White Paper

Accton Tests Programmable Switch-Server for Virtualized Networks

Accton CSP-9550 switch-server uses Intel® Xeon® Scalable processors for high-performance compute, and Barefoot Tofino™ P4 switch for programmable connectivity. Tests show near wire-speed throughput with low CPU utilization.

Authors

Charles Tso
Accton

Joyce Huang
Accton

Vincent Cho
Intel

Accton
Making Partnership Work

Table of Contents

Executive Summary

Network modernization is a real, mission-critical endeavor that is central to the goals of cloud service providers, enterprises, and telecom companies. This network evolution is driven by the dramatic increase in the number of connected internet of things devices, the emergence of fast 5G wireless networks, and increased data volumes from streaming video and cloud computing.

To deal with this expanded network scale, network operators need to dramatically simplify network operations and lower operational costs using the latest in network equipment, network automation, and software-defined networking applications. The Accton CSP-9550 is a new breed of switch-server that combines programmable switching capability with high-performance compute into a new network appliance that meets the needs of these high-capacity networks.

Background

To keep up with the significant growth in data in network, data centers and communications service providers (CommSPs) are turning to network functions virtualization (NFV) and containerization. Virtualized services can be run on Intel® architecture-based servers, which can reduce costs dramatically and can also enhance the operator's ability to scale services up or down based on demand. 5G networks are already making heavy use of virtualized network technology for service agility and cost effectiveness.

A newer trend that is bringing flexibility to networks is switch programmability. Legacy switch ASICs were fixed in their support of networking protocols and switching functionality. It could take several years for a switch ASIC to support a new protocol, and certain packet actions that are optimally done at the network ingress have to be forwarded to separate server for processing. In 2014, the P4 programming language was proposed as an alternative to fixed function switching. A programmable P4 switch ASIC supports both new networking protocols and can run packet processing applications such as firewalls and load balancers. The P4.org standards body, a project of the Open Networking Foundation, is driving the progress of the P4 community, which currently has more than 100 members.

These trends have changed the switch-server landscape (see Figure 1) as legacy top of rack (ToR) switches have evolved to the switch server, a combination of a programmable switch and compute functionality that offers comprehensive capabilities for packet processing that can be used in data centers or at the network edge.

Accton is a pioneer in this emerging market with the availability of its CSP-9550, a switch-server that utilizes Intel® Xeon® Scalable processors for high-performance compute and the Barefoot Tofino™ P4 switch for programmable connectivity.

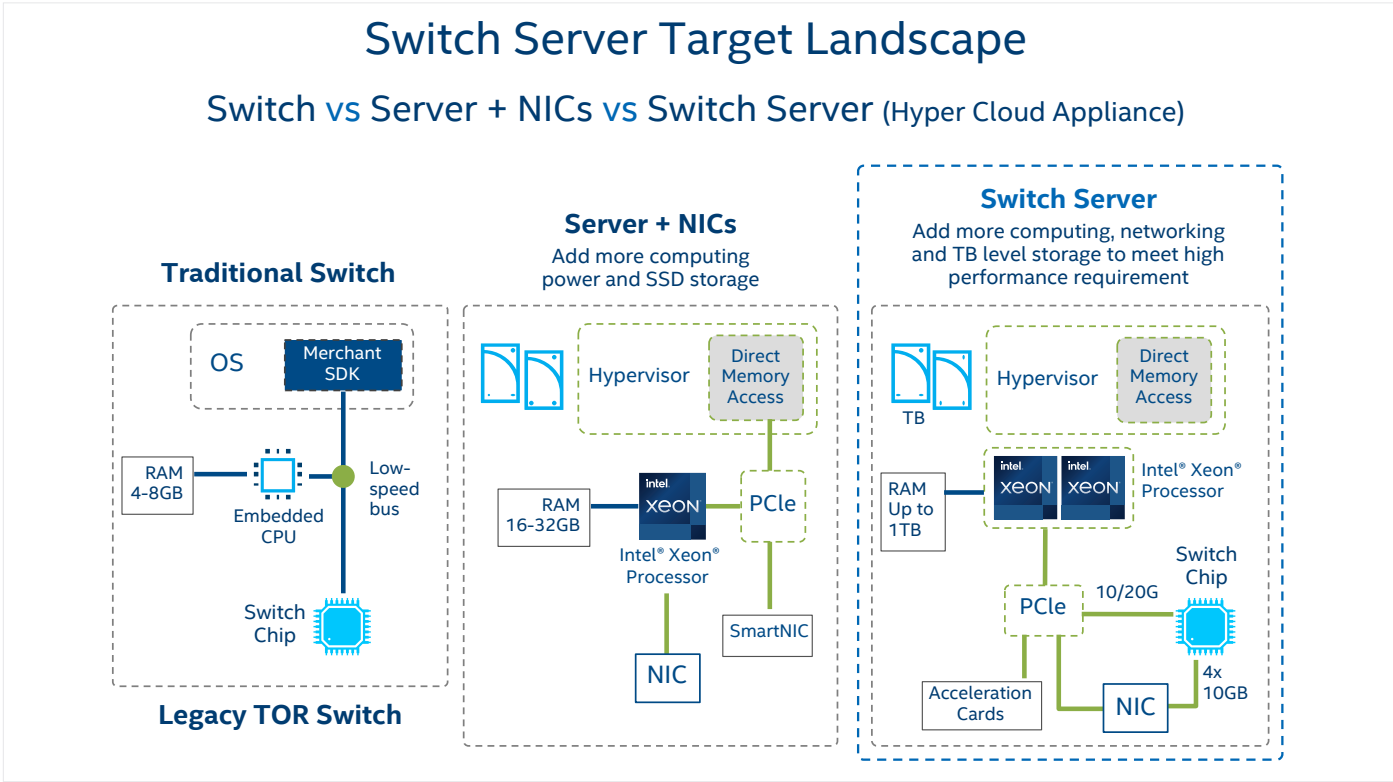

Figure 1. Switch-server evolution

Accton CSP-9550

Accton is a leading provider of open hardware network fabric for cloud data centers, and the Accton CSP-9550 is part of the company's family of open cloud data center switching and compute solutions. The Accton CSP 9550 is the company's first hyper cloud appliance, integrating computing, storage, and networking into a single device that can be configured for use cases such as the following:

- Real-time network visibility and security (in-band network telemetry)
- Software-defined WAN (SD-WAN)
- Network packet brokers
- Application delivery controller
- 4G/5G virtual radio access network
- Cloud orchestration
- Virtualized data security

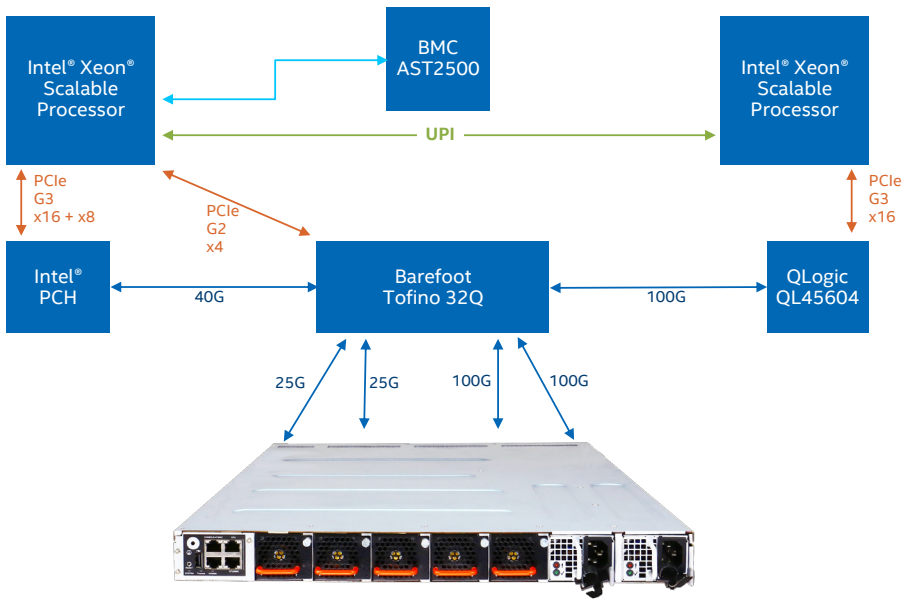

Figure 2. Accton CSP-9550 high-level block diagram

As seen in the block diagram in Figure 2, the 1RU high Accton CSP-9550 server is based on a dual-socket design that can be configured with a number of Intel® Xeon® Scalable processors. These processors offer a highly advanced compute core designed into a broad portfolio of balanced CPUs to enable communications networks to power through compute-hungry workloads and to scale to meet the dynamic performance requirements increasingly seen in cloud deployments.

With up to 28 cores and significant increases in memory bandwidth (six memory channels) and I/O bandwidth and throughput (48 PCIe lanes), this platform delivers notable efficiencies in deep packet inspection workloads and packet processing for virtual network functions.

The high-capacity Accton CSP-9550 design supports 12 DDR4 DIMM slots per CPU, with local storage options including two SATA III and two U.2 NVMe sockets (maximum of two 2.5" devices). The hardware platform can operate using any supported operating system and virtualization infrastructure, or can be custom developed based on application requirements.

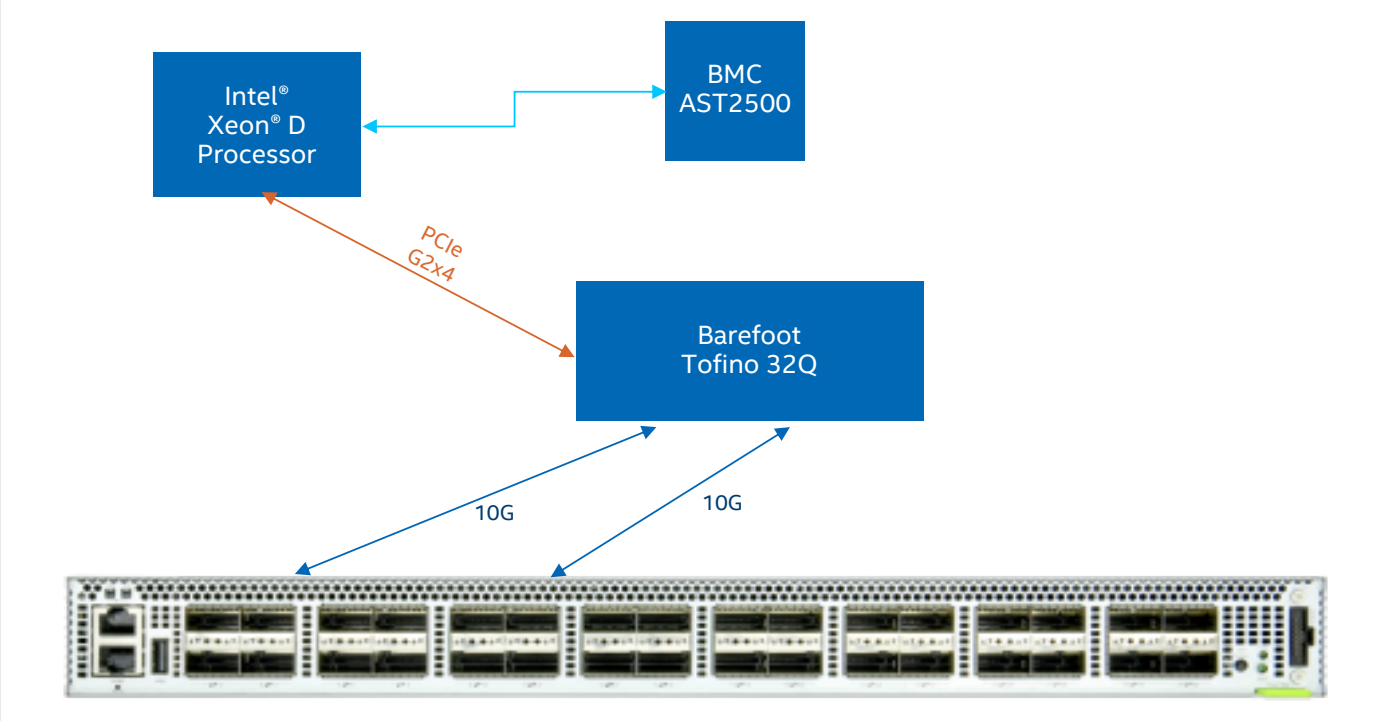

Figure 3. Wedge100BF-32X high-level block diagram

Wedge100BF-32X

Accton's other P4 programmable switch is the EdgeCore Wedge100BF-32X (see Figure 3). In a comparison test for this paper (see below), the Wedge100BF-32X was configured as a top-of-rack (ToR) switch that is typically used in a high-performance data center. The Wedge100BF-32X comes in a 1RU form factor and features the Barefoot Tofino switch ASIC, although no programmability was utilized in the testing. The switch comes configured with 32 x QSFP28 ports, but only two 10 Gbps network connections were used in the tests.

Programmable Switching

The programmable switching built into the Accton CSP-9550 is based on the Barefoot Tofino switch and is configured to support 48x 25 GbE (SFP28) and 8x 100 GbE (QSFP28) network ports.

The Barefoot Tofino switch supports up to 25 100 GbE interfaces, but is available in four configurations with capacities that range from 1.9 Tbps to 6.5 Tbps. The P4 programmability of the Barefoot Tofino switch is delivered using the switch ASIC's Protocol Independent Switch Architecture (PISA). PISA is based on a programmable packet parsing logic and match-action forwarding engine. The Figure 4 diagram shows the match-action engine, where Accton CSP-9550 users can program their switches using Barefoot P4 Studio, a complete suite of tools for development, debugging, and optimization of P4 applications. In addition, the device and abstraction APIs built into Barefoot P4 Studio allow developers to easily integrate their P4 applications with the local or remote control plane. These tools and APIs enable OEMs, cloud service providers, telecom operators, and ecosystem partners to build highly differentiated fit-for-purpose networking solutions.

The Accton CSP-9550 allows for linkage between the switch and the server, enabling virtualized applications to share the flexibility of the programmable switching, along with the high-performance CPU.

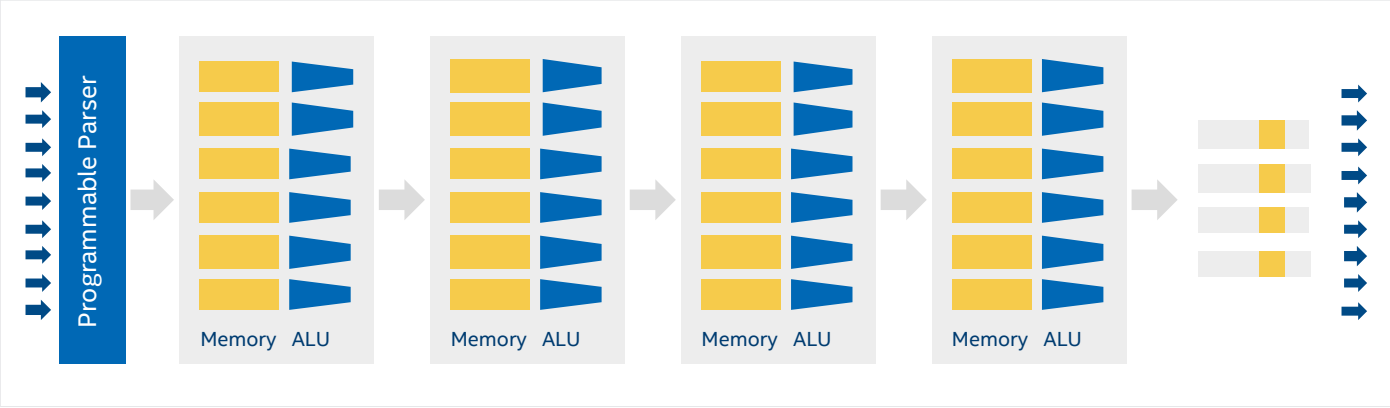

Figure 4. Programmable packet parser and match action forwarding engine of Barefoot Tofino P4 programmable switch

Accton CSP-9550 Performance Testing

To demonstrate the performance of the Accton CSP-9550, the company set up a series of tests to evaluate the data throughput performance between the CPU and the media access control (MAC) on the Ethernet network interface controller (NIC) at speeds up to 100 Gbps.

The tests were configured to display throughput by individual port, total switch throughput, file transmission performance, and the CPU loading for each interface when operating at maximum throughput.

The test results demonstrate the switch-server can operate at near wire-speed without utilizing a significant portion of the CPU power.[1] This leaves bandwidth and compute available to run applications in the switch, such as deep packet inspection (DPI). Because of the programmability of Accton CSP-9550, DPI can be deployed as a P4 program running in the switch so that packets are processed in the switch and transferred to the CPU for processing before being transferred back to the data plane to be forwarded to their destination.

Test Setup

The test connectivity for both switches is shown in Figure 5 and Figure 6. For the Accton CSP-9550, direct channels were set up between the NIC and the CPU on a 100 Gbps connection and on four 10GKR connections. For the Wedge100BF-32X, the same connections were set up using two 10GKR links.

iPerf, a popular open source network performance measurement and tuning tool, was used to test bandwidth capability. Turbostat was used to measure the CPU loading.

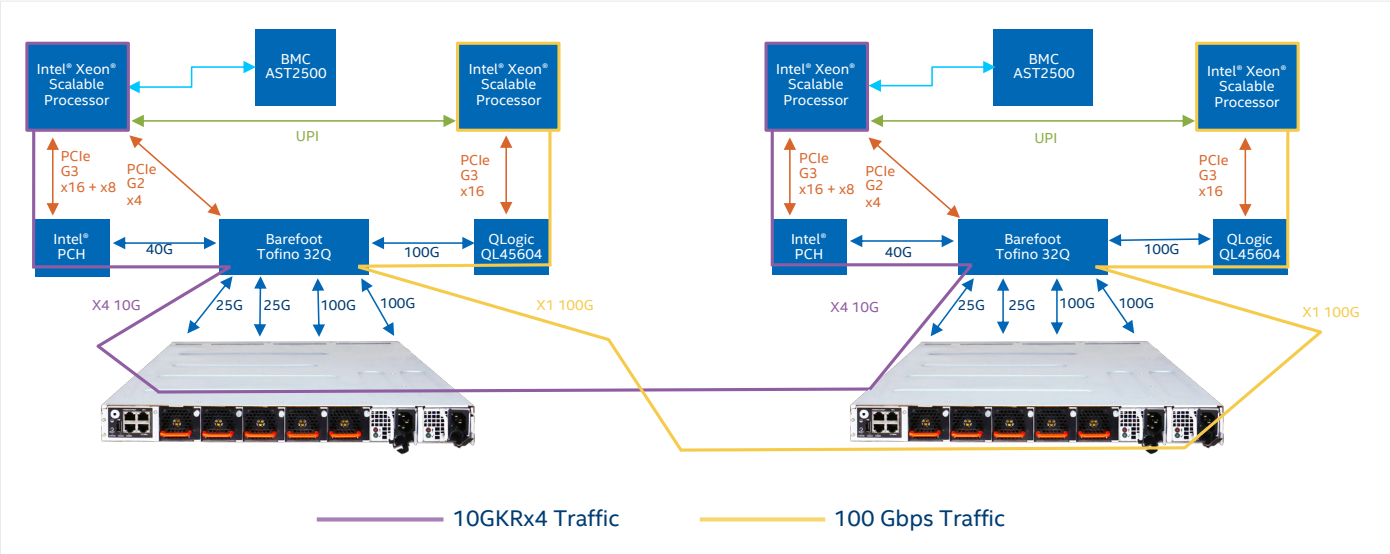

Figure 5. CSP-9550 test configuration

[1] Testing done by Accton in May 2020. System 1 is the dual-socket Accton CSP-9550 switch-server powered by the 28-core Intel® Xeon® Gold 6258R processor (microcode: MBF50657_05002f00). Programmable switch ASIC is the Barefoot Tofino. Both Intel® Hyper-Threading Technology and Intel® Turbo Boost Technology were turned on and the BIOS was PurleyCrb_0ACLA050. The system featured 192 GB of RAM comprising 24 8 GB DDR4-2666 MHz memory modules. System storage totaled 128 GB via a SanDisk X600 2.5" SATA SSD. Networking was provided by an Intel® Ethernet Controller I210. Operating system was the SONiC201911 (kernel: 4.9.0-9-2-amd64). Compilers used in the test were gcc (Debian 6.3.0-18+deb9u1) 6.3.0 20170516 and g++ (Debian 6.3.0-18+deb9u1) 6.3.0 20170516. Both iperf version: 2.0.9 (1 June 2016) pthreads and turbostat libraries were used for the testing, with set_irq_affinity.sh also in use.

System 2 is the Accton Wedge100BF-32X 32-port switch powered by the Barefoot Tofino switch ASIC with control plane processing provided by a 4-core Intel® Pentium® Processor D1517 (microcode: 07000012). Intel® Hyper-Threading Technology was turned on and Intel® Turbo Boost Technology was turned off, and the BIOS was R1.00.E2. The system featured 8 GB of RAM comprising 1 8 GB SO-DIMM DDR4-2133 MHz memory module. System storage totaled 128 GB via a Transcend 128 G M.2 SSD. Operating system was Open Network Linux (kernel: 4.14.151-OpenNetworkLinux). Compilers used in the test were gcc (Debian 6.3.0-18+deb9u1) 6.3.0 20170516 and g++ (Debian 6.3.0-18+deb9u1) 6.3.0 20170516. Both iperf version: 2.0.9 (1 June 2016) pthreads and turbostat libraries were used for the testing, with set_irq_affinity.sh also in use.

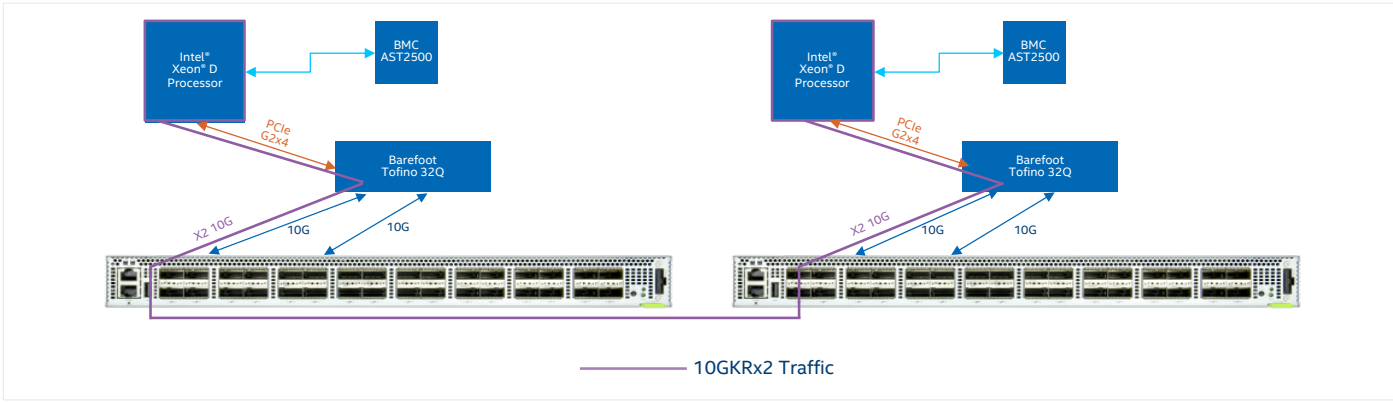

Figure 6. Wedge100BF-32X test configuration.

Test Results

Figure 7 and Figure 8 show the results of the throughput tests of each individual interface. In these tests, data was collected on the performance over time from each NIC on both systems being tested. Test time started at 60 seconds and was increased by 60-second intervals until it reached 600 seconds. The tests showed very high throughput of 94.2 Gbps on the 100 Gbps connection; throughput held steady until the 480-second mark, dropping to 89.6 Gbps by the 600-second mark. Performance of the four 10 Gbps links in the CSP-9550 held steady at 37 Gbps throughout the tests. For the Wedge100BF-32X, performance of the two 10 Gbps links remained steady at 18.8 Gbps across the entire test time frame.

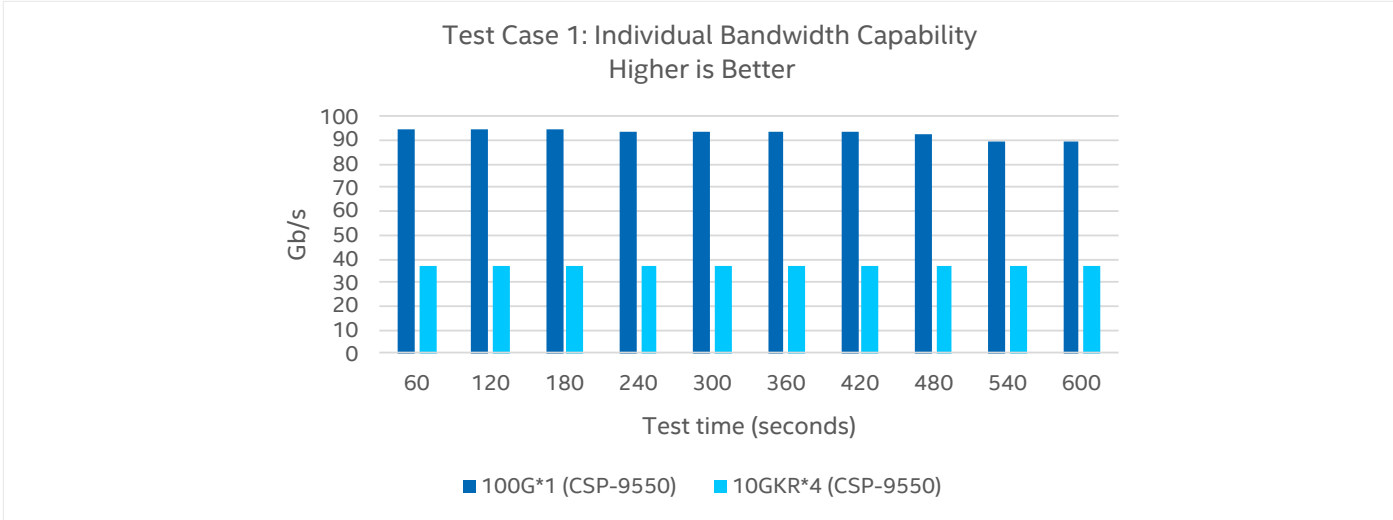

Figure 7. CSP9550 Individual bandwidth test results for 100Gx1 and 10GKRx4

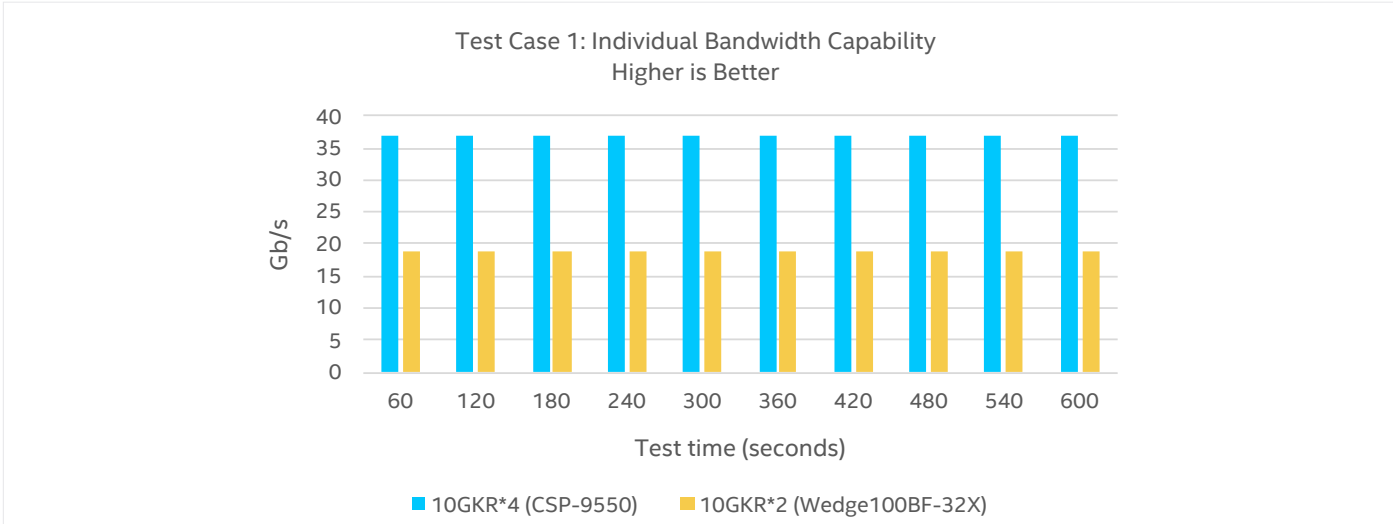

Figure 8. Individual bandwidth test results for CSP9550 (10GKRx4) and Wedge100BF-32X (10KGKRx2)

The second test case compared the sum bandwidth capacity of the CSP-9550 and the Wedge100BF-32X over the same time period, marked by 60-second intervals. As seen in Figure 9 the CSP-9550 initially delivered 123.2 Gbps out of the 140 Gbps potential of its 100 Gbps NIC and four 10 Gbps links. That performance held constant until the 360-second mark when it started to decline slightly, eventually ending up at 116.6 Gbps at the 600-second mark. For the Wedge100BF-32X, the total bandwidth of its two 10 Gbps NICs was consistent at 18.8 Gbps throughout the entire test time.

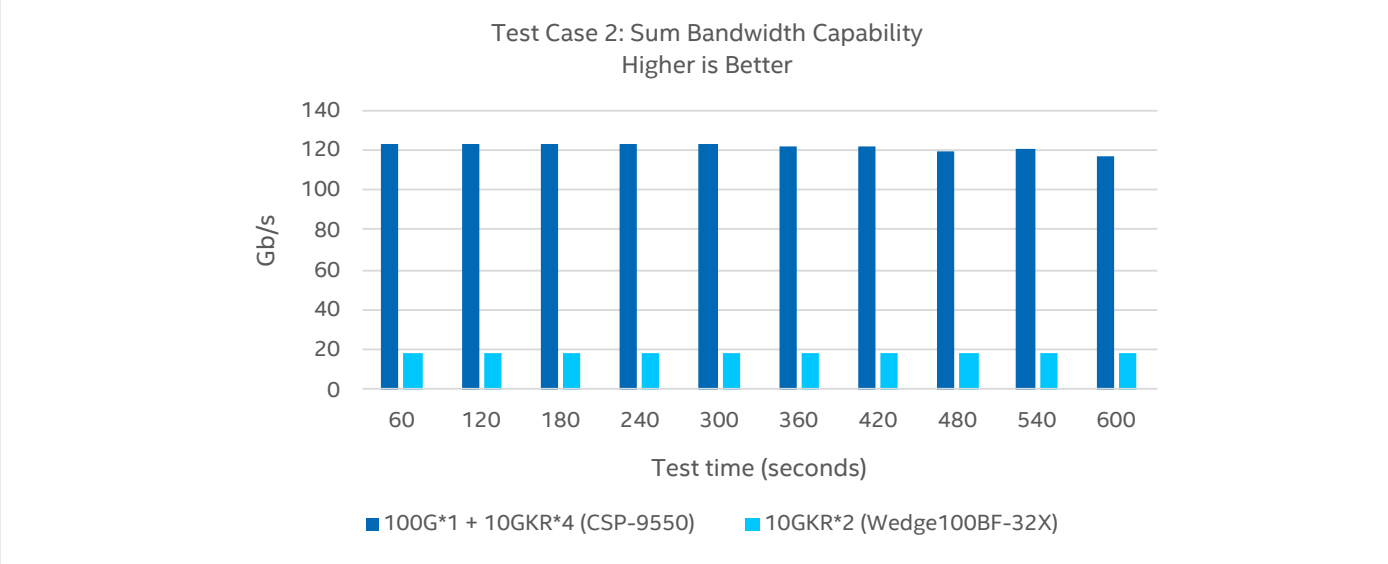

Figure 9. Sum bandwidth performance

The third test case showed the transfer time in seconds for a range of files that increased in size from 80 GB to 800 GB. Figure 10 shows these transfer times for each NIC on both systems, with a focus on the 10 Gbps connections.

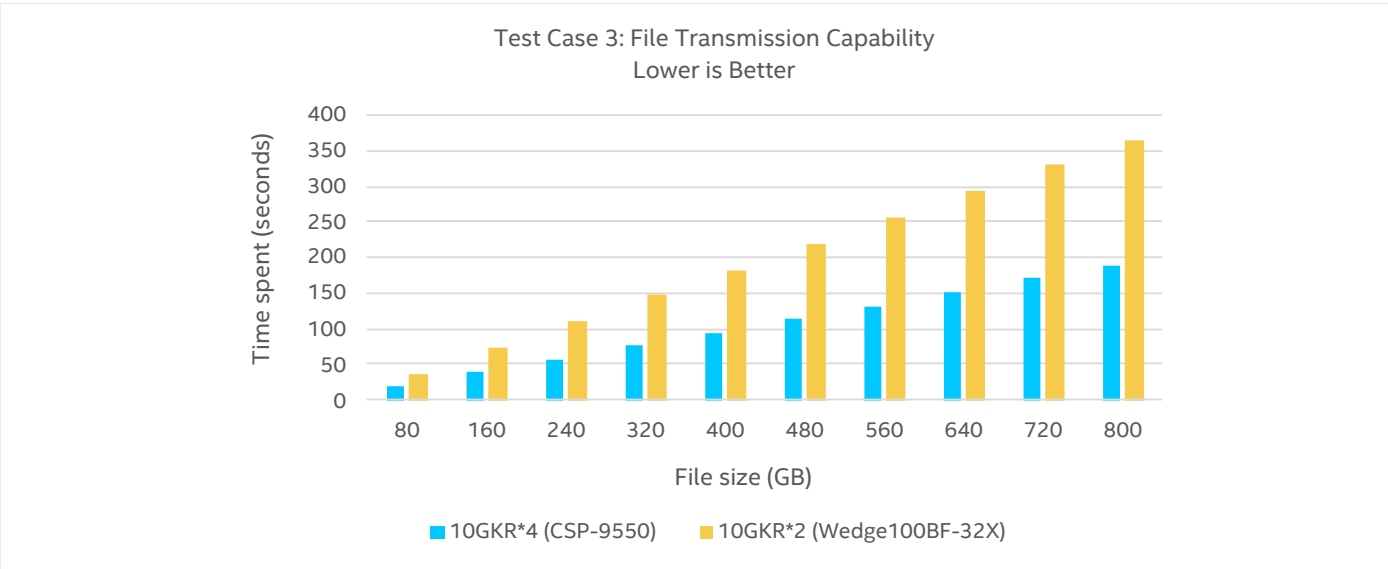

Figure 10. File transmission test results

Test case four demonstrated the CPU loading on the CSP-9550 and Wedge100BF-32X when each of the ports transmitted at line rate. Figure 11 shows the CPU load on both platforms when the two 10 Gbps NICs are engaged. Under these conditions, the CPU load on the CSP-9550 is between 0.75% and 0.78%, while the load on the Wedge100BF-32X is between 21.96% and 22.03%. This difference is explained by the higher performance delivered by the Intel Xeon Gold 6258R processor used in the CSP-9550.

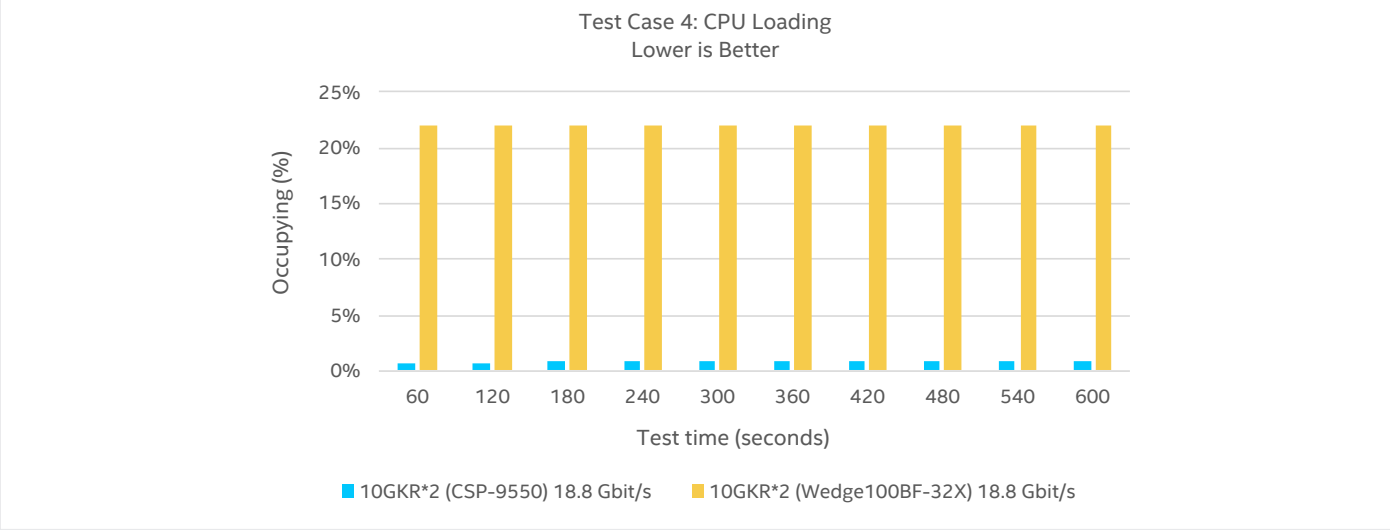

Figure 11. CPU loading test results

Conclusion

Accton's CSP-9550 switch server utilizes the Intel Xeon Scalable processors to serve as a right-sized compute appliance suitable for applications in hyperscale data centers to mid-size enterprise network infrastructure. In the tests described in this whitepaper, Accton has demonstrated significant bandwidth advantages in data plane traffic with low CPU load. With this performance, and by utilizing the programmability of Barefoot Tofino switch ASIC, the Accton CSP-9550 provides a highly flexible hardware solution with a programmable data plane that cloud service providers or communications service providers can trust in networking applications.

Learn More

Accton Technology Corporation: https://www.accton.com, jesse_ding@accton.com

Intel® Network Builders: http://networkbuilders.intel.com

Appendix: Additional Testing Contributors at Accton

Shone Wu

Will Chen

Norman Chou

Farley Hsu

Alvin Feng

Weichen Chen

Zoey Chen

Jianxin Yeh

Sheldon Chang

Tiejun Jin

0121/DO/H09/PDF ♻ Please Recycle 343871-001US