STAC Update for Fast Data

Peter Lankford
Founder and Director, STAC

peter.lankford@STACresearch.com

HPE Trade & Match Server with SLES and Mellanox

- SUT ID: HPE170814
  - HPE ProLiant XL170r Gen9 Servers
  - 1 x Intel E5-1680 v3 (Haswell) processor
  - Custom HPE tools to enable over-clocking
  - HPE Ethernet 10/25Gb 2-port 640SFP28
  - 25 GbE
  - SUSE Linux ES 12 SP2
  - Mellanox VMA 8.4.3
- Tested using STAC-N1
  - Measures the performance of a host network stack (server, OS, drivers, host adapter)
  - Round-trip software timestamping
  - Market data style workload
  - Matrix of message rates and message sizes

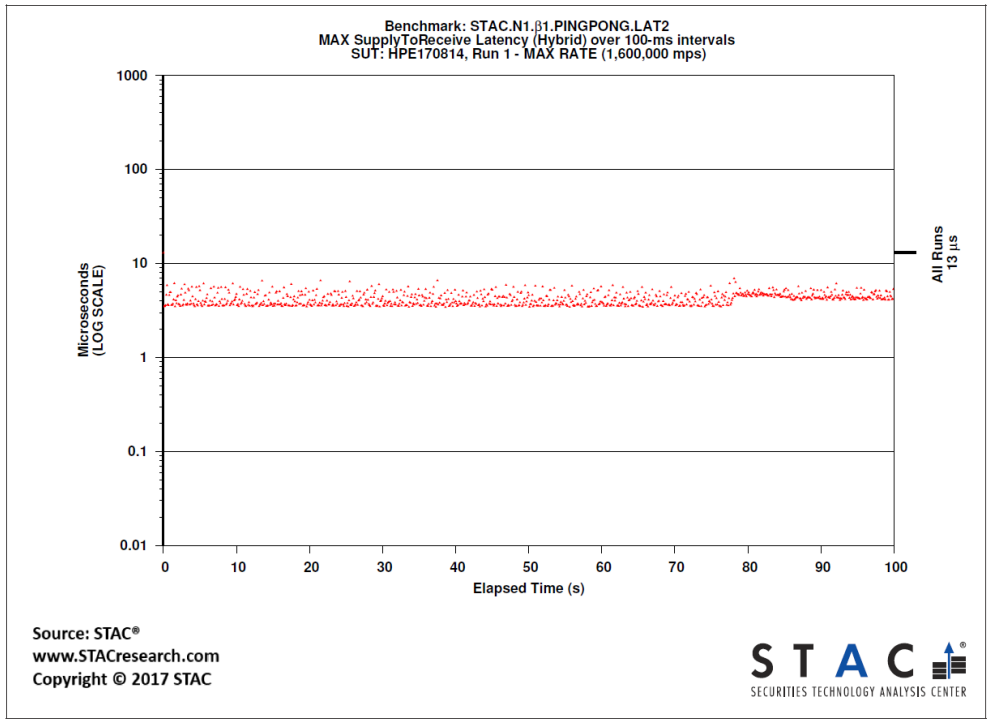
www.STACresearch.com/**HPE**

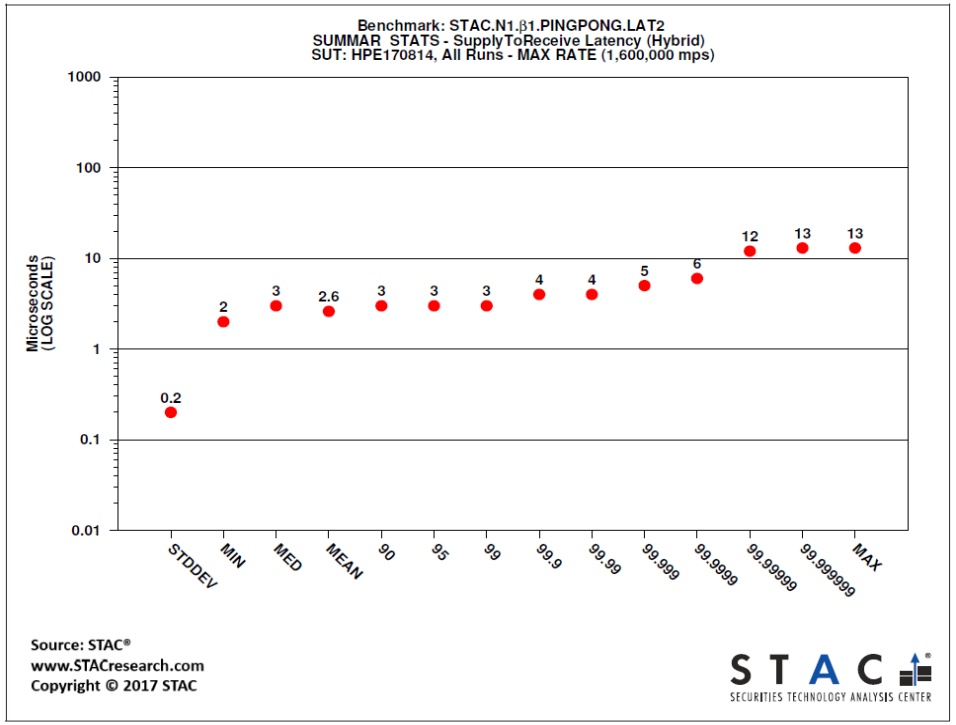
STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

HPE Trade & Match Server with SLES and Mellanox

Compared to all other public STAC-N1 reports of Ethernet-based SUTs:

- Lowest mean latency at both the base rate and the highest rate tested.
- Lowest max latency at the base rate.
- Lowest max latency at or above 1 million messages per second.
- Highest max rate reported: 1.6 million messages per second.
- 99.9999th percentile latency of just 5 µsec at the base rate and 6 µsec at 1.6 million mps





STAC
SECURITIES TECHNOLOGY ANALYSIS CENTER

Introducing STAC-T0

- Proposed specifications developed in consultation with Council members
- First use: testing an FPGA-base stack
  - Solarflare + LDA firmware + AlphaData/Xilinx FPGA + Penguin Computing server
  - Will review results a few slides later
- Please feed back

STAC-T0 goals

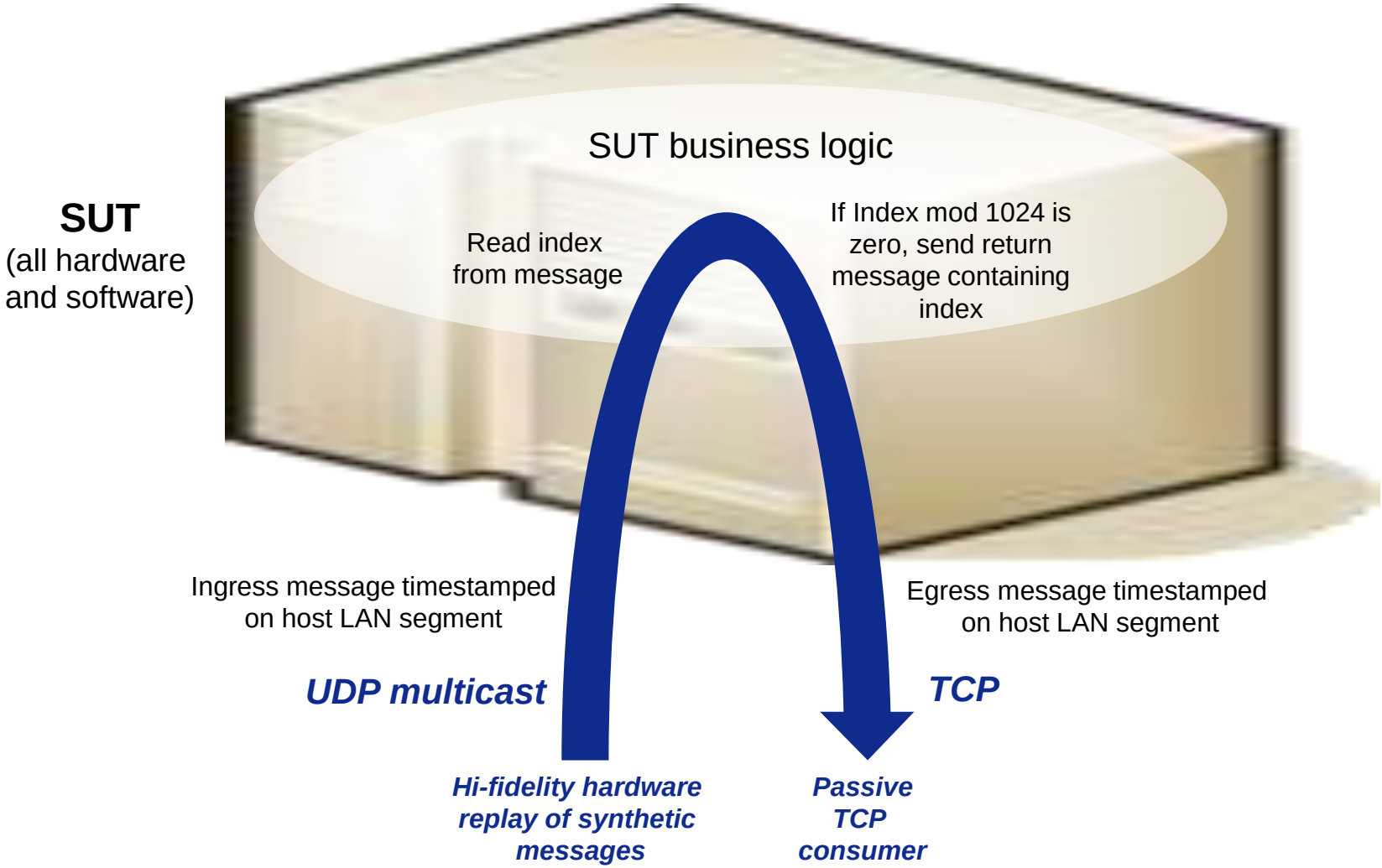
- Isolate network I/O latency
  - Latency that cannot be squeezed out by business logic
- Do not co-mingle I/O latency with market-specific logic
- Maximize accuracy
- Work with any trading platform
  - Software with kernel bypass
  - FPGA with IP cores (PHY, TOE, etc.)
- Exercise tick to trade pattern
  - UDP in, TCP out
- Can measure any software/hardware technology stack. E.g.:

Don’t we have that already?

- Not really
- STAC-N1
  - Software timestamps – Resulting latency metrics have too much jitter for many cases
  - Requires STAC software on the SUT. Can’t work with FPGA.
- STAC-T1
  - Wire-to-wire (UDP in, TCP out) and can measure any technology stack
  - However, includes market-specific logic that obscures underlying TCP/IP networking
    - Market data decode
    - Order generation
    - Product functionality (e.g., handling ACKs, Fills)
  - Requires SUT to implement market-specific logic it might not otherwise support
  - Logistical obstacle to benchmarking
  - Have to update when protocols change
  - “Race to the bottom” on market-specific functionality

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

STAC-T0 – High level



- Like a more universal and less complex STAC-T1
- Like a more universal and more accurate STAC-N1

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

Details: packets

- Two UDP ingress packet structures (one structure in a given stream):
  - Structure A: 503 bytes including headers (507B including FCS)
  - Structure B: 64 bytes including headers (68B including FCS)
- Index in ingress packet is an unpredictable uint64
  - Stands in for price, side, size, other relevant info
  - Begins in a fixed location within each packet structure
- Algorithm:
  - Send TCP message whenever Index is 0 mod 1024 (decimal)
  - Requires SUT to read entire Index but spend very little time evaluating
- TCP egress structure:
  - 118 bytes including headers (64 byte payload)
  - Index in specific location
  - Other payload bytes are static

Details: timestamping & benchmarks

- Timestamps
  - t_{LI} = time of the last bit of the ingress frame
  - t_{FI} = time of the first bit of the ingress frame
  - t_{IND} = time of the last bit of the Index
    - This is the first moment that the SUT has the information necessary to trigger an order
    - Harder to figure out in real life
  - t_{LO} = time of the last bit of the egress packet
  - t_{FO} = time of the first bit of the egress packet
  - Monitors generally timestamp just one bit within a each packet
    - Timestamps of other bits can be inferred from theoretical serialization delay
    - E.g., if the timestamper reports t_{FO}, then $t_{FI} = t_{FO}$ - serialization delay of (bits in packet - 1)
- Benchmarks
  - STAC-T0.ACTIONABLE.LAT: $t_{FO} - t_{IND}$
  - STAC-T0.FILO.LAT: $t_{LO} - t_{FI}$
  - STAC-T0.LIFO.LAT: $t_{FO} - t_{LI}$
  - STAC-T0.FIFO.LAT: $t_{FO} - t_{FI}$

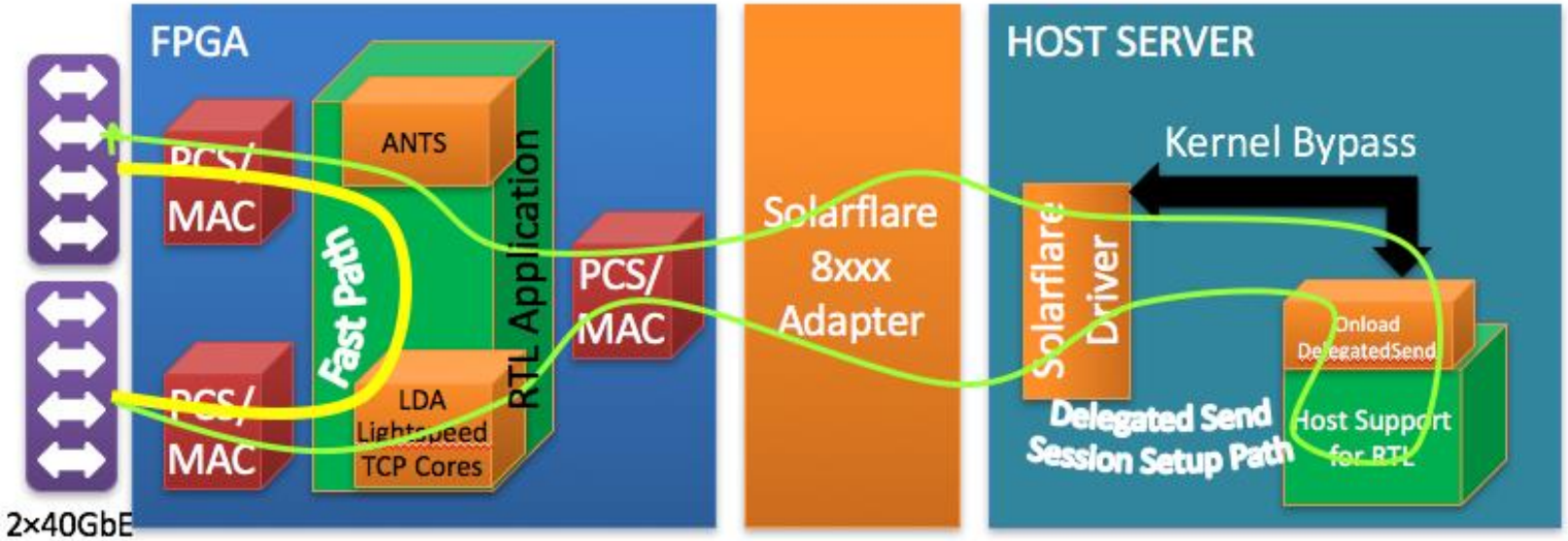
STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

Notes

- STAC-T0.ACTIONABLE.LAT
  - How long it takes the SUT to start sending the egress frame once it has all the algorithmically relevant information
- STAC-T0.FILO.LAT ($t_{LO} - t_{FI}$) is the same measurement as STAC-T1
  - Most inclusive latency but includes a lot of serialization delay
- STAC-T0.FIFO.LAT, STAC-T0.LIFO.LAT are familiar from networking

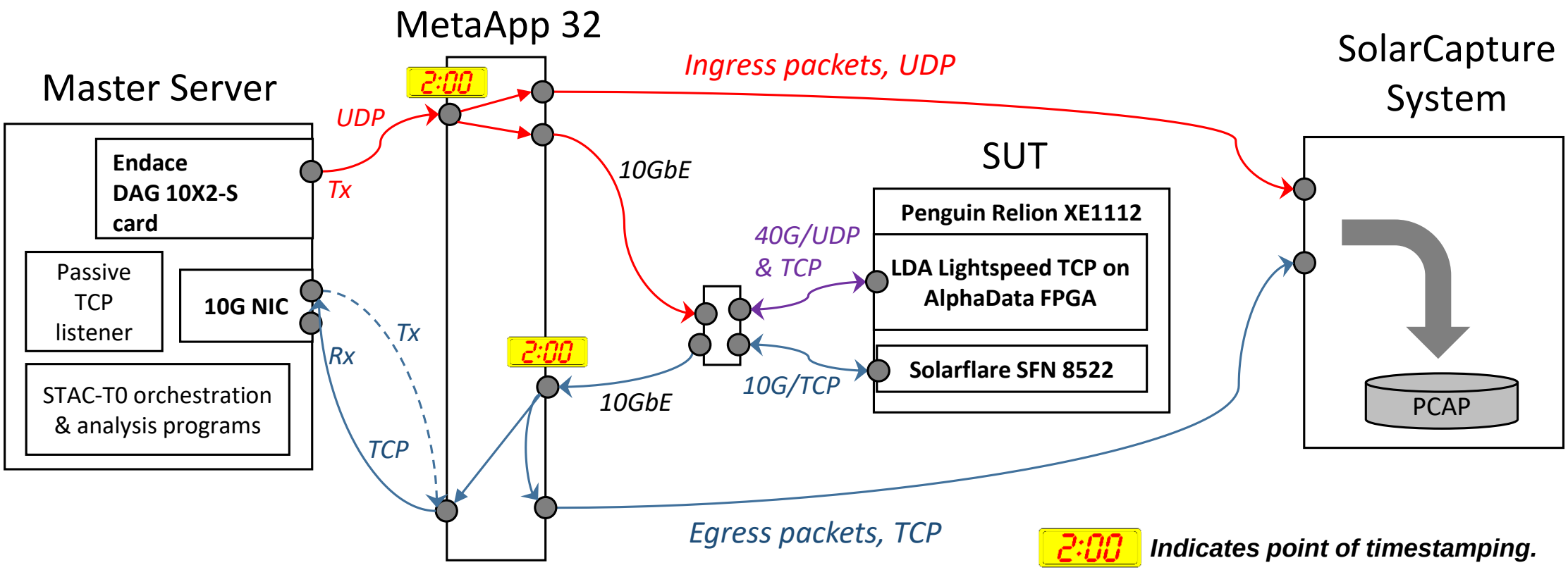
STAC-T0 on Solarflare / LDA / AlphaData / Penguin

- SUT ID: SFC170920
- Penguin Computing Relion® XE1112 server with dual Intel® Xeon® Scalable processors
- Solarflare Flareon Ultra SFN8722 dual-port 10 GbE SFP+ adapter with ANTS
- LDA Technologies LightSpeed TCP and MAC/PCS (FPGA Cores)
- Alpha Data ADM-PCIE-KU3, a Xilinx FPGA NIC with dual QSFP ports



STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

Test setup



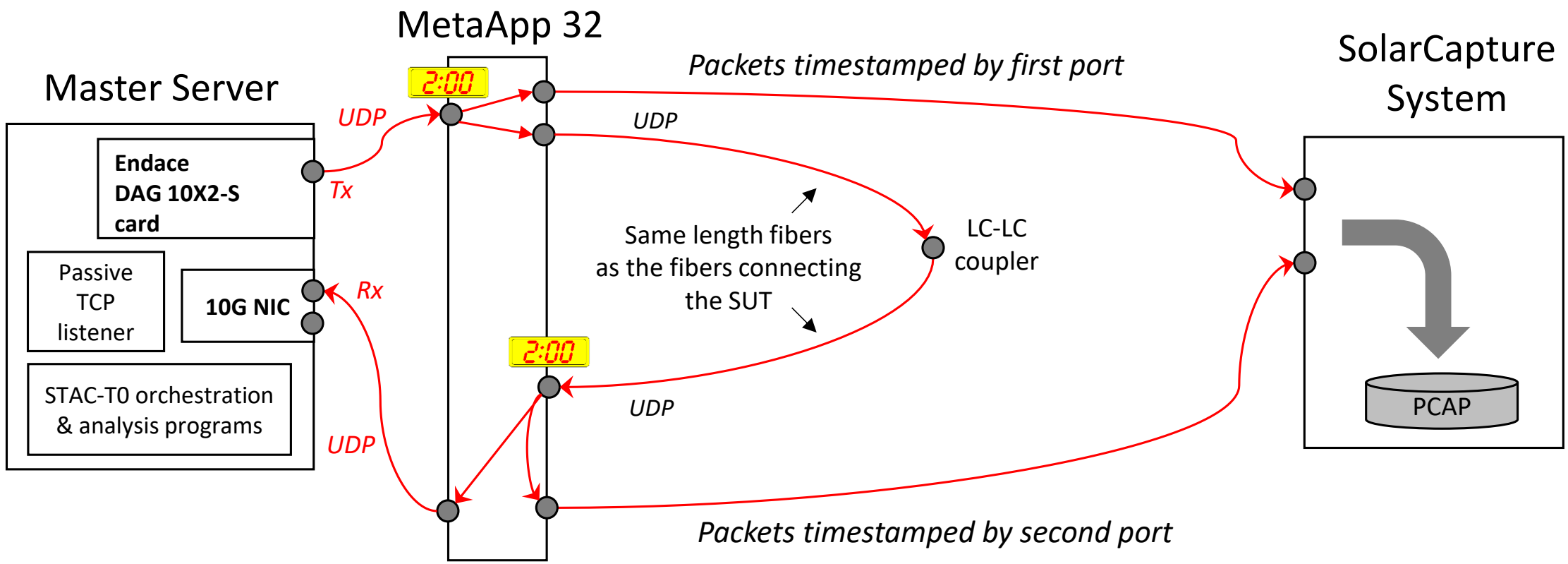
2:00 ***Indicates point of timestamping.***

TCP Tx from Master Server is not captured. Full Tx path not shown.

Management network not shown.

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

Calibrating the test harness



- Skew: 37ns
- Uncertainty: +/- 2ns

STAC
SECURITIES TECHNOLOGY ANALYSIS CENTER

Results highlights

- At 1Gbps with 68-byte frames, max Actionable Latency = 98 nanoseconds
- At 1Gbps with 507-byte frames, max Actionable Latency = 109 nanoseconds
- At 1Gbps with 507-byte frames, max LIFO latency = -65 nanoseconds
  - Negative because the SUT was able to send simulated orders before the entire ingress packet was received
- Max of each latency metric did not vary by more than 2 nanoseconds
  - From the very low rate to the highest rate reported
  - Across both message sizes
  - This is within the measurement uncertainty of the test harness (+/- 2 ns).

www.STACresearch.com/**SFC170831**

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

Interested?

- Go to www.STACresearch.com/nio
- Click “Enable me!”

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER