STAC Update:
Historical time series stacks

Michel Debiche
Director of Analytics Research, STAC

michel.debiche@STACresearch.com

STAC-M3

- Performance benchmarks for enterprise tick analytics
  - Language/DBMS neutral
  - Developed by banks and hedge funds

www.STACresearch.com/m3

- Workload:
  - Synthetic data modeled on NYSE TAQ
  - Mix of I/O- and compute-intensive operations (read-heavy)
  - Simulates concurrent access with varying number of users

- 3 Suites
  - Baseline (Antuco): Limited dataset forced to access storage
  - Scale (Kanaga): Large dataset
  - Small in-memory (Shasta): Limited dataset not forced to access storage

STAC ®
SECURITIES TECHNOLOGY ANALYSIS CENTER

STAC-M3 Suites

Suite	Computational Complexity	I/O	Input Data Size	Pre-load data?	Concurrency
Baseline (Antuco)	Low to Moderate	Variable	Small (~4 TB)	No	Varies
Small in-memory (Shasta)	Low to Moderate	Very low	Small (~4 TB)	Yes	Varies
Scale (Kanaga)	Low to Moderate	Variable	Large (>50 TB), Variable	Yes	Varies

STAC ®
SECURITIES TECHNOLOGY ANALYSIS CENTER

STAC-M3 Suites

Suite	Computational Complexity	I/O	Input Data Size	Pre-load data?	Concurrency
Baseline (Antuco)	Low to Moderate	Variable	Small (~4 TB)	No	Varies
Small in-memory (Shasta**)**	Low to Moderate	Very low	Small (~4 TB)	Yes	Varies
Scale (Kanaga)	Low to Moderate	Variable	Large (>50 TB), Variable	Yes	Varies

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

STAC-M3 Database Scaling for Kanaga Suite

STAC-M3 Kanaga dataset size by 'year,' using kdb+
(Triangles = annual, Squares = cumulative)

Note: Values indicate size of the dataset itself and may not completely reflect database or filesystem overhead, particularly for years prior to 2011.
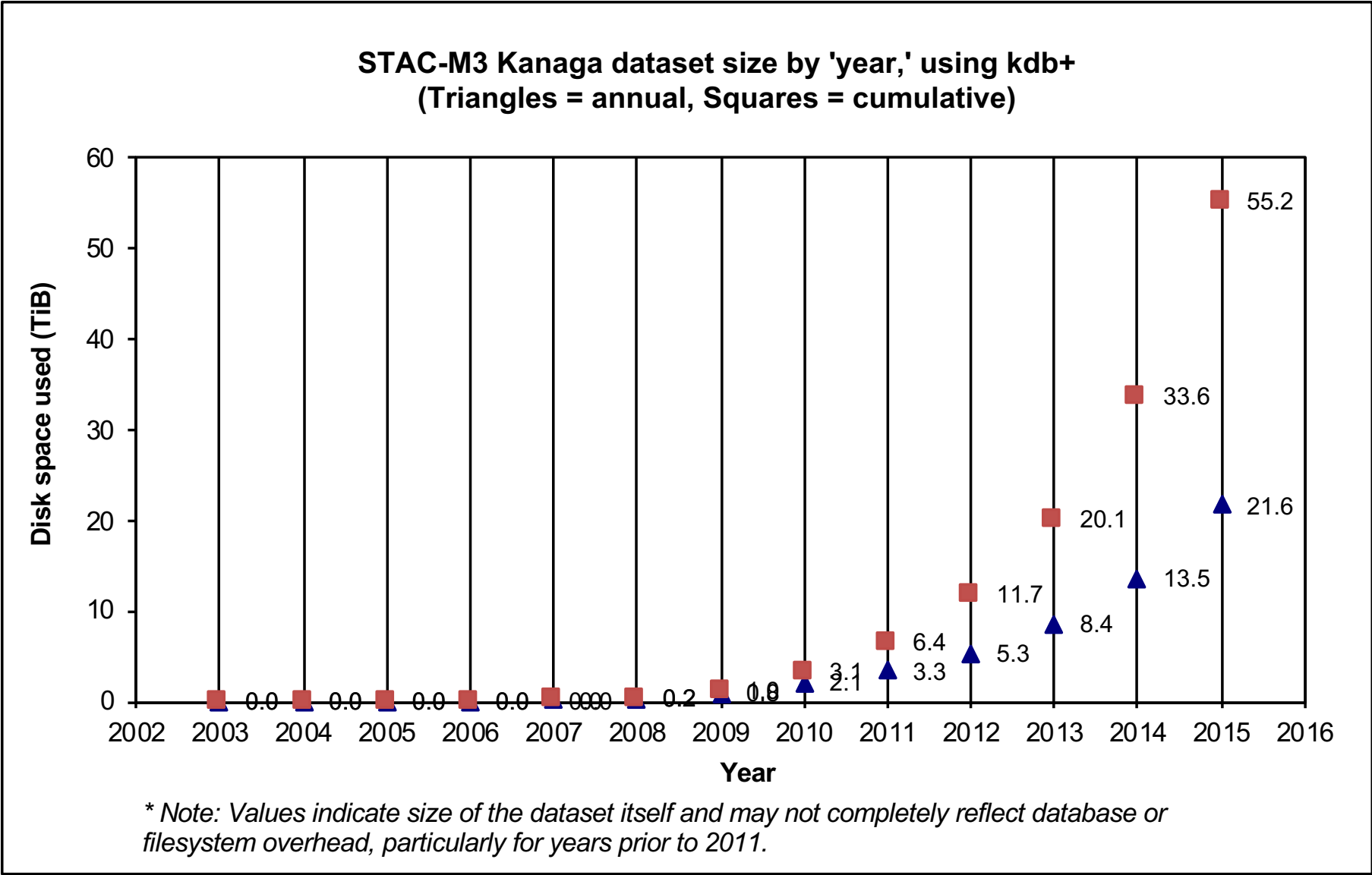
STAC-M3 (continued)

- Wide range of implementations

  - Databases: kdb+, eXtremeDB

  - Clustered file systems, parallel file systems, NFS, flash arrays, NVME over Fabric, direct-attached SSD, NAND and post-NAND Flash (e.g. Optane)

  - Single database server, database cluster (bare metal and cloud)

www.STACresearch.com/**m3**

S T A C ®
SECURITIES TECHNOLOGY ANALYSIS CENTER

STAC-M3 / kdb+ / Dell EMC R640 and PowerScale F200

- ## SUT ID: KDB200914

- ## Stack:

  - ### kdb+ 3.6

  - ### 9 x Dell EMC PowerEdge R640 servers

  - ### 2 x 18-core Intel® Xeon® Gold 6254 @ 3.10GHz

  - ### 3-node Dell EMC PowerScale F200 All-Flash Scale-Out cluster

  - ### 12 x 960 GB SSD

- ## STAC-M3 Antuco suite

www.STACresearch.com/KDB200914

STAC
SECURITIES TECHNOLOGY ANALYSIS CENTER

STAC-M3 Antuco results (baseline suite)

The Dell EMC based solution involving 9 database servers accessing networked flash storage:

- Had the highest storage efficiency of any publicly reported solution involving kdb+ 3.6 (STAC-M3.v1.1.STORAGE.EFF)
  - 166% for this SUT vs. 149% for all others
- was 7.7x the speed of a solution involving a parallel file system with 14 database servers and 18 storage servers (SUT ID KDB200401) in the single-user NBBO operation (STAC-M3.β1.1T.NBBO.TIME)

www.STACresearch.com/KDB200914

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

The PowerScale F200 based solution was faster than a solution using Dell EMC's larger flash storage appliance (SUT ID KDB190430) in 2 of 17 mean-response time 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)

www.STACresearch.com/KDB200914

STAC

SECURITIES TECHNOLOGY ANALYSIS CENTER

STAC-M3 / kdb+ / DDN EXAScaler & AI400X Flash appliances

- SUT ID: KDB200915

- Stack:
  - kdb+ 3.6
  - 15 x Intel® S2600BPB servers
  - 1 x 20-core Intel® Xeon® Gold 6138
  - 2 x DDN AI400X storage appliances
  - 48 x 3.8TB NVME SSD total
  - DDN EXAScaler Parallel filesystem v5.1.1

- STAC-M3 Antuco and Kanaga suites

www.STACresearch.com/KDB200915

STAC
SECURITIES TECHNOLOGY ANALYSIS CENTER

- The DDN EXAScaler and AI400X based solution outperformed all publicly disclosed results in all year-high bid throughput benchmarks (STAC-M3.β1.1T.*.BPS)

www.STACresearch.com/KDB200915

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

Versus a solution based on a parallel filesystem

Versus a kdb+ solution based on a parallel filesystem with 14 db servers & 18 storage servers (KDB200401), the DDN solution was faster in 4 of 17 Antuco mean-response time benchmarks, including:

- 9.4x speedup in NBBO (STAC-M3.β1.1T.NBBO.TIME)

- 46% faster in 10-user market snapshot (STAC-M3.β1.10T.MKTSNAP.TIME)

- 15% faster in 10-user aggregate stats (STAC-M3.β1.10T.STATS-AGG.TIME)

- 3% faster in 100-user interval stats (STAC-M3.β1.100T.STATS-UI.TIME)

www.STACresearch.com/KDB200915

STAC
SECURITIES TECHNOLOGY ANALYSIS CENTER

Versus a solution based on networked flash storage (Antuco)

Versus a kdb+ solution based on networked flash storage with 4 database servers (KDB190430), the DDN solution:

- was faster in 14 of 17 Antuco mean-response time benchmarks, including:
  - 12x speedup in 10-user market snapshot (STAC-M3.β1.10T.MKTSNAP.TIME)
  - 6x speedup or greater in all four user counts doing intervalized stats (STAC-M3.β1.*.STATS-UI.TIME)
  - 5x speedup in 10-user aggregate stats (STAC-M3.β1.10T.STATS-AGG.TIME)

www.STACresearch.com/KDB200915

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

Versus a kdb+ solution based on networked flash storage with 4 database servers (KDB190430), the DDN solution:
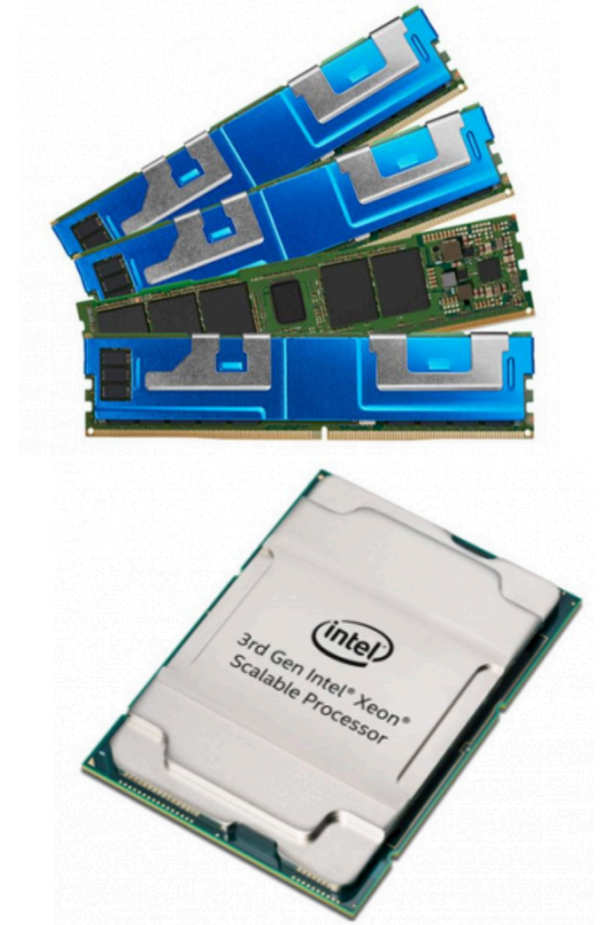
- was faster in 20 of 24 Kanaga mean-response time benchmarks, including:
  - 19x speedup or greater in all four 10-user market snapshots (STAC-M3.β1.10T.YR[n]-MKTSNAP.TIME)
  - 3x speedup or greater in year-high bid for all four years (STAC-M3.β1.1T.[n]YRHIBID.TIME)

www.STACresearch.com/KDB200915

STAC
SECURITIES TECHNOLOGY ANALYSIS CENTER

- SUT ID: KDB201016

- Stack:
  - kdb+ 4.0
  - Lenovo SR860 V2 server
  - 4 x Intel® Xeon Platinum 8380HL (Cooper Lake)
  - 24 x 512 GiB Intel® PMem 200 Series
  - 1 x 8TB Intel SSD DC P4510 (3DNAND)

- STAC-M3 Antuco & Kanaga (2-yr.)

- Unaudited results in the STAC Vault

www.STACresearch.com/KDB201016

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER