Improving Load Balancer Performance: NGINX Plus Acceleration with Solarflare OpenOnload and Flareon Ultra 40GbE Server I/O Adapter

SolarflareWhitePaper

Executive Summary

The Solarflare Flareon™ SFN7142Q 40GbE Server I/O Adapter with Solarflare OpenOnload® can deliver up to a 7x the connection rate performance over the Intel XL710 40GbE Adapter with its kernel driver using NGINX Plus as a load balancer at 1K Byte packet sizes. At 10K and 100K Byte packets, the SFN7142Q Adapter with OpenOnload delivers up to 5x the connection rate over the Intel XL710 Adapter. These performance increases are accomplished by using the OpenOnload high-performance, open source, user-level networking stack. OpenOnload enables the NGINX Plus load balancer to achieve higher performance by bypassing the operating system's kernel and leveraging the advanced functionality of Solarflare server I/O adapters.

When we look at the performance as a function of how many CPU cores are used, the Solarflare 7142Q with OpenOnload continues to scale linearly with increasing connection rates, while the Intel XL710 performance with its kernel driver remains relatively flat or peaking at 4-5 CPU cores. OpenOnload unlocks the power of the CPU cores and increases NGINX Plus HTTP and TCP connection rates. In short, the NGINX Plus load balancer application running with Solarflare optimized network I/O enables both increased connection rate performance and CPU core scalability across all packet sizes .

Introduction: Scope and Purpose

A load balancer distributes incoming Ethernet packets, e.g., HTTP and TCP connection requests, over a number of compute resources – often a cluster of servers – according to a specific algorithm. The algorithm often looks to optimize the cluster resources, maximize connection rate and throughput, minimize response time, and avoid overload of any single compute resource.

For many Web, CDN (Content Delivery Network), cloud and enterprise customers, increasing their load balancer performance is a constant area of focus. Beyond algorithm optimization, these improvements can include increasing the network and server performance, and enabling the delivery of higher application performance. Software load balancers inherently receive a performance boost with each server technology improvement, however additional improvements can be realized through the use high-performance network adapters and application acceleration middleware such as the Solarflare Flareon 10/40 Gb Ethernet server adapters and the OpenOnload kernel bypass user-level networking stack especially when they are used with software load balancers including the NGINX Plus solution from NGINX, Inc.

NGINX is the world's most popular open source web server for high traffic sites, powering over 100 million properties. NGINX Plus includes enterprise-ready features such as HTTP and TCP load balancing, session persistence, healh checks, advanced monitoring, and management to give you the freedom to innovate without being constrained by infrastructure.[1]

[1] *http://www.nginx.com/products Retrieved 2016-01-28*

sales@solarflare.com

US 1.949.581.6830 x2930

UK +44 (0)1223 477171

HK +852 2624-8868

www.solarflare.com

SolarflareWhitePaper

This paper investigates the use case where a software load balancer running NGINX Plus handles many concurrent connections from multiple clients targeting a cluster of two web servers at varying payload sizes. The performance benchmark shows the performance improvement with a Solarflare SFN7142Q 40GbE with OpenOnload over an Intel XL710 40GbE network adapter using its Linux net driver.

OpenOnload

Solarflare OpenOnload is a Linux-based, open source, high-performance application accelerator middleware. It is an implementation of TCP and UDP over IP which is dynamically linked into the address space of user-mode applications, and granted direct (but safe) access to the network adapter hardware. The result is that data can be transmitted to and received from the network directly by the application, without involvement of the operating system, using a technique called "kernel bypass."

Transitioning into and out of the kernel from a user-space application is a relatively expensive operation: the equivalent of hundreds or thousands of instructions. With conventional networking such a transition is required every time the application sends and receives data. Kernel bypass avoids disruptive events such as system calls, context switches and interrupts and so increases the efficiency with which a processor can execute application code. This also directly reduces the host processing overhead, typically by a factor of two, leaving more CPU time available for application processing. The effect is most pronounced for applications which are network intensive including:

- Web-caching, load-balancing and Memcached applications
- HTTP web serving, web e-commerce servers
- Content Delivery Networks (CDNs)
- High-bandwidth video-streaming
- HPC (High Performance Computing)
- Market-data and trading applications

Benchmark Setup

Load Balancer

- Dell R630 server with two Intel Xeon E5-2620 v3 CPUs (2.40 GHz, 6 cores w/HT)
- 64 GB DDR4 SD-RAM at 1867 MHz
- Red Hat Enterprise Linux 7.0, kernel version: 3.10.0-123.el7.x86_64
- NGINX Plus v1.7.117
  - Running in upstream-keepalive and downstream-nonkeepalive modes
  - Upstream keepalive connections set to 4000
- Network adapters:
  - Solarflare SFN7142Q, full-feature firmware
  - Intel XL710
  - One 40GbE port upstream; one 40GbE port downstream
- Network adapter driver/firmware versions:
  - Solarflare Onload: OpenOnload-201502-u1/ v4.7.0.1009
  - Intel XL710: i40e 1.2.37

SOLARFLARE®

sales@solarflare.com

US 1.949.581.6830 x2930

UK +44 (0)1223 477171

HK +852 2624-8868

www.solarflare.com

SolarflareWhitePaper

Clients

- Eight Dell R210 servers each with one Intel Xeon E3-1230 CPU (3.20 GHz, 4 cores, no HT)
- 16 GB DDR3 SD-RAM at 1333 MHz
- ApacheBench (`ab`)[2] version 2.3 (patched: see Appendix I)
- Network adapters: Solarflare SFN6122F, Driver version OpenOnload-201502-u1
- The kernel networking stack handled the traffic on the client machines.

Web Servers

- Two Dell R630 server with two Intel Xeon E5-2620 v3 CPUs (2.40 GHz, 6 cores w/HT)
- 64 GB DDR4 SD-RAM at 1867 MHz
- Red Hat Enterprise Linux 7.0, kernel version: 3.10.0-123.el7.x86_64
- Nginx web server 1.7.7 (`SO_REUSEPORT`[3] was enabled for Nginx workers)
- Network adapters:
  - Solarflare SFN7002F and SFN7142Q, ultra-low latency firmware
  - Intel XL710
- Network adapter driver/firmware versions:
  - Solarflare Onload: OpenOnload-201502-u1/ v4.7.0.1009
  - Intel XL710: i40e 1.2.37

In each test, four `ab` instances were spawned on each client host, configured to attempt an equal number of concurrent connections varying according to the specific test. The sum of these attempted connections across all `ab` instances across all the client hosts is denoted in this paper by "concurrency". The `ab` instances were run without keep-alive, meaning that each `ab http` request established a new TCP connection.

Connectivity

A high port density top of the rack 40GbE switch was used to connect the load balancer, the clients and the web servers. This switch has four 40GbE ports and forty-eight 10GbE SFP+ ports. The clients and web server side networks used static routes for isolation; alternatively, VLANs or separate physical networks could have been used. The network interfaces on the web servers were configured with single IP addresses, whereas sixteen IP addresses were configured on each client interface to avoid TCP port-space exhaustion.

Methodology

The benchmark's goal was to maximize the connection rate that NGINX Plus can handle as a function of server configuration and packet size, i.e., the number of CPU cores used was artificially restricted in order to demonstrate the comparative efficiency of OpenOnload versus the Linux kernel for 1K, 10K and 100K Byte packets. Note that while the results are indicative of performance gains, absolute gains will be dependent on actual workloads. Because OpenOnload presents a standard BSD sockets API, NGINX Plus required no modification.

[2] *http://httpd.apache.org/docs/2.2/programs/ab.html Retrieved 2015-03-17.*

[3] *http://forum.nginx.org/read.php?29,241470 Retrieved 2015-03-18.*

SOLARFLARE®

sales@solarflare.com

US 1.949.581.6830 x2930

UK +44 (0)1223 477171

HK +852 2624-8868

www.solarflare.com

SolarflareWhitePaper

Benchmarking Parameters

In order to model a realistic load balancer scenario, it was necessary to run the ab clients with a fair level of concurrency. If the concurrency level was set too high, the server load increased beyond its capacity and the aim of the benchmark to assess processing efficiency became confounded by other factors, e.g. TCP connection retries, etc. It was decided to run the tests at a concurrency of 192.

Results

Benchmark experiments were conducted on the NGINX Plus load balancer with 40GbE Solarflare SFN7142Q and Intel XL710 adapters. **Figures 1, 2 and 3** show the performance results at 1K Byte, 10K Byte, and 100K Byte packet sizes, respectively.

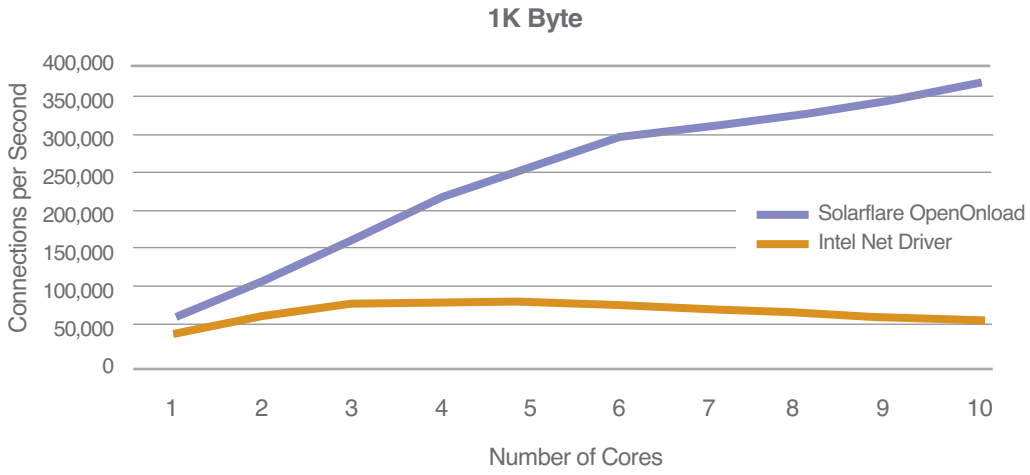

Figure 1. NGINX Plus 1K Byte Performance: Cores and Connections per Second.

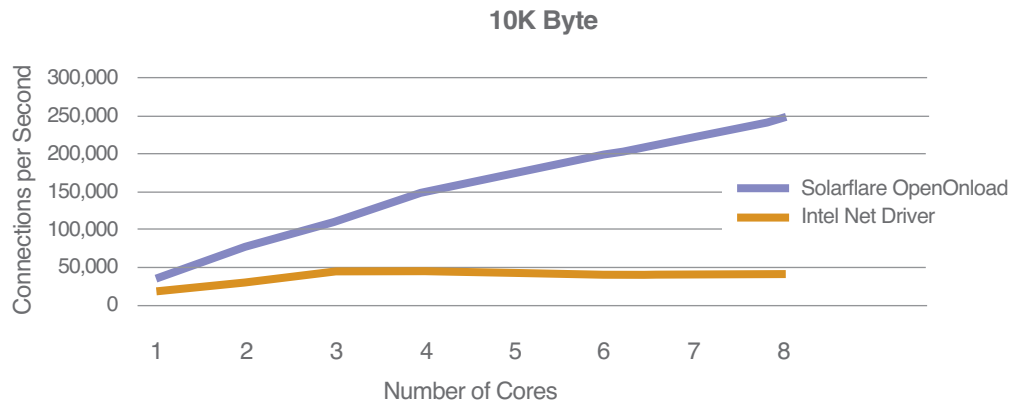

Figure 2. NGINX Plus 10K Byte Performance: Cores and Connections per Second.

sales@solarflare.com

US 1.949.581.6830 x2930

UK +44 (0)1223 477171

HK +852 2624-8868

www.solarflare.com

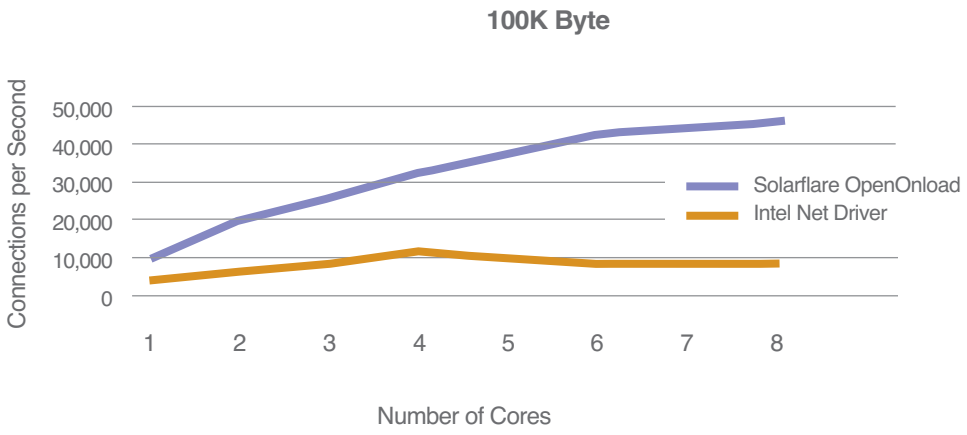

Figure 3. NGINX Plus 100K Byte Performance: Cores and Connections per Second.

The results in **Figure 1, 2, and 3** demonstrate that NGINX Plus with a Solarflare Flareon® Ultra SFN7142Q dual-port 40 Gbps adapter and OpenOnload delivers increased load balancer performance of up to 7 times for 1K Byte packets at 8 cores and 5 times for 10K and 100K Byte packets at 7 cores when compared against the Intel XL710 40Gbps adapter and the Intel Net driver. The results also demonstrate that NGINX Plus connection rate with a Solarflare SFN7142Q adapter and OpenOnload scales linearly with the number of CPU cores while the Intel XL710 adapter performance remains relatively flat across increased CPU cores.

Conclusions

With OpenOnload, an industry standard server running with the NGINX Plus software load balancer and Solarflare SFN7142Q 40GbE server adapters scales in performance with each additional CPU core and can achieve a connection rate improvement of 7 times over competitive solutions.

Bottom line: software load balancers including NGINX Plus can realized substantial performance gains with Solarflare OpenOnload and 40GbE server I/O adapters.

Appendix I. ApacheBench modification

To overcome some of ApacheBench limitations, its sources have been modified to achieve a number of features:

- Allow synchronizing multiple `ab` instances.
- Enable use of array of ip addresses per `ab` instance.
- Do not stop despite connection failures.
- Preset expected response size.
- Provide number of successful requests.

SF-115762-CD Issue 1
Improving NGINX_Plus_WP_021916
Copyright © 2016
Solarflare Communications, Inc.
All rights reserved.

SOLARFLARE®

sales@solarflare.com

US 1.949.581.6830 x2930

UK +44 (0)1223 477171

HK +852 2624-8868

www.solarflare.com