STAC Update:
Big data benchmarks

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 / Vexata / kdb+

- SUT ID: KDB170421
- Stack:
  - Software: kdb+ 3.5 / CentOS 7.2
  - Servers: 4 x white boxes with 2 x 22-core Broadwells
  - Storage: 1 x 76 TiB Vexata VX-100 Flash Array
  - Interconnect: Fibre Channel
- Baseline tests (Antuco)
- Scale tests (Kanaga)

www.STACresearch.com/vexata

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

Highlights – baseline suite (Antuco)

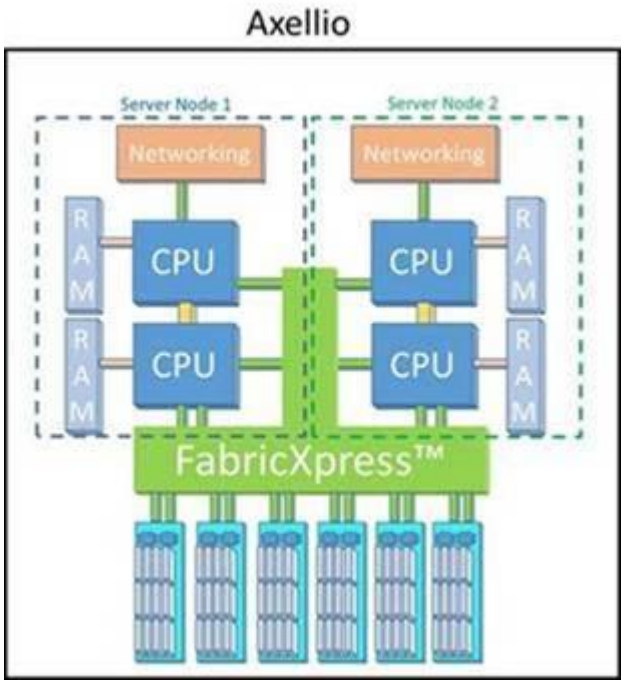
- Compared to all publicly benchmarked solutions:
  - Set new records in 8 of 17 mean response-time benchmarks
- Compared to the most recent record-setting system (SUT ID KDB160930), which used a PCIe-connected flash array and 4 servers with the same processors as this Fibre Channel-connected Vexata solution, the Vexata solution:
  - Was faster in 14 of 17 benchmarks
  - Including all of KDB160930's previous records
  - Was 68% faster in STAC-M3.β1.1T.NBBO.TIME
  - Was 47% faster in the STAC-M3.β1.1T.YRHIBID.TIME

Highlights – volume & user scaling (Kanaga)

- Compared to all publicly reported STAC-M3 Kanaga results to date
  - Set new records in 14 of 24 mean response-time benchmarks
  - Set new records for mean throughput in all of the multi-year high-bid benchmarks
    - Reaching 36.88 GB/sec in STAC-M3.β1.5T.YRHIBID.MBPS
- Compared to SUT ID KDB160930:
  - Was faster in 14 of 24 benchmarks
  - Was 72% faster in STAC-M3.β1.1T.5YRHIBID.TIME
    - The most intense sequential read workload

STAC-M3 / X-IO Axellio

- SUT ID: KDB170420
- Results are in the STAC Vault
  - www.STACresearch.com/KDB170420
- Stack:
  - Software: kdb+ 3.5 / CentOS 7.3 / ext4
  - 1 x X-IO Axellio Edge Computing System
- Both baseline and scale tests



STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

STAC-M3 / Intel Optane

- SUT IDs: KDB170521, KDB170522
- Results are in the STAC Vault
  - www.STACresearch.com/KDB170522
- Stack:
  - Software: kdb+ 3.4 / CentOS 7.3 / ext4
  - Server: White box with 2 x 22-core Broadwells
  - Storage:
    - KDB170521: 4 x Intel SSD DC P4800X (Optane), 375GB
    - KDB170522: 4 x Intel SSD DC P3700, 2TB (NAND flash)
- Not a complete STAC-M3 database
- Not official benchmarks
- Still a useful comparison

STAC-A3

- Recap:
  - Workloads that emulate real-world backtesting jobs
    - Range of parallelism and IO/compute intensity
  - Measure speed, scalability, efficiency of any architecture
- Research agenda:
  - Shared storage (e.g., parallel FS) and shared-nothing (e.g., HDFS)
  - Drive insight & improvements at both software and hardware level
- Spark implementation now available:
  - STAC-A3 Pack for Spark rev B
  - Scala implementation originally by Cloudera and Intel
  - Enhanced by Levyx (moved to data frames, changed data layout)

www.STACresearch.com/**backtesting**

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

STAC-A3 / Spark with Levyx Xenon on GCP

- SUT ID: SPRK170603
- Stack:
  - STAC-A3 Pack for Spark rev B
  - Apache Spark 1.6.1
  - Levyx Xenon 3.2.0 software
  - Google Cloud Platform 5-node cluster
    - n1-standard-64, 3TB local SSD per node
- Results (STAC Report forthcoming shortly):
  - STAC-A3.β1.SWEEP.MAX60 = 200 simulations
    - Max simulations using 50 instruments that could be completed in 60 minutes
  - STAC-A3.β1.SWEEP.SPEED = 14.08 simulations per minute
    - SWEEP.MAX60 divided by the elapsed time of the SWEEP.MAX60 sequence
  - Max instrument simulations/second in scaling tests = 12.6
    - 100 simulations on 200 instruments in ~26.6 minutes

STAC-A3 / Spark with Levyx Xenon on GCP

- Compared to SUT ID INTC141220-VI
  - Hadoop Streaming/Python (CDH 5.4.8)
  - 14 x Dell PowerEdge servers
  - Direct-attached spindle-based storage

	INTC141220-VI	SPRK170603	Improvement
STAC-A3.ß1.SWEEP.MAX60	25	200	8x
STAC-A3.β1.SWEEP.SPEED	0.43	14.08	32.7x
Max instrument simulations/second	2.5	12.6	5.0x

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

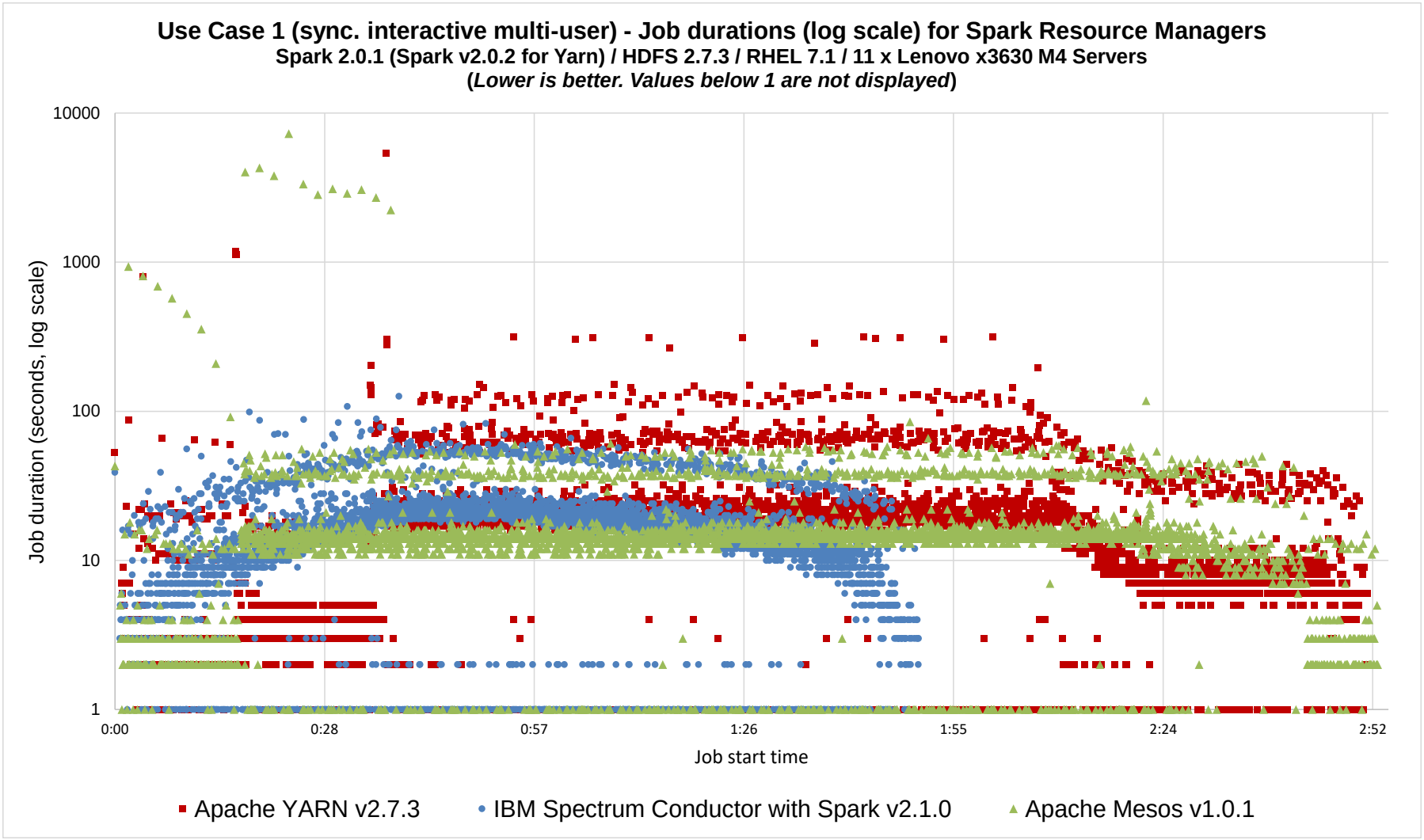
Custom benchmark: Spark resource managers

- Spark supports many resource managers
- Last year, we tested three:
  - YARN
  - Mesos
  - IBM Platform Conductor for Spark
- Used first version of open source benchmark by IBM
  - Spark Multi-User Benchmark (SMB-1)
  - Exercises the resource manager
  - Single job type: small sort
  - Synchronous submission pattern, simulating interactive use case.

Custom benchmark: Spark resource managers

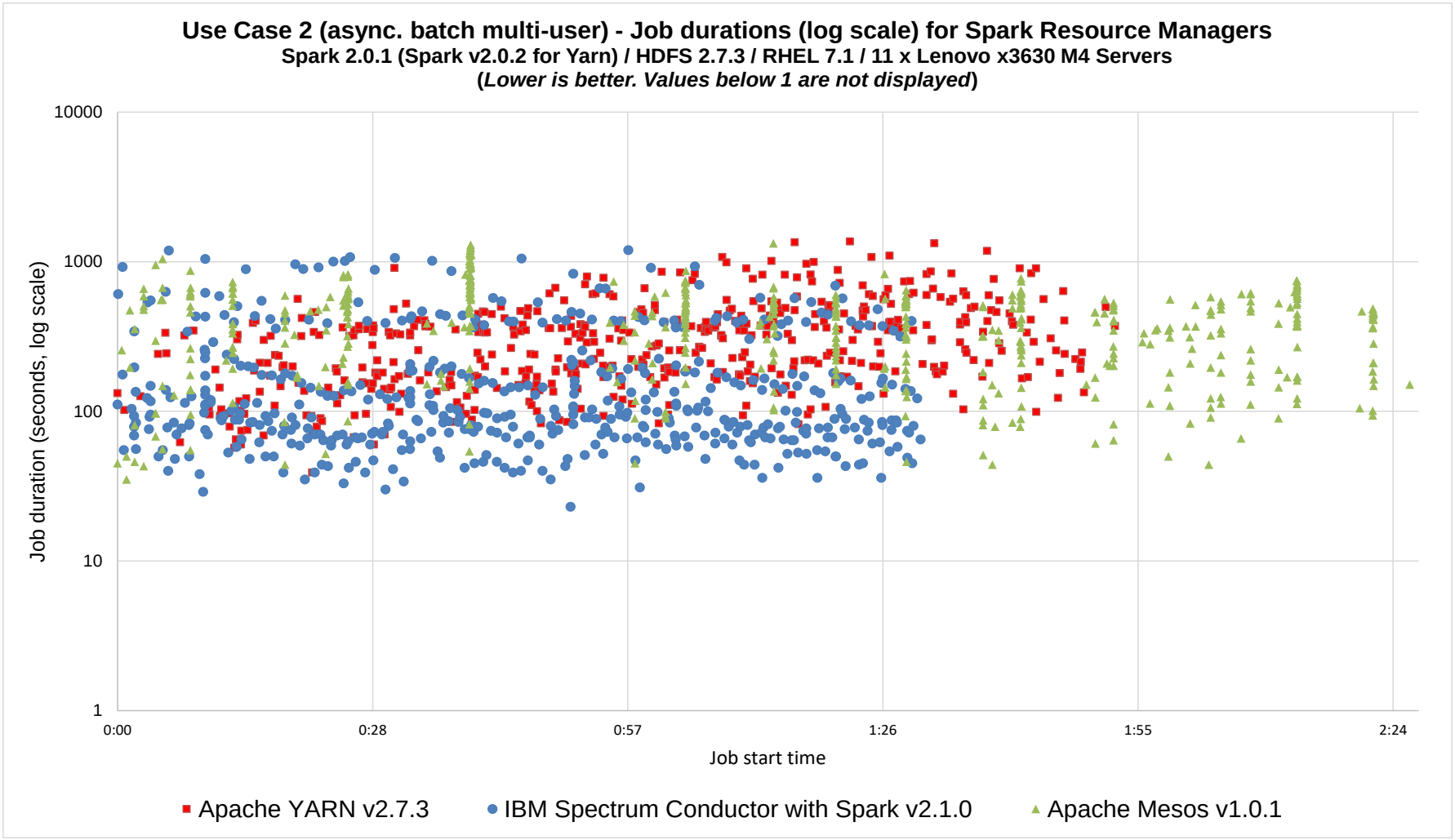
- This year, we tested them again
- Used second version of open source benchmark by IBM
  - Spark Multi-User Benchmark (SMB-2)
- Multiple job types:
  - Synchronous interactive jobs
  - Asynchronous batch jobs
- Four use cases:
  - Synchronous interactive multi-user
  - Asynchronous batch multi-user
  - Mixed multi-user (interactive & batch)
  - Mixed multi-tenant (interactive & batch with prioritization)

Results - Visualizations



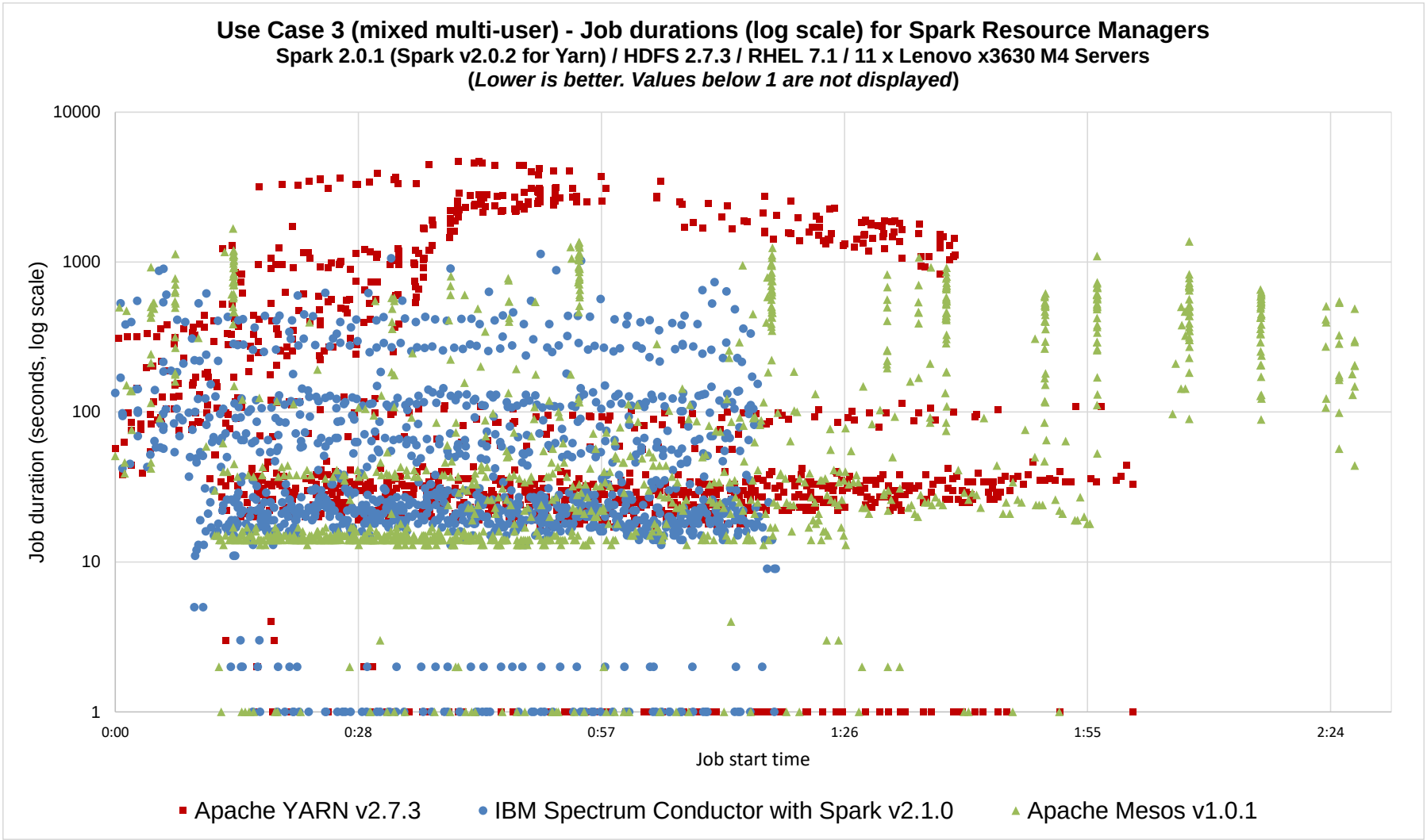
STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

