


**www.phoronix-test-suite.com**

# new tests

Intel Core i9-13900K testing with a ASUS PRIME Z790-P WIFI (0602 BIOS) and AMD Radeon RX 6700 XT 12GB on Ubuntu 22.04 via the Phoronix Test Suite.

**Automated Executive Summary**

*b had the most wins, coming in first place for 33% of the tests.*

*Based on the geometric mean of all complete results, the fastest (a) was 1.019x the speed of the slowest (d). b was 0.984x the speed of a, c was 0.999x the speed of b, d was 0.998x the speed of c.*

*The results with the greatest spread from best to worst included:*

*KeyDB (Test: LPOP - Parallel Connections: 100) at 1.43x*
*KeyDB (Test: LPOP - Parallel Connections: 500) at 1.399x*
*KeyDB (Test: LPOP - Parallel Connections: 900) at 1.389x*
*KeyDB (Test: SET - Parallel Connections: 100) at 1.172x*
*KeyDB (Test: LPOP - Parallel Connections: 50) at 1.169x*
*KeyDB (Test: SADD - Parallel Connections: 100) at 1.168x*
*KeyDB (Test: HMSET - Parallel Connections: 100) at 1.125x*
*KeyDB (Test: SADD - Parallel Connections: 900) at 1.109x*
*KeyDB (Test: SET - Parallel Connections: 500) at 1.106x*

*KeyDB (Test: LPUSH - Parallel Connections: 500) at 1.102x.*

# Test Systems:

## a

## b

## c

## d

Processor: Intel Core i9-13900K @ 4.00GHz (24 Cores / 32 Threads), Motherboard: ASUS PRIME Z790-P WIFI (0602 BIOS), Chipset: Intel Device 7a27, Memory: 32GB, Disk: 1000GB Western Digital WDS100T1X0E-00AFY0, Graphics: AMD Radeon RX 6700 XT 12GB (2855/1000MHz), Audio: Realtek ALC897, Monitor: ASUS VP28U, Network: Intel Device 7a70

OS: Ubuntu 22.04, Kernel: 6.0.0-060000-generic (x86_64), Desktop: GNOME Shell 42.4, Display Server: X Server 1.21.1.3 + Wayland, OpenGL: 4.6 Mesa 22.3.0-devel (git-8945375 2022-10-04 jammy-oibaf-ppa) (LLVM 14.0.6 DRM 3.48), Vulkan: 1.3.230, Compiler: GCC 11.2.0, File-System: ext4, Screen Resolution: 3840x2160

Kernel Notes: Transparent Huge Pages: madvise
Compiler Notes: --build=x86_64-linux-gnu --disable-vtable-verify --disable-werror --enable-bootstrap --enable-cet --enable-checking=release --enable-clocale=gnu --enable-default-pie --enable-gnu-unique-object --enable-languages=c,ada,c++,go,brig,d,fortran,objc,obj-c++,m2 --enable-libphobos-checking=release --enable-libstdcxx-debug --enable-libstdcxx-time=yes --enable-link-serialization=2 --enable-multiarch --enable-multilib --enable-nls --enable-objc-gc=auto --enable-offload-targets=nvptx-none=/build/gcc-11-gBFGDP/gcc-11-11.2.0/debian/tmp-nvptx/usr,amdgcn-amdhsa=/build/gcc-11-gBFGDP/gcc-11-11.2.0/debian/tmp-gcn/usr --enable-plugin --enable-shared --enable-threads=posix --host=x86_64-linux-gnu --program-prefix=x86_64-linux-gnu- --target=x86_64-linux-gnu --with-abi=m64 --with-arch-32=i686 --with-build-config=bootstrap-lto-lean --with-default-libstdcxx-abi=new --with-gcc-major-version-only --with-multilib-list=m32,m64,mx32 --with-target-system-zlib=auto --with-tune=generic --without-cuda-driver -v
Processor Notes: Scaling Governor: intel_pstate powersave (EPP: balance_performance) - CPU Microcode: 0x10e - Thermald 2.4.9
Python Notes: Python 3.10.6
Security Notes: itlb_multihit: Not affected + l1tf: Not affected + mds: Not affected + meltdown: Not affected + mmio_stale_data: Not affected + retbleed: Not affected + spec_store_bypass: Mitigation of SSB disabled via prctl + spectre_v1: Mitigation of usercopy/swapgs barriers and __user pointer sanitization + spectre_v2: Mitigation of Enhanced IBRS IBPB: conditional RSB filling PBRSB-eIBRS: SW sequence + srbds: Not affected + tsx_async_abort: Not affected

| | a | b | c | d |
|---|---|---|---|---|
| **Neural Magic DeepSparse - N.D.C.o.b.u.o.l - A.M.S (items/sec)** | 17.0977 | **17.0299** | 17.1322 | **17.2573** |
| **Normalized** | 99.08% | 98.68% | 99.28% | 100% |
| **Standard Deviation** | 0.9% | | | |
| **Neural Magic DeepSparse - N.D.C.o.b.u.o.l - S.S.S (items/sec)** | 15.1510 | 15.1488 | **15.1431** | **15.1564** |
| **Normalized** | 99.96% | 99.95% | 99.91% | 100% |
| **Standard Deviation** | 0.2% | | | |

# new tests

| | | | | |
|---|---|---|---|---|
| **Neural Magic DeepSparse - N.S.A.8.P.Q.B.B.U - A.M.S (items/sec)** | 201.8858 | 202.8721 | 202.7414 | 203.7202 |
| **Normalized** | 99.1% | 99.58% | 99.52% | 100% |
| **Standard Deviation** | 1.9% | | | |
| **Neural Magic DeepSparse - N.S.A.8.P.Q.B.B.U - S.S.S (items/sec)** | 97.0266 | 96.6004 | 97.6971 | 97.0456 |
| **Normalized** | 99.31% | 98.88% | 100% | 99.33% |
| **Standard Deviation** | 0.3% | | | |
| **Neural Magic DeepSparse - N.Q.A.B.b.u.S.1.P - A.M.S (items/sec)** | 62.4727 | 62.0325 | 62.4335 | 62.9368 |
| **Normalized** | 99.26% | 98.56% | 99.2% | 100% |
| **Standard Deviation** | 0.4% | | | |
| **Neural Magic DeepSparse - N.Q.A.B.b.u.S.1.P - S.S.S (items/sec)** | 44.4092 | 44.3471 | 44.2034 | 44.3803 |
| **Normalized** | 100% | 99.86% | 99.54% | 99.93% |
| **Standard Deviation** | 0.3% | | | |
| **Neural Magic DeepSparse - C.D.Y.C - A.M.S (items/sec)** | 100.5956 | 102.1854 | 100.922 | 101.5808 |
| **Normalized** | 98.44% | 100% | 98.76% | 99.41% |
| **Standard Deviation** | 0.8% | | | |
| **Neural Magic DeepSparse - C.D.Y.C - S.S.S (items/sec)** | 67.8527 | 68.4113 | 68.2268 | 69.2308 |
| **Normalized** | 98.01% | 98.82% | 98.55% | 100% |
| **Standard Deviation** | 0.8% | | | |
| **Neural Magic DeepSparse - C.C.R.5.I - A.M.S (items/sec)** | 220.2242 | 223.1123 | 222.5816 | 220.629 |
| **Normalized** | 98.71% | 100% | 99.76% | 98.89% |
| **Standard Deviation** | 1.1% | | | |
| **Neural Magic DeepSparse - C.C.R.5.I - S.S.S (items/sec)** | 125.7040 | 126.9919 | 126.7132 | 126.8482 |
| **Normalized** | 98.99% | 100% | 99.78% | 99.89% |
| **Standard Deviation** | 0.5% | | | |
| **Neural Magic DeepSparse - N.T.C.D.m - A.M.S (items/sec)** | 146.8362 | 146.7638 | 148.9785 | 146.7079 |
| **Normalized** | 98.56% | 98.51% | 100% | 98.48% |
| **Standard Deviation** | 0.2% | | | |
| **Neural Magic DeepSparse - N.T.C.D.m - S.S.S (items/sec)** | 84.2139 | 84.3354 | 84.6155 | 84.7035 |
| **Normalized** | 99.42% | 99.57% | 99.9% | 100% |
| **Standard Deviation** | 0.4% | | | |
| **Neural Magic DeepSparse - C.S.9.P.Y.P - A.M.S (items/sec)** | 22.4583 | 22.572 | 22.4829 | 22.8771 |
| **Normalized** | 98.17% | 98.67% | 98.28% | 100% |
| **Standard Deviation** | 0.3% | | | |
| **Neural Magic DeepSparse - C.S.9.P.Y.P - S.S.S (items/sec)** | 17.9198 | 17.9723 | 17.8257 | 17.902 |
| **Normalized** | 99.71% | 100% | 99.18% | 99.61% |
| **Standard Deviation** | 0.4% | | | |
| **Neural Magic DeepSparse - N.T.C.B.b.u.S - A.M.S (items/sec)** | 74.6763 | 73.963 | 75.2056 | 74.4384 |
| **Normalized** | 99.3% | 98.35% | 100% | 98.98% |
| **Standard Deviation** | 0.3% | | | |

# new tests

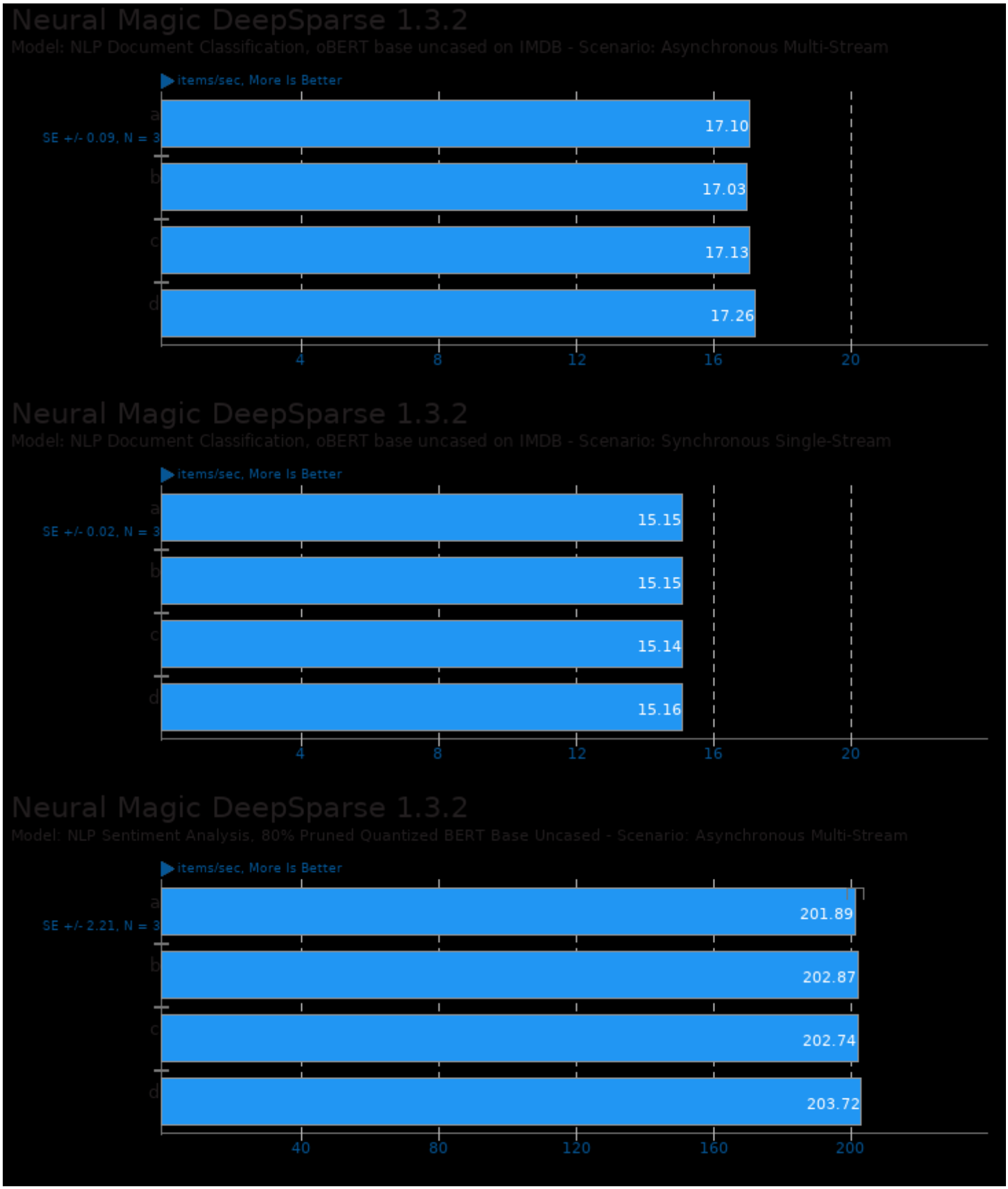
| | | | | |
|---|---|---|---|---|
| **Neural Magic DeepSparse - N.T.C.B.b.u.S - S.S.S (items/sec)** | 42.4124 | **42.4107** | **42.5922** | 42.4206 |
| **Normalized** | 99.58% | 99.57% | 100% | 99.6% |
| **Standard Deviation** | 0.3% | | | |
| **Neural Magic DeepSparse - N.T.C.B.b.u.c - A.M.S (items/sec)** | 17.0593 | **17.2641** | **17.0024** | 17.2405 |
| **Normalized** | 98.81% | 100% | 98.48% | 99.86% |
| **Standard Deviation** | 0.6% | | | |
| **Neural Magic DeepSparse - N.T.C.B.b.u.c - S.S.S (items/sec)** | 15.1861 | **15.2008** | **15.1247** | 15.1758 |
| **Normalized** | 99.9% | 100% | 99.5% | 99.84% |
| **Standard Deviation** | 0.5% | | | |
| **KeyDB - GET - 50 (Reqs/sec)** | **2301457** | **2203905** | 2258764 | 2273451 |
| **Normalized** | 100% | 95.76% | 98.14% | 98.78% |
| **Standard Deviation** | 0.9% | | | |
| **KeyDB - SET - 50 (Reqs/sec)** | 2051262 | **2074172** | **2007790** | 2013693 |
| **Normalized** | 98.9% | 100% | 96.8% | 97.08% |
| **Standard Deviation** | 2.2% | | | |
| **KeyDB - GET - 100 (Reqs/sec)** | **1526251** | **1482228** | 1488317 | 1484384 |
| **Normalized** | 100% | 97.12% | 97.51% | 97.26% |
| **Standard Deviation** | 1.5% | | | |
| **KeyDB - GET - 500 (Reqs/sec)** | 1493933 | **1558846** | **1476887** | 1553519 |
| **Normalized** | 95.84% | 100% | 94.74% | 99.66% |
| **Standard Deviation** | 5.1% | | | |
| **KeyDB - GET - 900 (Reqs/sec)** | 1563299 | **1582379** | 1527464 | **1510802** |
| **Normalized** | 98.79% | 100% | 96.53% | 95.48% |
| **Standard Deviation** | 2.2% | | | |
| **KeyDB - LPOP - 50 (Reqs/sec)** | **2355769** | 2032603 | **2014748** | 2016942 |
| **Normalized** | 100% | 86.28% | 85.52% | 85.62% |
| **Standard Deviation** | 1.9% | | | |
| **KeyDB - SADD - 50 (Reqs/sec)** | **2240005** | 2210043 | 2215134 | **2186557** |
| **Normalized** | 100% | 98.66% | 98.89% | 97.61% |
| **Standard Deviation** | 1.1% | | | |
| **KeyDB - SET - 100 (Reqs/sec)** | 1277181 | 1180777 | **1105999** | **1296008** |
| **Normalized** | 98.55% | 91.11% | 85.34% | 100% |
| **Standard Deviation** | 0.9% | | | |
| **KeyDB - SET - 500 (Reqs/sec)** | 1242079 | **1314786** | **1188354** | 1307531 |
| **Normalized** | 94.47% | 100% | 90.38% | 99.45% |
| **Standard Deviation** | 1.1% | | | |
| **KeyDB - SET - 900 (Reqs/sec)** | 1227965 | **1200999** | 1209599 | **1305926** |
| **Normalized** | 94.03% | 91.97% | 92.62% | 100% |
| **Standard Deviation** | 3.2% | | | |
| **KeyDB - HMSET - 50 (Reqs/sec)** | **770419** | **827089** | 822829 | 781067 |
| **Normalized** | 93.15% | 100% | 99.48% | 94.44% |
| **Standard Deviation** | 2.3% | | | |
| **KeyDB - LPOP - 100 (Reqs/sec)** | **1548102** | **1082814** | 1133556 | 1110667 |
| **Normalized** | 100% | 69.94% | 73.22% | 71.74% |
| **Standard Deviation** | 4.7% | | | |
| **KeyDB - LPOP - 500 (Reqs/sec)** | **1541363** | 1142413 | 1166752 | **1101928** |
| **Normalized** | 100% | 74.12% | 75.7% | 71.49% |
| **Standard Deviation** | 3.9% | | | |
| **KeyDB - LPOP - 900 (Reqs/sec)** | **1572283** | **1132093** | 1171454 | 1168661 |
| **Normalized** | 100% | 72% | 74.51% | 74.33% |
| **Standard Deviation** | 5.4% | | | |

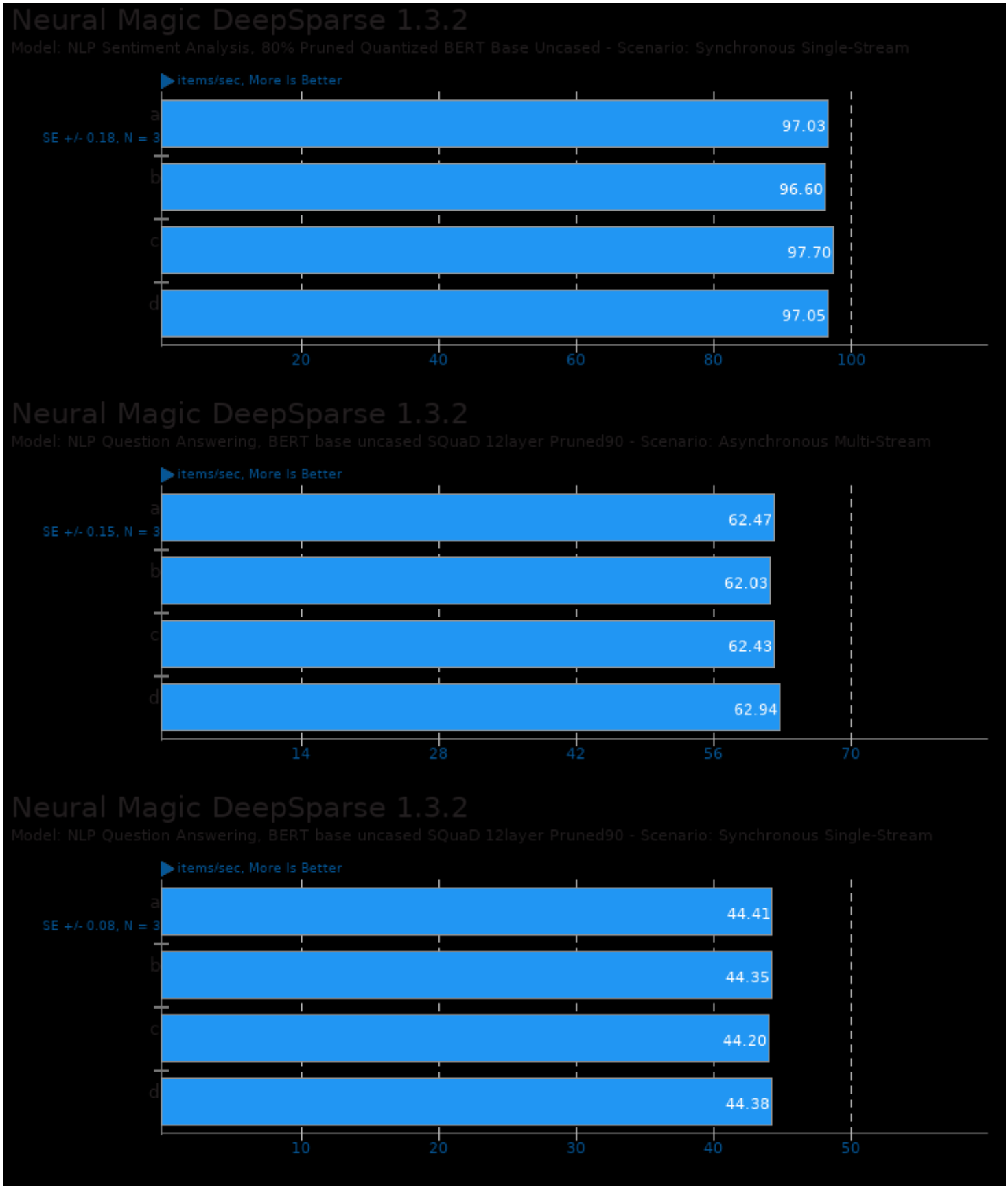
# new tests

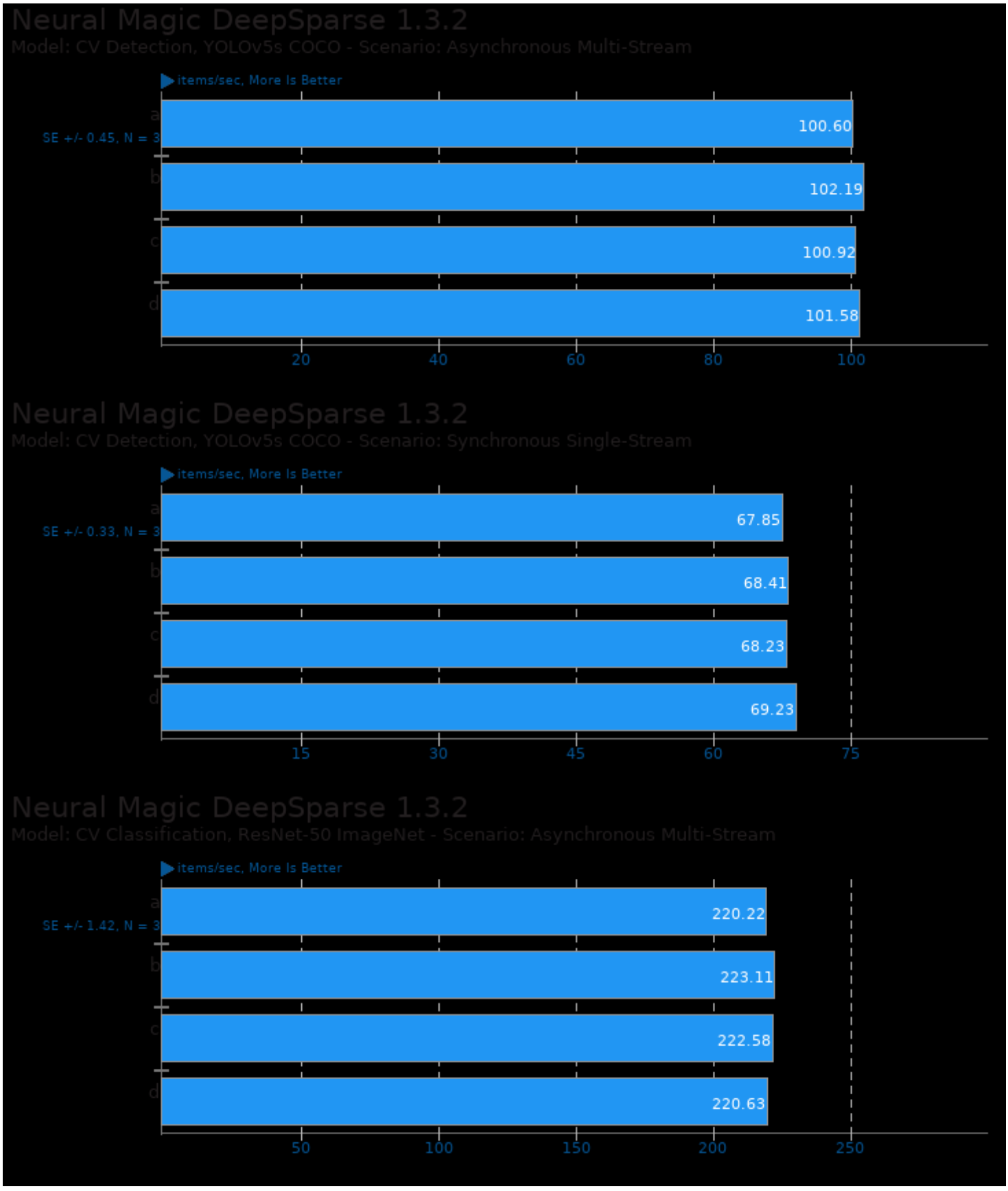
| | | | | |
|---|---|---|---|---|
| **KeyDB - LPUSH - 50 (Reqs/sec)** | 1885807 | 1849523 | **1964096** | **1816596** |
| **Normalized** | 96.01% | 94.17% | 100% | 92.49% |
| **Standard Deviation** | 2.3% | | | |
| **KeyDB - SADD - 100 (Reqs/sec)** | 1301147 | 1326330 | **1443335** | **1235605** |
| **Normalized** | 90.15% | 91.89% | 100% | 85.61% |
| **Standard Deviation** | 2.3% | | | |
| **KeyDB - SADD - 500 (Reqs/sec)** | **1369793** | 1365262 | **1294096** | 1296243 |
| **Normalized** | 100% | 99.67% | 94.47% | 94.63% |
| **Standard Deviation** | 5.1% | | | |
| **KeyDB - SADD - 900 (Reqs/sec)** | 1318362 | **1402721** | 1331771 | **1264958** |
| **Normalized** | 93.99% | 100% | 94.94% | 90.18% |
| **Standard Deviation** | 3.8% | | | |
| **KeyDB - HMSET - 100 (Reqs/sec)** | 811193 | 838926 | **858045** | **762637** |
| **Normalized** | 94.54% | 97.77% | 100% | 88.88% |
| **Standard Deviation** | 3.4% | | | |
| **KeyDB - HMSET - 500 (Reqs/sec)** | 540612 | **575162** | 544526 | **514398** |
| **Normalized** | 93.99% | 100% | 94.67% | 89.44% |
| **Standard Deviation** | 8.4% | | | |
| **KeyDB - HMSET - 900 (Reqs/sec)** | 405829 | 385410 | **373871** | **405920** |
| **Normalized** | 99.98% | 94.95% | 92.1% | 100% |
| **Standard Deviation** | 7% | | | |
| **KeyDB - LPUSH - 100 (Reqs/sec)** | 1079455 | **1009102** | 1053253 | **1161009** |
| **Normalized** | 92.98% | 86.92% | 90.72% | 100% |
| **Standard Deviation** | 6.8% | | | |
| **KeyDB - LPUSH - 500 (Reqs/sec)** | 972929 | 932940 | **1020158** | **925806** |
| **Normalized** | 95.37% | 91.45% | 100% | 90.75% |
| **Standard Deviation** | 4.7% | | | |
| **KeyDB - LPUSH - 900 (Reqs/sec)** | **913353** | **988729** | 967361 | 934195 |
| **Normalized** | 92.38% | 100% | 97.84% | 94.48% |
| **Standard Deviation** | 1.5% | | | |
| **Neural Magic DeepSparse - N.D.C.o.b.u.o.I - A.M.S (ms/batch)** | **693.7534** | **697.4591** | 696.4586 | 693.8845 |
| **Normalized** | 100% | 99.47% | 99.61% | 99.98% |
| **Standard Deviation** | 0.4% | | | |
| **Neural Magic DeepSparse - N.D.C.o.b.u.o.I - S.S.S (ms/batch)** | 65.9984 | 66.0079 | **66.0331** | **65.9748** |
| **Normalized** | 99.96% | 99.95% | 99.91% | 100% |
| **Standard Deviation** | 0.2% | | | |
| **Neural Magic DeepSparse - N.S.A.8.P.Q.B.B.U - A.M.S (ms/batch)** | **59.3747** | 59.0448 | 59.1069 | **58.8037** |
| **Normalized** | 99.04% | 99.59% | 99.49% | 100% |
| **Standard Deviation** | 1.9% | | | |
| **Neural Magic DeepSparse - N.S.A.8.P.Q.B.B.U - S.S.S (ms/batch)** | 10.3007 | **10.3459** | **10.2299** | 10.2988 |
| **Normalized** | 99.31% | 98.88% | 100% | 99.33% |
| **Standard Deviation** | 0.3% | | | |
| **Neural Magic DeepSparse - N.Q.A.B.b.u.S.1.P - A.M.S (ms/batch)** | 191.4059 | **192.5598** | 191.3873 | **189.9321** |
| **Normalized** | 99.23% | 98.64% | 99.24% | 100% |
| **Standard Deviation** | 0.4% | | | |
| **Neural Magic DeepSparse - N.Q.A.B.b.u.S.1.P - S.S.S (ms/batch)** | **22.5114** | 22.5427 | **22.616** | 22.5261 |
| **Normalized** | 100% | 99.86% | 99.54% | 99.93% |
| **Standard Deviation** | 0.3% | | | |

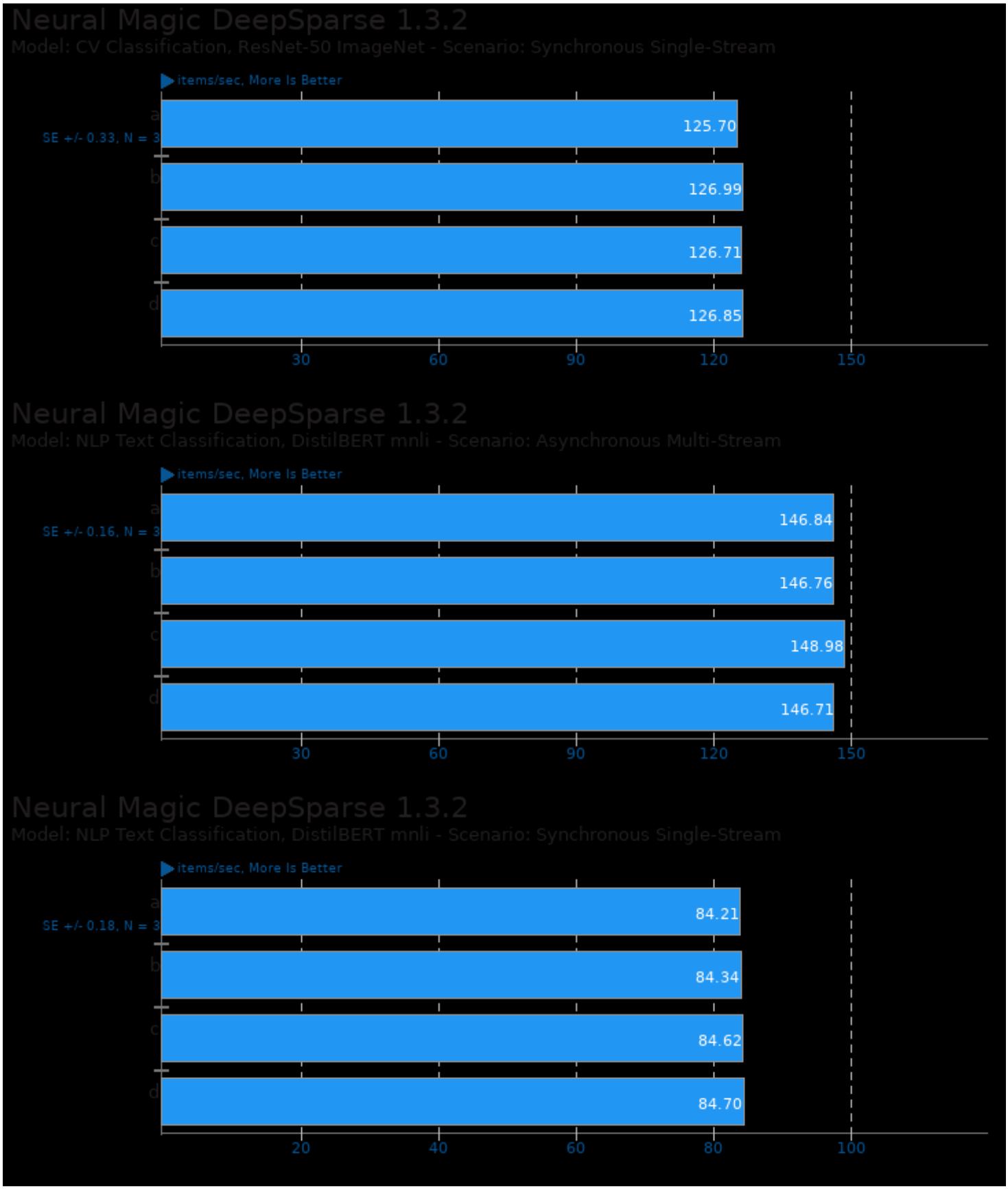
# new tests

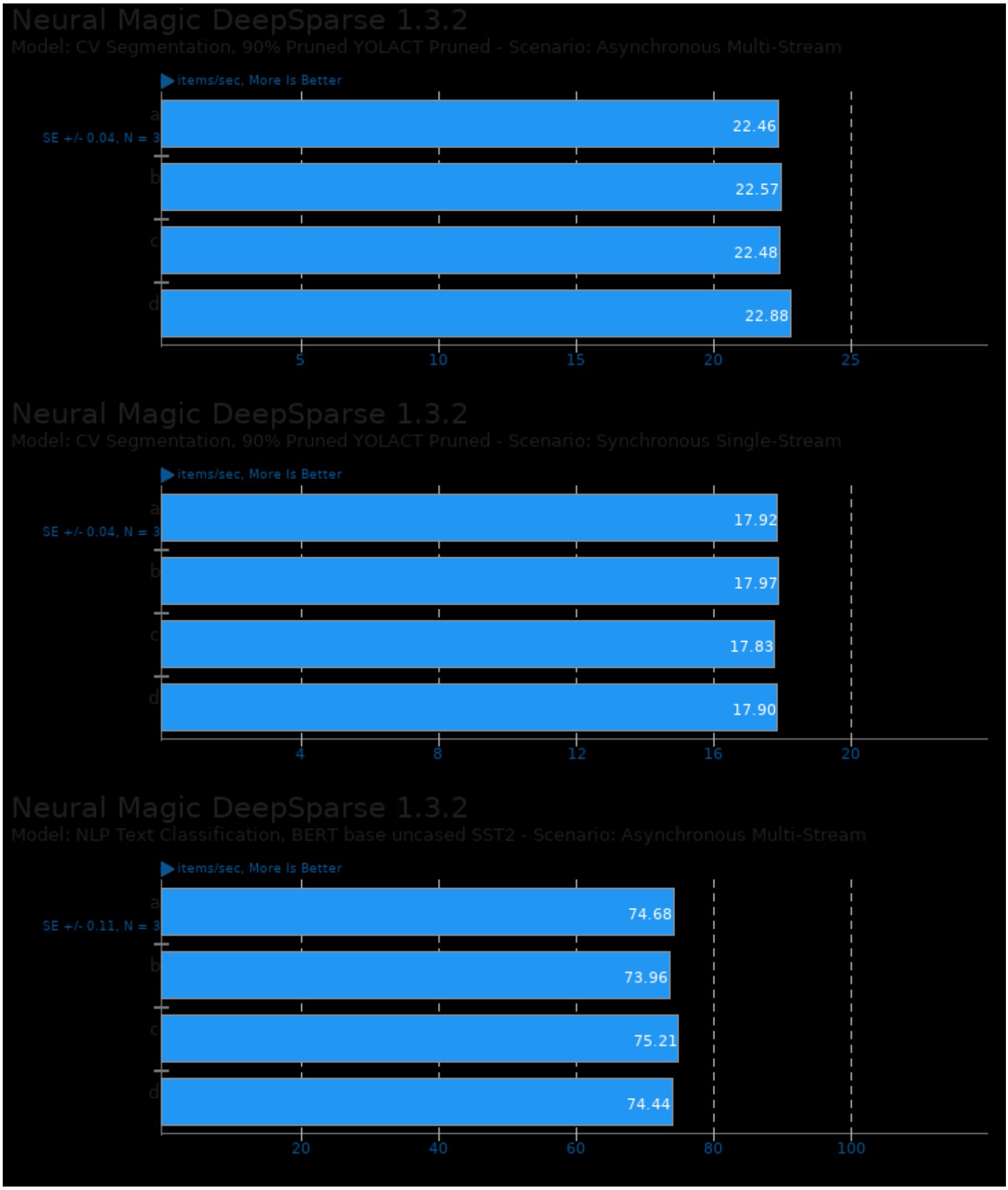
| | | | | |
|---|---|---|---|---|
| **Neural Magic DeepSparse - C.D.Y.C - A.M.S (ms/batch)** | 118.9735 | 117.1808 | 118.6662 | 117.9462 |
| **Normalized** | 98.49% | 100% | 98.75% | 99.35% |
| **Standard Deviation** | 0.9% | | | |
| **Neural Magic DeepSparse - C.D.Y.C - S.S.S (ms/batch)** | 14.7283 | 14.607 | 14.6466 | 14.4335 |
| **Normalized** | 98% | 98.81% | 98.55% | 100% |
| **Standard Deviation** | 0.8% | | | |
| **Neural Magic DeepSparse - C.C.R.5.I - A.M.S (ms/batch)** | 54.4516 | 53.7491 | 53.8481 | 54.3784 |
| **Normalized** | 98.71% | 100% | 99.82% | 98.84% |
| **Standard Deviation** | 1.1% | | | |
| **Neural Magic DeepSparse - C.C.R.5.I - S.S.S (ms/batch)** | 7.9509 | 7.8702 | 7.8877 | 7.8787 |
| **Normalized** | 98.99% | 100% | 99.78% | 99.89% |
| **Standard Deviation** | 0.5% | | | |
| **Neural Magic DeepSparse - N.T.C.D.m - A.M.S (ms/batch)** | 81.6431 | 81.6896 | 80.4175 | 81.7629 |
| **Normalized** | 98.5% | 98.44% | 100% | 98.35% |
| **Standard Deviation** | 0.2% | | | |
| **Neural Magic DeepSparse - N.T.C.D.m - S.S.S (ms/batch)** | 11.8712 | 11.8541 | 11.8147 | 11.8028 |
| **Normalized** | 99.42% | 99.57% | 99.9% | 100% |
| **Standard Deviation** | 0.4% | | | |
| **Neural Magic DeepSparse - C.S.9.P.Y.P - A.M.S (ms/batch)** | 531.4908 | 530.1764 | 528.6955 | 523.9449 |
| **Normalized** | 98.58% | 98.82% | 99.1% | 100% |
| **Standard Deviation** | 0.1% | | | |
| **Neural Magic DeepSparse - C.S.9.P.Y.P - S.S.S (ms/batch)** | 55.7950 | 55.6313 | 56.0887 | 55.85 |
| **Normalized** | 99.71% | 100% | 99.18% | 99.61% |
| **Standard Deviation** | 0.4% | | | |
| **Neural Magic DeepSparse - N.T.C.B.b.u.S - A.M.S (ms/batch)** | 160.4842 | 161.8553 | 159.443 | 161.0202 |
| **Normalized** | 99.35% | 98.51% | 100% | 99.02% |
| **Standard Deviation** | 0.3% | | | |
| **Neural Magic DeepSparse - N.T.C.B.b.u.S - S.S.S (ms/batch)** | 23.5749 | 23.5757 | 23.4753 | 23.57 |
| **Normalized** | 99.58% | 99.57% | 100% | 99.6% |
| **Standard Deviation** | 0.3% | | | |
| **Neural Magic DeepSparse - N.T.C.B.b.u.c - A.M.S (ms/batch)** | 694.2030 | 693.6961 | 695.2334 | 693.8329 |
| **Normalized** | 99.93% | 100% | 99.78% | 99.98% |
| **Standard Deviation** | 0.5% | | | |
| **Neural Magic DeepSparse - N.T.C.B.b.u.c - S.S.S (ms/batch)** | 65.8471 | 65.7826 | 66.1119 | 65.8907 |
| **Normalized** | 99.9% | 100% | 99.5% | 99.84% |
| **Standard Deviation** | 0.5% | | | |

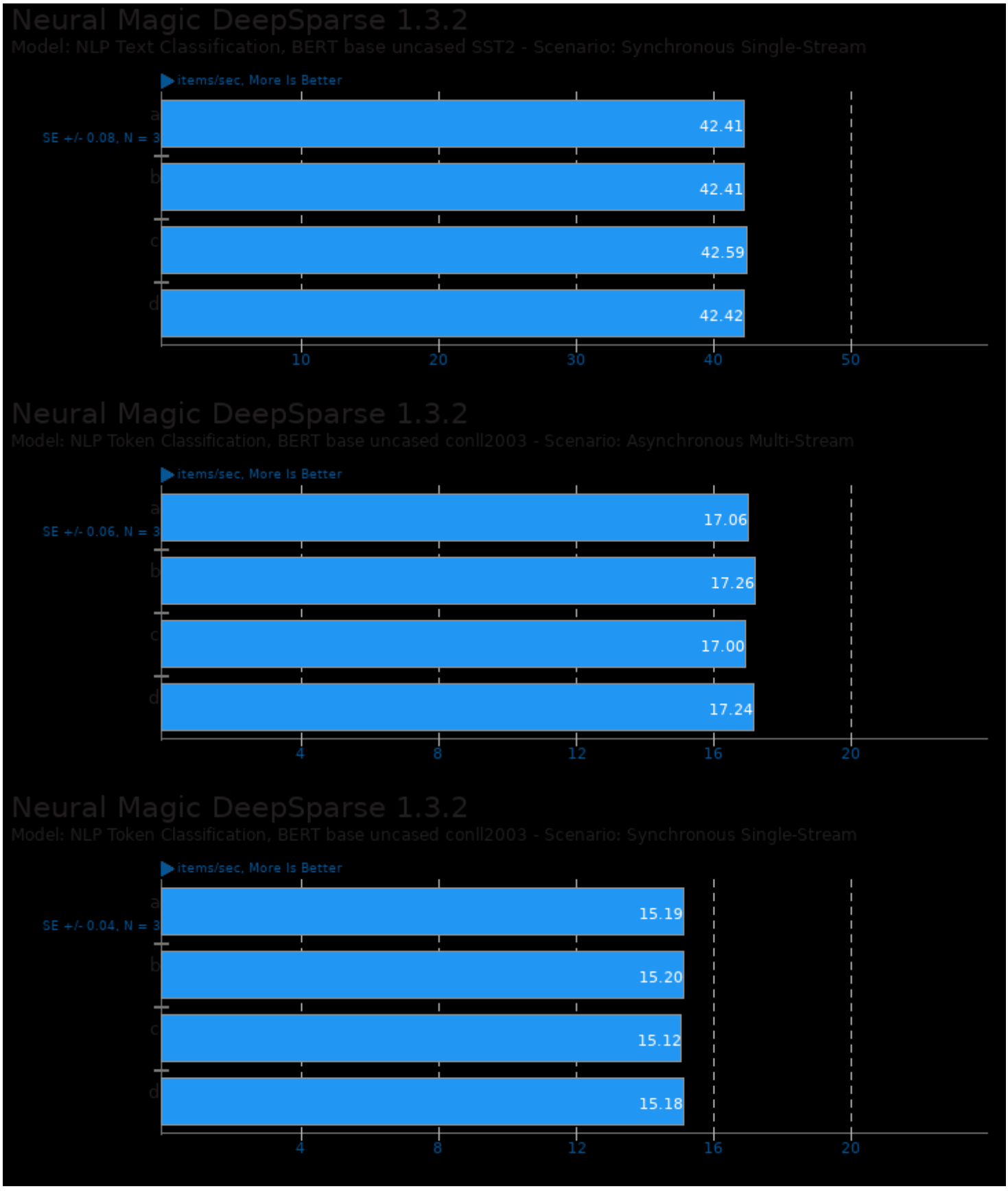

Neural Magic DeepSparse 1.3.2
Model: NLP Document Classification, oBERT base uncased on IMDB - Scenario: Asynchronous Multi-Stream
items/sec, More Is Better
SE +/- 0.09, N = 3
a
17.10
b
17.03
c
17.13
d
17.26
4
8
12
16
20
Neural Magic DeepSparse 1.3.2
Model: NLP Document Classification, oBERT base uncased on IMDB - Scenario: Synchronous Single-Stream
items/sec, More Is Better
SE +/- 0.02, N = 3
a
15.15
b
15.15
c
15.14
d
15.16
4
8
12
16
20
Neural Magic DeepSparse 1.3.2
Model: NLP Sentiment Analysis, 80% Pruned Quantized BERT Base Uncased - Scenario: Asynchronous Multi-Stream
items/sec, More Is Better
SE +/- 2.21, N = 3
a
201.89
b
202.87
c
202.74
d
203.72
40
80
120
160
200

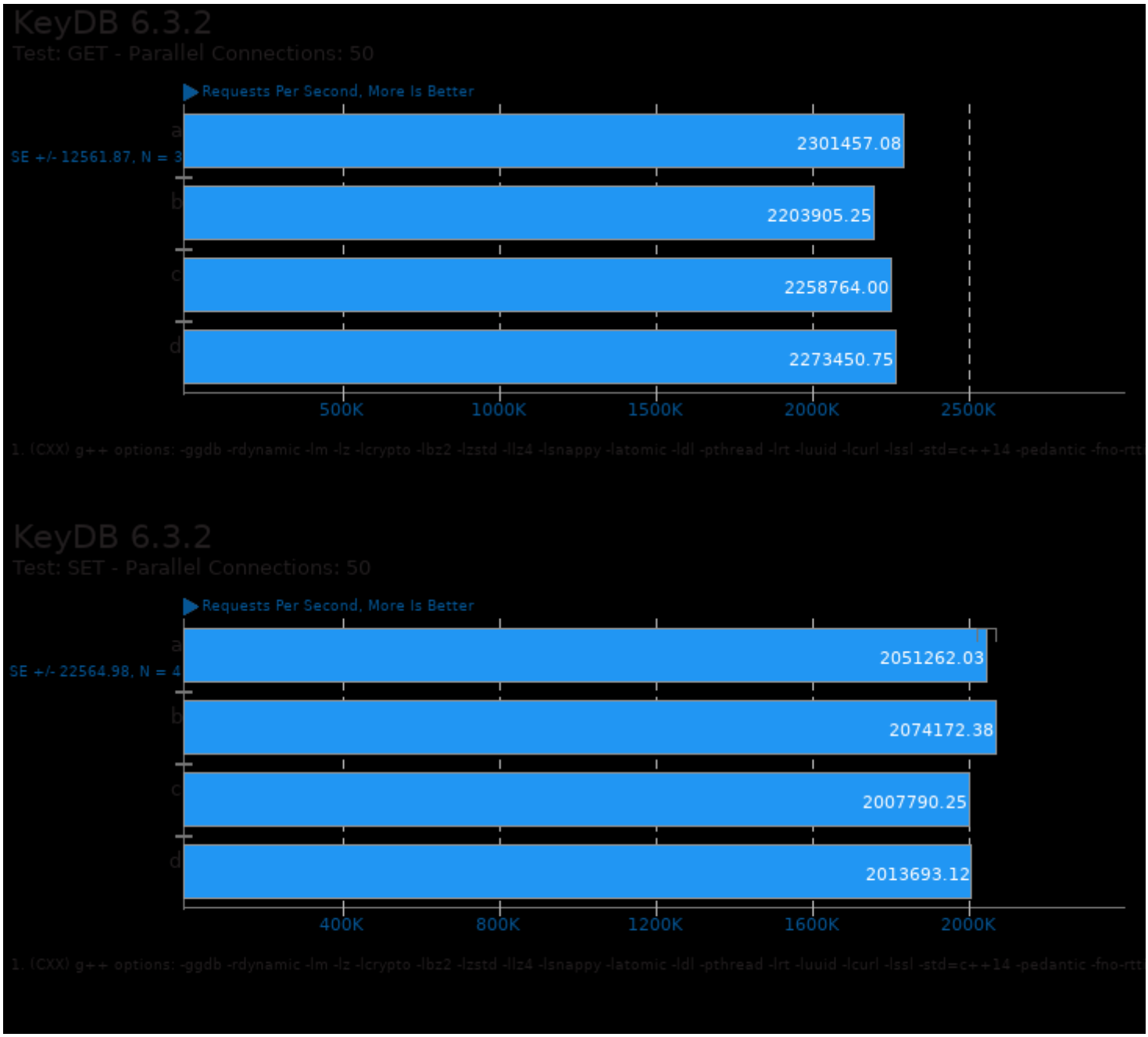

Neural Magic DeepSparse 1.3.2
Model: NLP Sentiment Analysis, 80% Pruned Quantized BERT Base Uncased - Scenario: Synchronous Single-Stream
items/sec, More Is Better
SE +/- 0.18, N = 3
a
97.03
b
96.60
c
97.70
d
97.05
20
40
60
80
100
Neural Magic DeepSparse 1.3.2
Model: NLP Question Answering, BERT base uncased SQuaD 12layer Pruned90 - Scenario: Asynchronous Multi-Stream
items/sec, More Is Better
SE +/- 0.15, N = 3
a
62.47
b
62.03
c
62.43
d
62.94
14
28
42
56
70
Neural Magic DeepSparse 1.3.2
Model: NLP Question Answering, BERT base uncased SQuaD 12layer Pruned90 - Scenario: Synchronous Single-Stream
items/sec, More Is Better
SE +/- 0.08, N = 3
a
44.41
b
44.35
c
44.20
d
44.38
10
20
30
40
50

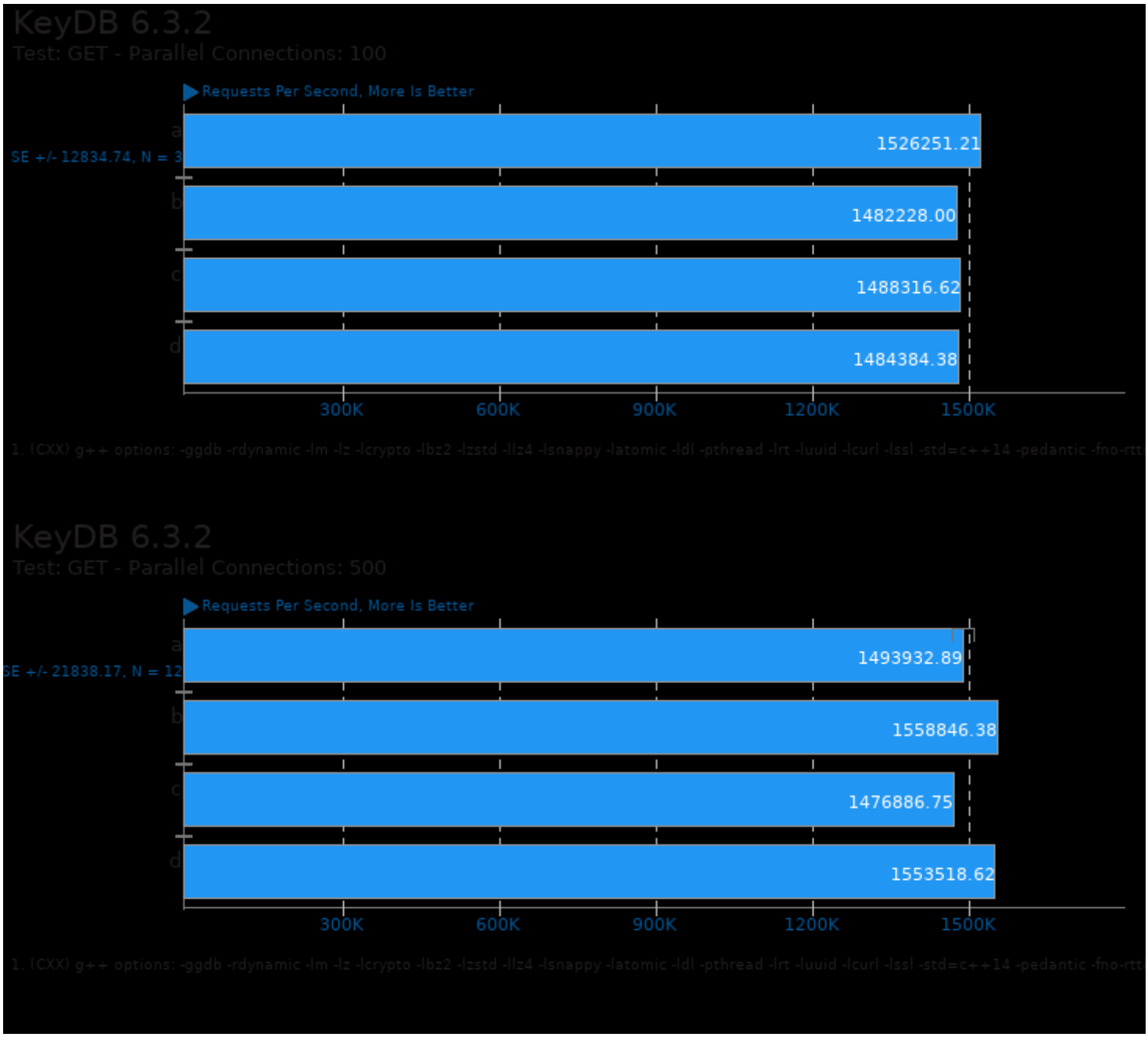

Neural Magic DeepSparse 1.3.2
Model: CV Detection, YOLOv5s COCO - Scenario: Asynchronous Multi-Stream
items/sec, More Is Better
SE +/- 0.45, N = 3
a
100.60
b
102.19
c
100.92
d
101.58
20
40
60
80
100
Neural Magic DeepSparse 1.3.2
Model: CV Detection, YOLOv5s COCO - Scenario: Synchronous Single-Stream
items/sec, More Is Better
SE +/- 0.33, N = 3
a
67.85
b
68.41
c
68.23
d
69.23
15
30
45
60
75
Neural Magic DeepSparse 1.3.2
Model: CV Classification, ResNet-50 ImageNet - Scenario: Asynchronous Multi-Stream
items/sec, More Is Better
SE +/- 1.42, N = 3
a
220.22
b
223.11
c
222.58
d
220.63
50
100
150
200
250

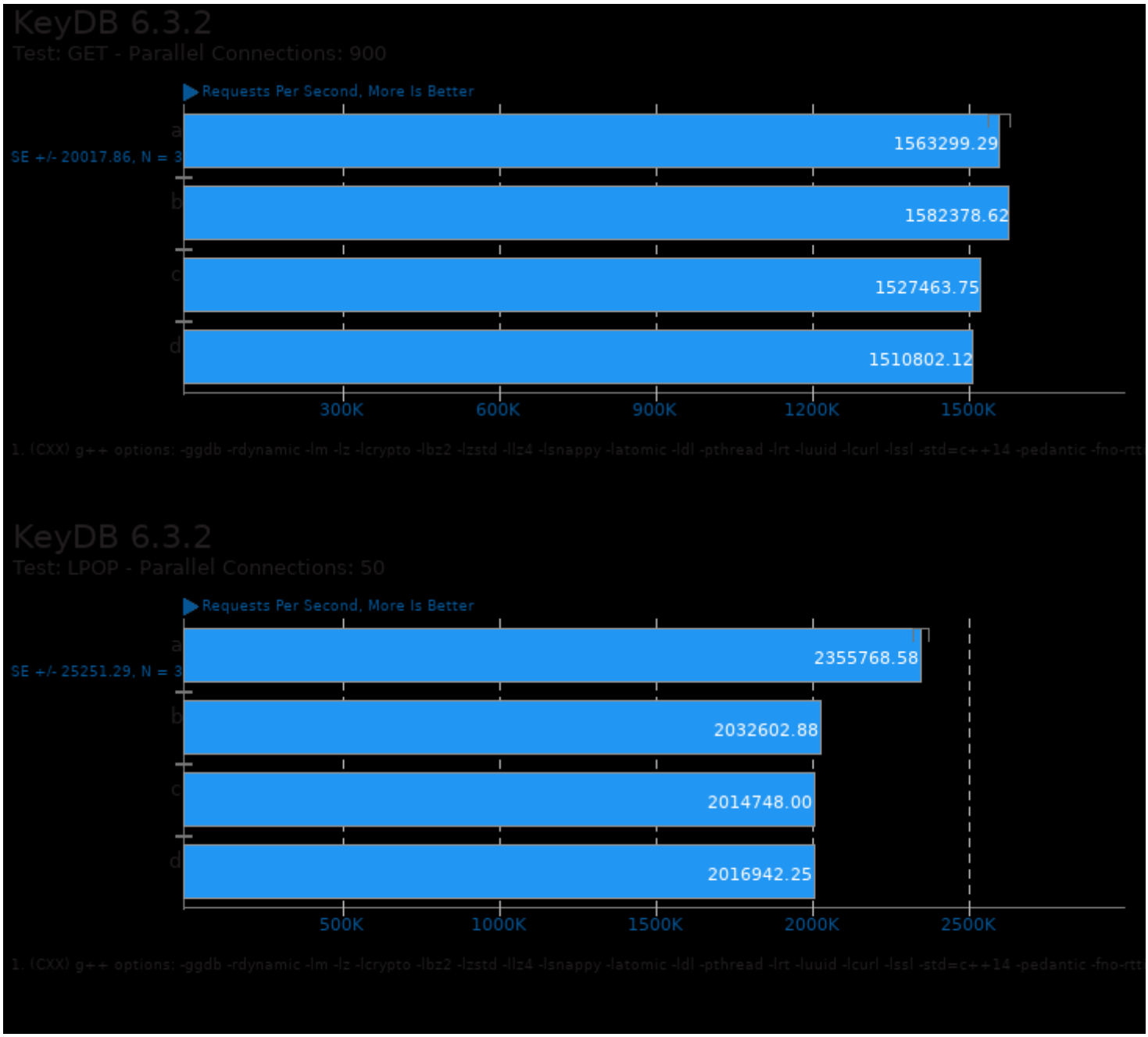

Neural Magic DeepSparse 1.3.2
Model: CV Classification, ResNet-50 ImageNet - Scenario: Synchronous Single-Stream
items/sec, More Is Better
SE +/- 0.33, N = 3
a
125.70
b
126.99
c
126.71
d
126.85
30
60
90
120
150
Neural Magic DeepSparse 1.3.2
Model: NLP Text Classification, DistilBERT mnli - Scenario: Asynchronous Multi-Stream
items/sec, More Is Better
SE +/- 0.16, N = 3
a
146.84
b
146.76
c
148.98
d
146.71
30
60
90
120
150
Neural Magic DeepSparse 1.3.2
Model: NLP Text Classification, DistilBERT mnli - Scenario: Synchronous Single-Stream
items/sec, More Is Better
SE +/- 0.18, N = 3
a
84.21
b
84.34
c
84.62
d
84.70
20
40
60
80
100

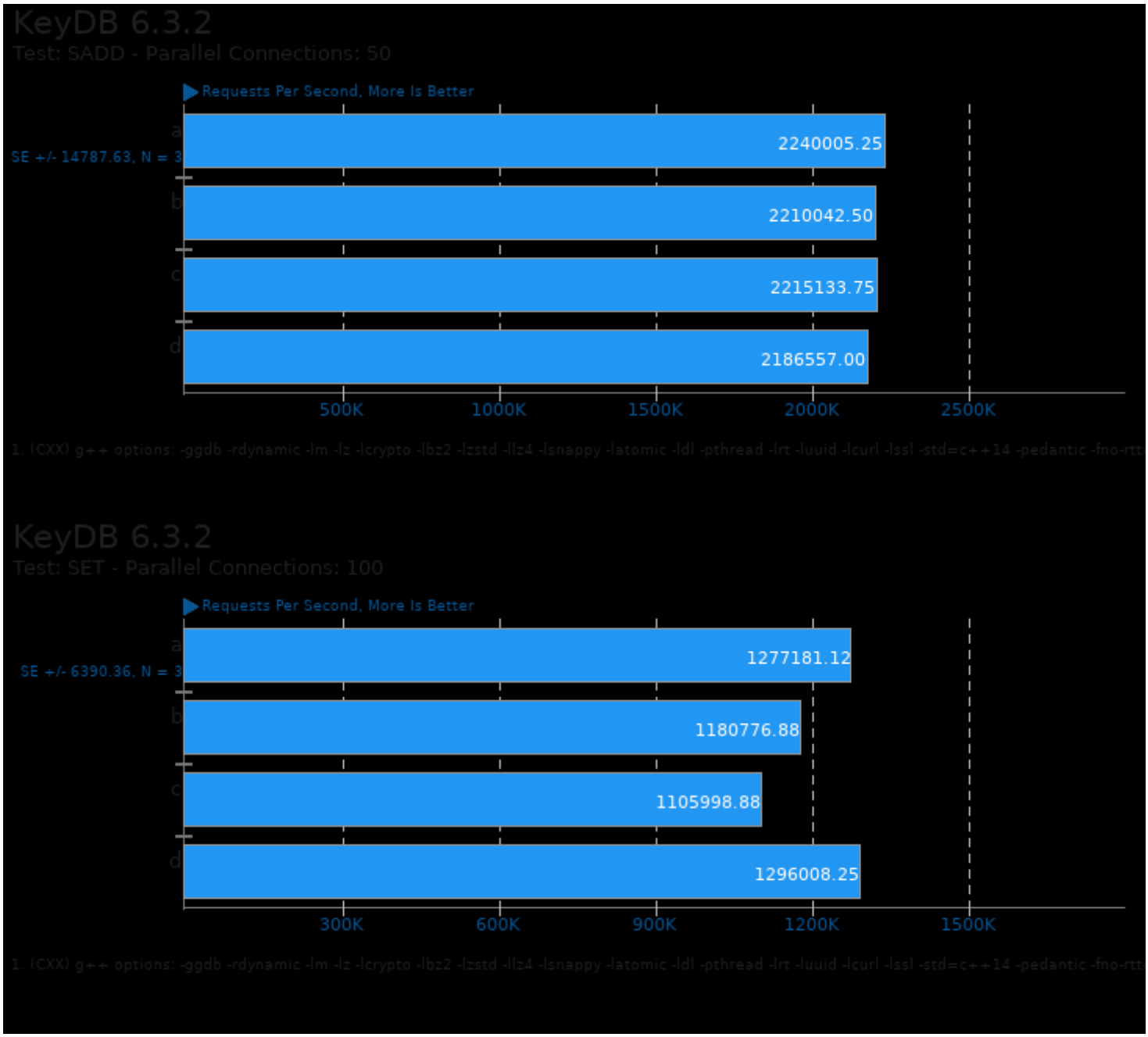

Neural Magic DeepSparse 1.3.2
Model: CV Segmentation, 90% Pruned YOLACT Pruned - Scenario: Asynchronous Multi-Stream
items/sec, More Is Better
SE +/- 0.04, N = 3
a
22.46
b
22.57
c
22.48
d
22.88
5
10
15
20
25
Neural Magic DeepSparse 1.3.2
Model: CV Segmentation, 90% Pruned YOLACT Pruned - Scenario: Synchronous Single-Stream
items/sec, More Is Better
SE +/- 0.04, N = 3
a
17.92
b
17.97
c
17.83
d
17.90
4
8
12
16
20
Neural Magic DeepSparse 1.3.2
Model: NLP Text Classification, BERT base uncased SST2 - Scenario: Asynchronous Multi-Stream
items/sec, More Is Better
SE +/- 0.11, N = 3
a
74.68
b
73.96
c
75.21
d
74.44
20
40
60
80
100

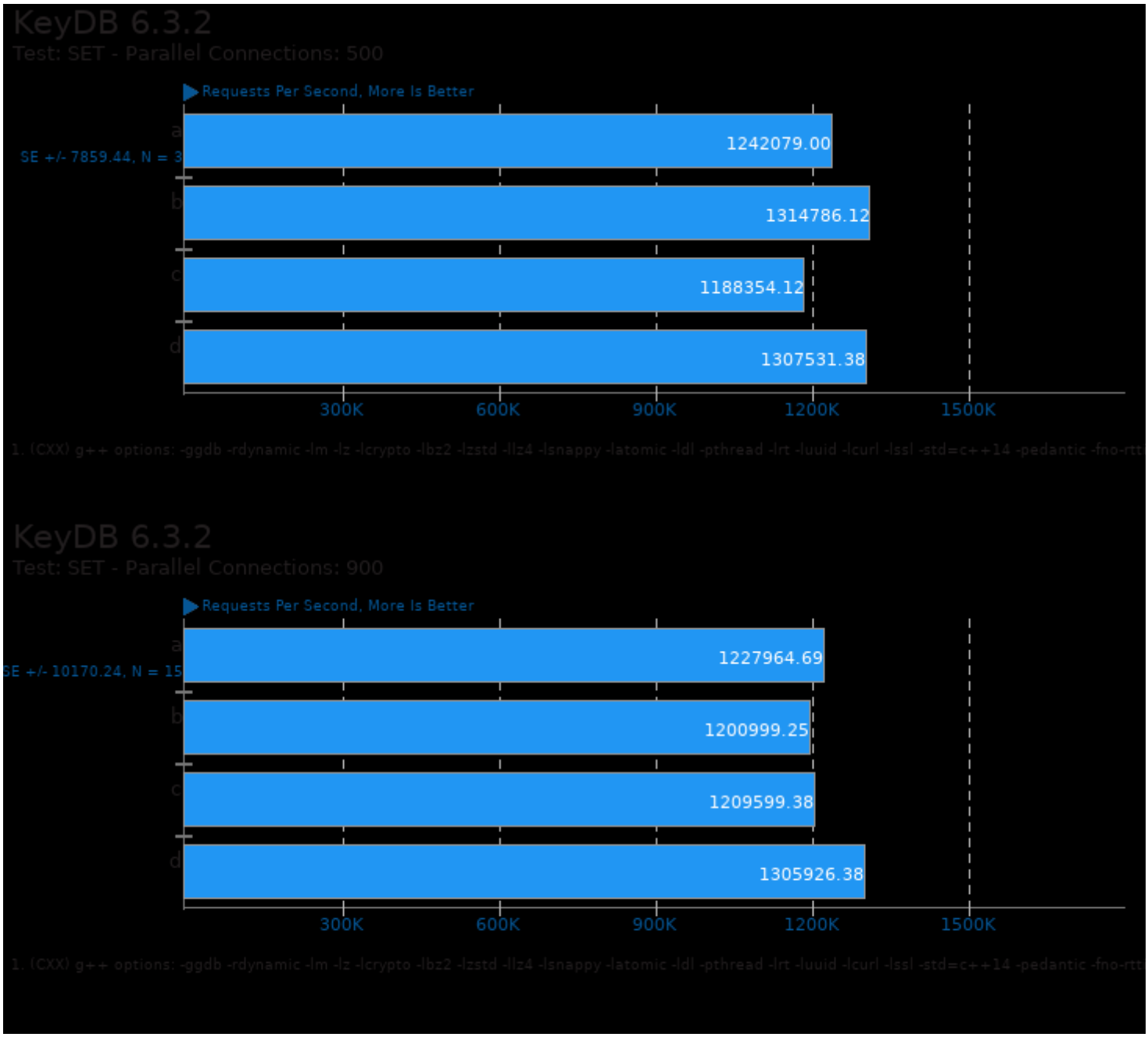

Neural Magic DeepSparse 1.3.2
Model: NLP Text Classification, BERT base uncased SST2 - Scenario: Synchronous Single-Stream
items/sec, More Is Better
SE +/- 0.08, N = 3
a
42.41
b
42.41
c
42.59
d
42.42
10
20
30
40
50
Neural Magic DeepSparse 1.3.2
Model: NLP Token Classification, BERT base uncased conll2003 - Scenario: Asynchronous Multi-Stream
items/sec, More Is Better
SE +/- 0.06, N = 3
a
17.06
b
17.26
c
17.00
d
17.24
4
8
12
16
20
Neural Magic DeepSparse 1.3.2
Model: NLP Token Classification, BERT base uncased conll2003 - Scenario: Synchronous Single-Stream
items/sec, More Is Better
SE +/- 0.04, N = 3
a
15.19
b
15.20
c
15.12
d
15.18
4
8
12
16
20

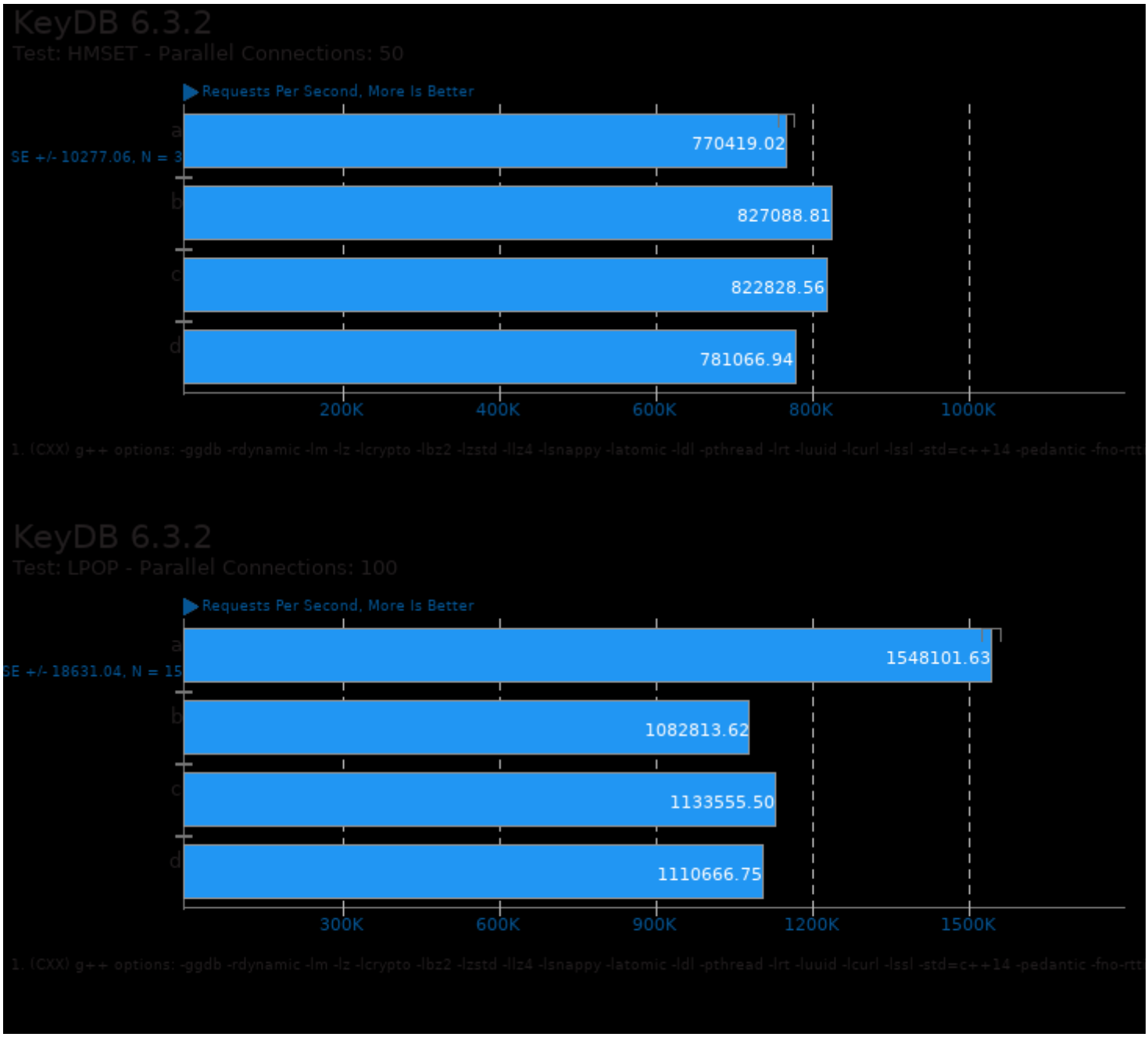

KeyDB 6.3.2
Test: GET - Parallel Connections: 50
Requests Per Second, More Is Better
SE +/- 12561.87, N = 3
a
2301457.08
b
2203905.25
c
2258764.00
d
2273450.75
500K
1000K
1500K
2000K
2500K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtt
KeyDB 6.3.2
Test: SET - Parallel Connections: 50
Requests Per Second, More Is Better
SE +/- 22564.98, N = 4
a
2051262.03
b
2074172.38
c
2007790.25
d
2013693.12
400K
800K
1200K
1600K
2000K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtt

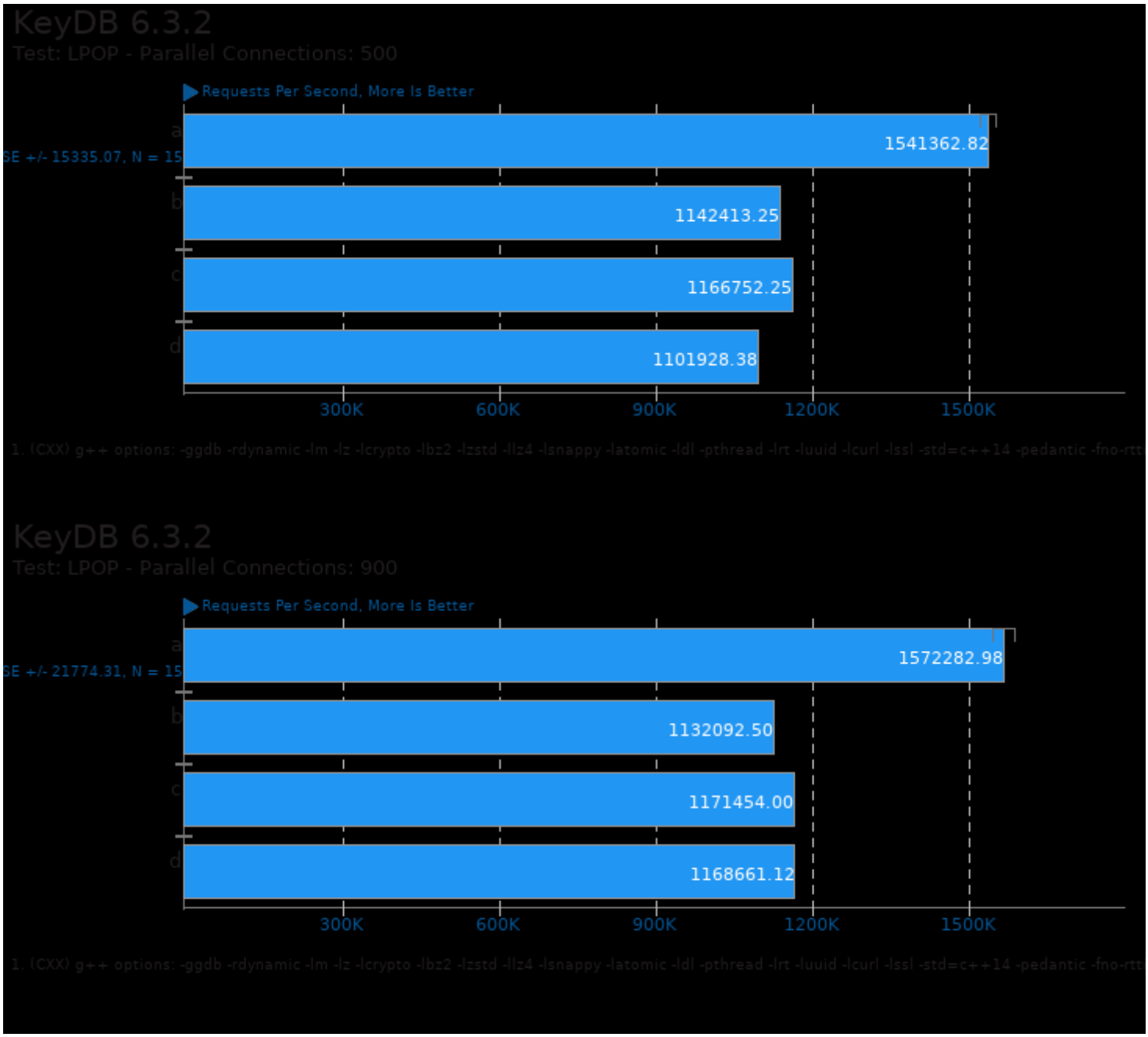

KeyDB 6.3.2
Test: GET - Parallel Connections: 100
Requests Per Second, More Is Better
SE +/- 12834.74, N = 3
a
1526251.21
b
1482228.00
c
1488316.62
d
1484384.38
300K
600K
900K
1200K
1500K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtt
KeyDB 6.3.2
Test: GET - Parallel Connections: 500
Requests Per Second, More Is Better
SE +/- 21838.17, N = 12
a
1493932.89
b
1558846.38
c
1476886.75
d
1553518.62
300K
600K
900K
1200K
1500K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtt

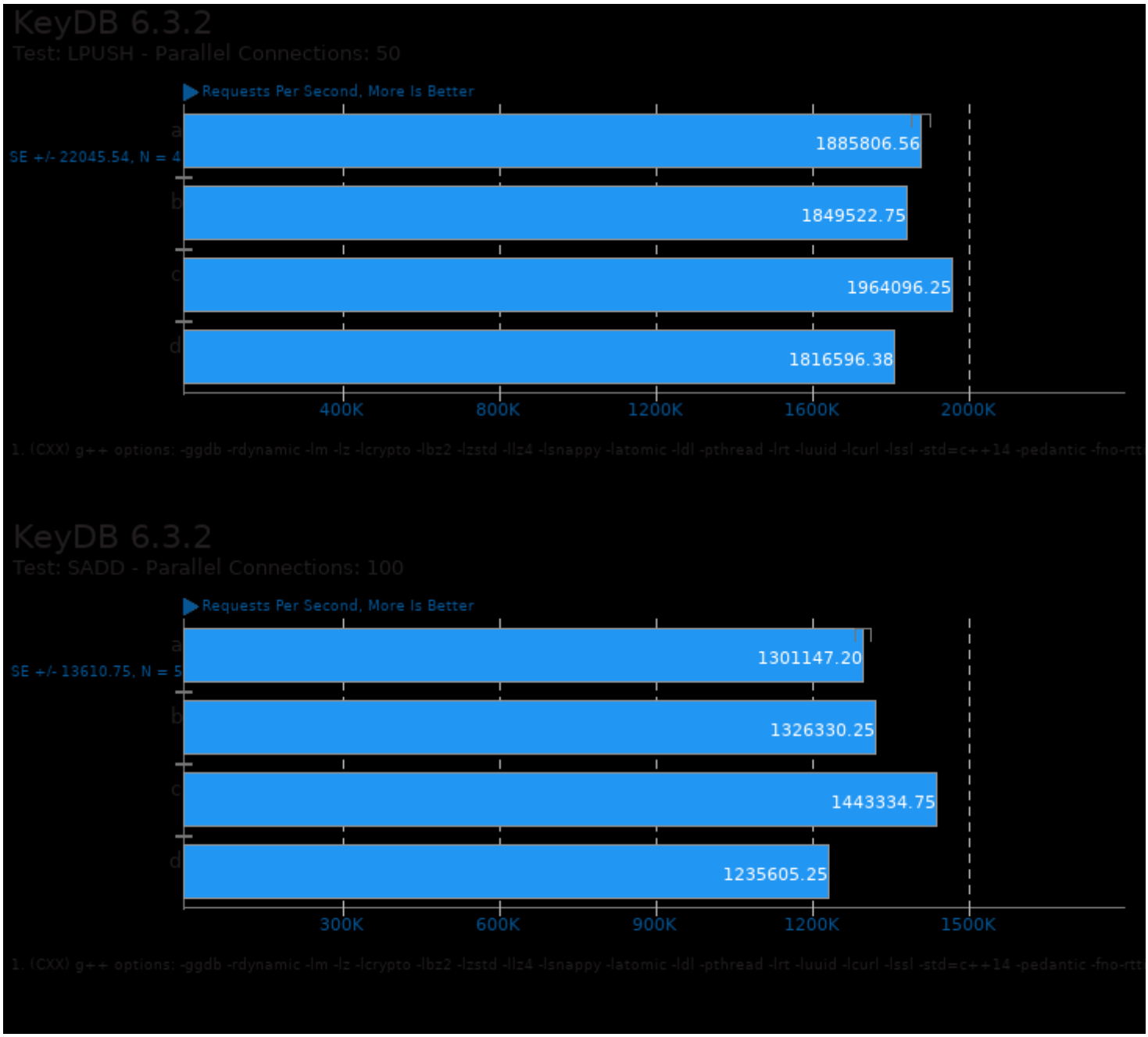

KeyDB 6.3.2
Test: GET - Parallel Connections: 900
Requests Per Second, More Is Better
SE +/- 20017.86, N = 3
a
b
c
d
1563299.29
1582378.62
1527463.75
1510802.12
300K
600K
900K
1200K
1500K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtt
KeyDB 6.3.2
Test: LPOP - Parallel Connections: 50
Requests Per Second, More Is Better
SE +/- 25251.29, N = 3
a
b
c
d
2355768.58
2032602.88
2014748.00
2016942.25
500K
1000K
1500K
2000K
2500K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtt

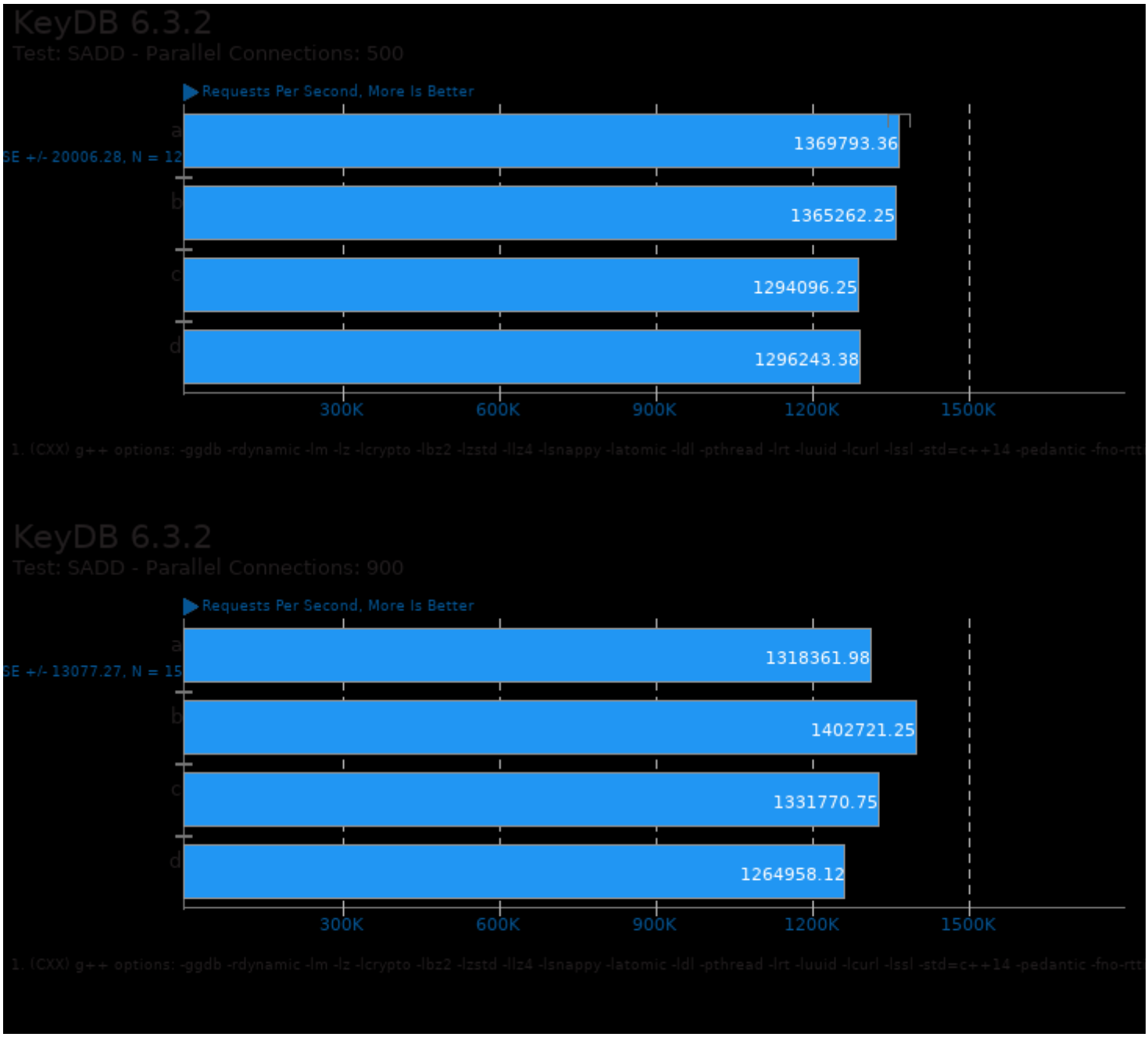

KeyDB 6.3.2
Test: SADD - Parallel Connections: 50
Requests Per Second, More Is Better
SE +/- 14787.63, N = 3
a
2240005.25
b
2210042.50
c
2215133.75
d
2186557.00
500K
1000K
1500K
2000K
2500K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtt
KeyDB 6.3.2
Test: SET - Parallel Connections: 100
Requests Per Second, More Is Better
SE +/- 6390.36, N = 3
a
1277181.12
b
1180776.88
c
1105998.88
d
1296008.25
300K
600K
900K
1200K
1500K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtt

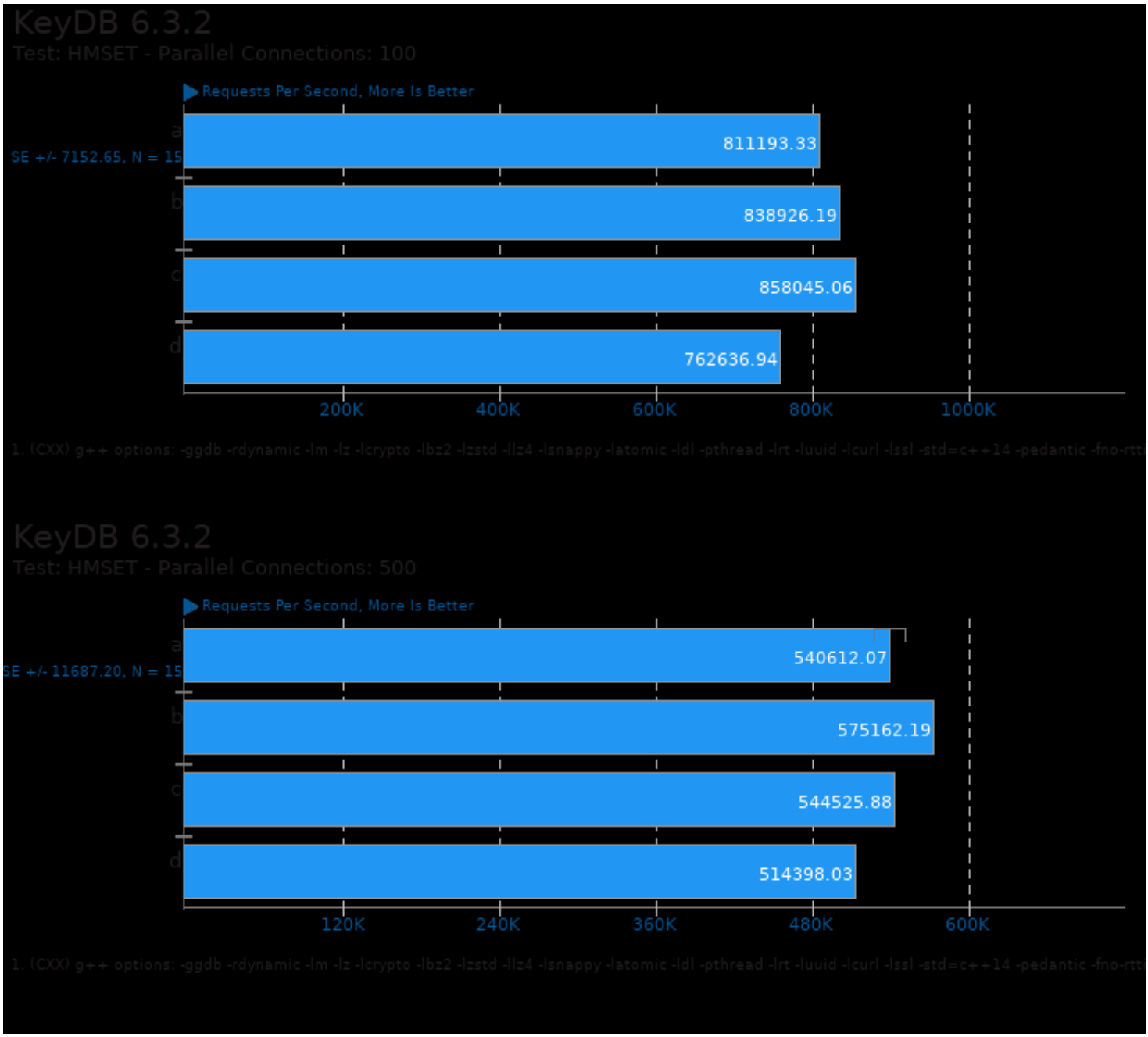

KeyDB 6.3.2
Test: SET - Parallel Connections: 500
Requests Per Second, More Is Better
SE +/- 7859.44, N = 3
a
1242079.00
b
1314786.12
c
1188354.12
d
1307531.38
300K
600K
900K
1200K
1500K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtt
KeyDB 6.3.2
Test: SET - Parallel Connections: 900
Requests Per Second, More Is Better
SE +/- 10170.24, N = 15
a
1227964.69
b
1200999.25
c
1209599.38
d
1305926.38
300K
600K
900K
1200K
1500K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtt

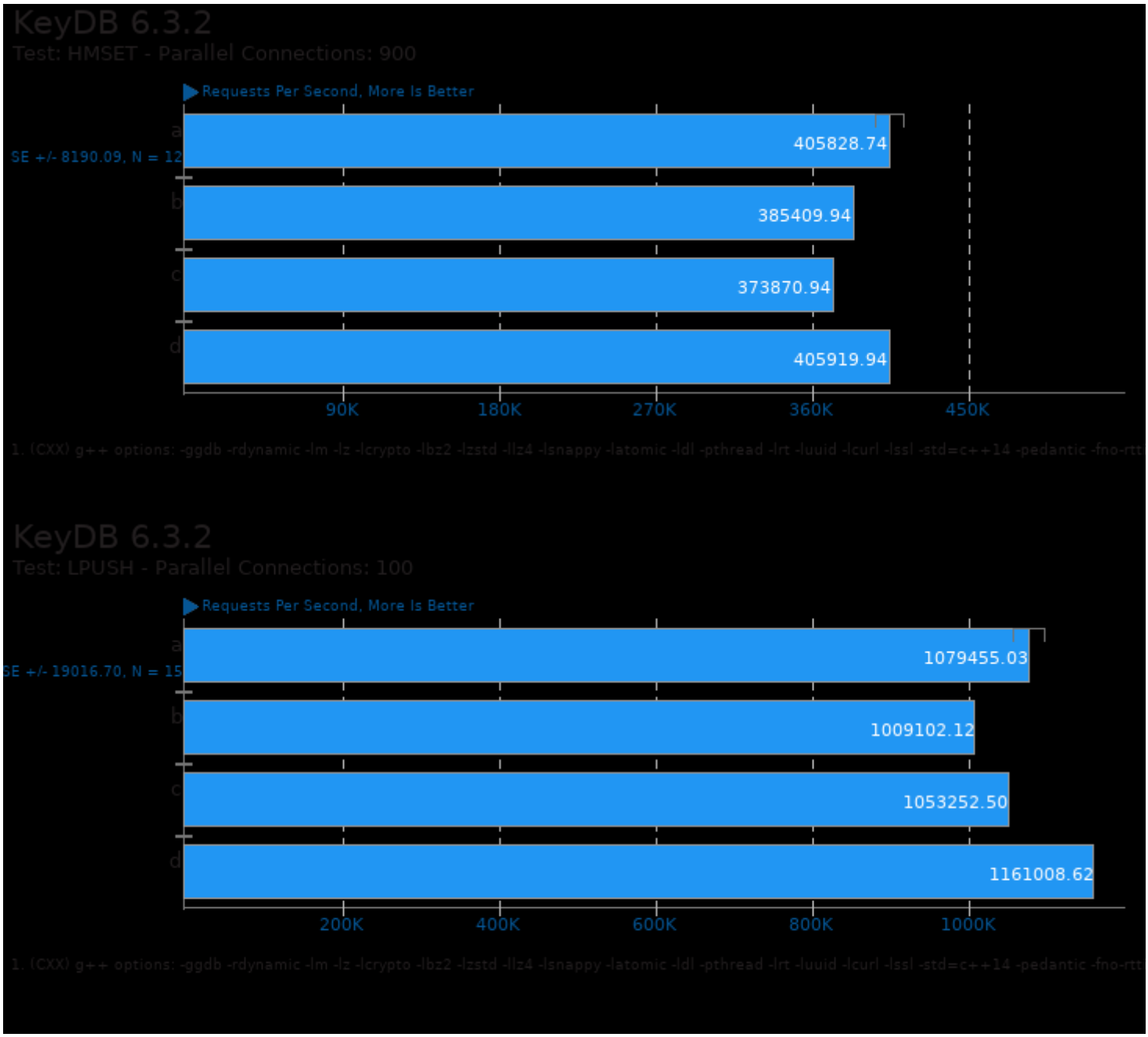

KeyDB 6.3.2
Test: HMSET - Parallel Connections: 50
Requests Per Second, More Is Better
SE +/- 10277.06, N = 3
a
770419.02
b
827088.81
c
822828.56
d
781066.94
200K
400K
600K
800K
1000K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtt
KeyDB 6.3.2
Test: LPOP - Parallel Connections: 100
Requests Per Second, More Is Better
SE +/- 18631.04, N = 15
a
1548101.63
b
1082813.62
c
1133555.50
d
1110666.75
300K
600K
900K
1200K
1500K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtt

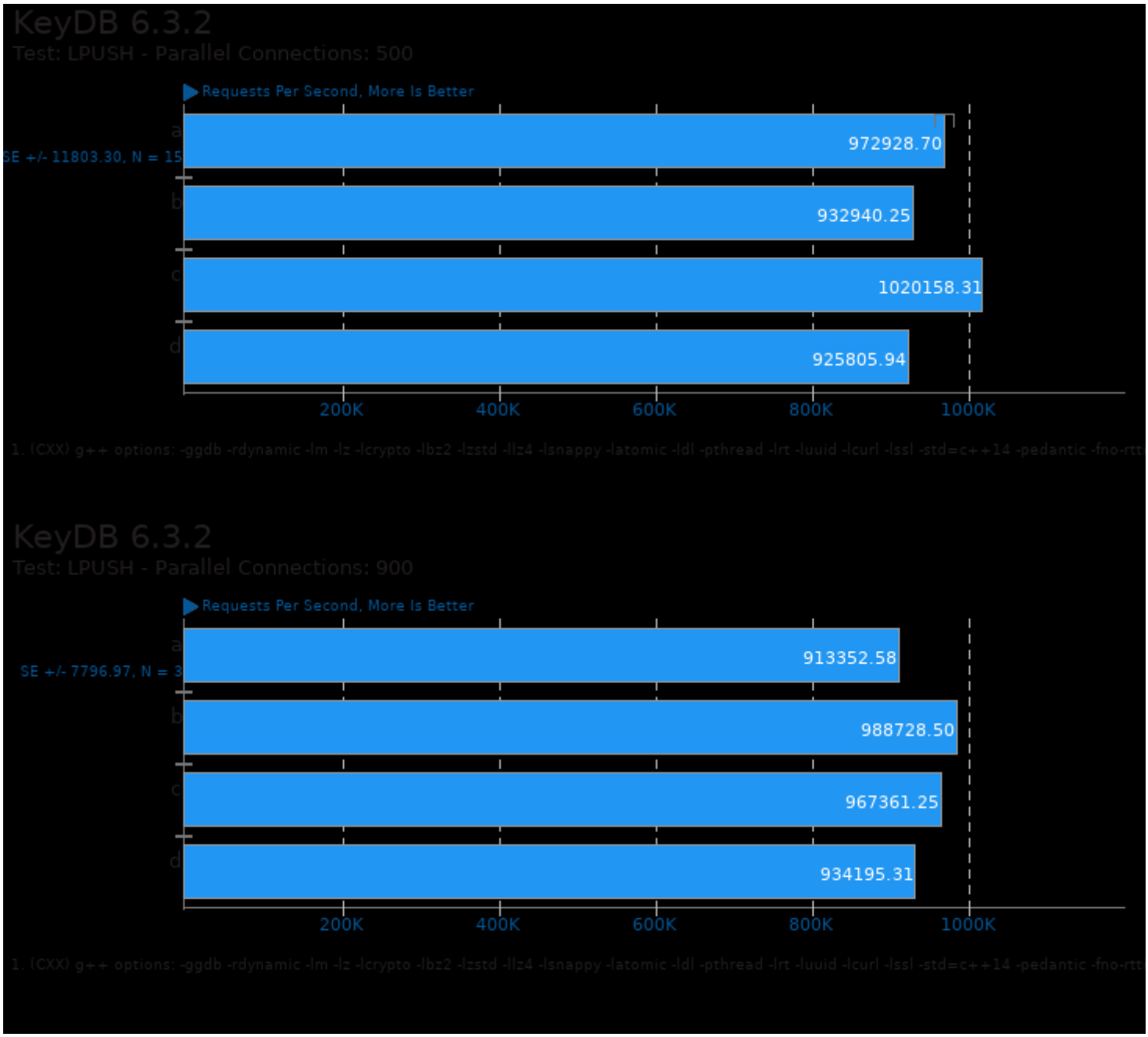

KeyDB 6.3.2
Test: LPOP - Parallel Connections: 500
Requests Per Second, More Is Better
SE +/- 15335.07, N = 15
a
b
c
d
1541362.82
1142413.25
1166752.25
1101928.38
300K
600K
900K
1200K
1500K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtt
KeyDB 6.3.2
Test: LPOP - Parallel Connections: 900
Requests Per Second, More Is Better
SE +/- 21774.31, N = 15
a
b
c
d
1572282.98
1132092.50
1171454.00
1168661.12
300K
600K
900K
1200K
1500K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtt

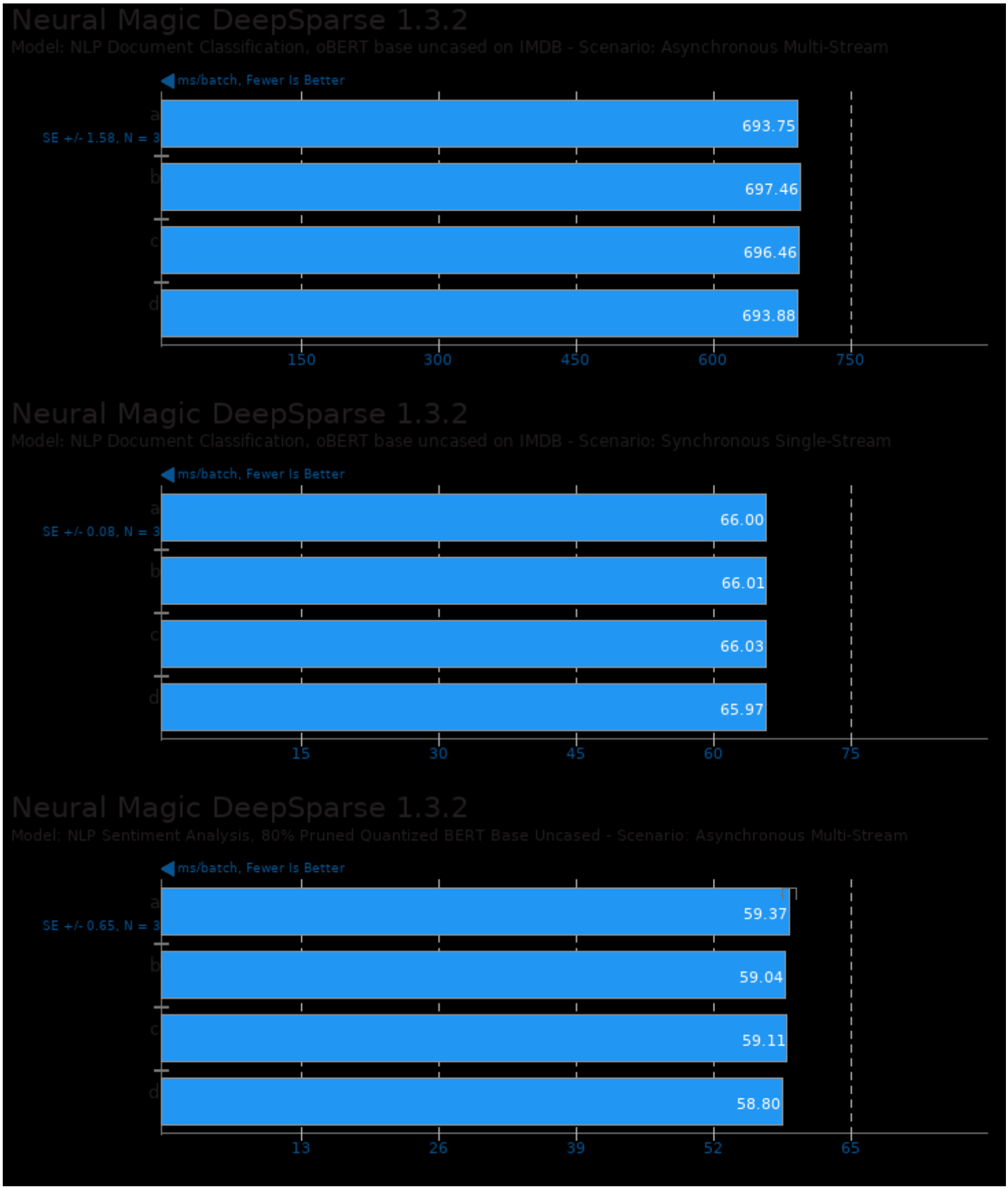

KeyDB 6.3.2
Test: LPUSH - Parallel Connections: 50
Requests Per Second, More Is Better
SE +/- 22045.54, N = 4
a
b
c
d
1885806.56
1849522.75
1964096.25
1816596.38
400K
800K
1200K
1600K
2000K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtt
KeyDB 6.3.2
Test: SADD - Parallel Connections: 100
Requests Per Second, More Is Better
SE +/- 13610.75, N = 5
a
b
c
d
1301147.20
1326330.25
1443334.75
1235605.25
300K
600K
900K
1200K
1500K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtt

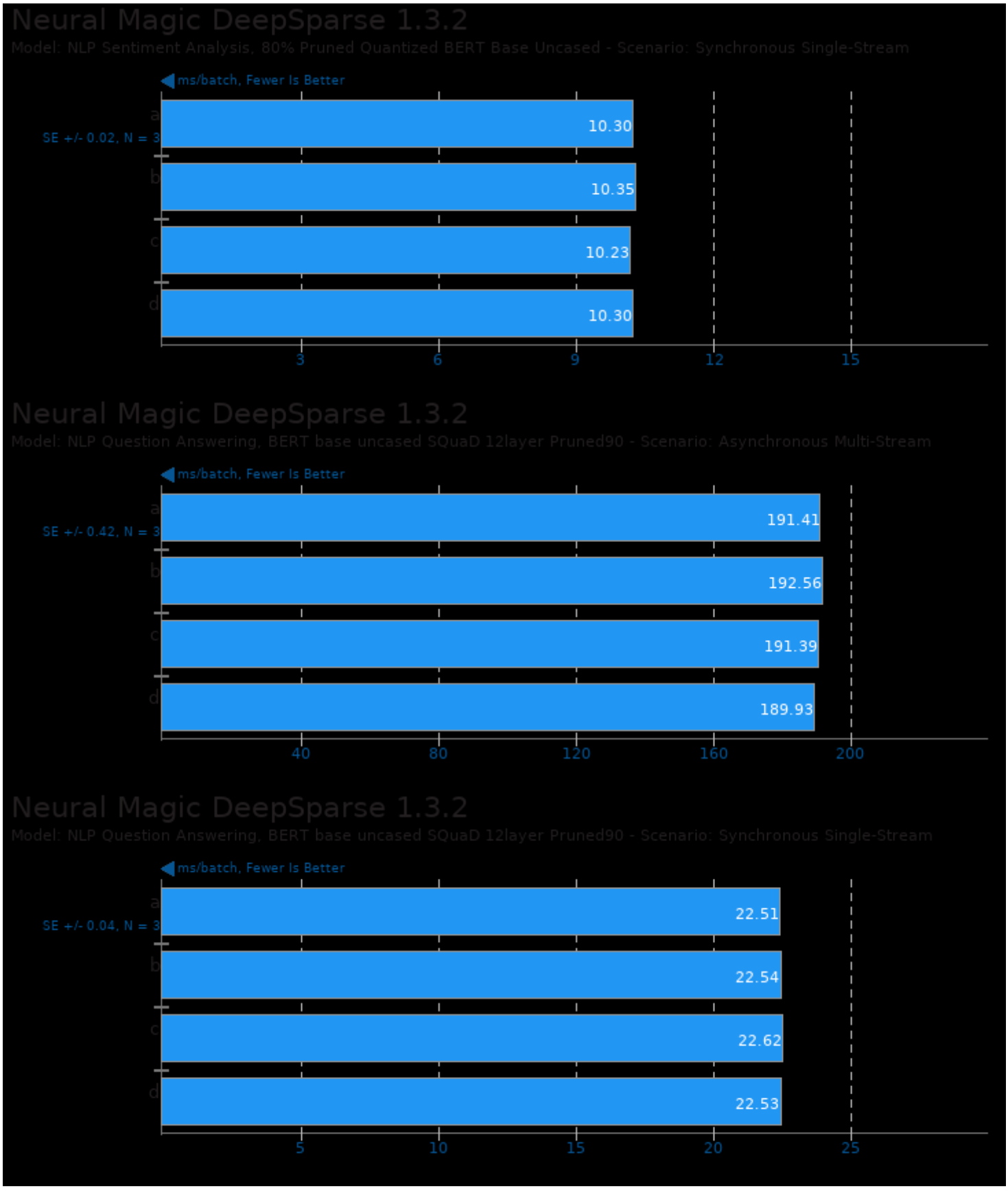

KeyDB 6.3.2
Test: SADD - Parallel Connections: 500
Requests Per Second, More Is Better
a
SE +/- 20006.28, N = 12
1369793.36
b
1365262.25
c
1294096.25
d
1296243.38
300K
600K
900K
1200K
1500K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtt
KeyDB 6.3.2
Test: SADD - Parallel Connections: 900
Requests Per Second, More Is Better
a
SE +/- 13077.27, N = 15
1318361.98
b
1402721.25
c
1331770.75
d
1264958.12
300K
600K
900K
1200K
1500K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtt

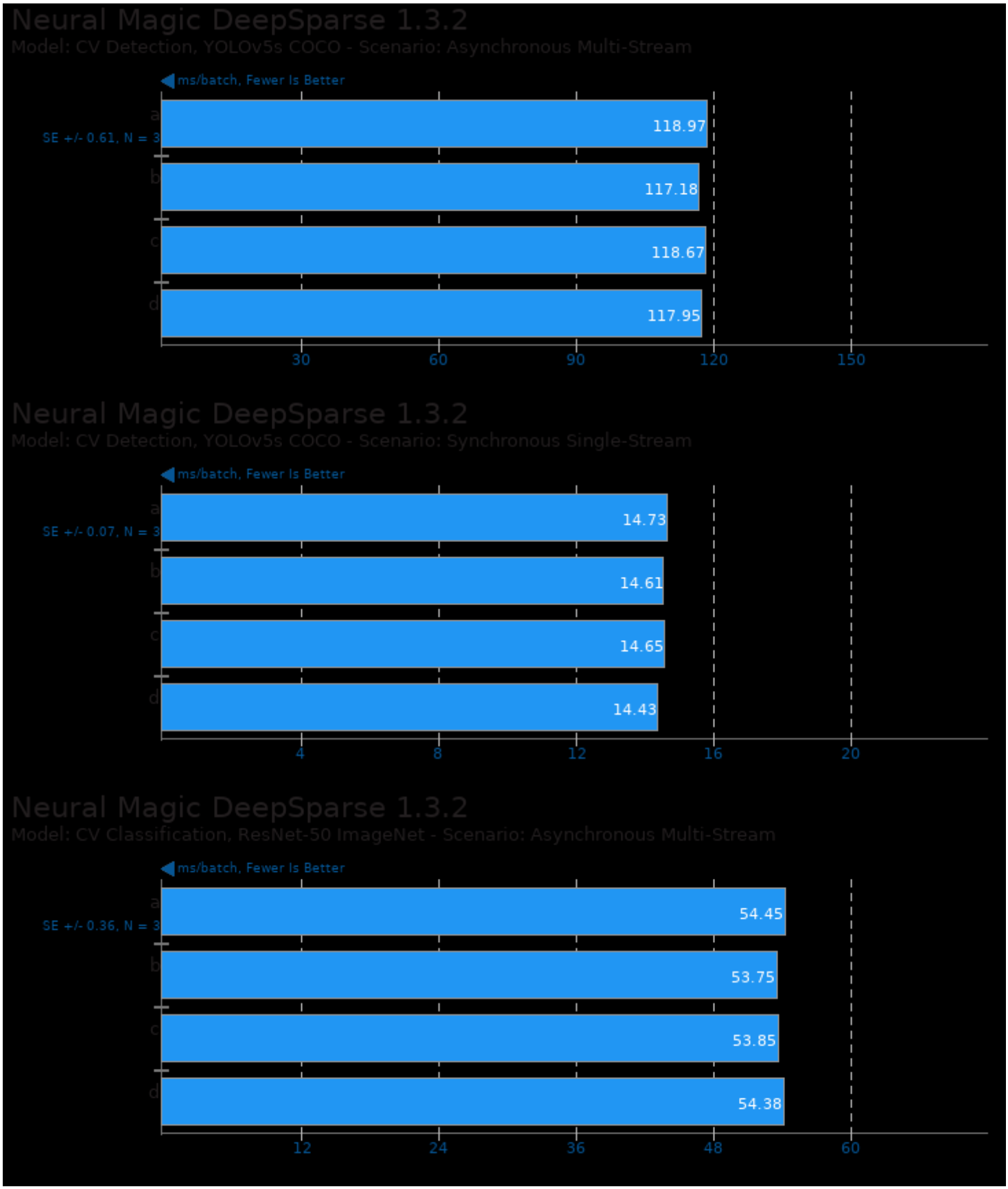

KeyDB 6.3.2
Test: HMSET - Parallel Connections: 100
Requests Per Second, More Is Better
a
b
c
d
811193.33
838926.19
858045.06
762636.94
200K
400K
600K
800K
1000K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtti
KeyDB 6.3.2
Test: HMSET - Parallel Connections: 500
Requests Per Second, More Is Better
a
b
c
d
540612.07
575162.19
544525.88
514398.03
120K
240K
360K
480K
600K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtti

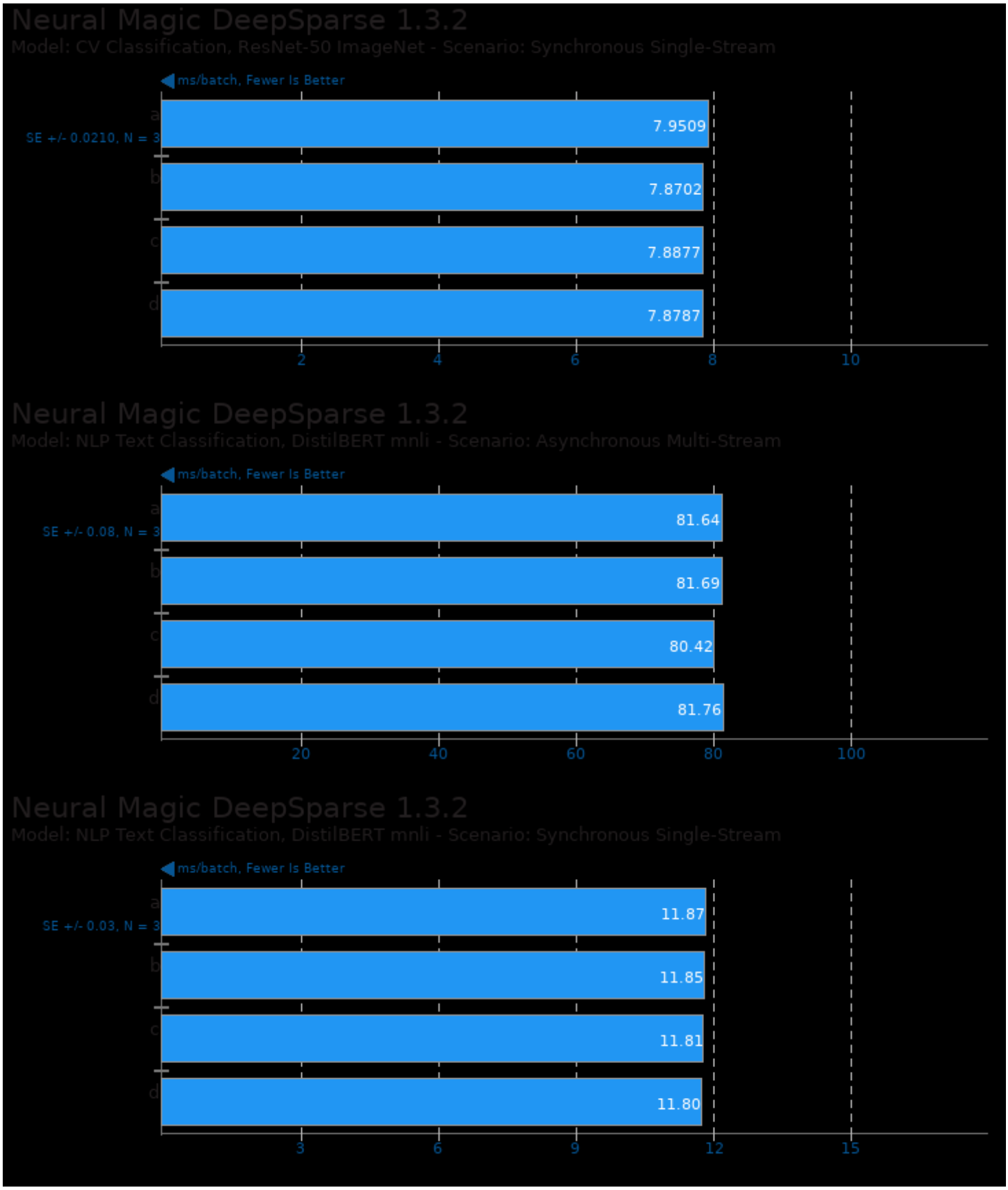

KeyDB 6.3.2
Test: HMSET - Parallel Connections: 900
Requests Per Second, More Is Better
SE +/- 8190.09, N = 12
405828.74
385409.94
373870.94
405919.94
90K
180K
270K
360K
450K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtt
KeyDB 6.3.2
Test: LPUSH - Parallel Connections: 100
Requests Per Second, More Is Better
SE +/- 19016.70, N = 15
1079455.03
1009102.12
1053252.50
1161008.62
200K
400K
600K
800K
1000K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtt

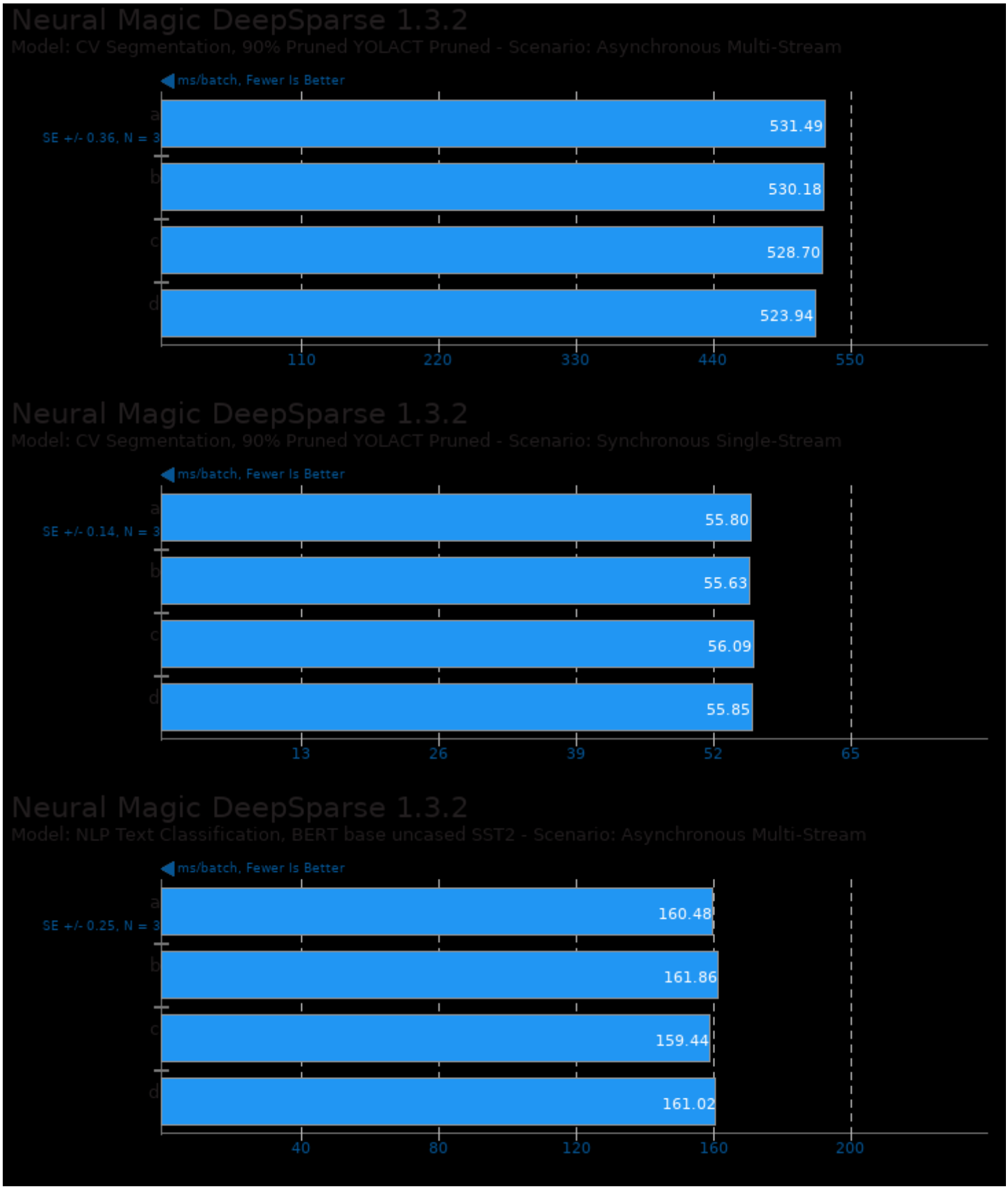

KeyDB 6.3.2
Test: LPUSH - Parallel Connections: 500
Requests Per Second, More Is Better
SE +/- 11803.30, N = 15
a
b
c
d
972928.70
932940.25
1020158.31
925805.94
200K
400K
600K
800K
1000K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtt
KeyDB 6.3.2
Test: LPUSH - Parallel Connections: 900
Requests Per Second, More Is Better
SE +/- 7796.97, N = 3
a
b
c
d
913352.58
988728.50
967361.25
934195.31
200K
400K
600K
800K
1000K
1. (CXX) g++ options: -ggdb -rdynamic -lm -lz -lcrypto -lbz2 -lzstd -llz4 -lsnappy -latomic -ldl -pthread -lrt -luuid -lcurl -lssl -std=c++14 -pedantic -fno-rtt

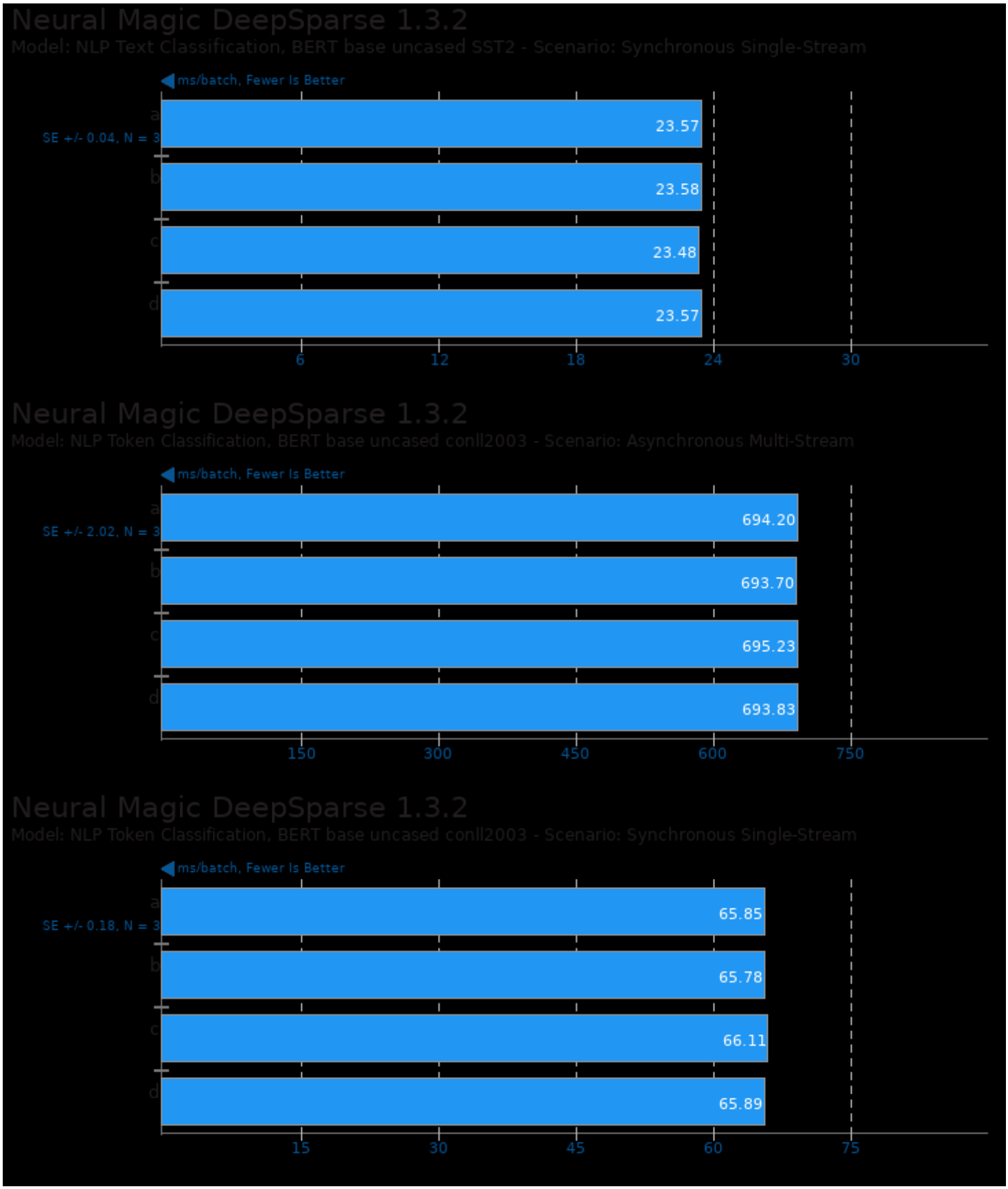

Neural Magic DeepSparse 1.3.2
Model: NLP Document Classification, oBERT base uncased on IMDB - Scenario: Asynchronous Multi-Stream
ms/batch, Fewer Is Better
SE +/- 1.58, N = 3
a
693.75
b
697.46
c
696.46
d
693.88
150
300
450
600
750
Neural Magic DeepSparse 1.3.2
Model: NLP Document Classification, oBERT base uncased on IMDB - Scenario: Synchronous Single-Stream
ms/batch, Fewer Is Better
SE +/- 0.08, N = 3
a
66.00
b
66.01
c
66.03
d
65.97
15
30
45
60
75
Neural Magic DeepSparse 1.3.2
Model: NLP Sentiment Analysis, 80% Pruned Quantized BERT Base Uncased - Scenario: Asynchronous Multi-Stream
ms/batch, Fewer Is Better
SE +/- 0.65, N = 3
a
59.37
b
59.04
c
59.11
d
58.80
13
26
39
52
65


Neural Magic DeepSparse 1.3.2
Model: NLP Sentiment Analysis, 80% Pruned Quantized BERT Base Uncased - Scenario: Synchronous Single-Stream
ms/batch, Fewer Is Better
SE +/- 0.02, N = 3
a
10.30
b
10.35
c
10.23
d
10.30
3
6
9
12
15
Neural Magic DeepSparse 1.3.2
Model: NLP Question Answering, BERT base uncased SQuaD 12layer Pruned90 - Scenario: Asynchronous Multi-Stream
ms/batch, Fewer Is Better
SE +/- 0.42, N = 3
a
191.41
b
192.56
c
191.39
d
189.93
40
80
120
160
200
Neural Magic DeepSparse 1.3.2
Model: NLP Question Answering, BERT base uncased SQuaD 12layer Pruned90 - Scenario: Synchronous Single-Stream
ms/batch, Fewer Is Better
SE +/- 0.04, N = 3
a
22.51
b
22.54
c
22.62
d
22.53
5
10
15
20
25


Neural Magic DeepSparse 1.3.2
Model: CV Detection, YOLOv5s COCO - Scenario: Asynchronous Multi-Stream
ms/batch, Fewer Is Better
SE +/- 0.61, N = 3
a
118.97
b
117.18
c
118.67
d
117.95
30
60
90
120
150
Neural Magic DeepSparse 1.3.2
Model: CV Detection, YOLOv5s COCO - Scenario: Synchronous Single-Stream
ms/batch, Fewer Is Better
SE +/- 0.07, N = 3
a
14.73
b
14.61
c
14.65
d
14.43
4
8
12
16
20
Neural Magic DeepSparse 1.3.2
Model: CV Classification, ResNet-50 ImageNet - Scenario: Asynchronous Multi-Stream
ms/batch, Fewer Is Better
SE +/- 0.36, N = 3
a
54.45
b
53.75
c
53.85
d
54.38
12
24
36
48
60


Neural Magic DeepSparse 1.3.2
Model: CV Classification, ResNet-50 ImageNet - Scenario: Synchronous Single-Stream
ms/batch, Fewer Is Better
SE +/- 0.0210, N = 3
a
7.9509
b
7.8702
c
7.8877
d
7.8787
2
4
6
8
10
Neural Magic DeepSparse 1.3.2
Model: NLP Text Classification, DistilBERT mnli - Scenario: Asynchronous Multi-Stream
ms/batch, Fewer Is Better
SE +/- 0.08, N = 3
a
81.64
b
81.69
c
80.42
d
81.76
20
40
60
80
100
Neural Magic DeepSparse 1.3.2
Model: NLP Text Classification, DistilBERT mnli - Scenario: Synchronous Single-Stream
ms/batch, Fewer Is Better
SE +/- 0.03, N = 3
a
11.87
b
11.85
c
11.81
d
11.80
3
6
9
12
15


Neural Magic DeepSparse 1.3.2
Model: CV Segmentation, 90% Pruned YOLACT Pruned - Scenario: Asynchronous Multi-Stream
ms/batch, Fewer Is Better
SE +/- 0.36, N = 3
a
531.49
b
530.18
c
528.70
d
523.94
110
220
330
440
550
Neural Magic DeepSparse 1.3.2
Model: CV Segmentation, 90% Pruned YOLACT Pruned - Scenario: Synchronous Single-Stream
ms/batch, Fewer Is Better
SE +/- 0.14, N = 3
a
55.80
b
55.63
c
56.09
d
55.85
13
26
39
52
65
Neural Magic DeepSparse 1.3.2
Model: NLP Text Classification, BERT base uncased SST2 - Scenario: Asynchronous Multi-Stream
ms/batch, Fewer Is Better
SE +/- 0.25, N = 3
a
160.48
b
161.86
c
159.44
d
161.02
40
80
120
160
200




*This file was automatically generated via the Phoronix Test Suite benchmarking software on Friday, 29 March 2024 09:50.*