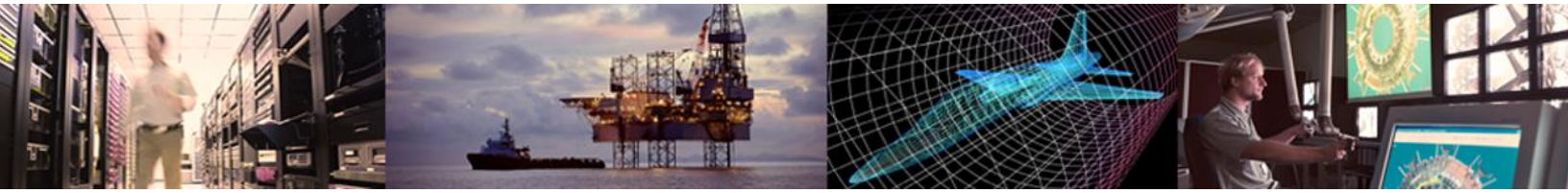
Low Latency Test Report

Ultra-Low Latency 10GbE Switch and Adapter Testing

Bruce Tolley, PhD, Solarflare

Testing conducted by Solarflare and Fujitsu shows that ultra low latency application-level performance can be achieved with commercially available 10G Ethernet Switch and Server Adapter Products



ABSTRACT

Solarflare, the pioneer in high-performance, low-latency 10GbE server networking and application acceleration middleware solutions and Fujitsu recently completed switch to server latency and message rate testing. The test found that the Solarflare® SFN5122F server adapter and Solarflare OpenOnload® application acceleration middleware in combination with the Fujitsu XG2600 10GbE Switch achieved 4.5 microsecond mean latency over 10GbE in a TCP latency test with 64 byte packet size typical of market data applications. The Fujitsu switch is one of the lowest latency switches on the market, demonstrating mean latency in this test of 427 nanoseconds. With back to back servers, the server adapter achieved impressive minimum TCP latencies as low at 4.1 microseconds. The latency of the overall system was also very deterministic with 99% of the messages being delivered with a latency of less than 4.5 microseconds at near zero jitter. Solarflare and Fujitsu measured the performance of TCP and UDP messaging using Solarflare developed benchmarks with commercially available products: the Solarflare SFN5122F 10 Gigabit server adapters and Fujitsu 10 Gigabit switch. The test platform used servers and processors typically found in use by financial firms today.

THE NEED FOR LOW LATENCY IN AUTOMATED, REAL-TIME TRADING

The rapid expansion of automated and algorithmic trading has increased the critical role of network and server technology in market trading, first in the requirement for low latency and second in the need for high throughput in order to process the high volume of transactions. Given the critical demand for information technology, private and public companies that are active in electronic markets continue to invest in their LAN and WAN networks and in the server infrastructure that carries market data and trading information.

In some trading markets, firms can profit from less than one millisecond of advantage over competitors, which drives them to search for sub-millisecond optimizations in their trading systems. The spread of automated trading across geographies and asset classes, and the resulting imperative to exploit arbitrage opportunities based on latency, has increased the focus on if not created an obsession with latency.

With this combination of forces, technologists, IT and data center managers in the financial services sector are constantly evaluating new technologies that can optimize performance. One layer of the technology stack that receives continuous scrutiny is messaging, i.e., the transmission of information from one process to another, over networks with specialized home-grown or commercial messaging middleware.

The ability to handle predictably the rapid growth of data traffic in the capital markets continues to be a major concern. As markets become more volatile, large volumes of traffic can overwhelm systems, increase latency unpredictably, and throw off application algorithms. Within limits, some algorithmic trading applications are more sensitive to the predictability of latency than they are to the mean latency. Therefore it is very important for the network solution stack to perform not just with low latency but with bounded, predictable latency. Solarflare and Fujitsu demonstrate in this paper that because of its low and predictable latency, a UDP multicast network built with 10 Gigabit Ethernet (10GigE) can become the foundation of messaging systems used in the financial markets.

FINANCIAL SERVICES AND OTHER APPLICATIONS THAT CAN TAKE ADVANTAGE OF LOW-LATENCY UDP MULTICAST

Messaging middleware applications were named above as one key financial services application that produce and consume large amounts of multicast data which can take advantage of low-latency UDP multicast. Other applications in the financial services industry that can take advantage of low-latency UDP multicast data include:

- Market data feed handler software that takes as input multicast data feeds and uses multicasting as the distribution mechanism
- Caching/data distribution applications that use multicast for cache creation or to maintain data state
- Any application that makes use of multicast and requires high packets per second (pps) rates, low data distribution latency, low CPU utilization, and increased application scalability

CLOUD NETWORKING AND BROADER MARKET IMPLICATIONS OF LOW LATENCY TO SUPPORT REAL-TIME APPLICATIONS

As stated above, the low-latency UDP multicast solution provided by Fujitsu switches and Solarflare server adapters can provide compelling benefit to any application that depends on multicast traffic where additional requirements exist for high throughput, low-latency data distribution, low CPU utilization, and increased application scalability. Typical applications that benefit from lower latency include medical imaging, radar and other data acquisition systems, and seismic image processing in oil and gas exploration. Yet moving forward, cloud networking is a market segment where requirements for throughput, low latency and real time application performance will also develop. The increasing deployment and build out of both public and private clouds will drive the increased adoption of social networking and Web 2.0 applications. These cloud applications will incorporate real-time media and video distribution and will need lower latency applications for both business to consumer (B2C) and business to business (B2B) needs. Perhaps more fundamentally, the need for real-time, high-speed analytics and search of large and often unstructured data sets will create demand for low latency and real time application response.

SOLARFLARE'S OPENONLOAD BENEFITS

Solarflare and Fujitsu measured the latency performance of messaging using Solarflare-developed benchmarks with commercially available products: the Solarflare SFN5122F SFP+ 10 Gigabit Ethernet server adapters, Solarflare OpenOnload application acceleration middleware, and the Fujitsu 10 Gigabit switch. A list of the hardware configurations and the benchmarks used is attached as an Appendix. The test platform used servers and processors typically found in use by financial firms today. The tests described below were run both switch to server adapter and server adapter to server adapter. The adapters were run in kernel mode and in OpenOnload mode.

OpenOnload is an open-source high-performance application acceleration middleware product. By improving the CPU efficiency of the servers, OpenOnload enables applications to leverage more server resources, resulting in dramatically accelerated application performance without changing the existing IT infrastructure. Using standard Ethernet, the solution combines state-of-the-art Ethernet switching and server technologies that dramatically accelerate applications. OpenOnload performs network processing at user-level and is binary-compatible with existing applications that use TCP/UDP with BSD sockets. It comprises a user-level shared library that implements the protocol stack, and a supporting kernel module.

FUNDAMENTAL FINDINGS

	size	mean	min	median	max	99%ile	stddev
tcp latency ool switch	**64**	4549	4408	4514	11982	5047	105
tcp latency b2b ool	**64**	4122	3963	4092	10957	4616	105

Exhibit 1: Half-Round Trip TCP Latency in Nano Seconds

Exhibit 1 summarizes results of TCP latency testing. The Fujitsu switch is an ultra-low latency switch contributing a mean latency of 427 nanoseconds to the system latency. In the testing for the 64 byte message sizes typical of market data messaging systems, very low latency was observed. The Solarflare server adapter in combination with the Fujitsu switch achieved mean latency of 4.5 microseconds. The Solarflare adapters back to back achieved an amazingly low latency of 4.1 microseconds. This latency was also very deterministic with 99% of the messages achieving a mean latency of less than 5.0 microseconds in the switch to server adapter configuration. By comparing the two values in the standard deviation column, the Fujitsu switch appears to contribute near zero jitter.

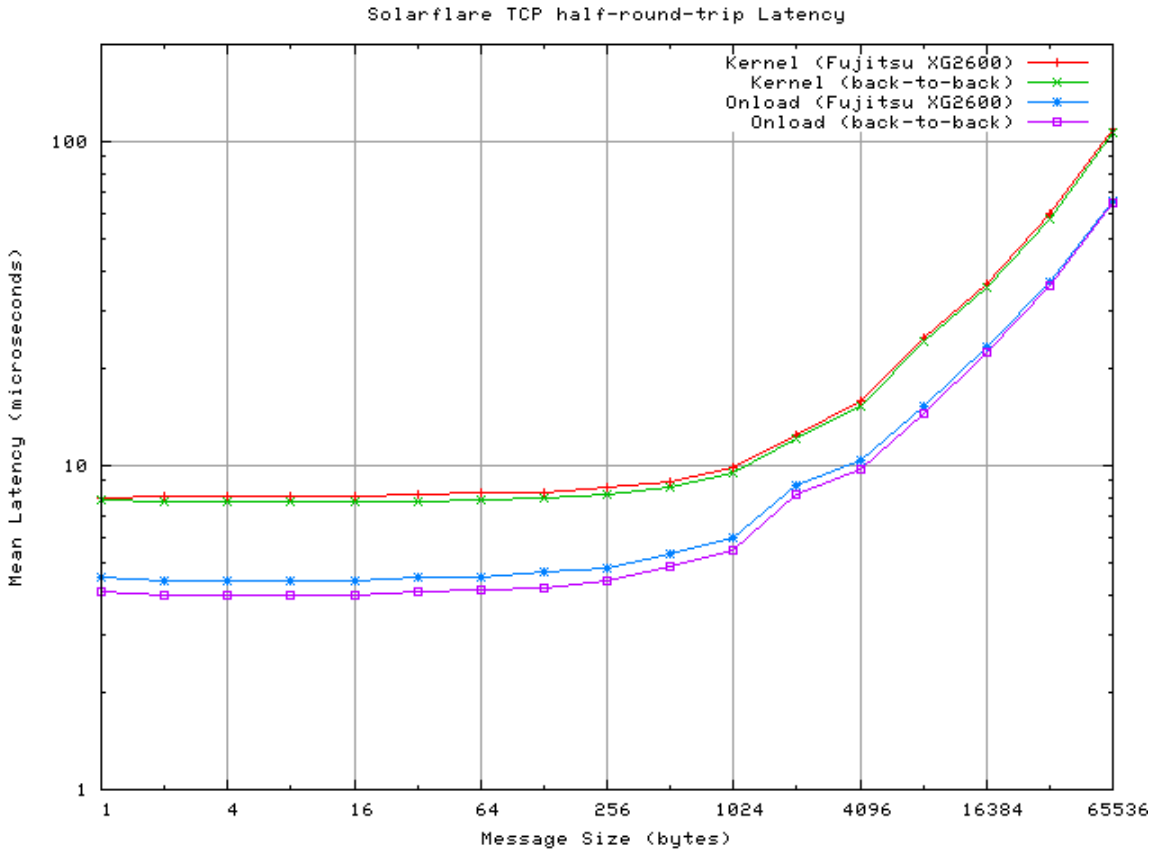


Exhibit 2: TCP Latency Performance vs. Message Size

Exhibit 2 above extends the data from Exhibit 1, where the x axis represents message size in bytes and the y axis represents TCP half round trip latency in microseconds. The data shows that the system demonstrated very low and deterministic latency from small to large message sizes up to 1500 bytes. The data plot also shows low latency in both kernel and OpenOnload mode. The Solarflare adapters back to back, achieved an amazingly minimum latency of 3.96 microseconds. This latency was very deterministic as demonstrated by the flatness of the curve as the message size approaches 1500 bytes. At a message size of 2048 bytes, mean latency was still low at 8.189 microseconds.

Exhibit 3 next page 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 1024 bytes. The data plot also shows very low latency in both kernel and OpenOnload mode. In OpenOnload mode with the switch and server adapter, minimum latencies go as low as 4.408 microseconds for 64 byte packet messages, and with the server adapters back to back, as low as 3.967 microseconds.

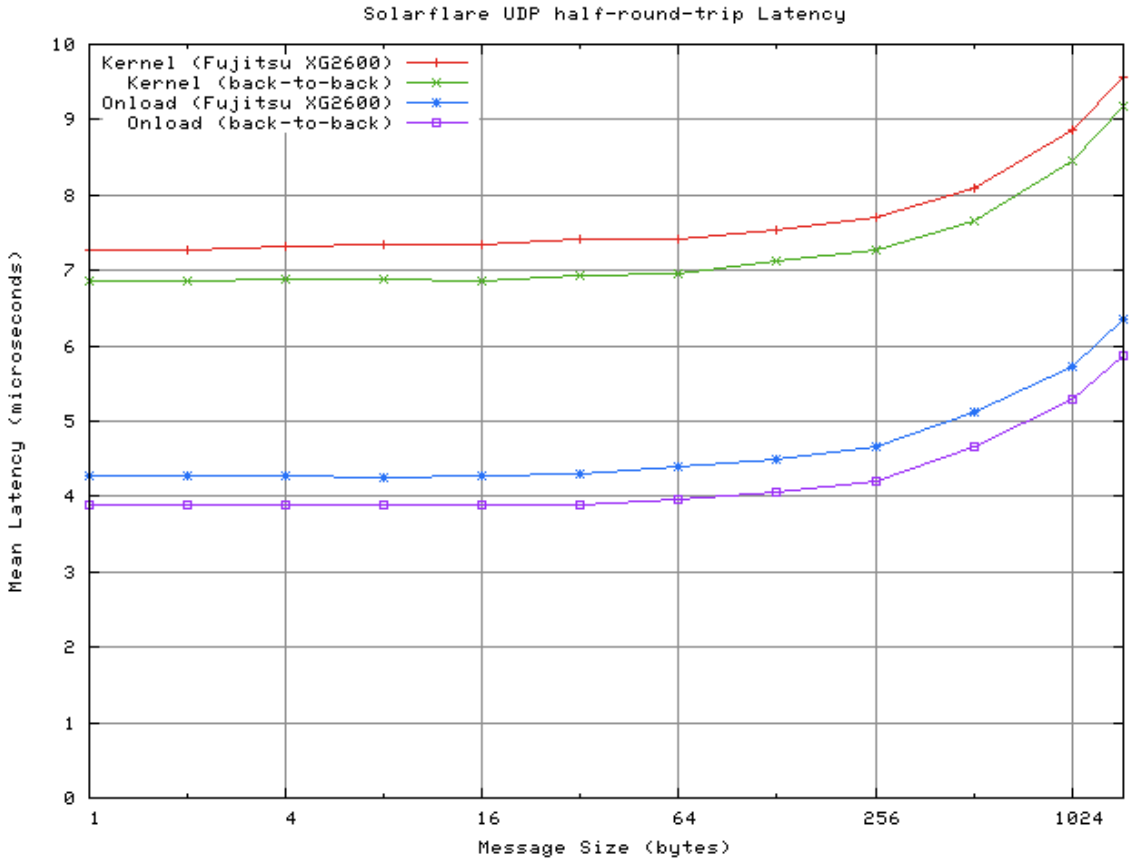


Exhibit 3: UDP Half Round Trip Latency

Exhibit 4 next page shows two plots of performance versus desired data rate of UDP multicast performance with and without OpenOnload. 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. The x axis shows the target message rate that the sender is trying to achieve. The y axis shows one-way latency (including a switch) and the achieved message rate. The kernel results are measured with Solarflare server adapters without OpenOnload. The plot combines results from three runs: kernel to kernel, OpenOnload to kernel, and OpenOnload to OpenOnload. The OpenOnload to kernel test is needed in order to fully stress the kernel receive performance.

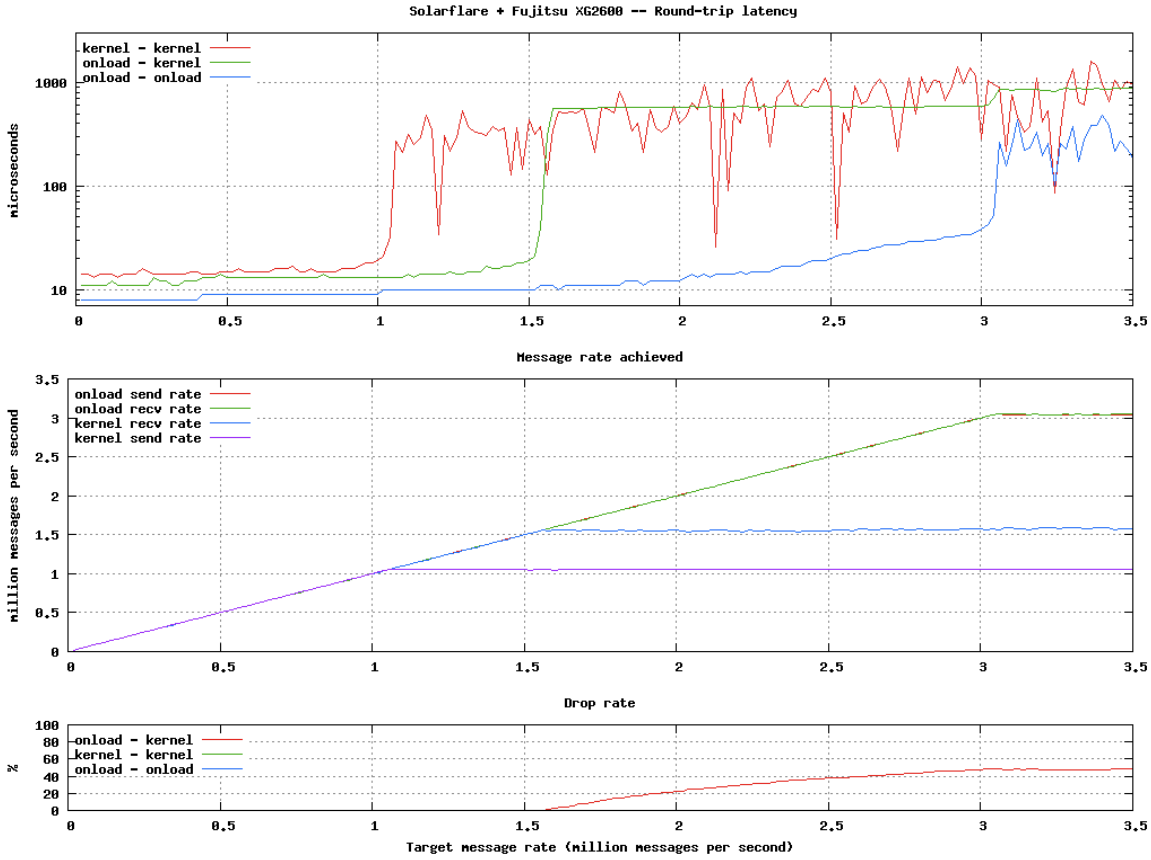


Exhibit 4: Message Rates Achieved with Upstream UDP

The top plot labeled Round Trip Latency shows the improved, deterministic low latency achieved with the Solarflare adapter, OpenOnload, and the Fujitsu switch. The y axis shows the round trip latency while the x axis shows the desired message rate in millions of messages per second at the receiver. With OpenOnload, not only is the system performing at much lower latency, but the latency is predictable and deterministic over the range of expected message rates. This is precisely the attribute desired in trading systems or any other application demanding real time performance.

The second plot in Exhibit 4, 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.

THE SOLARFLARE SOLUTION

The SFN5122F 10GbE SFP+ server adapter is the most frequently recommended 10Gb Ethernet server adapter for trading networks in New York, London and Chicago. In both kernel and OpenOnload modes, the adapter supports high frequency trading and HPC applications which demand very low latency and very high throughput. Tests were performed using the standard kernel TCP/IP stack as well as Solarflare OpenOnload mode.

OpenOnload is an open-source application acceleration middleware product from Solarflare. As Exhibit 5 shows, OpenOnload provides an optimized TCP/UDP stack in the application domain which can communicate directly with the

Solarflare server adapter. With OpenOnload, the adapter provides the application with protected, direct access to the network, bypassing the OS kernel, and hence reducing networking overheads and latency.

The typical TCP/UDP/IP stack resides as part of the kernel environment and suffers performance penalties due to context switching between the kernel and application layers, the copying of data between kernel and application buffers, and high levels of interrupt handling.

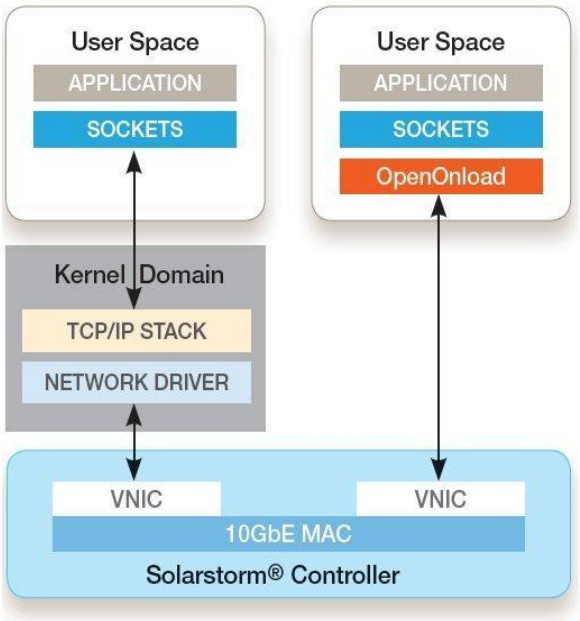


Exhibit 5: The Solarflare Architecture for OpenOnload

The kernel TCP/UDP/IP and OpenOnload stacks can co-exist in the same environment. This co-existence which is based on Solarflare's hybrid architecture allows applications that require a kernel-based stack to run simultaneously with OpenOnload. This coexistence feature was leveraged as part of the testing where the benchmarks were run through both the kernel and OpenOnload stacks in back to back fashion using the same build and without having to reboot the systems.

FUJITSU XG 2600 ULTRA-LOW-LATENCY 10GBE SWITCH

INDUSTRY LEADING LAYER 2 PERFORMANCE, SCALABILITY, AND HIGH AVAILABILITY

The Fujitsu XG 2600 offers industry-leading low latency in 10GbE switching. The XG 2600 is powered by Fujitsu's 4th generation Fujitsu AXEL-X2 ASIC, with measured chip-level performance as low as 300ns latency. The switch features 26-port SFP+ solution in compact 1RU high form factor and is highly reliable, with configurable airflow options (front-to-rear or rear-to-front airflow). With flexible SFP+ interfaces (Copper Twinax, SR and LR), it allows you to build the ideal network architecture for your requirements. The XG 2600 also features the lowest power consumption, fewer than 5 watts per port, which is best in its class. The Fujitsu XG 2600 10GbE switch is designed for high-performance systems.

CONCLUSIONS

- The findings analyzed in this white paper represent the results of testing of transmit latency of a configuration with the Solarflare server adapter with OpenOnload and the Fujitsu switch at transmission rates up to 3 million messages/second (mps). For the 64 byte message sizes typical of market data messaging systems, very low latency was observed:
- TCP latency Mean did not exceed 4.5 microseconds with the switch
- TCP latency Mean did not exceed 4.1 microseconds without the switch
- For TCP messaging traffic, 99th percentile did not exceed 5.0 microseconds with switch
- For UDP latency server to server, mean latency was as low at 3.9 microseconds

The system also demonstrated low jitter for both TCP and UDP traffic which delivers a very predictable messaging system. With Solarflare's 10GbE server adapter and OpenOnload application acceleration middleware, and Fujitsu's 10GbE switch, off-the-shelf 10GbE hardware can be used as the foundation of messaging systems for electronic trading with no need to re-write applications or use proprietary, specialized hardware.

By enabling financial trading customers to implement highly predictable systems, Solarflare's and Fujitsu's10GbE solutions provide competitive advantages and offer increased overall speeds, more accurate trading and higher profits. By leveraging the server adapter with OpenOnload, IT managers are able to build market data delivery systems designed to handle increasing message rates, while reducing message latency and jitter between servers.

SUMMARY

Solarflare Communications and Fujitsu have demonstrated performance levels with 10 Gigabit Ethernet that enables Ethernet to serve as the foundation of messaging systems used in the financial markets. Now, financial firms can use off-the-shelf Ethernet, TCP/IP, UDP and multicast solutions to accelerate market data systems without requiring the implementation of new wire protocols or changing applications. With off the shelf 10GbE gear, Solarflare's server adapter and the Fujitsu switch can be used as the foundation of messaging systems for electronic trading and the support of low-latency UDP multicast with no need to re-write applications or use proprietary, specialized hardware. IT and data center managers can deploy plain old Ethernet solutions today.

Moving forward, Solarflare Communications and Fujitsu also expect high performance 10G Ethernet solutions with low-latency UDP multicast to become an important technology component of public and private clouds that rely on real time media distribution for business to consumer and business to business applications.

ABOUT SOLARFLARE

Solarflare is the pioneer in high-performance, low-latency 10GbE server networking solutions. Our architectural approach combines hardware and software to deliver high-performance adapter products and application acceleration middleware for superior performance in a wide range of applications, including financial services, high performance computing (HPC), cloud computing, storage and virtualized data centers. Solarflare's products are used globally by many of the world's largest companies, and are available from leading distributors and value-added resellers, as well as from Dell and HP. Solarflare is headquartered in Irvine, California and has an R&D site in Cambridge, UK. For more information on Solarflare products, visit http://www.solarflare.com or contact productinfo@solarflare.com.

ABOUT FUJITSU FRONTECH NORTH AMERICA

Fujitsu Frontech North America Inc. provides market-focused IT solutions that enable customers to achieve their business objectives through integrated offerings for retail point of sale, self checkout solutions and kiosks, mobile, cloud computing, digital media solutions, biometric authentication technology, Ethernet switches, RFID tags and currency handling solutions. Working closely with a diverse network of partners for sales, systems integration, and managed services, FFNA delivers industry-specific solutions for the manufacturing, retail, healthcare, government, education, financial services, enterprise and communications sectors throughout North America. For more information, visit www.fujitsufrontechna.com.

ABOUT THE AUTHOR

Bruce Tolley is responsible for solutions marketing at Solarflare including technical, event, and partner marketing activities. Previously, he served as Solarflare's Vice President of Marketing, and earlier Director of Product Management. Prior to joining Solarflare, Bruce was a Senior Product Line Manager at Cisco Systems where he managed the Ethernet transceiver business that included product life cycle management and the launch of Metro Ethernet, 10 Gigabit, and 1000BASE-T switch solutions. Prior to Cisco, he served in various product and marketing management roles at 3Com Corporation. Formerly Study Group Chair of the IEEE 802.3aq 10GBASE-LRM standards project, Bruce is a frequent contributor to the IEEE 802.3 Ethernet standards projects. He is currently serving as Secretary and Director of the Ethernet Alliance. He is an alumnus of Selwyn College, Cambridge University and Tuebingen University, Germany and holds MA and Ph.D degrees from Stanford University and an MBA from Haas School of Business, UC Berkeley.

APPENDIX: LIST OF BENCHMARKS

Solarflare's test procedures are documented in its "Low Latency Quickstart Guide" available for download from the driver download page at www.solarflare.com.