

### STAC Update: Fast Data

Peter Nabicht President, STAC

peter.nabicht@STACresearch.com



# **FPGA Special Interest Group**



#### Current collaborations: 3 main projects

- RapidWright / RapidStream improvements, including
  - Common requirements, requests, and prioritized bugs
  - Collaborating with developers at AMD at a deeper level
- Language support
  - Jointly contribute to VHDL and SystemVerilog projects that check canonical language feature support in other tools
  - Use to convey of critical features to vendors
- Joint development of open-source Switch and/or NIC reference implementation
  - Exploring currently existing projects as starting points
  - Focus on the primary needs of trading firms



### Education

- Previously
  - Financial firms FPGA developers presented different build, test, and deploy pipelines
  - RapidWright project deep dive led by project engineers from AMD
- Upcoming
  - Tutorial for CXL for FPGA to CPU communication and impact on development from Intel

## **STAC-N1**

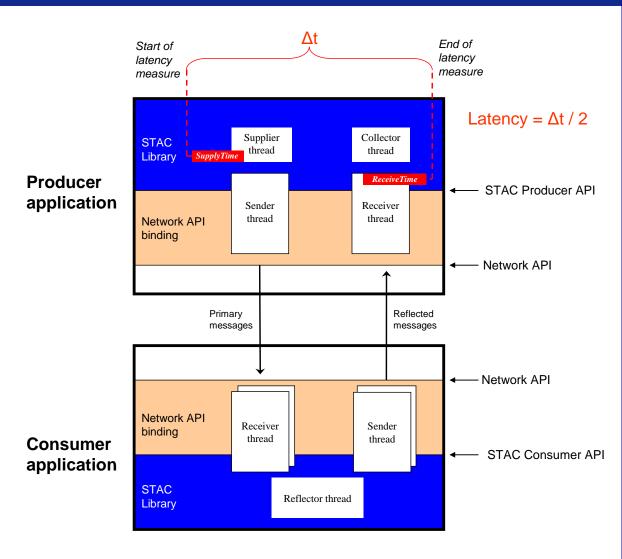
(network stacks)



### STAC-N1

- Measures the performance of a host network stack (server, OS, drivers, host adapter)
- Round-trip software timestamping
- Market data style workload
- Network API to network API
  - No middleware, feed handlers, etc.

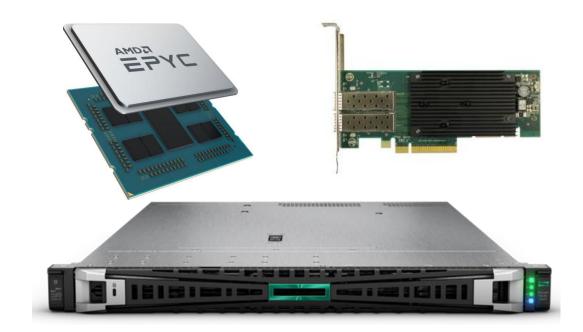






### STAC-N1 / UDP / AMD / HPE / XtremeScale / OpenOnload

- Looks at impact of generational updates, including AMD EPYC<sup>™</sup> Genoa
- Stack
  - STAC-N1 UDP-TCP binding
  - 2 x HPE ProLiant DL325 Gen11 servers
    - 1 x 32-core AMD EPYC<sup>™</sup> 9374F @ 3.85Ghz (4.3 GHz Boost)
    - AMD Xilinx XtremeScale<sup>™</sup> X2522-25G-PLUS Adapter
    - Red Hat Enterprise Linux 8.6
  - 25Gb (via cross-over cable, FEC off)

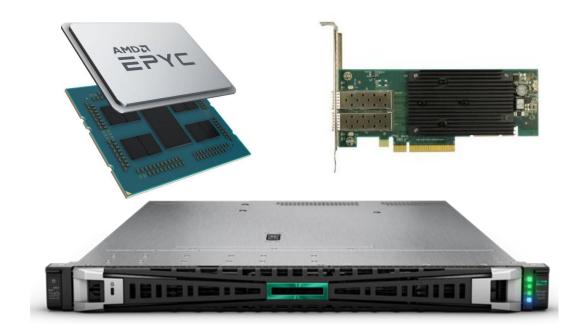






### Vs. all public results for UDP on non-overclocked servers

- The lowest mean and 99p latency for the base rate of 100k messages per second
  - STAC.N1.β1.PINGPONG.LAT1
- The highest maximum throughput tested of 1.5 million messages per second
  - STAC.N1. $\beta$ 1.PINGPONG.TPUT1
- The lowest 99p and maximum latency at the highest rate tested for SupplyToReceive latency
  - STAC.N1. $\beta$ 1.PINGPONG.LAT2

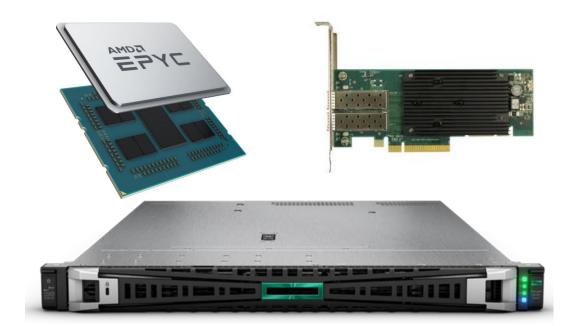






### Vs. prev. generation server, CPU, OS, and OpenOnload driver\*

- A maximum throughput tested that was 300,000 messages per second higher
  - (STAC.N1.β1.PINGPONG.TPUT1)
- A lower mean and maximum latency at the highest rate tested
  - (STAC.N1. $\beta$ 1.PINGPONG.LAT2







\*SUT ID: AMD221001