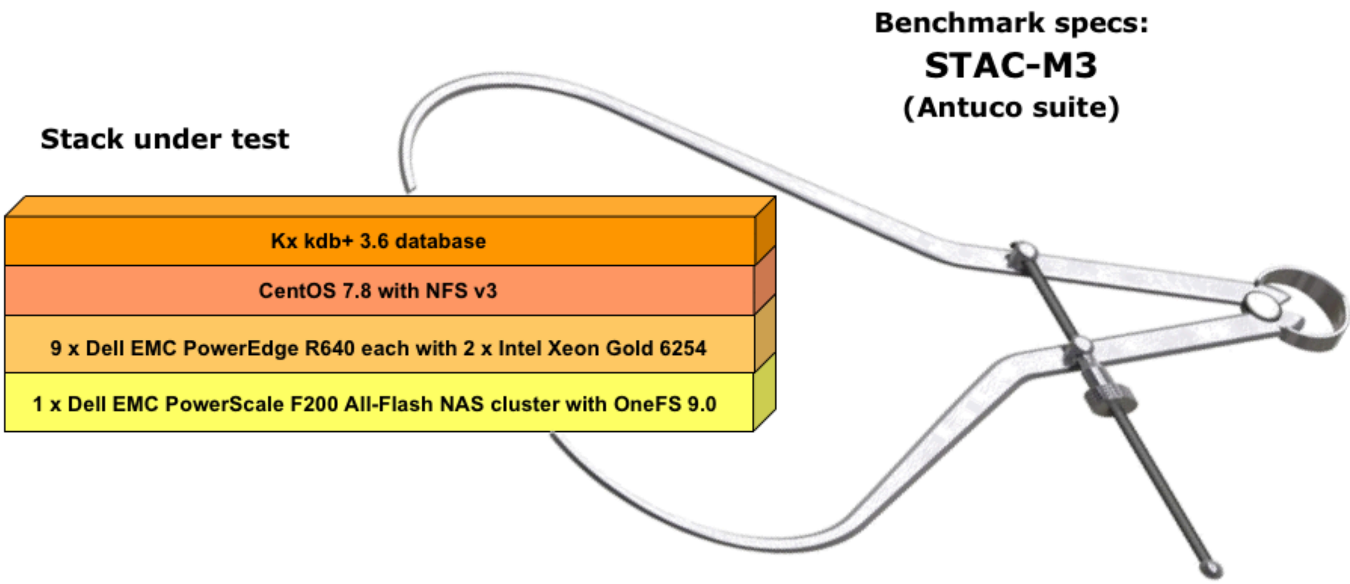
kdb+ 3.6 on 9 x Dell EMC PowerEdge R640 servers accessing a single 3-node Dell EMC PowerScale F200 All-Flash cluster

SUT ID: KDB200914

STAC-M3™ BENCHMARKS (Antuco Suite)

Test date: September 20, 2020
Draft 1.0, October 12, 2020



THESE TESTS FOLLOWED STAC BENCHMARK SPECIFICATIONS PROPOSED OR APPROVED BY THE STAC BENCHMARK COUNCIL (SEE WWW.STACRESEARCH.COM). BE SURE TO CHECK THE VERSION OF ANY SPECIFICATION USED IN A REPORT. DIFFERENT VERSIONS MAY NOT YIELD RESULTS THAT CAN BE COMPARED TO ONE ANOTHER.

This document was produced by the Securities Technology Analysis Center, LLC (STAC®), a provider of research and tools based on community-sourced benchmarks. To be notified of future reports or for more information, please visit www.STACresearch.com. Copyright © 2020, STAC. "STAC" and all STAC names are trademarks or registered trademarks of the Securities Technology Analysis Center, LLC. Other company and product names are trademarks of their respective owners.

Contents

About STAC 3
References 4
1. Summary 5
2. Vendor Commentary 6
3. Product background 6
4. Project participants and responsibilities 8
5. Contacts 8
6. Results status 9
7. Overview of the STAC-M3 Benchmark specifications 9
8. Specification particulars 16
8.1 Version 16
8.2 Limitations 16

Appendix A: STAC-M3 Antuco Results .... STAC-M3 Antuco Results - 1
Storage Efficiency .... STAC-M3 Antuco Results - 1
Light-Compute Benchmarks .... STAC-M3 Antuco Results - 1
Post-Trade Analytics Benchmarks .... STAC-M3 Antuco Results - 2
Research Analytics Benchmarks .... STAC-M3 Antuco Results - 3
NBBO Benchmark .... STAC-M3 Antuco Results - 4
Multi-day/Multi-User VWAB Benchmark .... STAC-M3 Antuco Results - 4
Chart view .... STAC-M3 Antuco Results – 4

About STAC

STAC® provides technology research and testing tools that are based upon community-source standards. STAC facilitates the STAC Benchmark™ Council (www.STACresearch.com/council), an organization of leading financial institutions and technology vendors that specifies standard ways to assess technologies used in finance. The Council is active in an expanding range of low-latency, big-compute, and big-data workloads.

STAC helps end-user firms relate the performance of new technologies to that of their existing systems by supplying them with STAC Benchmark reports as well as standards-based STAC Test Harnesses™ for rapid execution of STAC Benchmarks in their own labs. User firms do not disclose their results. Some STAC Benchmark results from vendor-driven projects are made available to the public, while those in the STAC Vault™ are reserved for qualified members of the Council (see www.STACresearch.com/vault).

To be notified when new STAC Reports™ become available, please sign up for free at www.STACresearch.com.

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

Disclaimer

This document was prepared by the Securities Technology Analysis Center (STAC®) at the request of Dell EMC. It is provided for your internal use only and may not be redistributed, retransmitted, or published in any form without the prior written consent of STAC. “STAC” and all STAC names are registered trademarks or trademarks of the Securities Technology Analysis Center LLC. All other trademarks in this document belong to their respective owners.

The test results contained in this report are made available for informational purposes only. Neither STAC nor the vendor(s) supplying the information in this report guarantee similar performance results. All information contained herein is provided on an “AS IS” BASIS WITHOUT WARRANTY OF ANY KIND. STAC explicitly disclaims any liability whatsoever for any errors or otherwise.

References

[1] Specifications used for this benchmark: STAC-M3 Benchmark Specifications, Antuco Suite, Rev Q – www.STACresearch.com/m3/antuco/revq . Accessible by qualified members of the STAC Benchmark Council.

[2] STAC Configuration Disclosure for this SUT: www.STACresearch.com/KDB200914. If you are unable to access these materials and would like to learn how to, please contact us at www.STACresearch.com/contact.

Copyright © 2020 STAC

1. Summary

STAC recently performed the baseline STAC-M3™ Benchmarks on a stack involving Kx's kdb+3.6 DBMS distributed across 9 Dell EMC PowerEdge R640 servers sharing a 3-node Dell EMC PowerScale F200 All-Flash Scale-Out NAS cluster.

STAC-M3 is the set of industry standard enterprise tick-analytics benchmarks for database software/hardware stacks that manage large time series of market data ("tick data"). This report highlights results from the baseline benchmark suite (code named Antuco).

In all, the STAC-M3 specifications deliver dozens of test results, which are presented through a variety of tables and visualizations in this report. Dell EMC chose to highlight the following:

This Dell EMC Isilon based solution involving 9 database servers accessing networked flash storage was/had:

- *Faster than a solution using Dell EMC's larger flash storage appliance (SUT ID KDB190430) in 2 of 17 mean-response time STAC-M3 Antuco benchmarks, including:*
  - *77% faster in the single-user VWAB operation (STAC-M3.v1.1T.VWAB-D.TIME)*
  - *33% faster in the 10-user market snapshot operation (STAC-M3.β1.10T.MKTSNAP.TIME)*

- *7.7x the speed of a solution involving a parallel file system with 14 database servers and 18 storage servers (KDB200401) in the single-user NBBO operation (STAC-M3.β1.1T.NBBO.TIME)*

- *Highest storage efficiency of any publicly reported solution involving kdb+ 3.6 (least storage used for the same database size)*
  - *166% for this SUT vs 149% for all others (STAC-M3.v1.1.STORAGE.EFF)*

Getting the most from these results

Any interested party can analyze public STAC Reports to compare the performance of different systems. However, members of the STAC Benchmark Council are able to put these reports to much greater use. Qualified members may:

- Read the detailed Configuration Disclosure [2] for the system tested in this report
- Read the detailed test specifications
- Access additional reports in the confidential STAC Vault™
- Obtain the materials to run the STAC-M3 Benchmarks on their own systems
- Discuss benchmarks, technologies, and related business issues with their peers.

To join the Council or upgrade your membership, please contact council@STACresearch.com.

2. Vendor Commentary

Dell EMC provided the following comments:

1. *Default Linux OS settings were used, i.e. no special tuning was done. The goal was to conduct the STAC benchmark tests using default settings on both the compute servers and on the single 3-node, 3U PowerScale F200 storage system.*

2. *Standard NFS protocol was used to access all data on the PowerScale storage system. The goal was to benchmark the system without requiring any special software, drivers or libraries on the Linux kdb+ servers to highlight the simplicity and performance of using PowerScale OneFS and the F200 storage cluster. Adding more F200 nodes is simple and would immediately scale I/O capacity. The numbers listed in this report are for a single 3-node, 3U PowerScale F200 cluster.*

3. *The distributed kdb+ configuration used to conduct the benchmark reflects a common type of deployment used in the industry. Client connections were all remote to the SUT to represent real-world kdb+ deployments.*

4. *With Dell EMC PowerScale, financial organizations and kdb+ administrators can effortlessly scale from tens of terabytes to tens of petabytes within a single file system, as a single volume, and with a single point of administration. Dell EMC PowerScale delivers high performance and high throughput during highly concurrent workloads without adding management complexity.*

5. *Caching was disabled on Isilon for these benchmark tests.*

3. Product background

This section provides a high level overview of the SUT in this report. A detailed STAC Configuration Disclosure for the SUT in this report is available to premium members of the STAC Benchmark Council at the same web page as this report [2]. That document provides the exact product version numbers, detailed tuning options, and other important information. Additional configuration details such as a SOS report may also be available, depending on the SUT platform.

The top of the stack under test was the benchmark implementation code (i.e., the STAC-M3 Clients and supporting scripts): the STAC-M3 Pack for kdb+ Rev 3.0 Antuco, Compatibility Rev E.

