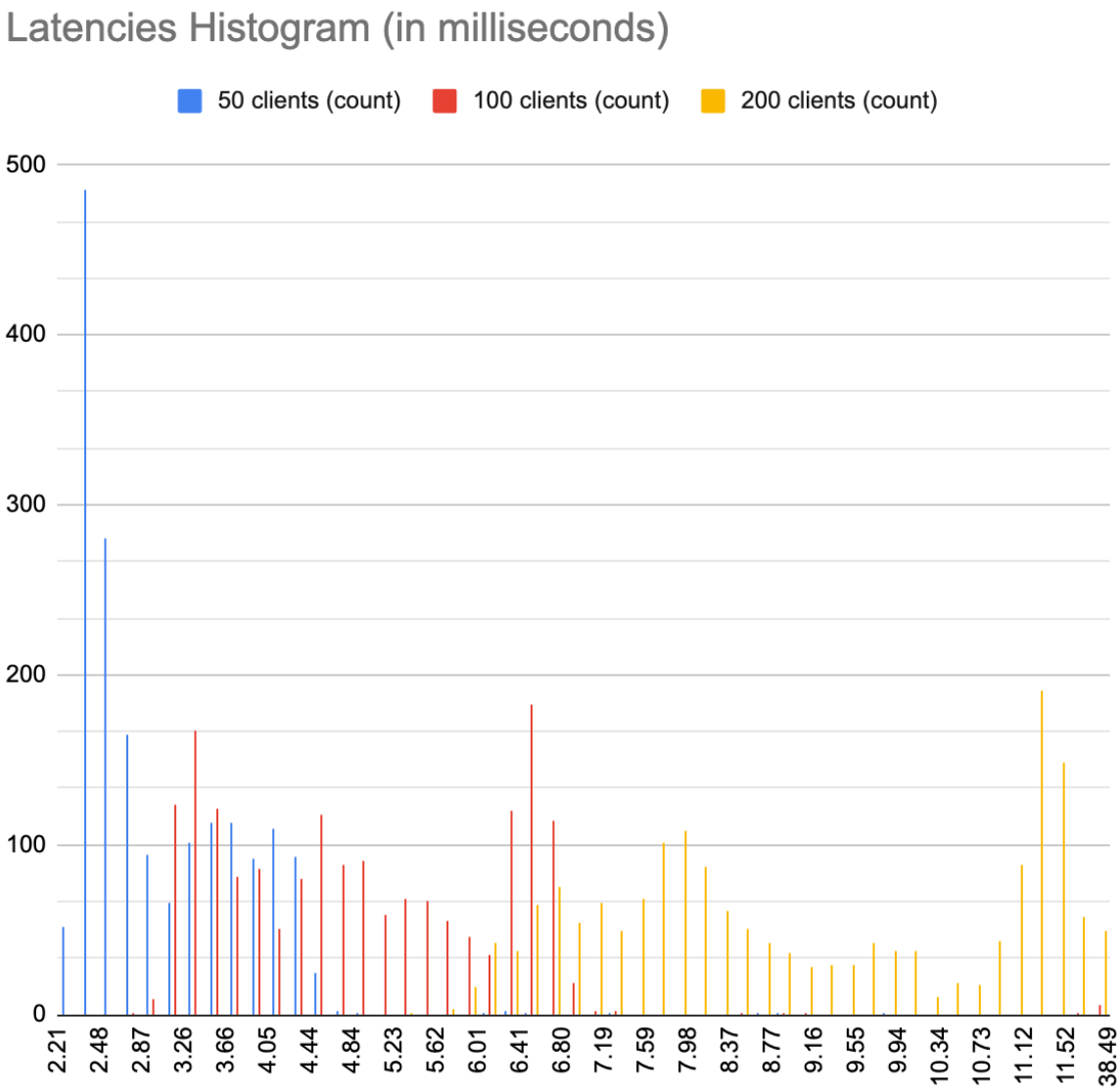

TPS Histogram
50 clients (count)
100 clients (count)
200 clients (count)
125
100
75
50
25
0
3274.00
4074.00
4874.00
5674.00
6474.00
7274.00
8074.00
8874.00
9674.00
10474.00
11274.00
12074.00
12874.00
13674.00
14474.00
15274.00
16074.00
16874.00
17674.00
18474.00
19274.00
20074.00
20874.00
21674.00
22474.00
23274.00
24074.00
24874.00
25674.00
26474.00
27274.00
28074.00
28874.00
29674.00
30474.00
31274.00
32074.00
32874.00
33674.00
34474.00
35274.00
36074.00
36874.00

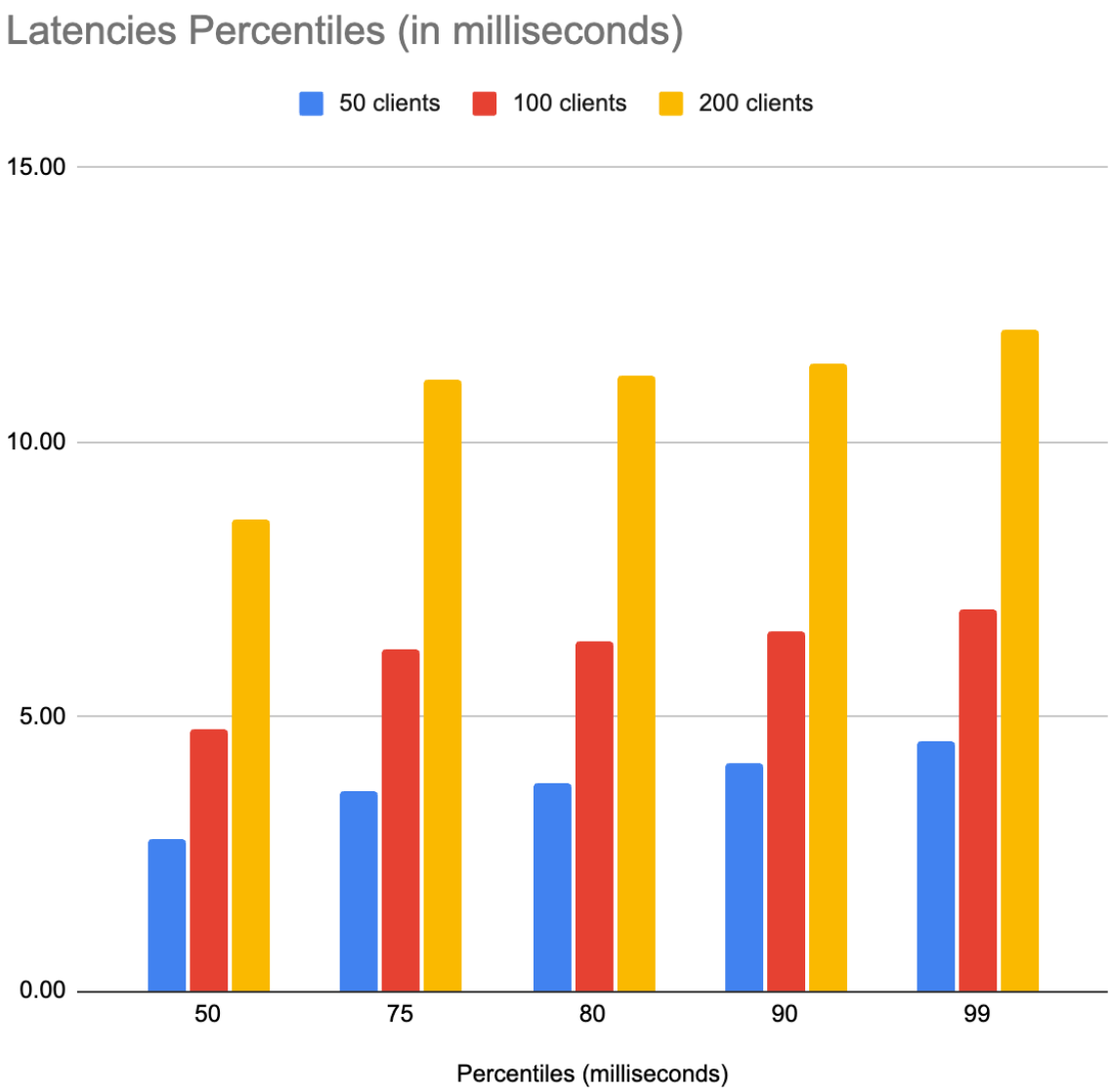
Transaction Latency

Clients	Average	Median	StdDev
50	3.06	2.78	0.76
100	4.928	4.758	1.389
200	9.082	8.596	2.191

Clients	50th	75th	80th	90th	99th
50	2.78	3.65	3.8	4.17	4.54
100	4.76	6.21	6.35	6.56	6.94
200	8.6	11.12	11.22	11.42	12.03




Latencies Histogram (in milliseconds)
50 clients (count)
100 clients (count)
200 clients (count)
500
400
300
200
100
0
2.21
2.48
2.87
3.26
3.66
4.05
4.44
4.84
5.23
5.62
6.01
6.41
6.80
7.19
7.59
7.98
8.37
8.77
9.16
9.55
9.94
10.34
10.73
11.12
11.52
38.49


Latencies Percentiles (in milliseconds)
50 clients
100 clients
200 clients
15.00
10.00
5.00
0.00
50
75
80
90
99
Percentiles (milliseconds)

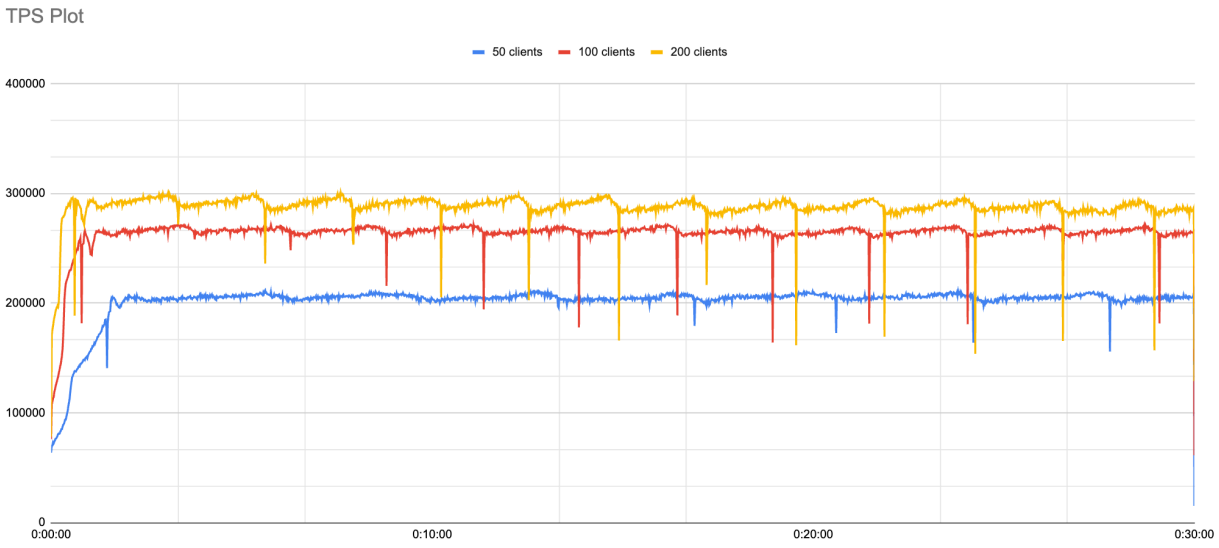
Select Benchmark

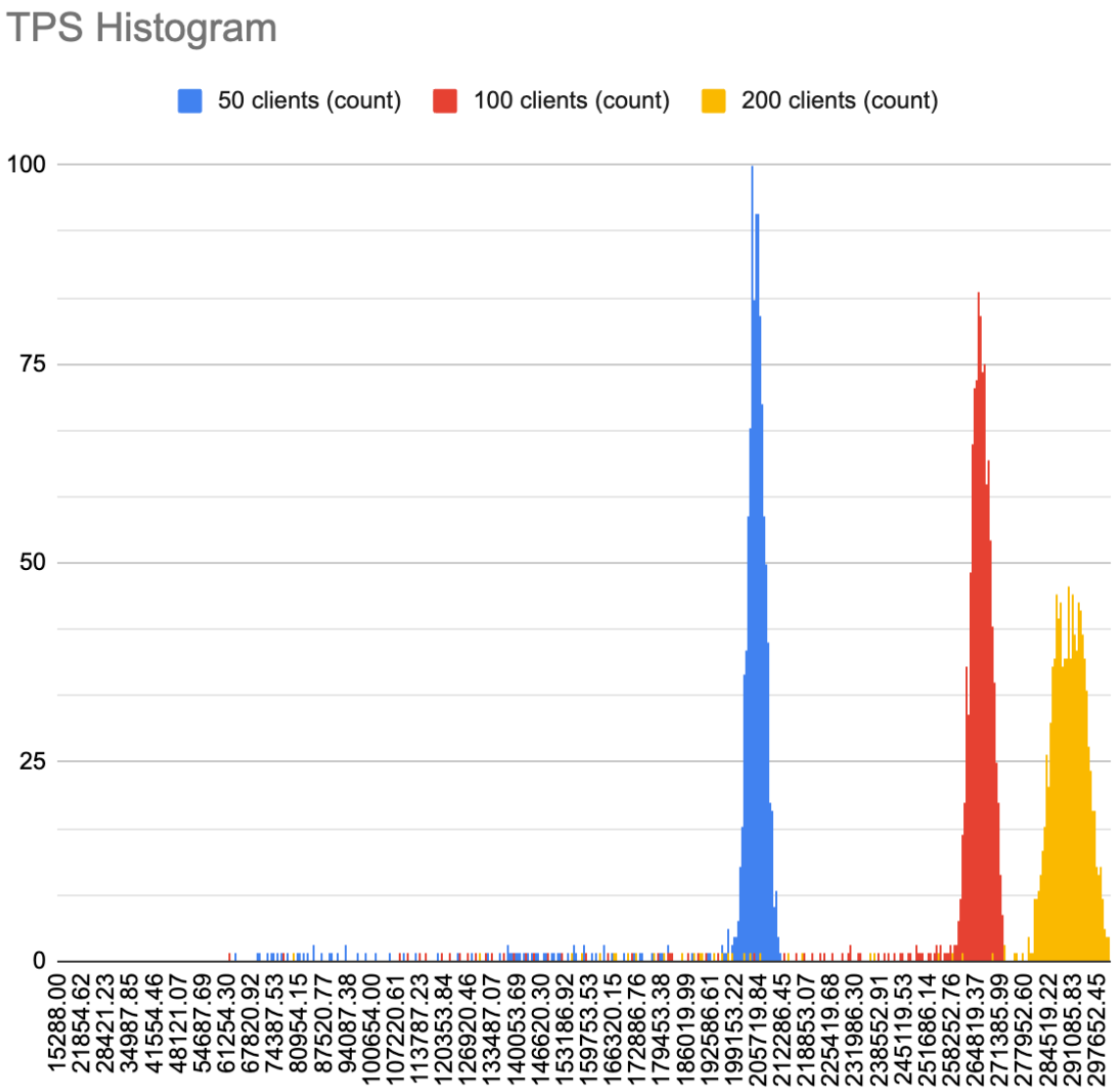
The select benchmark evaluates the system's ability to handle read-only operations — a crucial aspect for analytics, reporting, and read-heavy workloads.

The select benchmark was not disk-bound. Even at 16MB shared buffers, the disk only had a utilization of about 400,000 IOPS. When manually dropping the Linux caches as fast as possible, performance stabilized at about 190,000 TPS.

Transactions Per Second

Clients	Average	Median	StdDev
50	201447	205169	18677
100	262984	265522	17039
200	287553	289047	14225

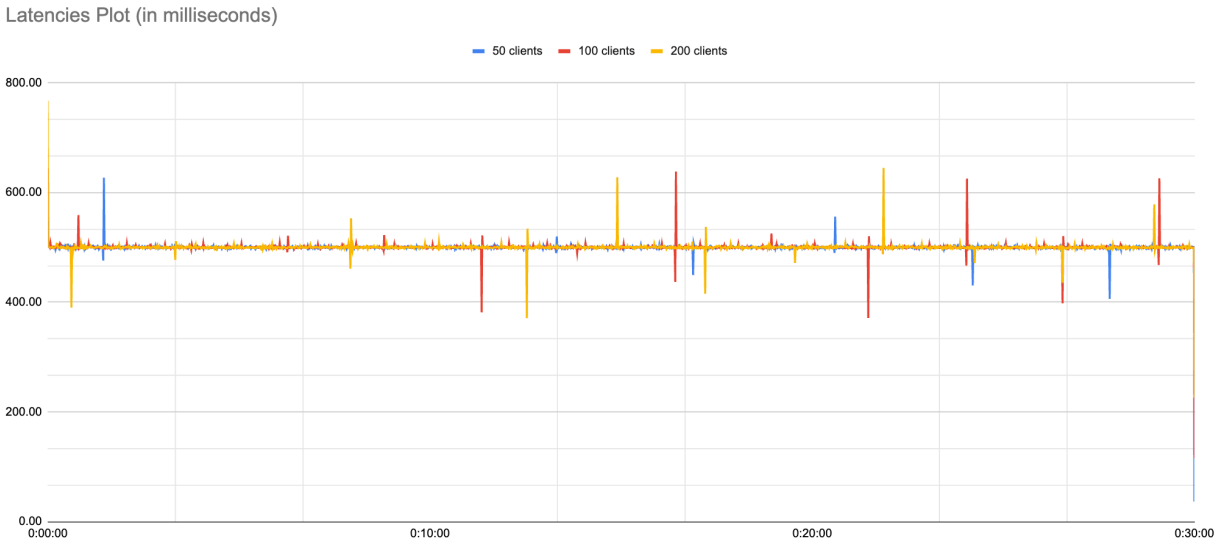


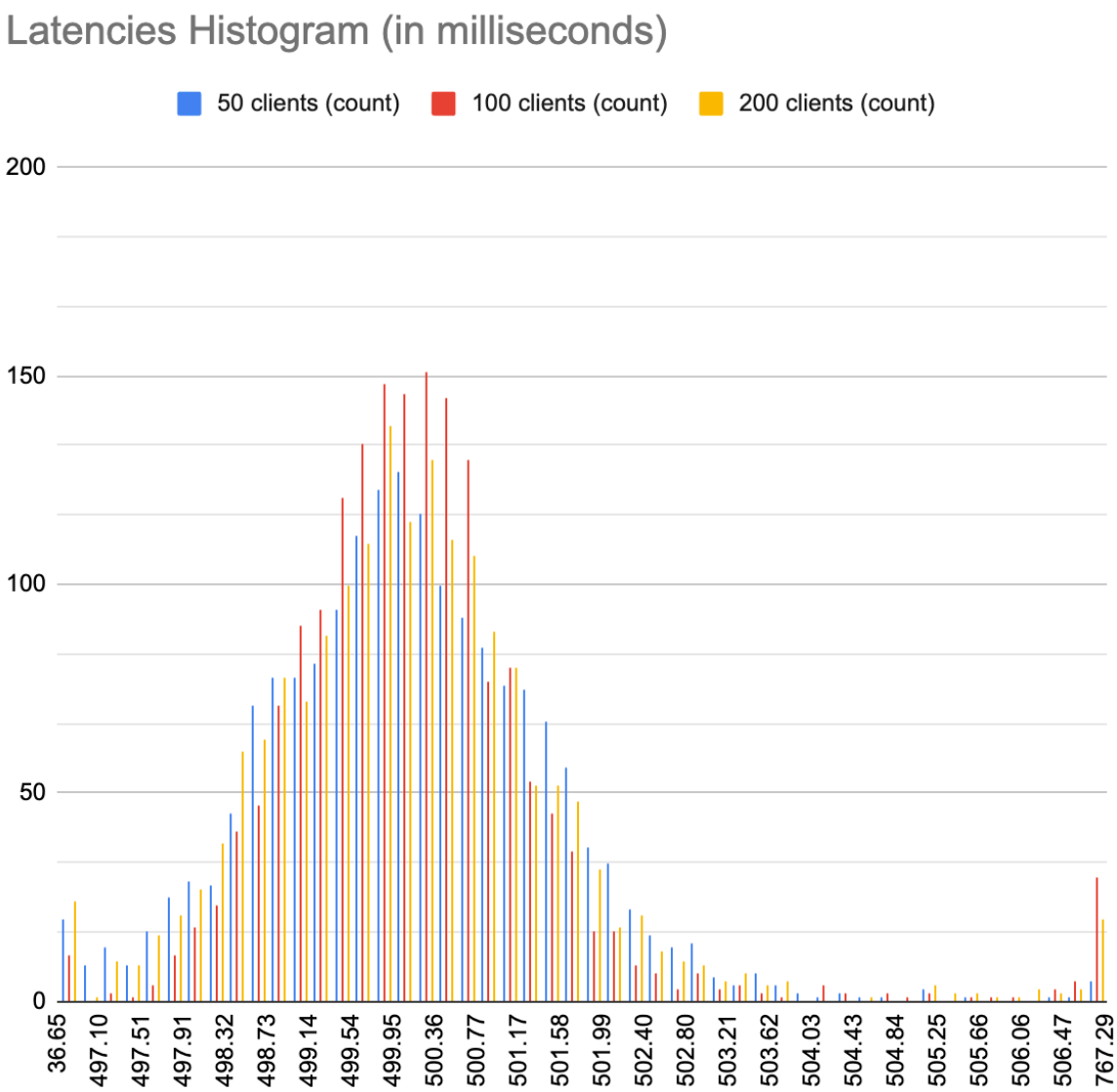

TPS Histogram
50 clients (count)
100 clients (count)
200 clients (count)
100
75
50
25
0
15288.00
21854.62
28421.23
34987.85
41554.46
48121.07
54687.69
61254.30
67820.92
74387.53
80954.15
87520.77
94087.38
100654.00
107220.61
113787.23
120353.84
126920.46
133487.07
140053.69
146620.30
153186.92
159753.53
166320.15
172886.76
179453.38
186019.99
192586.61
199153.22
205719.84
212286.45
218853.07
225419.68
231986.30
238552.91
245119.53
251686.14
258252.76
264819.37
271385.99
277952.60
284519.22
291085.83
297652.45

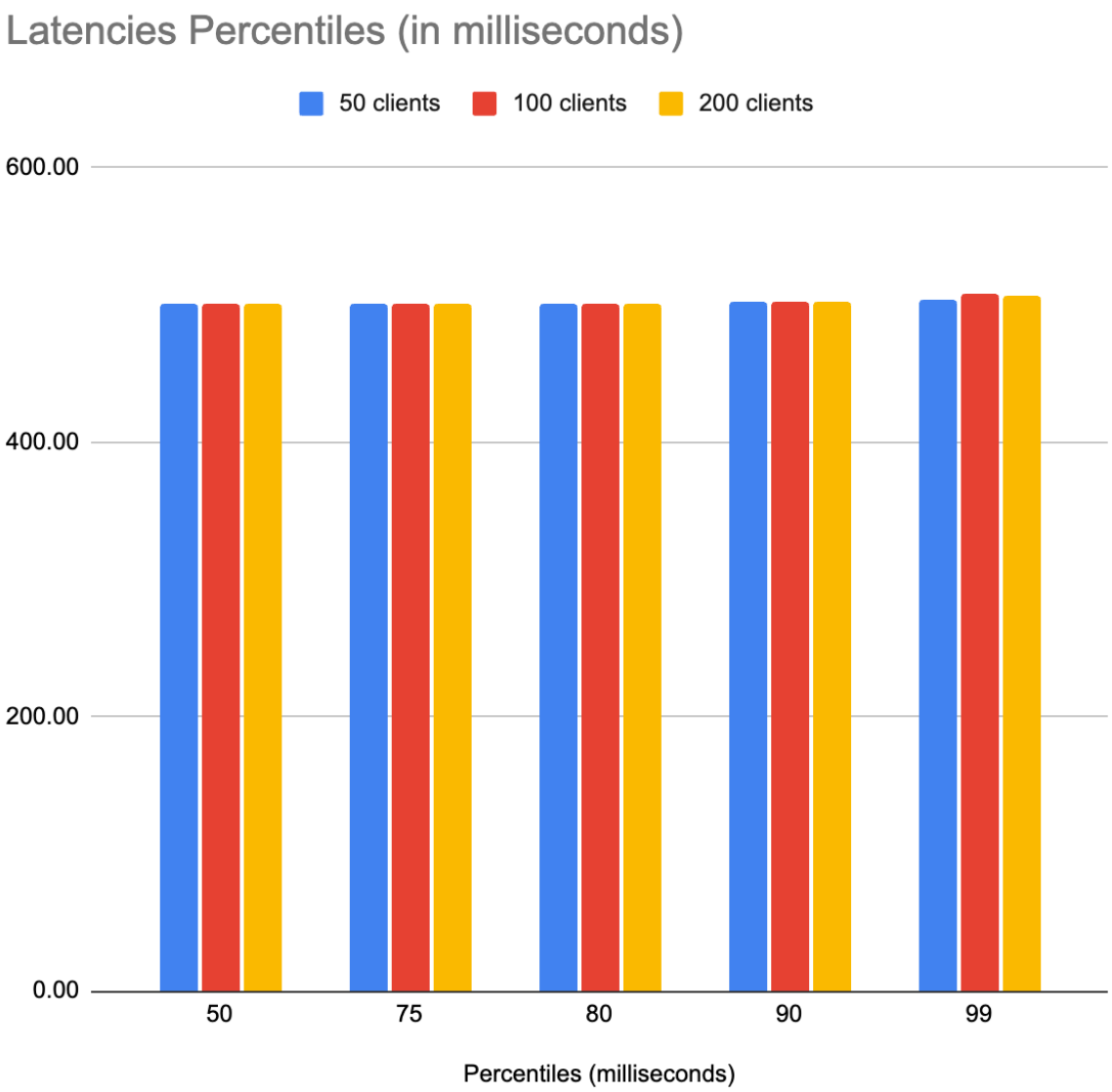
Transaction Latency

Clients	Average	Median	StdDev
50	499.85	500.06	11.95
100	500.10	500.05	12.39
200	500.09	500.01	11.63

Clients	50th	75th	80th	90th	99th
50	500.06	500.96	501.21	501.77	503.80
100	500.05	500.71	500.89	501.47	507.45
200	500.01	500.85	501.04	501.70	506.80




Latencies Histogram (in milliseconds)
50 clients (count)
100 clients (count)
200 clients (count)
200
150
100
50
0
36.65
497.10
497.51
497.91
498.32
498.73
499.14
499.54
499.95
500.36
500.77
501.17
501.58
501.99
502.40
502.80
503.21
503.62
504.03
504.43
504.84
505.25
505.66
506.06
506.47
767.29


Latencies Percentiles (in milliseconds)
50 clients
100 clients
200 clients
600.00
400.00
200.00
0.00
50
75
80
90
99
Percentiles (milliseconds)