Results - Visualizations



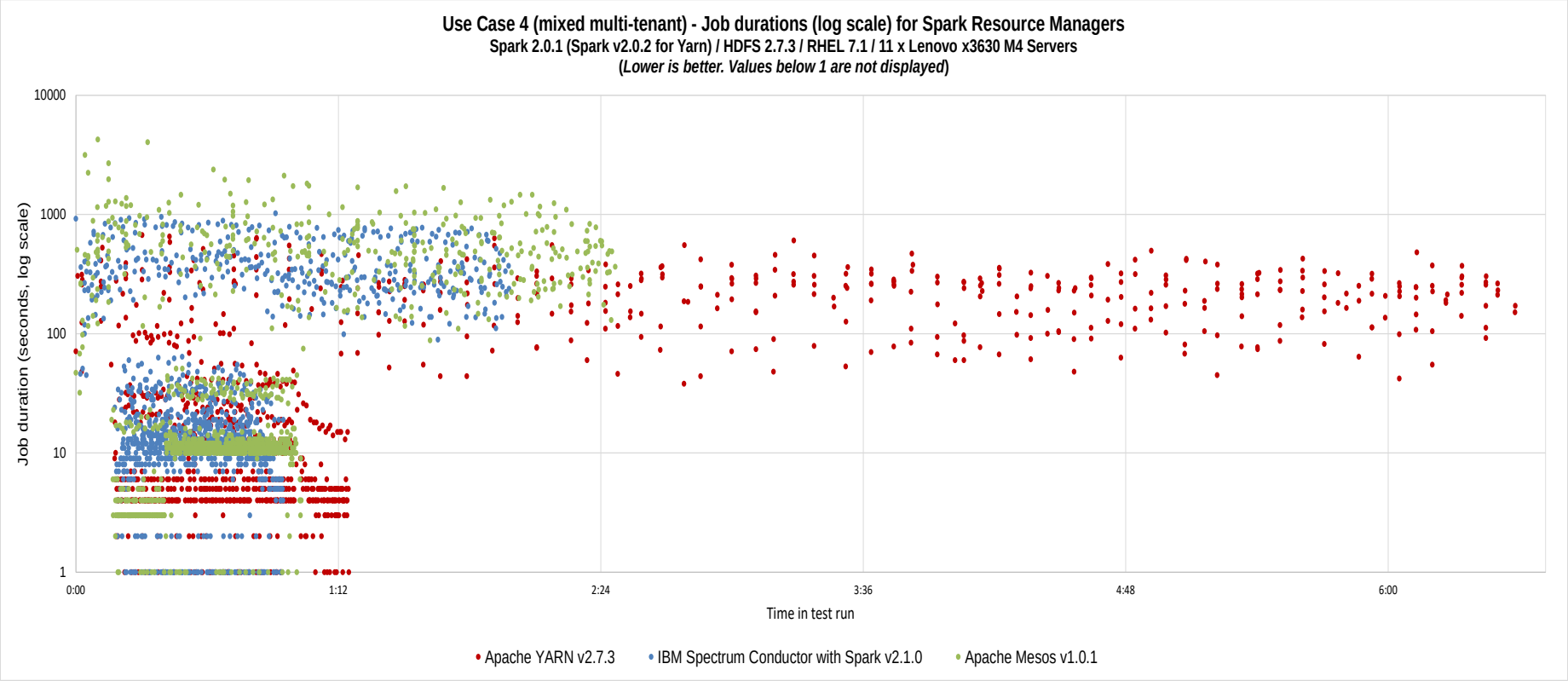
STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

