

Achieving Low, Consistent Latency in Market Data Distribution with Solace, Cisco and Solarflare

With more trading decisions and order execution processes being automated, it is critical to make sure information flows quickly and consistently through the front, mid and back office. The picture goes well beyond just raw speed, however, as market participants understand that consistency and predictability are just as important.

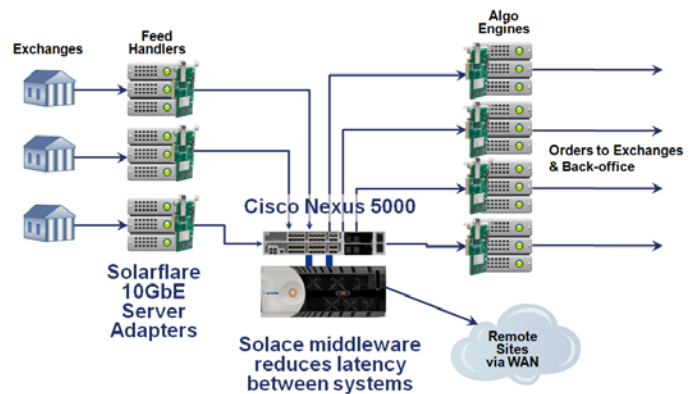
This paper describes how a system based on message routers from Solace Systems and 10GigE technology from Cisco Systems and Solarflare Communications can deliver lower and more consistent latency than software-based solutions. It also provides performance numbers including average and 99.9th percentile figures at a variety of rates. When routing 1,000,000 messages per second the platform exhibited average latency of 28 microseconds and 99.9th percentile latency of 32 microseconds. At five times that volume, a rate that represents substantial performance headroom for most firms, latency was still low with minimal jitter: averaging 39 microseconds with 99.9th percentile of 48.



Architectural Overview

In traditional ultra-low-latency architectures, messaging software runs on the publisher and subscriber servers, network protocol stacks are implemented in the operating system and network connectivity is via store-and-forward layer 2 switches that often use 1GigE links.

The architecture shown here removes the bottlenecks that exist in real-world, high-volume scenarios.



- The Solace message router implements message routing and filtering in purpose-built hardware using 10GigE links.
- A Cisco Nexus 5000 Series switch streams packets instead of performing store and forward.
- The Solarflare Adapter accelerates the process of moving Ethernet packets from the wire to the host CPU over PCIe.
- The Solarflare Enterprise OnLoad provides kernel bypass capabilities, so the TCP/IP stack runs in user space rather than in the kernel, which then avoids expensive interrupts and context switching.

Since message filtration and forwarding is in hardware, this configuration delivers remarkably low, consistent latency.

Test Configuration, Methodology and Results

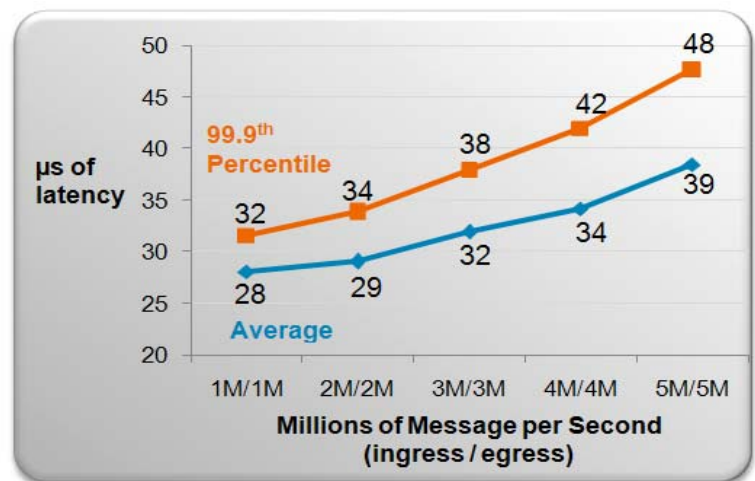
The tests were designed to approximate a real-world system and scenario, and independent testing by actual customers has validated the results, so firms can expect to see similar numbers in their own systems assuming similar parameters.

Latency measurements were taken from the time the publishing application calls the Solace API send function to when the subscribing application receives the message from the Solace API, so it is the latency a typical application would experience. The publisher and subscriber applications performing the measurements are located on separate CPU cores of the same server in order to ensure both accurate clock synchronization and accurate one-way latency measurements, rather than deducing one-way latency from a two-way latency measurement as the latencies are rarely symmetrical.

This test scenario does not result in microbursts of traffic that cause congestion in the network. In real-world environments, microbursts are prevalent, especially during periods of high-volatility. From a network infrastructure perspective, having adequate buffers to handle short-lived congestion is critical to prevent congestion collapse due to excessive packet drops.

Test System and Parameters

- Solace 3260 Content Router w/ 2x10 GigE NAB
- Cisco Nexus[®] 5010 Switch
- Solarflare Solarstorm SFN4112F 10GbE SFP+ server adapters
- Enterprise OpenOnload™
- Pub/Sub servers: Intel Xeon Nehalem 5550 quad core, 2.67 GHz running CentOS 5.2 X86 64 bit
- 100 byte messages with 12 byte topics
- 10 minute duration
- 10 publishers, 10 topics, 10 subscribers



Note that these results do not represent the maximum capabilities of each product, just their performance in this scenario.

Cisco Nexus 5000 Series Switches

Simplify data center transformation with the innovative, standards-based, high-performance Ethernet and Unified Fabric server access layer of Cisco Nexus 5000 Series Switches. Next-generation data centers are increasingly dense, multicore, and virtual machine intensive. The Cisco Nexus 5000 Series meets business, service, application, and operational requirements of such data centers. With Nexus 5000 Series Switches, you can:



- Consolidate the data center and protect investments in existing server, network, storage, and facilities assets
- Decrease the total cost of ownership by simplifying the data center infrastructure
- Increase business agility with easier, faster, and more pervasive data center virtualization
- Enhance business resilience with greater operational continuity
- Use existing operational models and administrative domains for easy deployment

Solarflare 10 GigE Server Adapters

Solarflare Communications is delivering the next level of high-performance 10 Gigabit Ethernet (10 GigE). The robust and power-efficient server adapter and silicon solutions make possible next-generation applications such as low-latency networking, cloud computing, server virtualization, and network convergence.



Solarflare Server Adapters provide the highest possible line-rate performance, excel in small message processing, and have demonstrated performance leadership in the most demanding application environments.

Solarflare Solarstorm 10 GigE server adapters provide the following benefits:

- The single-port SFN4112F offers 20 Gbps throughput (full-duplex) so full 10G bandwidth is available to critical applications. The dual-port SFN 5122F server adapters support PCIe 2.0 and offer 40 Gbps.
- Offloads for critical compute-intensive tasks to insure that minimum burden is placed on the server CPU, freeing up processor cycles for customer applications
- Fully compatible with Solarflare's OpenOnload™ software, which provides direct socket connections to applications

Solace 3260M Message Router

The Solace Message Router is a middleware appliance that enables organizations to increase the speed of their application infrastructure by 10 to 100 times, improve the manageability of their systems and reduce costs. It supports reliable messaging and non-persistent JMS, as well as guaranteed message delivery, persistent JMS and message caching. All of its capabilities and services are accessible and manageable through Solace's unified API and administration framework.



- **High Performance:** A hardware datapath with no OS handoffs or

context switching enables latency that is as consistent as it is low — just tens of microseconds even at very high rates and to the 99.9th percentile.

- **Low Complexity and TCO:** By supporting many messaging types and capabilities (reliable, guaranteed, persistent and non-persistent JMS, and message caching), and by offering the “out of the box” simplicity of a purpose-built appliance, the Solace Message Router is resource efficient in terms of rack space and power, as well as easy to deploy, operate, upgrade and scale.
- **Scalability:** Modular architecture enables the customization of features and capacity to meet specific requirements, and supports the addition of capacity or functionality over time without increasing the solution’s footprint.
- **High Availability:** Redundant components and connectivity provide built-in high-availability, and support for paired deployment provides out-of-the-box fault tolerance without requiring additional equipment or software.
- **Manageability:** The use of discrete TCP connections instead of multicast make it easy for administrators to find and fix faults thanks to visibility into client-specific metrics such as connections, queue depths and message rates, all with no impact on system performance. The Solace Message Router is based on flexible FPGA chips so administrators can upgrade them with a single command.