ARISTA Financial applications - Real world results

Overview - Sources of Trading Latency

Feeds

Feed Handlers

Ethernet Switch

Messaging Server

Algo Engines

Orders for Execution
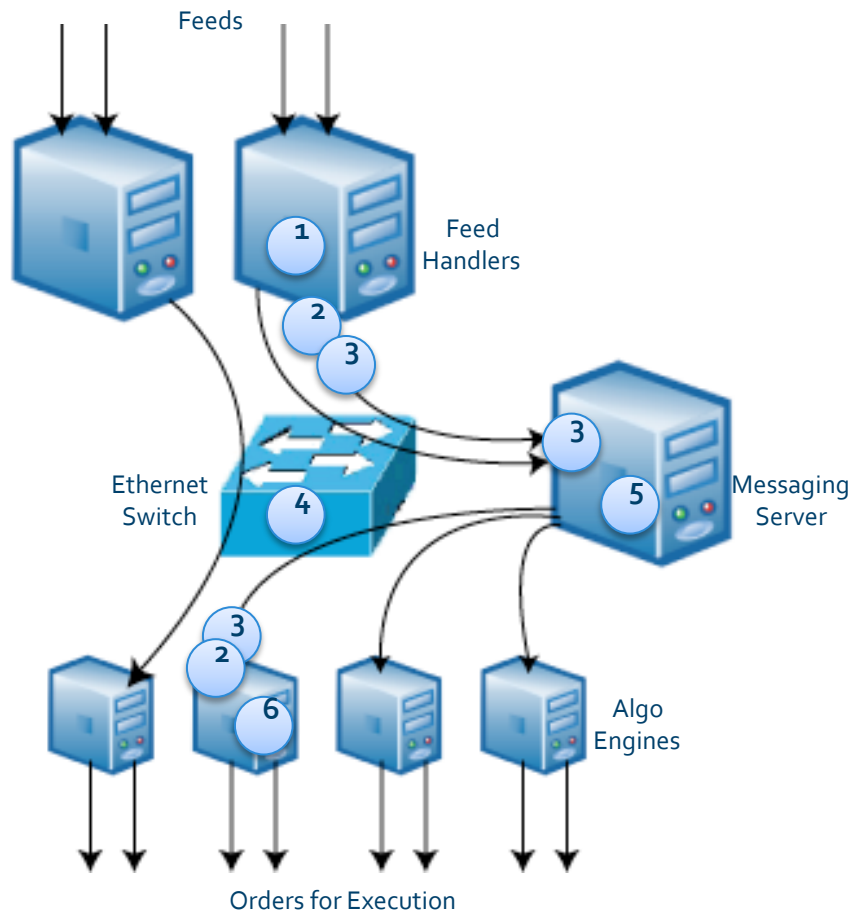
Sources of Latency

1. Feed handler/ticker plant
2. Messaging API
3. Network stack
4. Ethernet switch
5. Messaging middleware (server)
6. Algorithm performance
7. Transition to order execution

Other Latency Issues

High volumes:
- Jitter
- Context switching
- Network I/O

Physical distance:
- Speed of light

-Arista reduces latency at the heart of the trading network
-Arista ecosystem partners provide low latency end to end solutions

Overview - Arista insertion points and partners

Data feeds/Exchanges
NYSE · London Stock Exchange · NASDAQ OMX · BATS Making Markets Better · REUTERS · cme Chicago Mercantile Exchange

ARISTA

Ticker plants/Feed handlers
Custom UDP/TCP/RDMA · TERVELA · exegy · Celoxica · activ FINANCIAL

ARISTA

Messaging/Middleware Platforms
Custom UDP/TCP/RDMA · TERVELA · Solace Systems™ · TIBCO · 29 WEST INC.

ARISTA

Application servers/NICS
intel · Chelsio Communications Accelerate · Mellanox TECHNOLOGIES · Myricom · SOLARFLARE Communications

Arista attributes and features providing true VALUE

- LOW LATENCY

    -600ns (24 ports) - 1200ns(48 ports) for broadcast, unicast, and multicast traffic at 10G line rate – low 1G too!

- Multicast scalability

    -2048 IGMP groups in hardware, no pipelining or delay

- Predictable variance (jitter) for all frame sizes

    -~100ns, regardless of traffic type or frame size!

- Line rate, low latency mirroring

    -packet copy at the same rate/latency as the actual data

    -analyze and track execution in real time

- Future full L3 capabilities

    -BGP, PIM-SM to provide a "single device" exchange connectivity solutions in the collocation facility
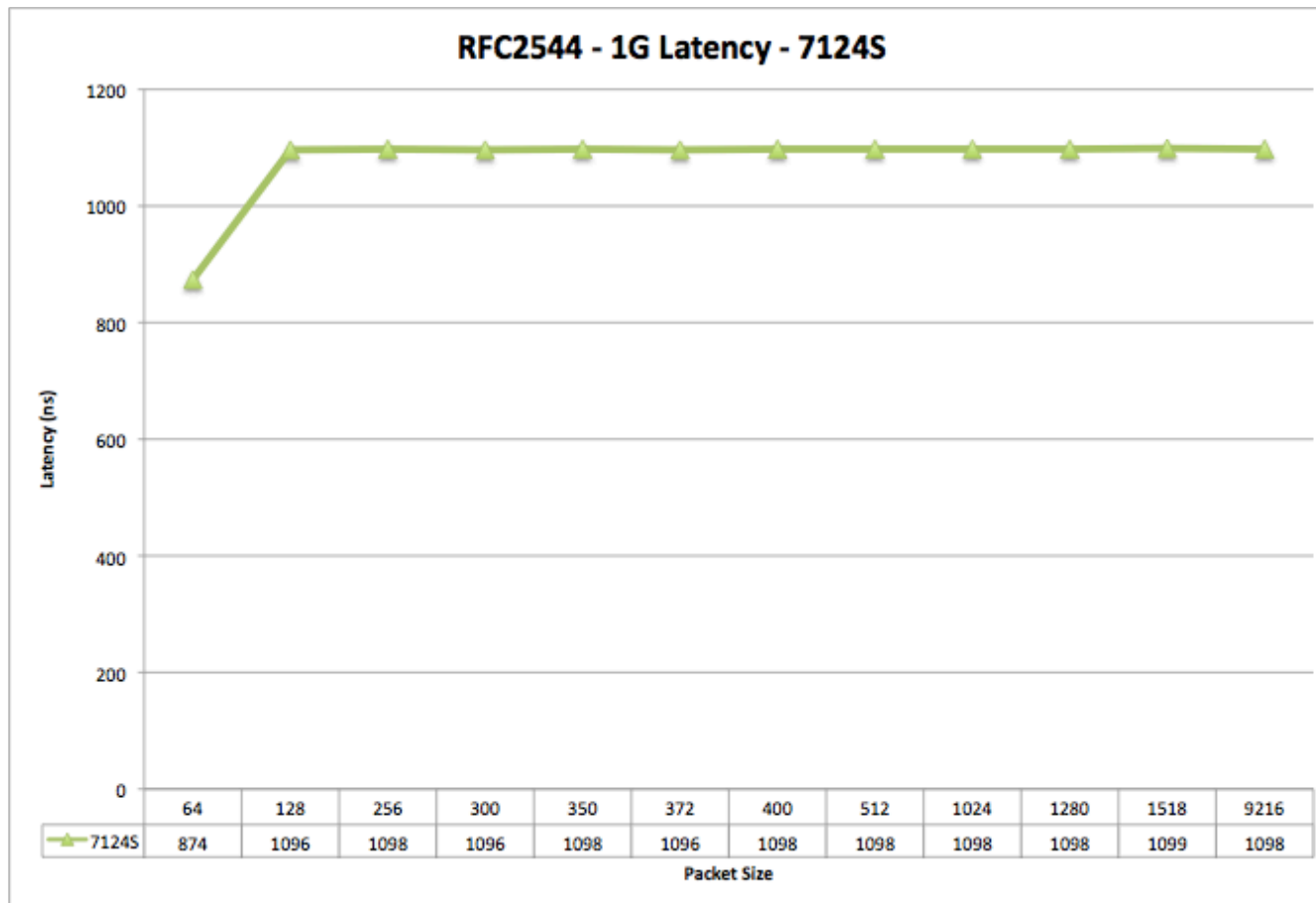
Arista 7100 Series- The lowest latency 10G switches!

Intelligent cut-through switching - the right solution for market data
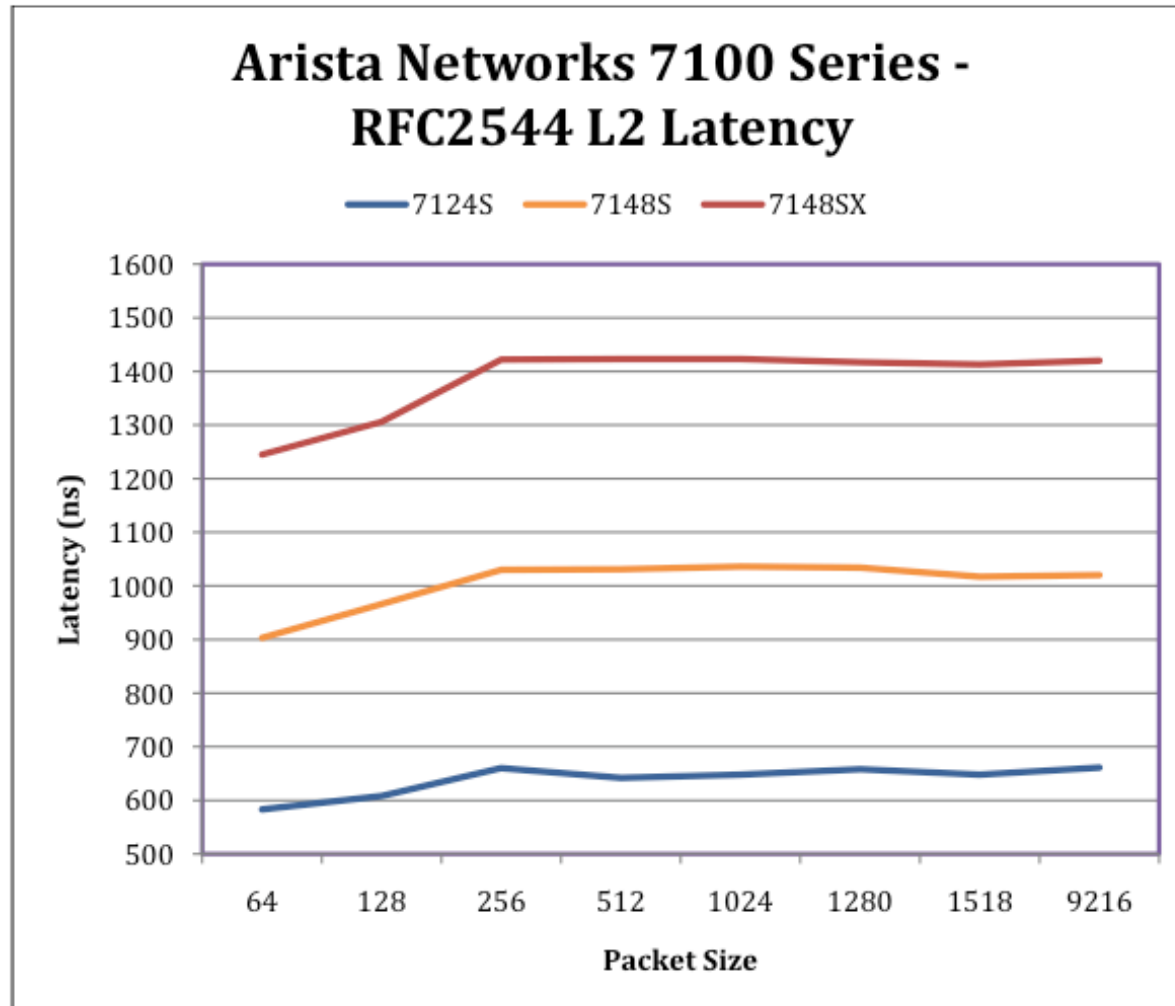
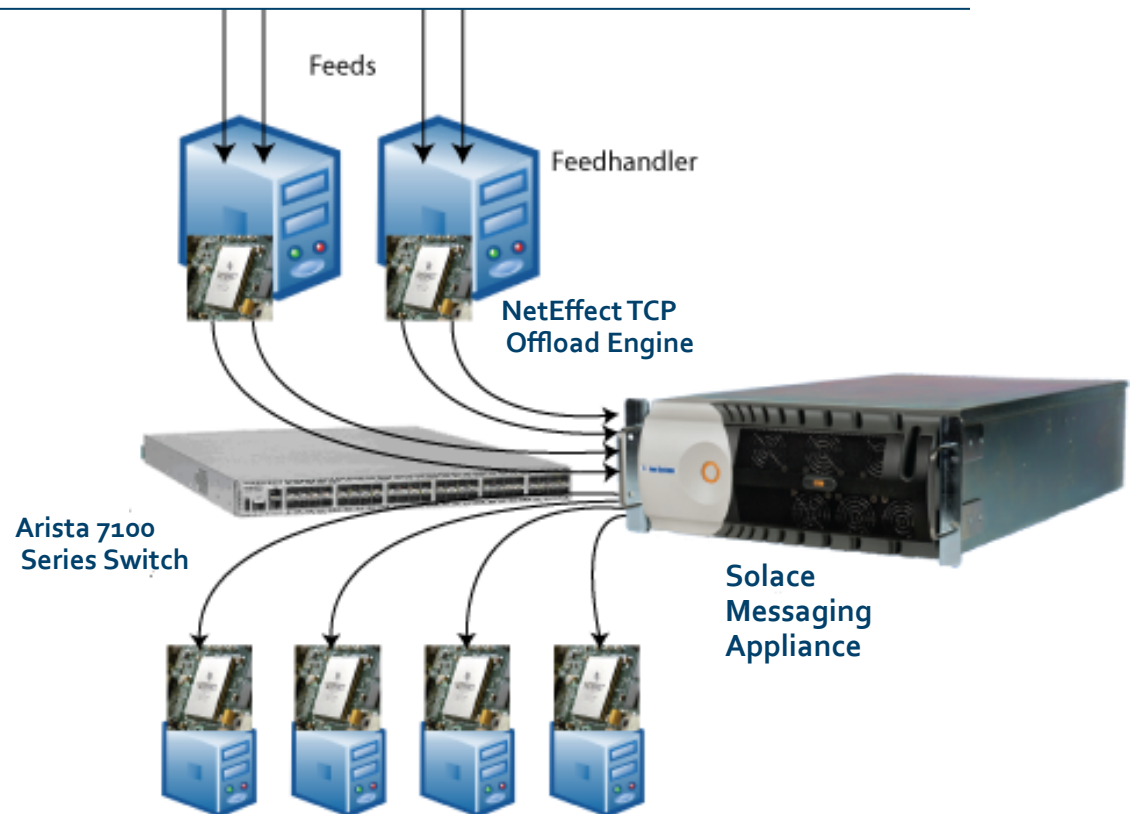# Arista 7100 Series- The lowest latency 1G switch too!

- Arista provides 1G and 10G on every interface
- ~1us 1G performance = latency reductions for exchange colo cross connects

RFC2544 - 1G Latency - 7124S

Packet Size	64	128	256	300	350	372	400	512	1024	1280	1518	9216
7124S	874	1096	1098	1096	1098	1096	1098	1098	1098	1098	1099	1098

Real world data – Solace Systems

Solace Systems + **ARISTA** + (intel) **NetEffect**

- **100% hardware data path for market data distribution**

- **500k msgs/sec**
  31 μsecs mean latency
  47 μsecs 99.9th percentile

- **1 million msgs/sec**
  35 μsecs mean latency
  55 μsecs 99.9th percentile

Feeds

Feedhandler

NetEffect TCP
Offload Engine

Arista 7100
Series Switch

Solace
Messaging
Appliance

Less than 7us end to end mean latency variance through the entire distribution path

Real world data – Solace Systems

- a total of 8 servers were used, traffic spread equally among the 8 ports
- each application "message" is in its own TCP segment - no message bundling
- actual throughput/mps on the Arista switch was 2x the test message rate shown
  - 1x from publisher to Solace, 1x Solace to subscriber
  - each message in is sent to exactly one subscribing application based on a subscription to the 12-byte topic
- rate = per-packet size of (100B msg payload + 12B topic + 12B msg overhead +40B TCP/IP + 8B Enet) * msgs/sec

1 Gigabit Ethernet Test; November 2008

- Solace 3260 Content Router w/ 8x1GbE NAB
- Arista DCS-7124S
- NetEffect NE020 1GbE single port TOE
- Message size: 100 bytes
- Topic size: 12 bytes
- Test duration: 10 minutes

Data Volume (messages/sec)	Mean Latency (µsecs)	Standard Deviation (µsecs)	99.9th % Latency (µsecs)
500,000	31	6	47
1 million	35	7	55