Results - Visualizations



STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

Results - Visualizations



STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

Results – By the numbers

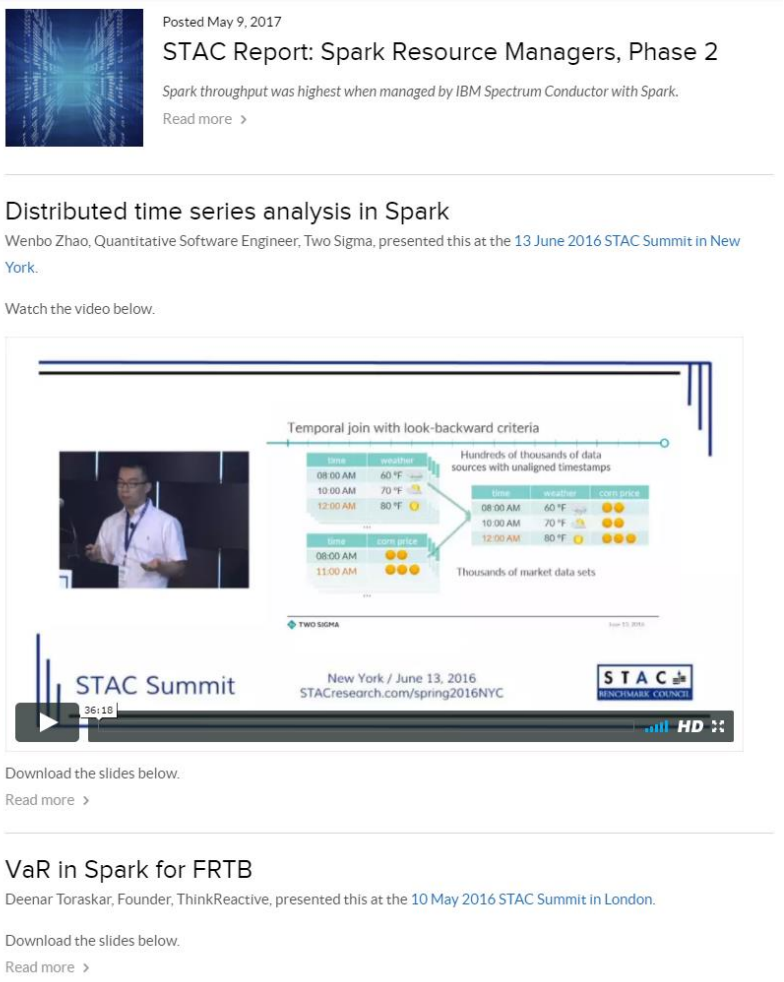
	Use Case 1: Synchronous interactive multi-user		Use Case 2: Asynchronous batch multi-user		Use Case 3: Mixed multi-user		Use Case 4: Mixed multi-tenant	
	YARN	Mesos	YARN	Mesos	YARN	Mesos	YARN	Mesos
Throughput advantage	1.6x	1.6x	1.3x	1.6x	1.5x	1.9x	3.2x	1.3x
RSD advantage	4.3x	10.9x	1.3x	2.0x	2.1x	2.6x	2.5x	2.7x

IBM Spectrum Conductor with Spark v2.1.0 (as patched) / Apache YARN v2.7.3 / Apache Mesos v1.0.1

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

Spark & STAC

- www.STACresearch.com/spark
- Single link is easy way to keep up with STAC & Spark



STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER