STAC Update:
Big data

Peter Lankford
Founder and Director, STAC

peter.lankford@STACresearch.com

STAC-M3

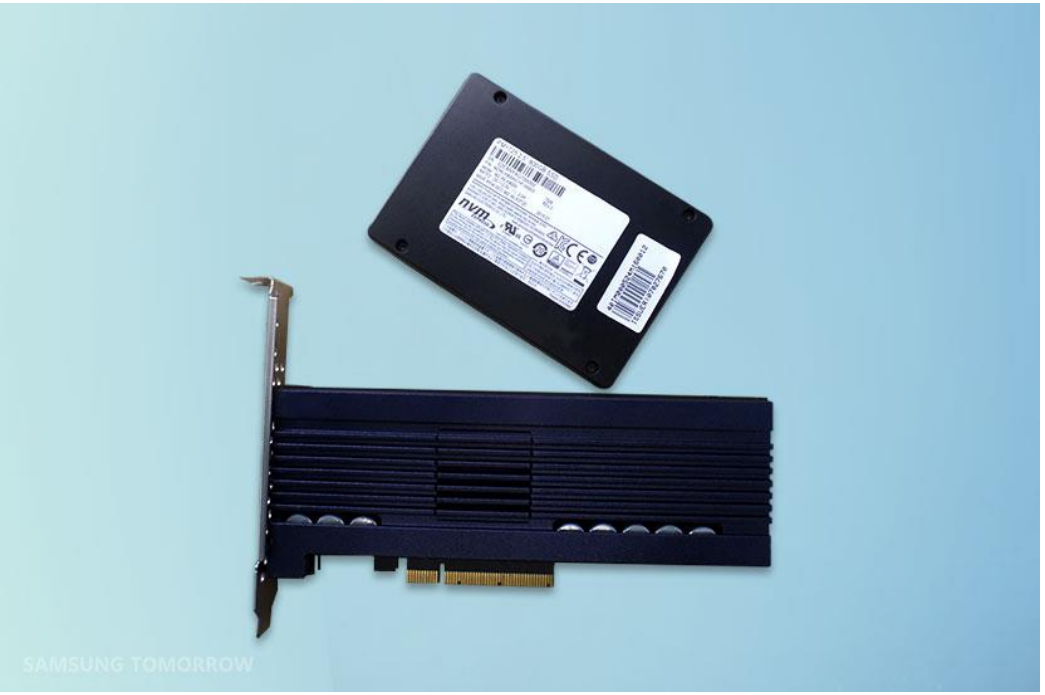
- Performance benchmarks for enterprise tick analytics
  - Language/DBMS neutral
  - Developed by banks
- Workload:
  - Synthetic data modeled on NYSE TAQ
  - Mix of I/O- and compute-intensive operations (read-heavy)
  - Scalable volume and number of users

www.STACresearch.com/m3

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

STAC-M3 / kdb+ / Dell 960 / Samsung PM1725-a SSD

- SUT ID: KDB180418b
- Stack:
  - Software: kdb+ 3.5 / RHEL 7 / xfs
  - Processor: 4 x Intel Gold 6154 (Skylake)
  - Server: 1 x Dell PowerEdge R960
  - Storage: 6x 1.6 TB PM1725a Samsung NVMe SSD
- Antuco suite (baseline benchmarks)



www.STACresearch.com/KDB180418b

STAC SECURITIES TECHNOLOGY ANALYSIS CENTER

Highlights

- In 11 of 17 mean response-time benchmarks, this solution performed the same as or better than a solution with 6 x 1.6TB 3D NAND flash drives from a competitor (KDB170703). Including:
  - 1.8x the performance in STAC-M3.β1.1T.MOHIBID.TIME and STAC-M3.β1.1T.WKHIBID.TIME
  - 1.7x the performance in STAC-M3.β1.100T.VWAB-12D-NO.TIME

STAC SECURITIES TECHNOLOGY ANALYSIS CENTER

STAC-M3 / kdb+ / Dell R960 / Samsung Z-SSD

- SUT ID: KDB180418a
- Stack:
  - Same system as above, but with 6 x 800GB Samsung SZ985 (Z-SSD)
- Antuco suite (baseline benchmarks)

www.STACresearch.com/KDB180418a

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

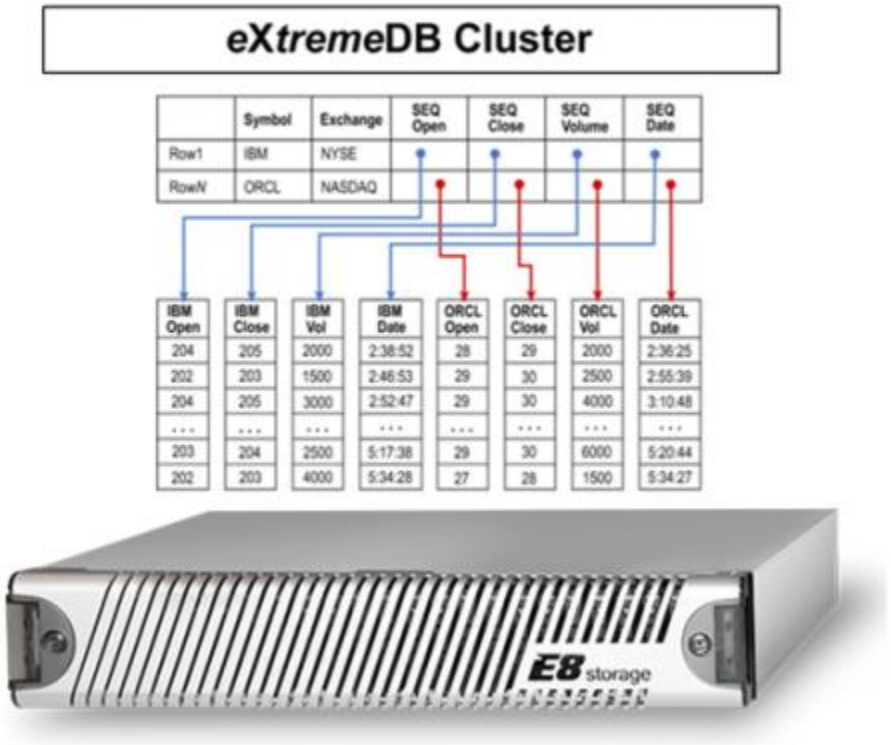
Highlights

- In 8 of 17 mean response-time benchmarks, this solution performed the same as or better than a solution with 6 x 750GB 3DXPoint drives from a competitor (KDB171010). Including:
  - 1.5x the performance in STAC-M3.β1.1T.STATS-UI.TIME
  - 1.4x the performance in STAC-M3.β1.10T.STATS-UI.TIME and STAC-M3.β1.10T.STATS-AGG.TIME

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

eXtremeDB & E8 Storage

- SUT ID: KDB180416
- Stack:
  - McObject eXtremeDB Financial Edition 8.0
  - E8 Storage E8-X24 array with Intel Optane D4800X NVMe SSD (NVMe-oF)
  - 6 x Supermicro SYS-2028TP-HTTR
  - CentOS 7.3.1611 with ext4
- Antuco suite (baseline benchmarks)



www.STACresearch.com/XTR180416

STAC SECURITIES TECHNOLOGY ANALYSIS CENTER

eXtremeDB & E8 Storage

- Faster mean response times than all previously published best results for 5 of the 17 operations
  - More consistent response times (lower standard deviation) than all previous best results for 7 of the 17 operations.
- Compared to the previously published best results for a non-McObject database (SUT ID: KDB171010), this SUT was:
  - 17.0x for the write benchmark (STAC-M3.β1.1T.WRITE.TIME)
  - 9.1x for the NBBO benchmark (STAC-M3.β1.1T.NBBO.TIME)
  - 2.4x for volume curves (STAC-M3.β1.10T.VOLCURV.TIME)

STAC-A3

- Recap:
  - Workloads that emulate real-world backtesting jobs
  - Measure speed, scalability, efficiency of any architecture
- Test harness hands the implementation jobs to execute
- Measures the throughput and efficiency of the SUT
- Currently defined algos are almost always bottlenecked on I/O

STAC
SECURITIES TECHNOLOGY ANALYSIS CENTER

Spark with Xenon on Vexata

- SUT ID: LEVX180608
- Stack:
  - Apache Spark 2.1.0
  - Levyx Xenon 3.2.0
  - 6 x Supermicro servers with 2 x 20-core Intel Xeon Scalable (Gold) & 512GB DRAM
  - Vexata VX-100F array with 64 x Intel P3700 SSD

report coming soon

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

Spark with Xenon on Vexata

- In STAC-A3.β1.SWEEP.SPEED2.BEST (instrument simulations/second on a basket of ideal size for this solution):
  - 69.5 instrument simulations per second using 6 servers VS
  - Previous record: 51.4 instrument simulations using 20 servers (SUT ID HPAT171029)
- In STAC-A3.β1.SWEEP.SPEED2.BASE (instrument simulations/second on a basket of 50 instruments)
  - This SUT was 69% faster than the previously published best results on 5 servers (SUT ID HPAT171028) despite having only 20% more servers

STAC
SECURITIES TECHNOLOGY ANALYSIS CENTER

Converged big data & big compute workloads

- Some workloads are both compute and data intensive
- Bottlenecks depend on architecture
- STAC-M3 has some of this, but working to introduce more
- First will be an additional algorithm for STAC-A3

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

What’s coming soon

- Variant that involves option pricing & Greeks
- Still trawls through huge amounts of data
- But has to do huge amounts of compute as it analyzes the data
- Also evaluating adding a portfolio optimization algo
- Stay tuned (and reach out if you’re interested)