10 Gigabit Ethernet Test; March 2009

- Solace 3260 Content Router w/ 2x10GbE NAB
- Arista DCS-7148SX
- NetEffect NE020 10GbE single port SPF+ TOE
- Message size: 100 bytes
- Topic size: 12 bytes
- Test duration: 10 minutes

Data Volume (messages/sec)	Mean Latency (µsecs)	Standard Deviation (µsecs)	99.9th % Latency (µsecs)
500,000	25	6	41
1 million	28	6	48
2 million	42	8	71

Real world data – 29West

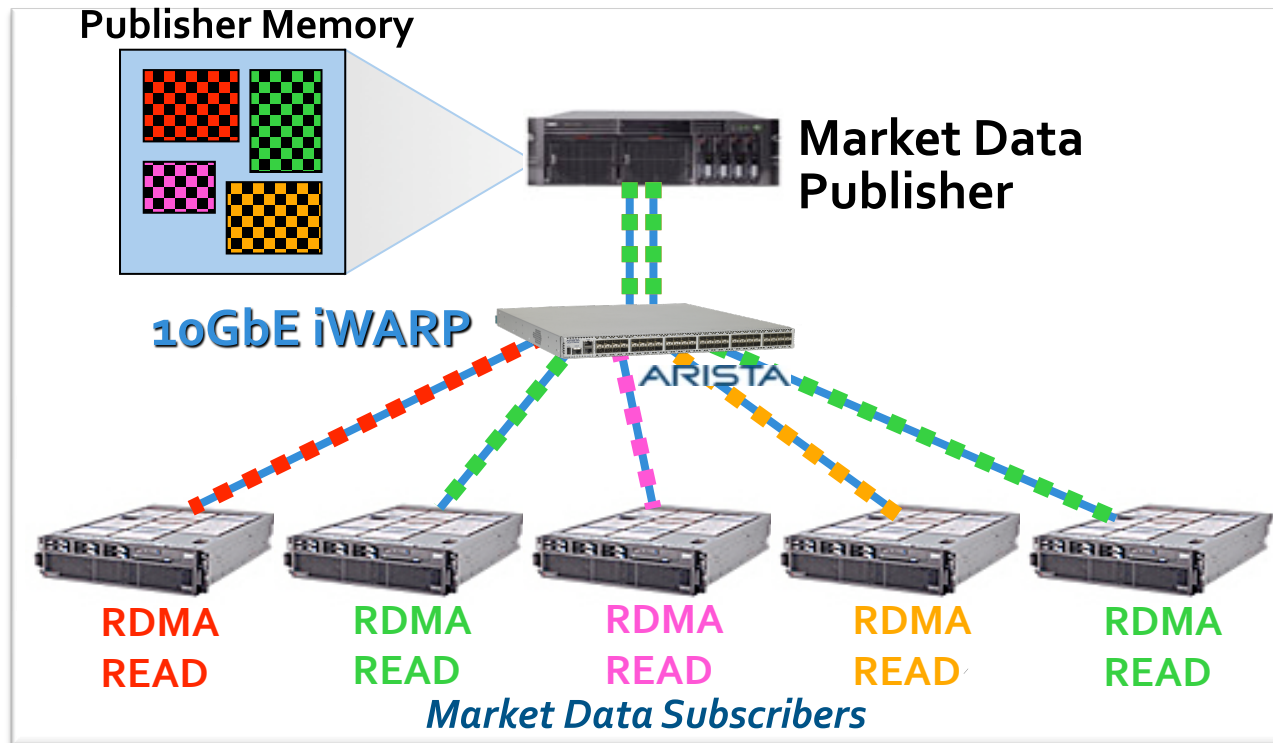**29 WEST INC. + ARISTA + intel**

Test Details:
- Two Dell Precision T3400 Servers with Q6600 CPUs running Linux Centos 5
- Intel NE020 10GbE NICS's (10G SR PHY)
- Arista 7124S 10 Gigabit switch
- The measurements were performed using version 3.3.6 of 29West, Inc. Latency Busters® Messaging.

Results:
- Payload delivery rates as high as 9.8 gbps for large messages
- Message delivery rates more than 2.3 million messages per second for small messages
- Average ping-pong latency of 36 microseconds
- Scalable performance as receivers are added when using UDP multicast
- Message delivery latency under 50 microseconds for rates up to 110,000 messages per second

ARISTA + **NYSE Technologies**

Publisher Memory

Market Data Publisher

10GbE iWARP

ARISTA

RDMA READ	RDMA READ	RDMA READ	RDMA READ	RDMA READ

Market Data Subscribers

- The test showcases NYSE Technologies' high performance middleware, Data Fabric™
- 1M, 100 Byte messages with RDMA, average latency of 26 usec!

Real world data - NYSE Technologies DF

Test Details:

- Arista 7124S switch

- single 8-core Intel® Xeon® 5500-based server

- 12 subscribers on 6 more similar servers

- NetEffect NE020 10 Gigabit Ethernet RDMA capable NIC

Test Results:

- 1 Million, 100 Byte messages per second

- 12 Million messages streaming through the fabric per second

- breakthrough, end-to-end, average one-second latency of 26 usec

- average one-second minimum latency at 1.8 usec

- average maximum one-second latency at 46.6 usec.

Real World Data – Solarflare

ARISTA + SOLARFLARE® Communications

Test Details:

Switch: Arista DCS-7124S 10 Gigabit switch
NIC: Solarflare Solarstorm SFN4112F SFP+ 10 Gigabit server adapters
Processor: 2 x Intel Xeon x5482 (3.2GHz)
RAM: 4 Gbytes
Chipset: Harpertown 32K L1, 6 Mb L2
OS: LinuxRed Hat Enterprise Linux Client release 5.1 (64 Bit)
Middleware: Solarflare OpenOnload

Benchmarks:
- Latency - udprtt- Generates Traffic and measures Round Trip latency with respect to different UDP payloads
- Message Throughput - udpstream - Generates traffic and measures message rate throughput
- Bandwidth - Udpsend/udpswallow - Generates traffic and measures bandwidth

Real World Data – Solarflare TCP Results

Exhibit 1: Half-Round Trip Latency in Nano Seconds

Network stack	Link	Min	Median	Mean	95th percentile	99th percentile
Kernel	Back to back	9801	10074	10185	10683	12075
Kernel	Switch	10233	10563	10745	11374	12704
Onload	Back to back	5517	5697	5884	6399	6498
Onload	Switch	6122	6327	6514	7030	7123

Exhibit 1 summarizes results of TCP latency testing. The Arista DCS7124 is a very low-latency switch contributing a mean latency of 560 nanoseconds to the system latency. In the testing for the 70 byte message sizes typical of market data messaging systems, very low latency was observed. The Solarstorm server adapter in combination with the Arista DCS 7124S switch achieved mean latency of 6.5 microseconds. The Solaflare adapters back to back, achieved an amazingly low mean latency of 5.9 microseconds. This latency was also very deterministic with 95% of the messages being delivered with a mean less than 7.0 microsecond and 99% of the messages with mean latency less than 7.1 microseconds in the switch to server adapter configuration.

Real World Data – Solarflare UDP Latency Results
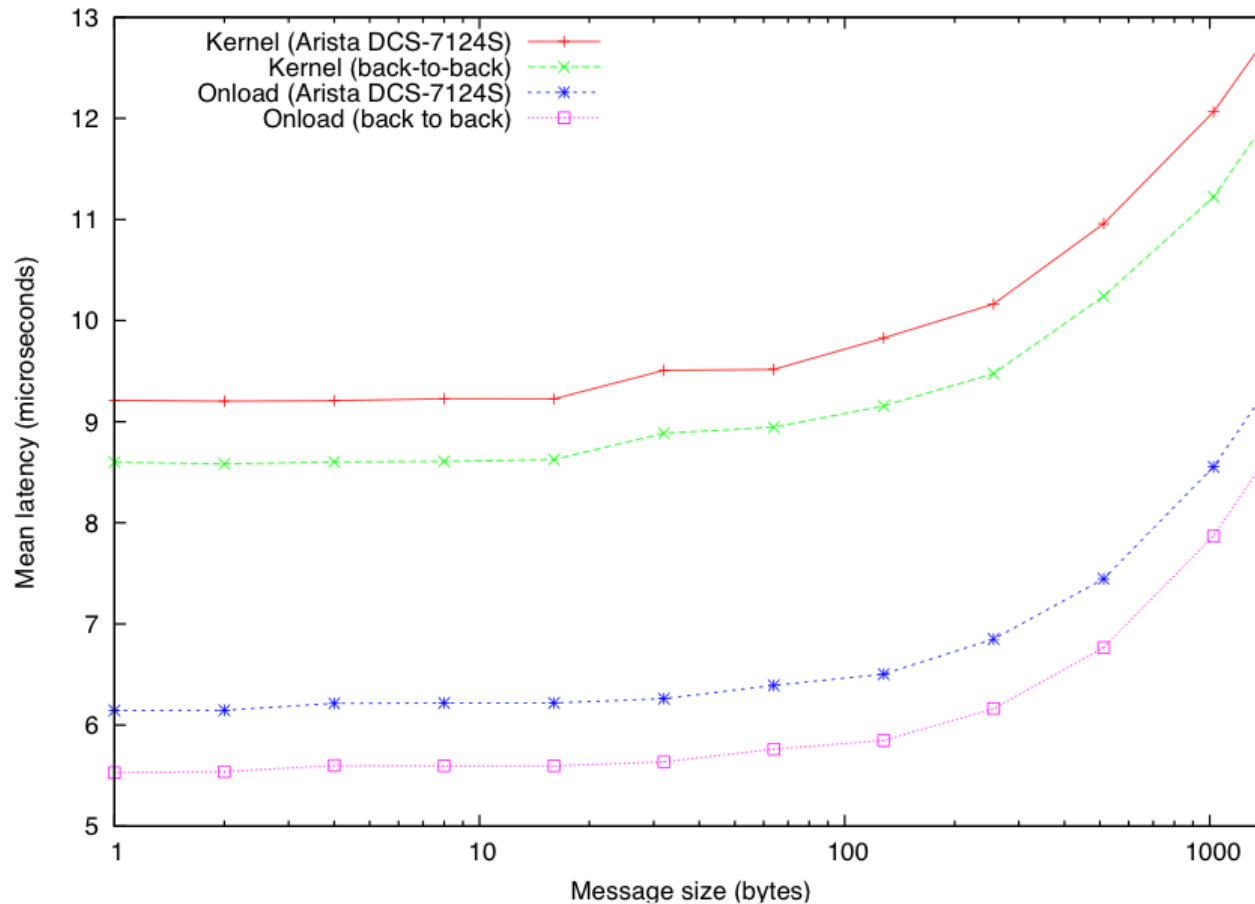
Exhibit 2: UDP Half Round Trip Latency

Exhibit 2 above plots UDP half round trip latency where the x axis represents message size in bytes and the y axis represents latency in microseconds. The data shows that the system demonstrated very low and deterministic latency from small up to very large message sizes of 1472 bytes In OpenOnload mode with the switch and server adapter, minimum latencies go as low as 6 microseconds.

Real World Data – Solarflare UDP Message Rate Results

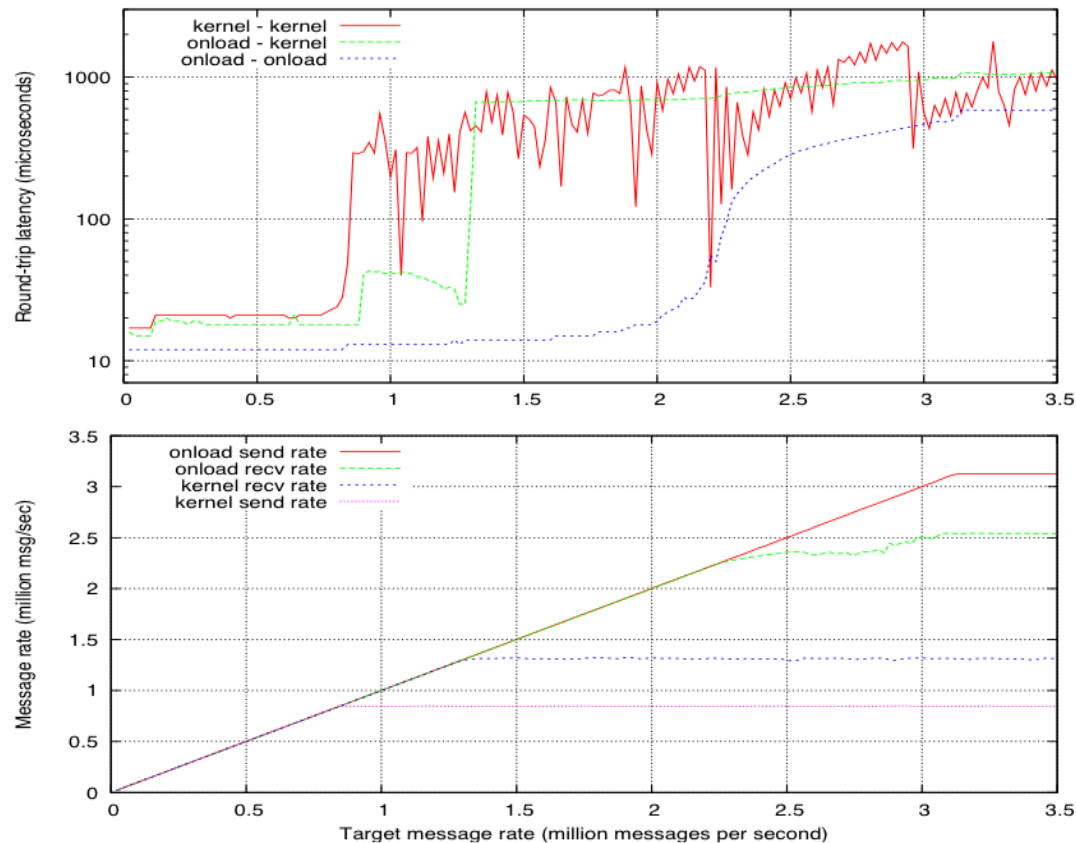Exhibit 3: Message Rates Achieved with Upstream UDP

Exhibit 3 - This test simulates a traffic pattern that is common in financial services applications. In the test, the system streams small messages from a sender to a receiver. The receiver reflects a small proportion of the messages back to the sender which the sender uses to calculate the round-trip latency. In the second plot in Exhibit 3, Message Rate Achieved shows the Solarflare OpenOnload system's ability to scale and perform as the message rate is increased. This is in contrast to the kernel stack where the greater CPU processing overheads of the stack limit performance as higher levels of load are put on the system.

Real World Data – Solarflare Results Summary

Test Summary:

The findings revealed the transmit latency of a configuration with the Solarflare Solarstorm server adapter with OpenOnload and the Arista DCS- 7124S at transmission rates up to 3 million messages/second (mps). For the 70 byte message sizes typical of market data messaging systems, very low latency was observed:

• Mean did not exceed 6.5 microseconds with switch
• Mean did not exceed 5.8 microseconds without switch
• 99th percentile did not exceed 7.1 microseconds with switch

The system demonstrated therefore very bounded jitter and very low UDP multicast latency which delivers very predictable messaging systems.

High volume Market Data Simulation – Details

ARISTA + IXIA

- **Test Details:**
- IXIA XM 12 Chassis
- IXIA 5.50 Sw
- Arista 7148SX switch

- *IXIA Port 1 -> Arista port 1*
- transmits 202 mcast streams
- 100 30Mb streams
- 100 50Mb streams
- 2 1000Mb streams
- 10Gb aggregate

- *IXIA 2 -> Arista Port 48*
- transmits 100 mcast streams
- 100 100Mb streams
- 10Gb aggregate
- IXIA ports 3-48 -> Arista ports 2-46

- *Arista ports 2-47 subscribe to 175 mcast streams each*
- 100 30Mb streams
- 50 50Mb streams
- 25 100Mb streams
- 1 1000Mb stream
- Frame size is random 64-527 bytes
- Ingress traffic is 20Gb
- Egress traffic is 414Gb, replicated by the Arista switch

- **Results:**
- Zero frame loss
- Mean latency of ~1.4us
- Minimum average of 89ons, maximum of 4us (3us variance with random frame sizes)
- The Arista 7148 process 8000+ IGMP messages in 8 seconds

More information

- **Solace Systems**
  - http://www.aristanetworks.com/en/Solace_Aristawp.pdf

- **29West**
  - http://www.aristanetworks.com/en/29WestArista_wp.pdf

- **Solarflare**
  - http://www.solarflare.com/technology/documents /
    AristaSolarflare_LowLatency10GigEthernet.pdf

- **Tervela**
  - http://www.aristanetworks.com/en/PR_20090602_01

- **IXIA RFC test reports**
  - http://www.aristanetworks.com/en/LatencyReport

- **Arista technology partners**
  - http://www.aristanetworks.com/en/Technology_Partners

- **Arista NIC partners**
  - http://www.aristanetworks.com/en/Host_Adapter_Solutions