Key products in the SUT included:

- Database:
  - kdb+ 3.6 running in distributed mode
- Database Server Nodes:
  - 9 x Dell EMC PowerEdge R640 servers, each with:
    - NFS v3
    - 2 x 18 core Intel® Xeon® Gold 6254 CPU @ 3.10GHz
    - 384GiB DRAM
- Storage:
  - 1 x 3-node Dell EMC PowerScale F200 All-Flash Scale-Out cluster
  - Dell EMC OneFS 9.0 storage cluster operating system
  - 12 x 960GB SSD total
  - 10TiB total physical capacity (8.3TiB usable)
- Network:

- Dell EMC PowerSwitch Z9332F-ON (database servers to storage cluster)
- Dell EMC PowerSwitch Z9100-ON (intra-storage cluster)

The STAC-M3 clients were hosted on a separate server not part of the SUT but identical to the database servers.

Servers in the SUT were configured to mitigate the full range of Spectre/Meltdown threats checked by the Spectre/Meltdown checker tool as of the test date. Details are available in the STAC Configuration Disclosure [2], along with the detailed Spectre/Meltdown tool output.

Dell EMC submitted the following information and claims about its products:

Dell EMC PowerScale F200 All-Flash Network Attached Storage (NAS) powered by OneFS

New PowerScale all-flash storage platforms - powered by the OneFS operating system - provide a powerful yet simple scale-out storage architecture to speed up access to massive amounts of unstructured data while dramatically reducing cost and complexity. They deliver high performance and efficiency for your most demanding unstructured data applications and workloads.
PowerScale F200: Provides the performance of flash storage in a cost-effective form factor to address the needs of a wide variety of workloads. Each node allows you to scale raw storage capacity from 3.84 TB to 15.36 TB per node and up to 3.8 PB of raw capacity per cluster. The F200 includes in-line compression and deduplication. The minimum number of PowerScale nodes per cluster is three while the maximum cluster size is 252 nodes.
The OneFS operating system provides the intelligence behind the highly scalable, high–performance modular storage solution that can grow with your business. With support for all-flash and NVMe, OneFS can help you accelerate processes and workflows while scaling easily to handle massive growth and providing the highest levels of data protection. This is all provided in a storage solution designed for unmatched ease of use.
With OneFS powered clusters consisting of PowerScale or Isilon nodes, you can eliminate storage silos, consolidate all your unstructured data, store petabytes of file data and analyze them in a data-first world. With up to 252 nodes in a cluster, you can scale both capacity and performance in a few minutes to meet your specific business needs—all without any additional IT burden.

Dell EMC PowerEdge R640

The PowerEdge R640 is a general purpose platform expandable up to 7.68TB of memory, up to twelve 2.5 inch drives, and flexible I/O options. The R640 can handle demanding workloads such as virtualization, dense private cloud, High Performance Computing (HPC) and software-defined storage.

The Dell EMC PowerEdge R640 is the ideal dual-socket, 1U platform for dense scale-out data center computing. The R640 combines density, performance and scalability to optimize application performance and data center density.

Compared to previous generations, the PowerEdge R640 offers faster processing power and advanced system management. The R640 system is a powerhouse 2S/1U rack system, which emphasizes performance and reliability in areas such as virtualization, power, thermal and systems management, and usability. It is designed for mid-size to large data centers that require high memory capacity and performance.

Kx submitted the following information and claims about its products:

Kx is part of First Derivatives (FD), a global technology and services provider with more than 20 years of experience working with some of the world's largest financial, technology, automotive, manufacturing and energy institutions. Kx Streaming Analytics, incorporating the kdb+ time-series database, is a high-performance, streaming analytics and operational intelligence platform that enables the real time analysis of any data – whether at speed, or at rest – allowing firms to make better business decisions 'in the moment' and solve complex problems that other data platforms cannot easily address. FD operates from

15 offices across Europe, North America and Asia Pacific, and employs more than 2,400 people worldwide. Solutions based on kdb+ are developed using q, an interactive declarative, list oriented, SQL-like language included in kdb+ at runtime.

The following diagram (figure 1) illustrates the kdb+ architecture used for Financial Services Market Data analytics, some elements of which were tested by STAC M3.

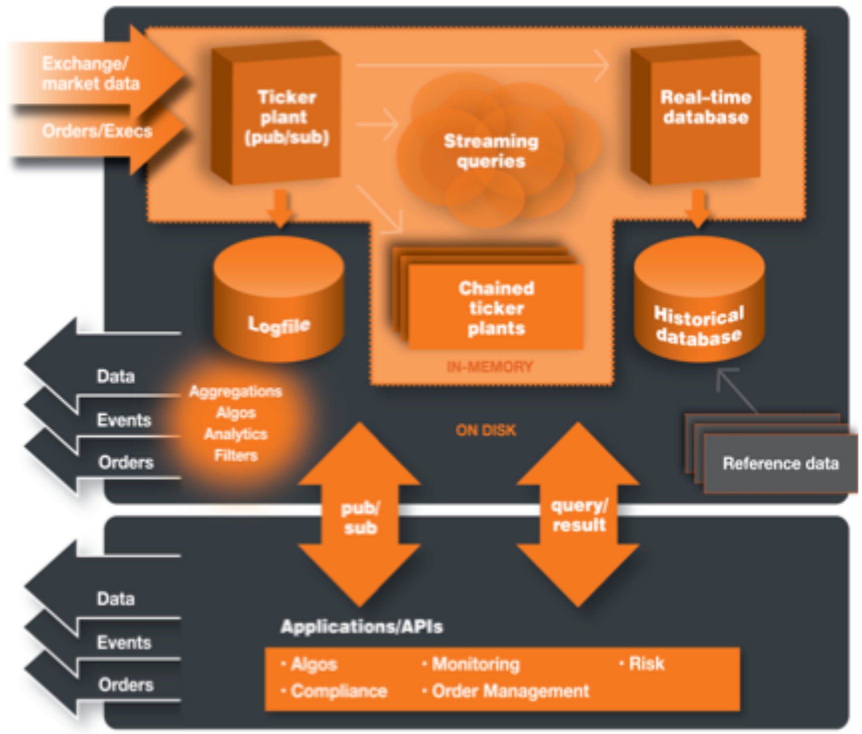


Figure 1: View of kdb+ system architecture

4. Project participants and responsibilities

The following firms participated in the project, with the associated responsibilities:

- Kx implemented the STAC-M3 STAC Pack using the STAC-M3 Benchmark specifications.
- Dell EMC supplied the storage hardware for the test and expertise around performance and tuning of the system; configured and optimized the full stack under test; supplied the lab/support, hardware and network infrastructure for the database server nodes; and sponsored the audit.
- STAC conducted the STAC-M3 Benchmark Audit, which included validating the database; inspecting any source-code revisions to the STAC Pack; validating the Operation results; executing the tests and documenting the results.

5. Contacts

- Dell EMC: Damien Mas, Analytics Solutions Architect, damien.mas@dell.com
- Kx Systems: Glenn Wright, gwright@kx.com, +44 7802 248372
- STAC: info@STACresearch.com

6. Results status

- These benchmark specifications were developed by the STAC-M3 Working Group of the STAC Benchmark Council.
- These test results were audited by STAC or a STAC-certified third party, as indicated in the Responsibilities section above. As such, they are official results. For details, see www.STACresearch.com/reporting.
- The vendors attest that they did not modify the SUT during the Audit.

7. Overview of the STAC-M3 Benchmark specifications

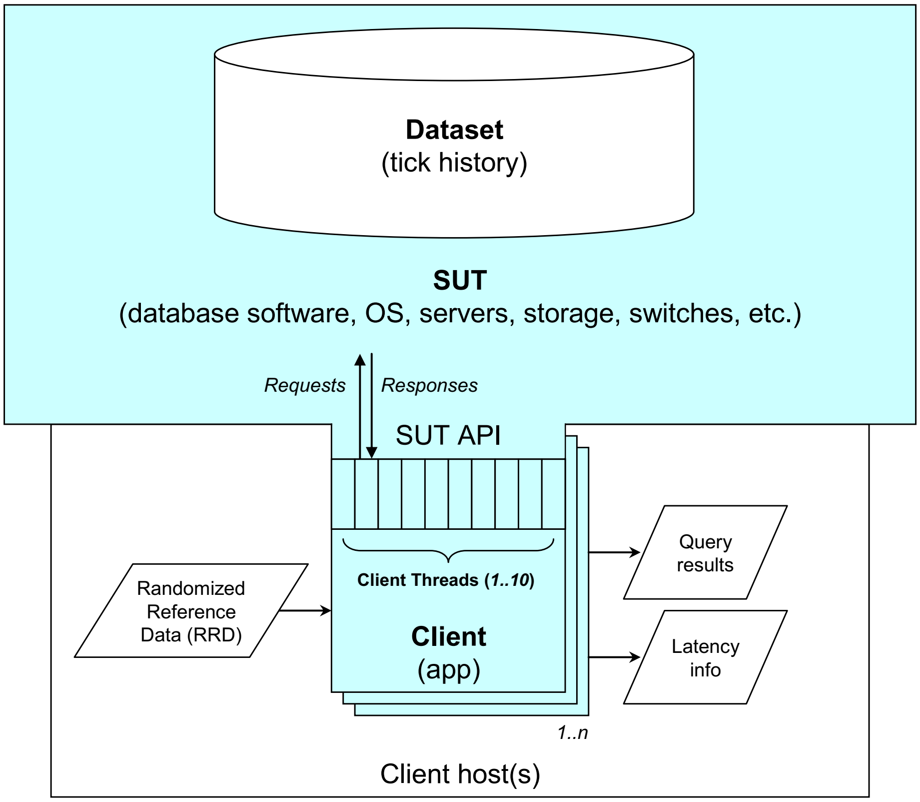
Analyzing time-series data such as tick-by-tick quote and trade histories is crucial to many trading functions, from algorithm development to risk management. But the domination of liquid markets by automated trading—especially high-frequency trading—has made such analysis both more urgent and more challenging. As trading robots try to outwit each other on a microsecond or sub-microsecond scale, they dish out quotes and trades in ever more impressive volumes. This places a premium on technology that can store and analyze that activity efficiently. For example, the faster an algorithm developer can backtest and discard a haystack of unprofitable ideas, the faster he will find the needle of a winning algorithm, leaving more time to exploit it in the market.

The STAC Benchmark Council has developed the STAC-M3 Benchmarks in order to provide a common basis for quantifying the extent to which emerging software, cloud, and hardware innovations improve the performance of the storage, retrieval, and analysis of market data time series ("tick") data.

STAC-M3 tests the ability of a complete solution stack of database software and infrastructure to perform a variety of operations on a large store of market data. The STAC-M3 Working Group designed these test specs to enable useful comparisons of entire solution stacks (i.e., to gauge the state of the art) as well as comparisons of specific stack layers while holding other layers constant. Comparisons can include (but are not limited to tick-database software products (typically columnar), storage architectures (including media, interconnects, and file systems), server products (including processors, chipsets, and memory), and cloud infrastructure (IaaS, DBaaS, etc.).

As shown below, the test setup for STAC-M3 consists of the "stack under test" (SUT) and client applications. No restrictions are placed on the architecture of the SUT or clients (though members of the STAC-M3 Working Group frequently provide input on architectures they would like to see tested). Threads within the clients take in Randomized Reference Data (RRD) such as dates and symbols, submit requests for the required operations, receive responses, and store the timings and results from these queries. Vendor-supplied code for the operations and response-time calculations are subjected to a combination of source-code inspection and empirical validation.

Copyright © 2020 STAC



Understanding the STAC-M3 Benchmark Suites

The STAC-M3 Working Group has developed three benchmark suites that address different testing needs, as explained in the table below.

Suite	Purpose	Dataset size*	Concurrent requests	Operations	Constraints related to memory and storage
Antuco	Using a limited dataset size for convenience, simulate performance that would be obtained with a larger real-world dataset residing mostly on non-volatile media. Study a broad range of read and write operations.	4.5 TB	1 to 100	Range of compute-bound and storage-bound analytics. A few operations involving writes.	• No pre-loading into memory • File system cache cleared at several points in test run
Shasta	Study performance across a broad range of operations for datasets that are relatively small in the real world. (While the dataset tested is the same size as in Antuco, there is no attempt to simulate the storage-access pattern of a larger	4.5 TB	1 to 100	Same as Antuco except operations involving writes are optional.	• Pre-loading into memory is allowed (most recent data first) • Caches not cleared during test run

	dataset.)				
Kanaga	Study performance on large datasets with large numbers of concurrent requests.	33 TB to 897 TB	1 to 450	A few storage-intensive queries.	• Pre-loading into memory is allowed (most recent data first) • Caches not cleared during test run • Storing certain data into faster storage tiers is allowed

* Reference size is based on a “standard” representation for each data type, making no allowance for optimizations or compression, nor for any overhead such as file headers, delimiters, indices, etc. Actual space requirements will vary by implementation and in practice tend to be smaller.

** Benchmark IDs that are identical except that one ends in ".TIME" and the other ends in ".LAT2" can be fairly compared. Prior to 2014, benchmarks in the STAC-M3 Antuco suite had two metrics: LAT1 (time to receive first result) and LAT2 (time to receive all results). Given that LAT1 and LAT2 results were identical for all systems reported from 2011 to 2013, LAT1 was eliminated in 2014. In addition, LAT2 was redesignated TIME in order to clarify that the measurement represents a response time at the application level and to avoid confusion with micro-level storage latency.

Datasets

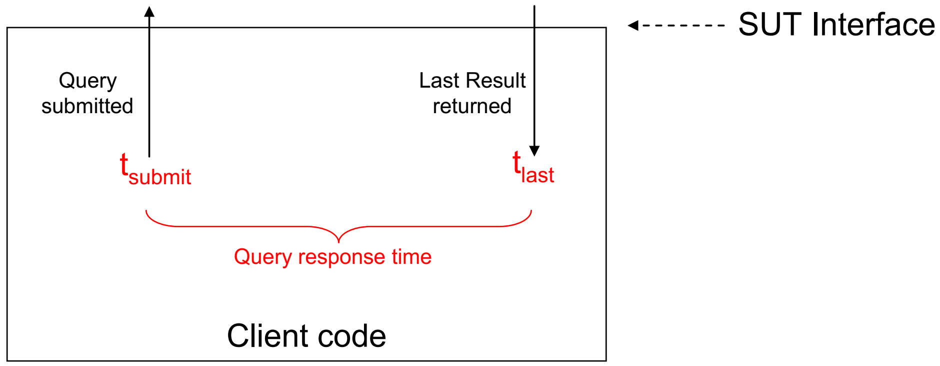
STAC-M3 draws from client experience with equities and FX use cases. The database is synthetic, modeled on NYSE TAQ data (US equities). While testing with real data is also desirable, synthetic data has three advantages that make it compelling for STAC-M3:

- Synthetic data allows us to control the database properties exactly, which in turn allows us to randomize elements of queries from project to project while keeping the resulting workload exactly the same (for example, we control how much volume is associated with each symbol).
- Synthetic data does not incur fee liability from a third party such as an exchange.
- Synthesizing the data makes it easy to scale the database to an arbitrarily large size.

The dataset consists of high-volume symbols and low-volume symbols in proportions based on observed NYSE data. The data volume per symbol in the baseline dataset was based on doubling the typical volume in NYSE TAQ in 1Q10. The resulting database is considerably smaller than databases in use at customer sites, but the benchmarks impose policies that force the database to access storage. This approach was the STAC-M3 Working Group’s way of minimizing the cost of running baseline benchmarks while still yielding results indicative of those that would occur with large databases.

Metrics

The key metric in STAC-M3 is query response time. This measurement is performed in the client. A client thread gets a local timestamp (t_{submit}) just before submitting a query. When it receives the complete results of the query (sorted appropriately), the client immediately gets a second timestamp (t_{last}). Query response time is (t_{last}) - (t_{submit}).



Timestamps and response time

Some of the I/O-focused benchmarks also measure the bytes read per second from persistent storage (i.e., excluding server cache), which is computed from the output of appropriate system utilities.

The algorithms in all benchmarks are defined so as to keep the result sets small. This ensures that network I/O between the test clients and server(s) is negligible compared to back-end processing times.

Test cases

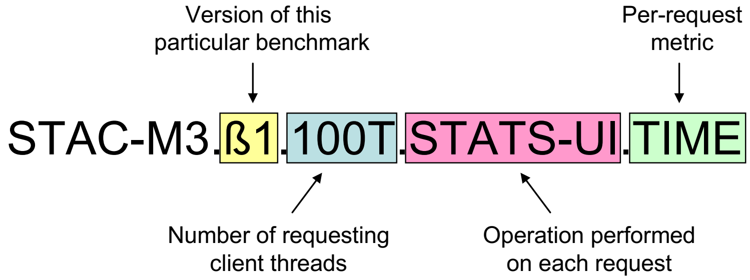
The tests in the baseline STAC-M3 suite (Antuco) are listed in the "STAC-M3 Antuco Benchmark Operations" table below. These benchmarks operate on baskets of instruments, accessing many fields of underlying tick data for both trades and quotes across varying time windows. The table classifies each test case as relatively heavy on I/O, compute, or both.

The tests require a client application that is written to a product API and is capable of submitting requests from 10 independent threads, each of which simulates a user. As detailed in the table, some of the benchmarks call for one client instance making requests from a single thread, while others call for one client using 10 threads, and still others require 10 clients each using 10 threads (100 total requesting threads). One set of benchmarks (using the STATS-UI operation) tests multi-user scaling by running with 1, 10, 50, and 100 client threads. In all cases, benchmark results refer to per-request response times. For example, the mean of 10T.MKTSNAP.TIME is the mean time to satisfy a market-snapshot request from one of the threads, not the total time to satisfy requests from all 10 client threads. (Note, however, that a single request typically requires access to multiple instruments, fields, dates, and/or times.)

The range of dates eligible for querying depends on the benchmark. For example, some algorithms operate on dates randomly chosen throughout the year, some stick to a recent date range, and some always run on the most recent date (see the "Input Date Range" column of the table). The purpose of this differentiation is to provide a "recency bias" for those workloads where such bias is observed in the real world, while preventing such bias for those workloads that do not exhibit it in the real world.

Benchmark identifiers

The STAC-M3 Report Card and accompanying charts identify each benchmark unambiguously, as follows:



In charts, the ID is sometimes decomposed, with part of it in the chart title or labels. Each individual STAC Benchmark™ specification has its own version number. The same version of a given spec may appear in multiple benchmark suites. Thus, the code names of the suites are irrelevant when making comparisons. Versioning individual specs enables the reader to compare a discrete result from this “stack under test” (SUT) to the corresponding result from another SUT. When making comparisons, be sure that the identifiers match exactly. If they do not, the benchmark results may not be capable of fair comparison.

Copyright © 2020 STAC

STAC-M3 Benchmarks in the Antuco Suite

The table below gives a brief overview of each test in this STAC-M3 suite. Version numbers of 1 or greater indicate benchmark specs that have been approved. Versions less than 1 are proposed by the STAC-M3 Working Group but not yet voted on by the full STAC Benchmark Council.

STAC-M3 Antuco Benchmark Operations

Root ID	Operation name	Ver	Number of requesting Client Threads	Algorithm performed on behalf of each requesting Client Thread	Algorithm I/O intensity	Algorithm compute intensity	Input date range*
VWAB-D	VWAB-Day	1	1	4-hour volume-weighted bid over one day for 1% of symbols (like VWAP but operating on quote data, so much higher input volume).	Heavy read	Light	Last 30 days
VWAB-12D-NO	VWAB-12DaysNoOverlap	1	100	4-hour volume-weighted bid over 12 days for 1% of symbols. No overlap in symbols among client threads.	Heavy read	Light	Full year
YRHIBID	Year High Bid	ß1	1	Max bid over the year for 1% of symbols.	Heavy read	Light	Full year
YRHIBID-2	Year High Bid Re-run	ß1	1	Re-run of YRHIBID (same symbols) without clearing the cache.	Heavy read†	Light	Full year
QTRHIBID	Quarter HighBid	ß1	1	Max bid over the quarter for 1% of symbols.	Heavy read	Light	Most recent quarter
MOHIBID	Month High Bid	ß1	1	Max bid over the month for 1% of symbols.	Heavy read	Light	Most recent month
WKHIBID	Week High Bid	ß1	1	Max bid over the week for 1% of symbols.	Heavy read	Light	Most recent week
STATS-AGG	Aggregate Stats	ß1	10	One set of basic statistics over 100 minutes for all symbols on one exchange. Each 100-minute range crosses a date boundary.	Heavy read	Heavy	Full year

Copyright © 2020 STAC

STATS-UI	Stats - Unpredictable Intervals	ß1	1, 10, 50, 100 (more optional)	Per-minute‡ basic statistics over 100 minutes for all high-volume symbols on one exchange. Each 100-minute range crosses a date boundary.	Heavy read	Heavy	Full year
MKTSNAP	Market Snapshot	ß1	10	Most recent trade and quote information for 1% of symbols as of a random time.	Heavy read	Heavy	Full year
VOLCURV	Volume Curves	ß1	10	Create an average volume curve (using minute intervals aligned on minute boundaries) for 10% of symbols over 20 days selected at random.	Light read	Heavy	Full year
THEOPL	Theoretical P&L	ß1	10	For a basket of 100 trades on random dates, find the future times at which 2X, 4X, and 20X the trade size traded in each symbol. Trade sizes cause up to 5 days of forward searching. Calculate the corresponding VWAP and total volume traded over those periods.	Light read	Heavy	Full year
NBBO	NBBO	ß1	1	Create the NBBO across all 10 exchanges for all symbols on the most recent day. Write to persistent storage.	Heavy read and write	Heavy	Most recent day
WRITE	Write	1	1	Write one day's quote data to persistent storage, following the same algorithm used to generate the randomized dataset used in the other Operations.	Heavy write	Light	n/a
STORAGE.EFF	Storage efficiency	1.1	n/a	Reference Size of the Dataset divided by size of the Dataset in the SUT format used for the performance benchmarks. Expressed as as percentage.	n/a	n/a	n/a

* In some cases, one or more dates at the end of the year were excluded from eligibility to prevent an algorithm that crosses days from running out of input data.
† Typically this will be reads from DRAM cache.
‡ In this case, interval start times are offset from minute boundaries by a consistent random amount per test run, so that the SUT cannot rely on pre-calculated minute statistics.

Copyright © 2020 STAC

8. Specification particulars

8.1 Version

This project followed the benchmark specifications in [1]. Qualified members of the STAC Benchmark Council can access these specifications and also download the programs used in this project in order to run the same tests on systems in the privacy of their own labs.

8.2 Limitations

- As discussed in the overview of the benchmark specifications, the STAC-M3 Antuco suite was designed to test operations on a limited amount of purely historical data. The STAC-M3 Kanaga suite involves larger amounts of historical data.
- As discussed in Section 1, the dataset used in this version of STAC-M3 is synthetic. The algorithm to generate the dataset creates random values for prices and sizes that can vary widely from tick to tick. In the real world, by contrast, there is significant correlation of successive prices (i.e., large differences from tick to tick are relatively rare). Compression algorithms often take advantage of this fact, such as by focusing on deltas between successive values. Hence, the storage efficiency of a SUT may be higher when working with real data than with the synthetic dataset of this version of STAC-M3.
- MKTSNAP is a random-access operation, and because systems tend to have areas of storage that differ in access time, the MKTSNAP response times can vary considerably. STAC-M3 requires a limited number of test runs, which means that the standard deviation of response times can be quite large relative to the mean. Thus, the mean MKTSNAP.TIME in a given year of data is, in general, not the best statistic to use from these tests (i.e., two systems with the same performance could get quite different mean response times simply by chance). Median and max are probably more instructive indicators, which is why these are used in Figure 11.
- Due to the architecture of the Dell EMC PowerScale F200 All-Flash Storage, the system utility iostat could not be used to directly measure bytes read from disk while performing STAC-M3 operations. The alternative used here was nfsstat, a utility that reports statistics from the NFS driver. See the STAC Notes in [2] for details.

Appendix A: STAC-M3 Antuco results

Below are the results from benchmarks in the Antuco suite of STAC-M3, in tabular and graphical forms.

Storage efficiency

Storage Efficiency	
The reference size of the dataset divided by the size of the dataset as stored by the SUT. The less storage space required, the higher the percentage.	
STAC-M3.v1.1.STORAGE.EFF	**166%**

Light-Compute Benchmarks

High Bid (1 Client Thread Requesting)

Return the high bid for a certain 1% of symbols over varying timeframes. Run the year-high bid a second time (YRHIBID-2) without clearing the cache.

	Response time (milliseconds)				
Spec ID	MEAN	MED	MIN	MAX	STDV
STAC-M3.ß1.1T.YRHIBID.TIME	760	763	739	782	14
STAC-M3.ß1.1T.YRHIBID-2.TIME	522	517	511	544	12
STAC-M3.ß1.1T.QTRHIBID.TIME	197	196	189	206	6
STAC-M3.ß1.1T.MOHIBID.TIME	137	138	131	143	4
STAC-M3.ß1.1T.WKHIBID.TIME	119	120	116	124	3
	Megabytes read per second*				
Spec ID	MEAN			MAX	
STAC-M3.ß1.1T.YRHIBID.MBPS	5,671			5,747	
STAC-M3.ß1.1T.YRHIBID-2.MBPS	n/a			n/a	
STAC-M3.ß1.1T.QTRHIBID.MBPS	5,531			5,763	
STAC-M3.ß1.1T.MOHIBID.MBPS	2,782			2,909	
STAC-M3.ß1.1T.WKHIBID.MBPS	759			776	

Megabytes read per second derived from nfsstat statistics. Not appropriate for STAC- M3.β1.YRHIBID-2.MBPS. Please see the Limitations section of the STAC Report or the STAC Notes in [2].

Write Test

Perform the Basic Data Generation Algorithm for 1 day's data.

	Response time* (milliseconds)				
Spec ID	MEAN	MED	MIN	MAX	STDV
STAC-M3.v1.1T.WRITE.TIME	49,014	48,661	48,567	50,243	629

* Time to write all results.

Copyright © 2020 Securities Technology Analysis Center LLC

Post-Trade Analytics Benchmarks

VWAB on 1 Day's Data
(1 Client Thread Requesting)

Return ~4-hour volume-weighted bid over a single day for certain 1% of symbols

	Response time (milliseconds)				
Spec ID	MEAN	MED	MIN	MAX	STDV
STAC-M3.v1.1T.VWAB-D.TIME	43	42	40	48	3

Theoretical P&L
(10 Client Threads Requesting)

For each of 10 Client Threads querying a unique set of 100 trades, find the amount of time until 2x, 4x, and 20x the size of each trade was traded in the market, and return the VWAP and total volume over those times intervals.

	Response time (milliseconds)				
Spec ID	MEAN	MED	MIN	MAX	STDV
STAC-M3.ß1.10T.THEOPL.TIME	310	329	84	524	123

Market Snapshot
(10 Client Threads Requesting)

To each of 10 Client Threads querying a unique date, time, and set of symbols (1% of the total symbols), return the price and size information for the latest quote and trade for each symbol.

	Response time (milliseconds)				
Spec ID	MEAN	MED	MIN	MAX	STDV
STAC-M3.ß1.10T.MKTSNAP.TIME	4,041	3,927	2,158	5,060	776

Copyright © 2020 Securities Technology Analysis Center LLC

Research Analytics Benchmarks

Volume Curves
(10 Client Threads Requesting)

To each of 10 Client Threads querying a unique set of 20 dates and set of symbols (10% of the total symbols), return the average proportion of volume traded in each minute interval for each symbol across the date set.

	Response time (milliseconds)				
Spec ID	MEAN	MED	MIN	MAX	STDV
STAC-M3.ß1.10T.VOLCURV.TIME	9,650	9,616	1,747	17,631	5,033

Aggregated Stats
(10 Client Threads Requesting)

For each of 10 Client Threads querying a unique exchange, date, and start time, return basic statistics calculated for the entirety of the 100-minute time range following the start time. Time ranges always cross a date boundary.

	Response time (milliseconds)				
Spec ID	MEAN	MED	MIN	MAX	STDV
STAC-M3.ß1.10T.STATS-AGG.TIME	35,451	35,542	5,547	70,643	18,886

Stats Over Unpredictable Intervals
(Variable Client Threads Requesting)

To each of some number of Client Threads querying a unique exchange, date, and start time, return basic statistics calculated for each minute interval in a 100-minute time range following the start time. Start times are offset from minute boundaries by a random amount. Time ranges always cross a date boundary. Tests must be run with 1, 10, 50, and 100 Client Threads. Tests with other numbers of Client Threads are optional.

	Response time (milliseconds)				
Spec ID	MEAN	MED	MIN	MAX	STDV
STAC-M3.ß1.1T.STATS-UI.TIME	1,918	1,506	1,485	2,555	516
STAC-M3.ß1.10T.STATS-UI.TIME	8,523	8,640	815	17,822	4,947
STAC-M3.ß1.50T.STATS-UI.TIME	30,681	30,660	797	60,858	17,200
STAC-M3.ß1.100T.STATS-UI.TIME	47,237	48,759	2,500	99,346	28,096

Copyright © 2020 Securities Technology Analysis Center LLC

NBBO Benchmark

NBBO

Calculate NBBO across all exchanges for all symbols on one day.

	Response time* (milliseconds)				
Spec ID	MEAN	MED	MIN	MAX	STDV
STAC-M3.ß1.1T.NBBO.TIME	24,653	24,421	24,060	25,567	587

* Time to write all results.

Multi-day/Multi-User VWAB Benchmark

VWAB for 12 Days with No Overlap in Interest (100 Client Threads Requesting)

To each of 100 Client Threads querying unique symbol sets, return 4-hour volume-weighted bid for 12

	Response time (milliseconds)				
Spec ID	MEAN	MED	MIN	MAX	STDV
STAC-M3.v1.100T.VWAB-12D-NO.TIME	25,749	25,426	491	47,941	13,413

Chart view

The charts that follow illustrate or elaborate on the results above:

- Figure 1 through Figure 4 plot the mean response time (TIME) benchmarks for all of the operations.
- Figure 5 and Figure 6 analyze the individual response-time observations for the multi-user/multi-day VWAB benchmark (STAC-M3.v1.100T.VWAB-12D-NO.TIME), first by sorting the results by response time, then by plotting them in a histogram.
- Figure 7 provides a more explicit look at multi-user scaling by plotting the response time for the intervalized statistics benchmark (STAC-M3.s1.[n]T.STATS-UI.TIME) against the number of simultaneously requesting client threads (n).
- Figure 8 and Figure 9 take the 100-client-thread case of Figure 7 and analyze the individual responsetime observations, first by sorting the results by response time, then by plotting them in a histogram.

Refer to the overview of the benchmark specifications below or the tables above for explanations of the benchmark IDs used in the charts.

The axes in the bar charts are fixed, so that results from this SUT may be visually compared to those of other SUTs. Because the results of future SUTs are unpredictable, the axes use a log scale.

Copyright © 2020 Securities Technology Analysis Center LLC

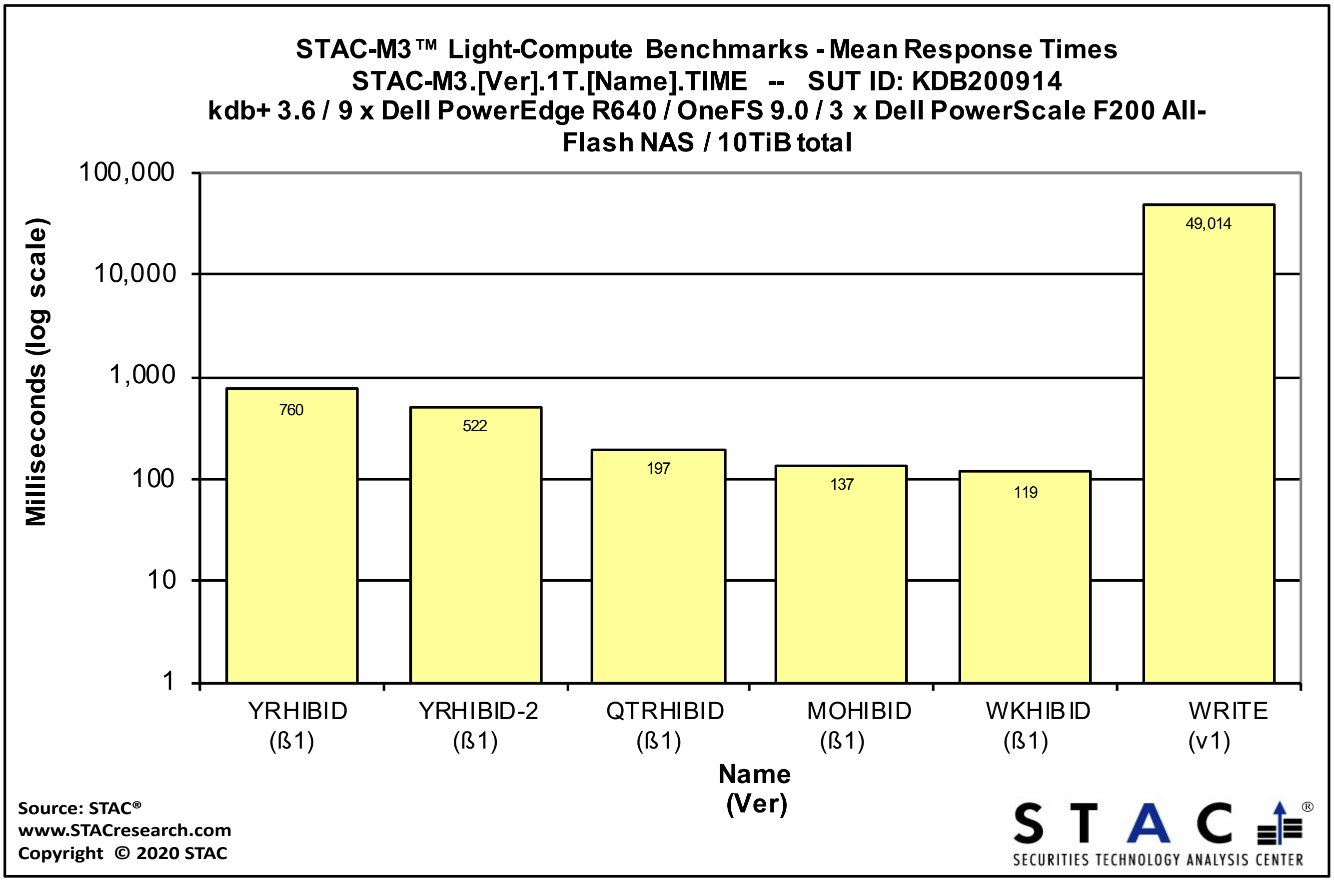

STAC-M3™ Light-Compute Benchmarks - Mean Response Times
STAC-M3.[Ver].1T.[Name].TIME -- SUT ID: KDB200914
kdb+ 3.6 / 9 x Dell PowerEdge R640 / OneFS 9.0 / 3 x Dell PowerScale F200 All-Flash NAS / 10TiB total
Milliseconds (log scale)
100,000
10,000
1,000
100
10
1
760
522
197
137
119
49,014
YRHIBID (ß1)
YRHIBID-2 (ß1)
QTRHIBID (ß1)
MOHIBID (ß1)
WKHIBID (ß1)
WRITE (v1)
Name (Ver)
Source: STAC®
www.STACresearch.com
Copyright © 2020 STAC
STAC® SECURITIES TECHNOLOGY ANALYSIS CENTER

Figure 1

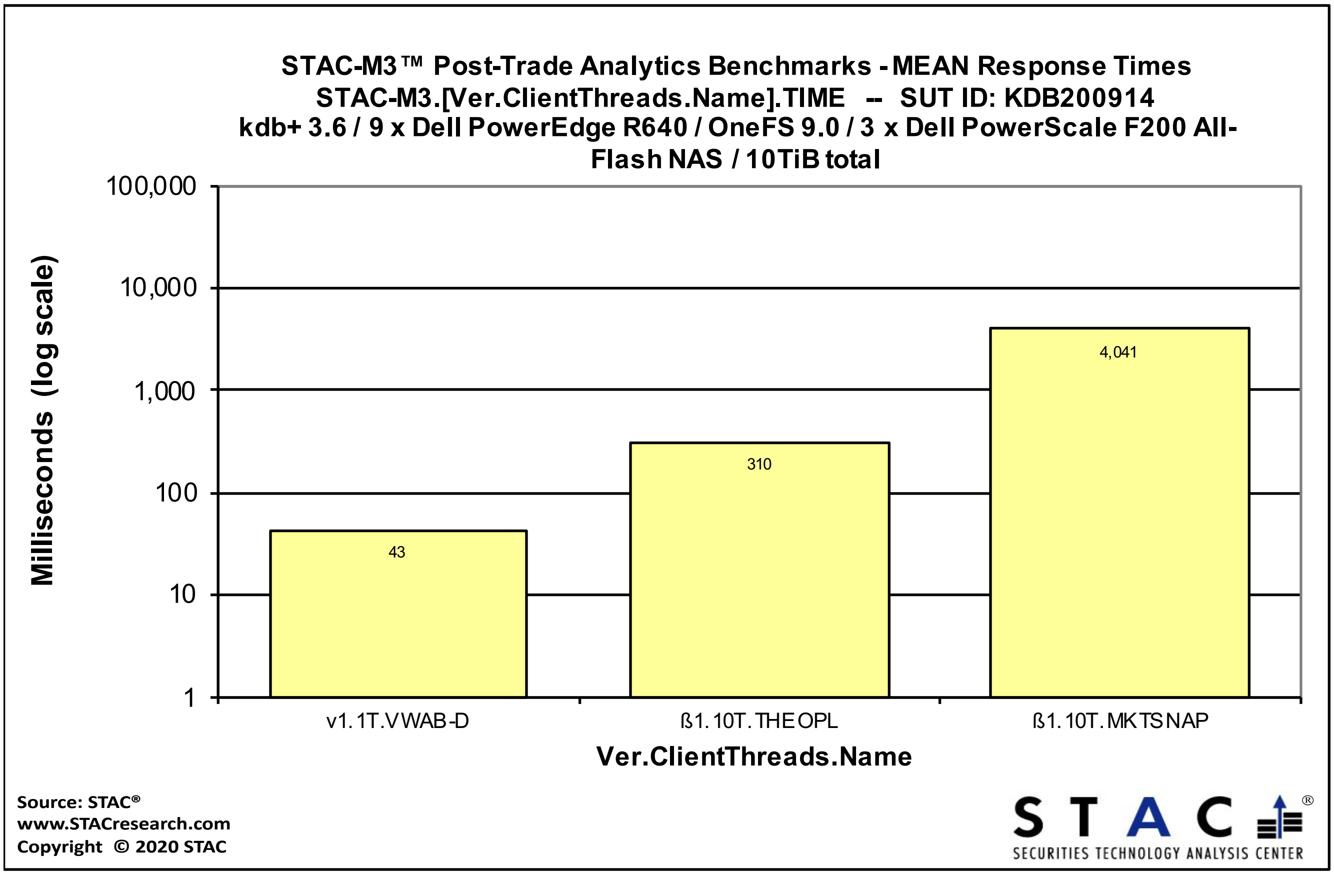

STAC-M3™ Post-Trade Analytics Benchmarks - MEAN Response Times
STAC-M3.[Ver.ClientThreads.Name].TIME -- SUT ID: KDB200914
kdb+ 3.6 / 9 x Dell PowerEdge R640 / OneFS 9.0 / 3 x Dell PowerScale F200 All-Flash NAS / 10TiB total
Milliseconds (log scale)
100,000
10,000
1,000
100
10
1
43
310
4,041
v1.1T.VWAB-D
ß1.10T.THEOPL
ß1.10T.MKTSNAP
Ver.ClientThreads.Name
Source: STAC®
www.STACresearch.com
Copyright © 2020 STAC
STAC
SECURITIES TECHNOLOGY ANALYSIS CENTER

Figure 2

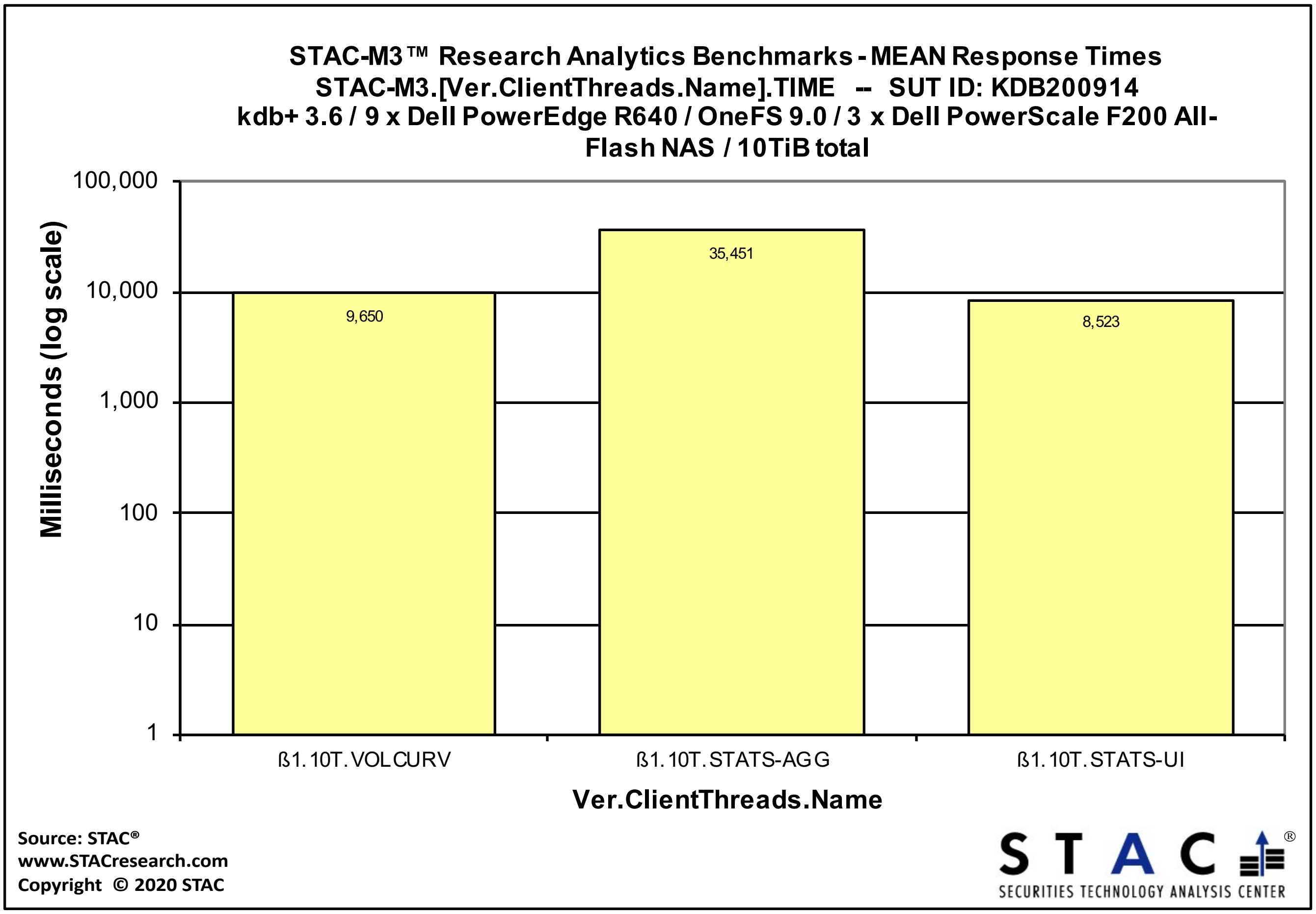

STAC-M3™ Research Analytics Benchmarks - MEAN Response Times
STAC-M3.[Ver.ClientThreads.Name].TIME -- SUT ID: KDB200914
kdb+ 3.6 / 9 x Dell PowerEdge R640 / OneFS 9.0 / 3 x Dell PowerScale F200 All-Flash NAS / 10TiB total
Milliseconds (log scale)
100,000
10,000
1,000
100
10
1
9,650
35,451
8,523
ß1.10T.VOLCURV
ß1.10T.STATS-AGG
ß1.10T.STATS-UI
Ver.ClientThreads.Name
Source: STAC®
www.STACresearch.com
Copyright © 2020 STAC
STAC
SECURITIES TECHNOLOGY ANALYSIS CENTER

Figure 3

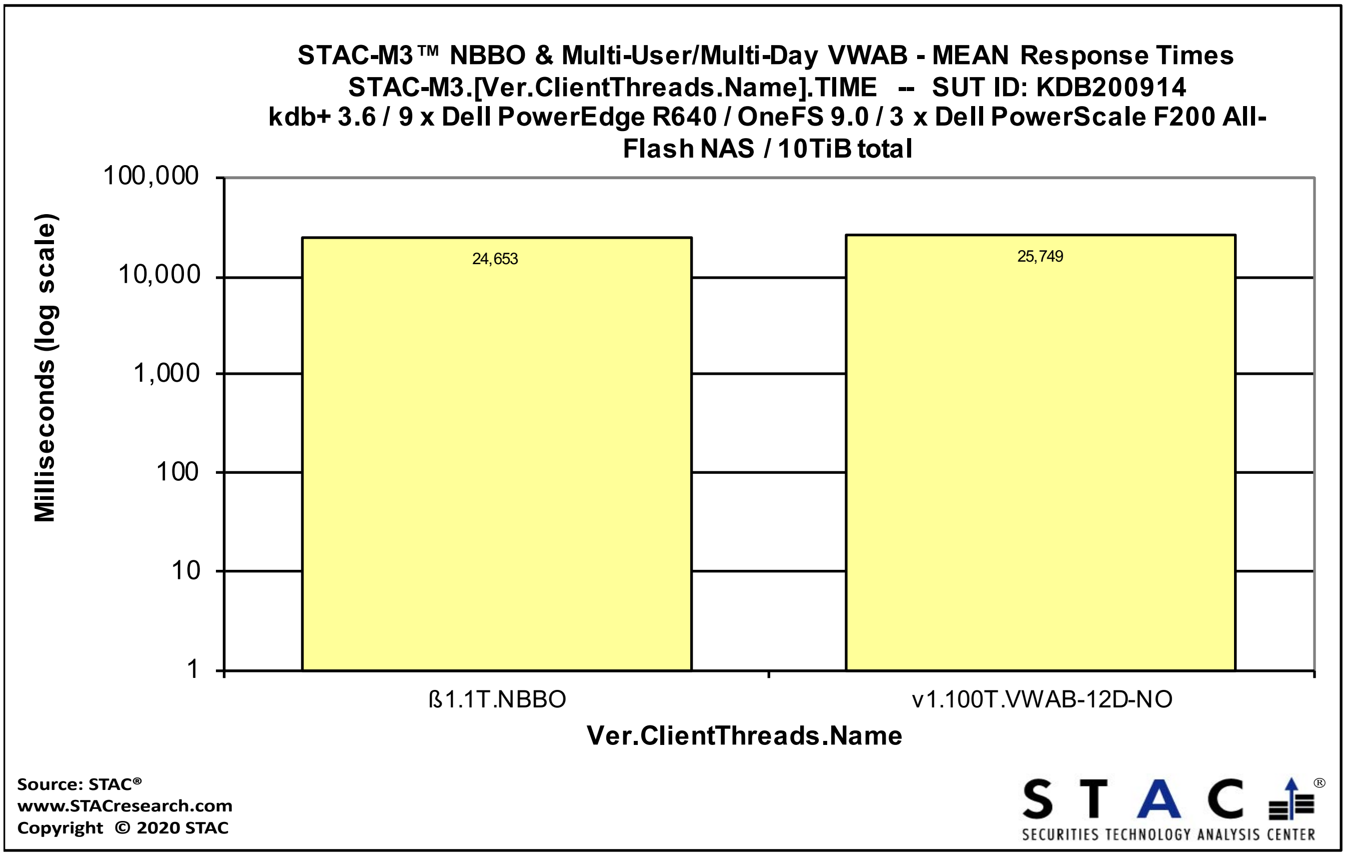

STAC-M3™ NBBO & Multi-User/Multi-Day VWAB - MEAN Response Times
STAC-M3.[Ver.ClientThreads.Name].TIME -- SUT ID: KDB200914
kdb+ 3.6 / 9 x Dell PowerEdge R640 / OneFS 9.0 / 3 x Dell PowerScale F200 All-Flash NAS / 10TiB total
Milliseconds (log scale)
100,000
10,000
1,000
100
10
1
24,653
25,749
ß1.1T.NBBO
v1.100T.VWAB-12D-NO
Ver.ClientThreads.Name
Source: STAC®
www.STACresearch.com
Copyright © 2020 STAC
SECURITIES TECHNOLOGY ANALYSIS CENTER

Figure 4

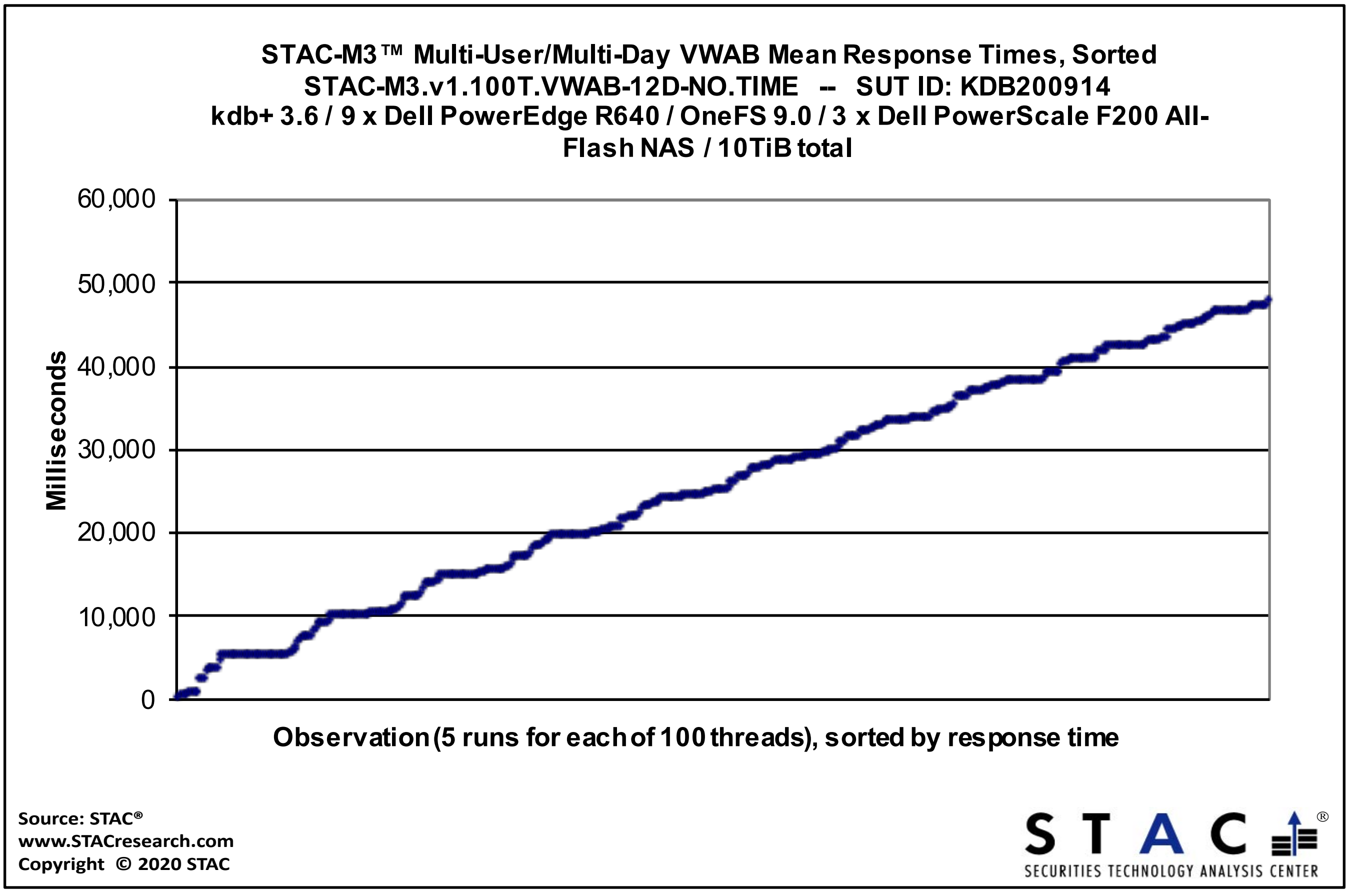

STAC-M3™ Multi-User/Multi-Day VWAB Mean Response Times, Sorted
STAC-M3.v1.100T.VWAB-12D-NO.TIME -- SUT ID: KDB200914
kdb+ 3.6 / 9 x Dell PowerEdge R640 / OneFS 9.0 / 3 x Dell PowerScale F200 All-Flash NAS / 10TiB total
60,000
50,000
40,000
30,000
20,000
10,000
0
Milliseconds
Observation (5 runs for each of 100 threads), sorted by response time
Source: STAC®
www.STACresearch.com
Copyright © 2020 STAC
STAC
SECURITIES TECHNOLOGY ANALYSIS CENTER

Figure 5

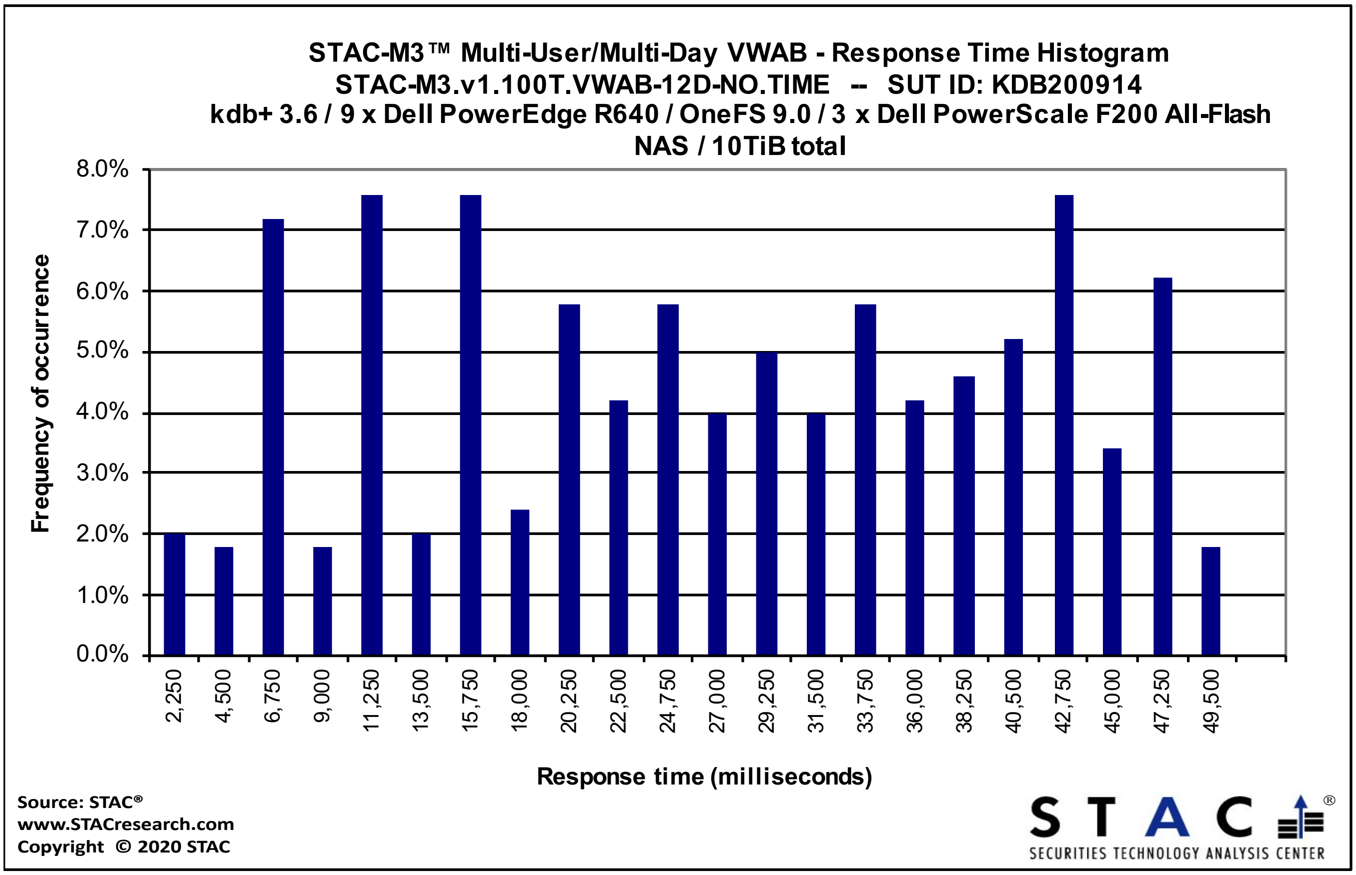

STAC-M3™ Multi-User/Multi-Day VWAB - Response Time Histogram
STAC-M3.v1.100T.VWAB-12D-NO.TIME -- SUT ID: KDB200914
kdb+ 3.6 / 9 x Dell PowerEdge R640 / OneFS 9.0 / 3 x Dell PowerScale F200 All-Flash NAS / 10TiB total
Frequency of occurrence
8.0%
7.0%
6.0%
5.0%
4.0%
3.0%
2.0%
1.0%
0.0%
2,250
4,500
6,750
9,000
11,250
13,500
15,750
18,000
20,250
22,500
24,750
27,000
29,250
31,500
33,750
36,000
38,250
40,500
42,750
45,000
47,250
49,500
Response time (milliseconds)
Source: STAC®
www.STACresearch.com
Copyright © 2020 STAC
STAC
SECURITIES TECHNOLOGY ANALYSIS CENTER

Figure 6

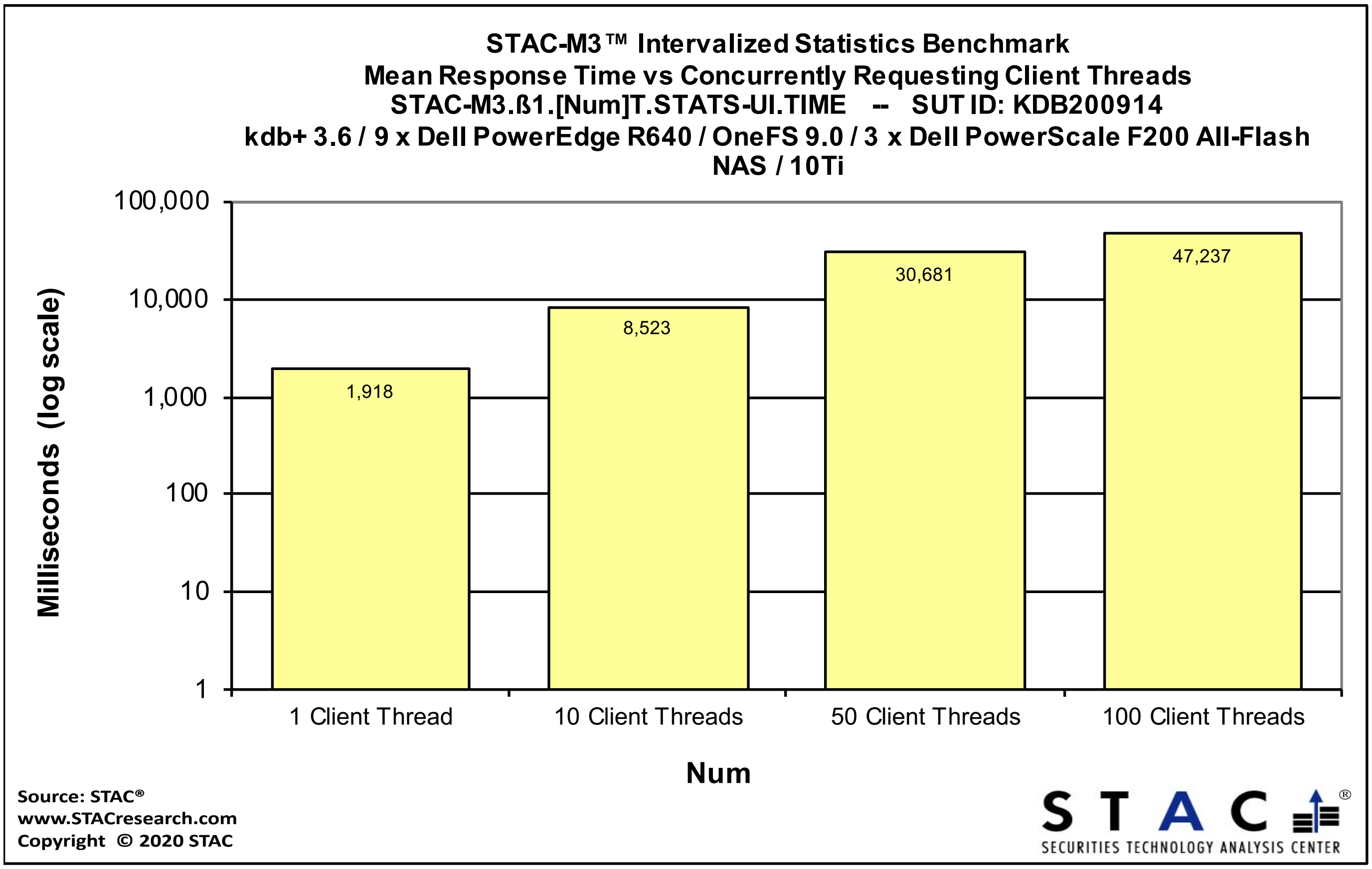

STAC-M3™ Intervalized Statistics Benchmark
Mean Response Time vs Concurrently Requesting Client Threads
STAC-M3.ß1.[Num]T.STATS-UI.TIME -- SUT ID: KDB200914
kdb+ 3.6 / 9 x Dell PowerEdge R640 / OneFS 9.0 / 3 x Dell PowerScale F200 All-Flash NAS / 10Ti
Milliseconds (log scale)
100,000
10,000
1,000
100
10
1
1,918
8,523
30,681
47,237
1 Client Thread
10 Client Threads
50 Client Threads
100 Client Threads
Num
Source: STAC®
www.STACresearch.com
Copyright © 2020 STAC
STAC
SECURITIES TECHNOLOGY ANALYSIS CENTER

Figure 7

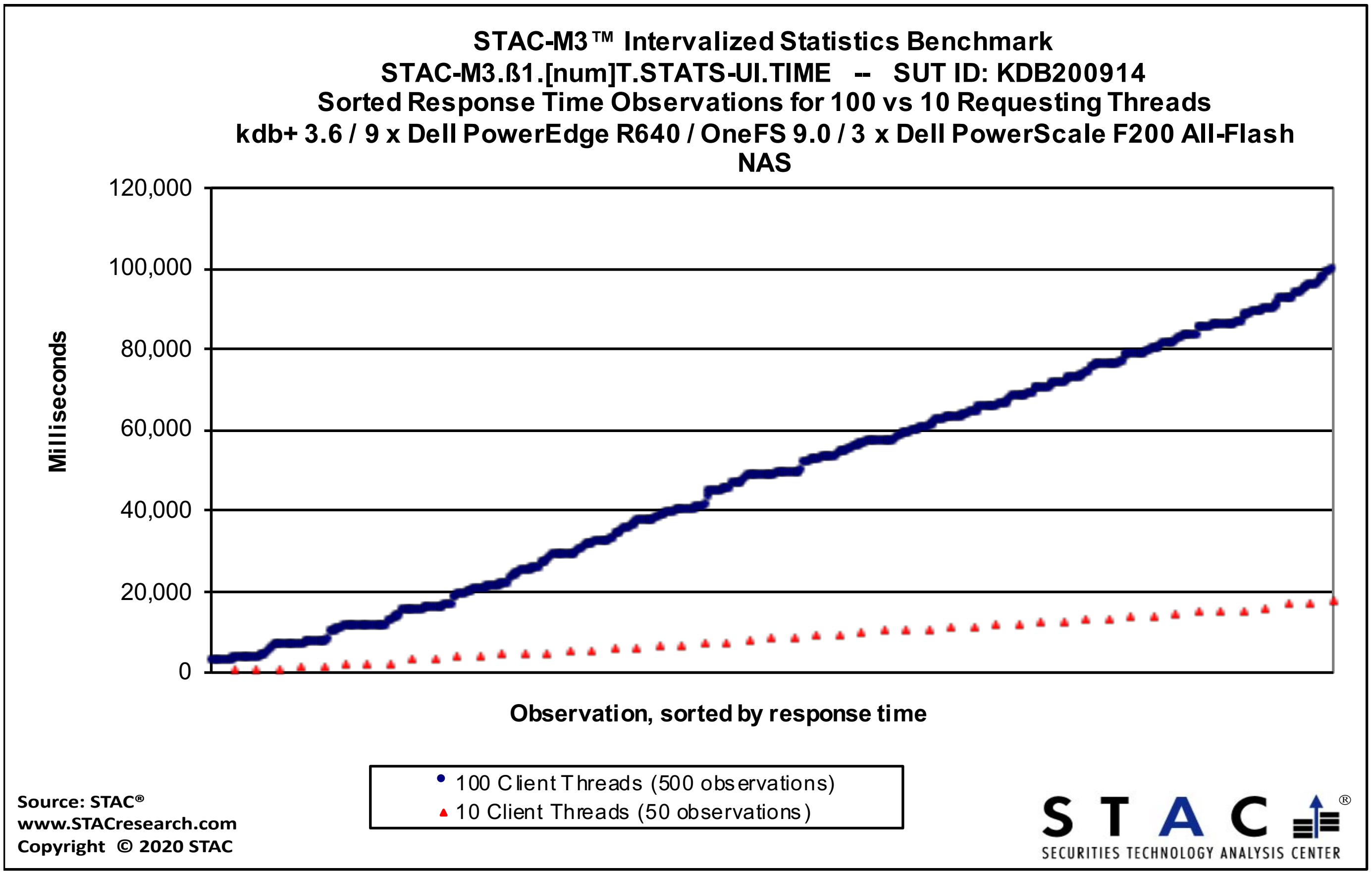

STAC-M3™ Intervalized Statistics Benchmark
STAC-M3.ß1.[num]T.STATS-UI.TIME -- SUT ID: KDB200914
Sorted Response Time Observations for 100 vs 10 Requesting Threads
kdb+ 3.6 / 9 x Dell PowerEdge R640 / OneFS 9.0 / 3 x Dell PowerScale F200 All-Flash NAS
Milliseconds
120,000
100,000
80,000
60,000
40,000
20,000
0
Observation, sorted by response time
100 Client Threads (500 observations)
10 Client Threads (50 observations)
Source: STAC®
www.STACresearch.com
Copyright © 2020 STAC
STAC
SECURITIES TECHNOLOGY ANALYSIS CENTER

Figure 8

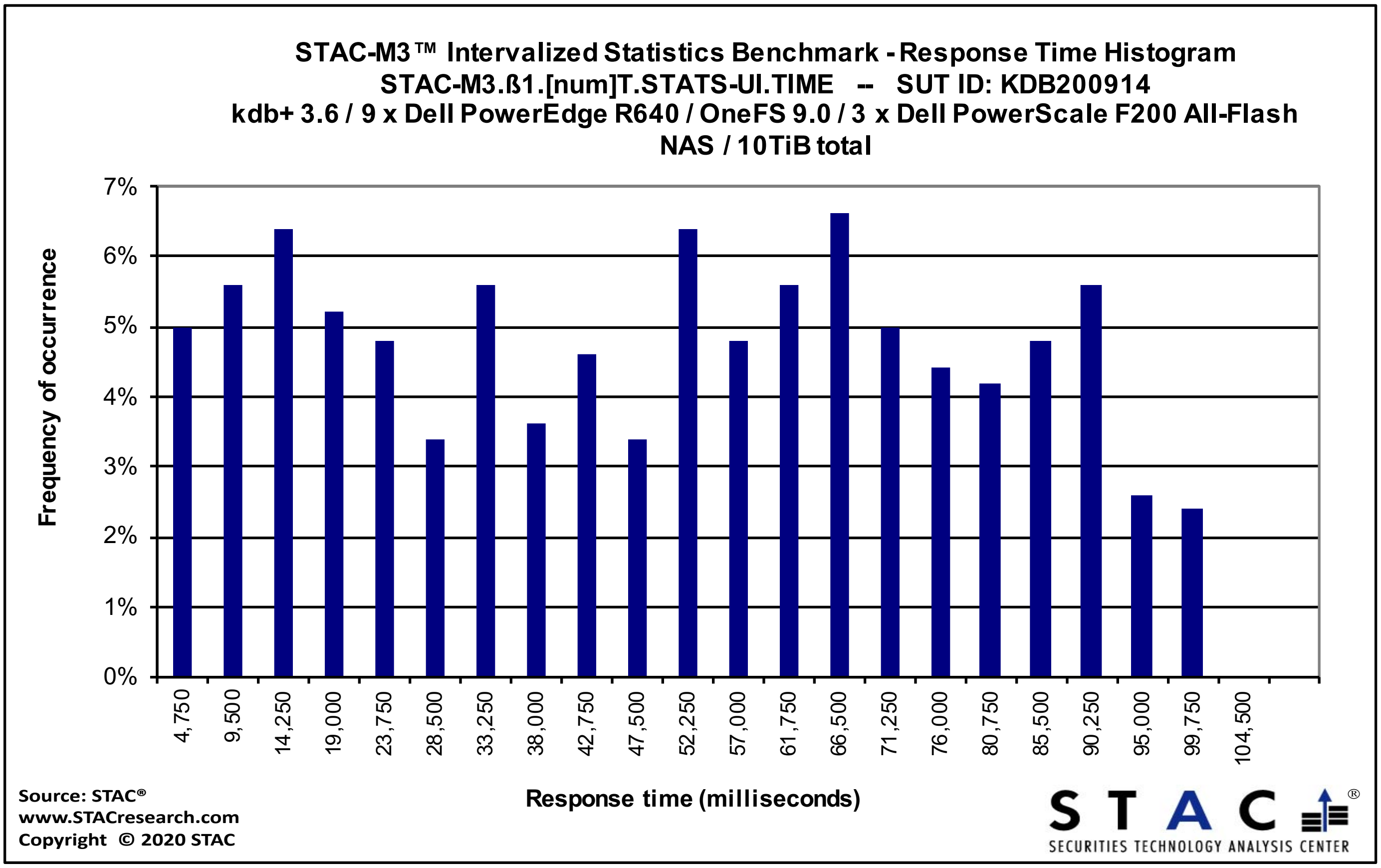

STAC-M3™ Intervalized Statistics Benchmark - Response Time Histogram
STAC-M3.ß1.[num]T.STATS-UI.TIME -- SUT ID: KDB200914
kdb+ 3.6 / 9 x Dell PowerEdge R640 / OneFS 9.0 / 3 x Dell PowerScale F200 All-Flash NAS / 10TiB total
Frequency of occurrence
7%
6%
5%
4%
3%
2%
1%
0%
4,750
9,500
14,250
19,000
23,750
28,500
33,250
38,000
42,750
47,500
52,250
57,000
61,750
66,500
71,250
76,000
80,750
85,500
90,250
95,000
99,750
104,500
Response time (milliseconds)
Source: STAC®
www.STACresearch.com
Copyright © 2020 STAC
STAC
SECURITIES TECHNOLOGY ANALYSIS CENTER

Figure 9