

The Ultra High-End Network Application Platform

Executive Summary

The technological momentum that networking is going through is at the same time exciting and challenging. On the one hand, the explosion of internet traffic drives the need to upgrade deployed interfaces to much higher bandwidths and this is increasing at an exponential pace. On the other hand, the advent of virtualization and software defined networking is transforming the overall network architecture and the relationship between the different network elements.

Both trends bring new complexities to the design of the underlying infrastructure and its ability to fulfil the promises that these new technologies hold: Terabit throughput and unprecedented flexibility. With the arrival of 100GbE interfaces, the CPU complex of platforms running data plane intensive applications struggles to keep up to speed with I/O throughput. The millions of packets to be processed by middle boxes turn hardware inefficiencies into bottlenecks. Valuable computing resources need to be dedicated to process application software and cannot be burned switching packets or performing packet processing tasks that are more cost effective when offloaded. The imminent arrival of virtualization to network appliances adds to this situation. The increasing number of network nodes and the new service chain quickly intensify inter-function traffic and move scalability to the top of the requirements list when building the new network infrastructure.

The FWA-6522C is Advantech's new concept of ultra high-end network platform to boost performance of applications migrating to 100GbE. The FWA-6522C makes the most of its four integrated Intel® Xeon® processors and its 400 Gbps of I/O by optimizing packet delivery from the network ports into the CPU sockets and into the virtual machines. This greater efficiency is achieved through the integrated Netronome FlowNIC technology that offloads flow processing and load balancing and accelerates Open vSwitch implementations for outstanding application performance. This paper describes this innovative platform architecture and shows some applications that can leverage the ultra high-end FWA-6522C appliance.



Hardware Platform – FWA-6522C

- Quad Intel® Xeon® Processor E5-4600 v3
- 32 x DDR4, ECC/REG DIMM, up to 1TB memory capacity
- Up to 400 Gbps of modular network I/O through a variety of GbE, 10GbE, 40GbE or 100GbE modules
- Netronome FlowNIC for L2-L7 stateful packet processing and load balancing
- Two internal PCIe slots for Intel® QuickAssist or other accelerators
- Compact 2U mechanical design with hot-swappable fans and PSU modules

Network Evolution

Traffic Growth

Networks are transforming at fast pace to cope with increasing internet traffic. IP traffic growth is mainly driven by two factors: the growing demand of video content delivered over the top and the increasing number of connected devices. According to Cisco Visual Networking Index (VNI), IP traffic will grow 3-fold from 2014 to 2019. In 2019, internet video traffic will be 80% of all consumer Internet traffic and the number of devices connected to IP networks will be three times as high as the global population.

To handle traffic growth, service providers and operators are constantly looking to deploy higher bandwidth network equipment. The IEEE 802.3 working group is picking up on this demand and responding positively. Ethernet bandwidth has grown about ten-fold every 5 years with 100GbE network ports now shipping. Next generation 400GbE feasibility has already been proven and the corresponding IEEE P802.3bs is expected to be approved early 2017.



Figure 1. Global IP Traffic Growth

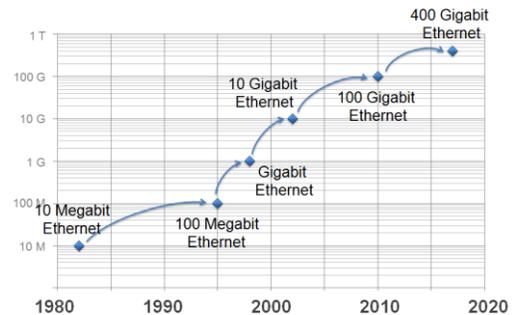


Figure 2. Ethernet Speed Evolution

Software-Defined Virtual Networking

Up until now, increasing throughput was the immediate answer from telecom equipment manufacturers to network traffic growth. But moving forward, the industry is anticipating a more profound transformation. The future virtualized and software defined network changes how services are provisioned and allows for a more flexible response to fluctuations in demand making a more efficient use of the infrastructure and opening up new business opportunities. Networks Functions Virtualization (NFV) and Software Defined Networking (SDN) are the disruptive technologies that enable this model and although they are independent, they can be coupled together to maximize benefits.

Virtualization techniques, the value of which has already been proven in the IT world, are now moving to the networking arena to replicate the advantage of decoupling software from hardware. The hypervisor implements the virtualization layer that abstracts the application from the infrastructure which is viewed as a pool of compute, network and storage resources. SDN separates the control and data plane, centralizing the network intelligence in a controller that manages white box switches implementing the forwarding function. Network administrators are no longer required to program thousands of devices and can remotely deploy network-wide policies down to the user level within an open software framework that leaves manufacturers' dependency behind.

These two new concepts also have a great impact on traditional development models since hardware and software life cycles no longer need to be tied together. Software is more flexible and less restrictive;

changes are easier to manage and development cycles are shorter. However, software performance still relies on the underlying hardware. The requirements for platforms running high-end network applications that have to process millions of packets per second are much different than those for hardware running enterprise IT software. At the high-end, equipment manufacturers need to find the right balance between standard server architectures and workload optimized platforms to achieve programmable flexibility without compromising performance. Support for industry leading SDN/NFV standards bodies is a requirement for systems to fulfil the OPEX savings promised by this new networking paradigm.

Security and the Need for Speed

When building the new infrastructure, two aspects are still a major concern: network security and quality of service. The applications that work to guarantee them sit in-line with traffic to capture and inspect 100% of the packets that cross the network with a twofold objective: detect security threats early and gather comprehensive network insights. In this scenario, the ability to classify and direct network flows efficiently becomes critical. Relegating these responsibilities to CPU cores will degrade throughput and consume processing cycles that could be put to better use. At network speeds of multiple 10Gbps, 40Gbps and

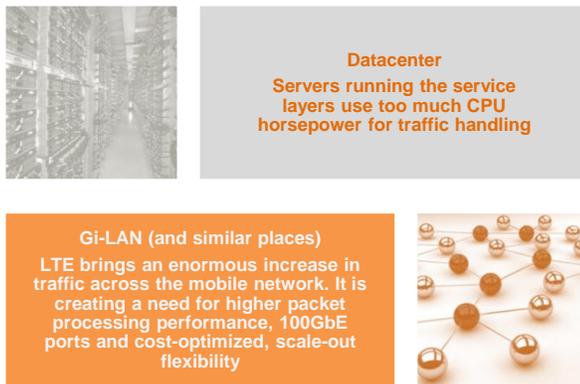


Figure 3. Hot Spots in the Network

100Gbps, these bottlenecks become exaggerated and can cripple a system that has not been designed to accommodate high traffic and processing performance. At the same time, trends in consumer behaviour and devices lead to an increased number of distinct flows. Today's smartphones, phablets and ultra-convertibles run an increasing number of apps where each app may create multiple control plane and data plane sessions, each representing a distinct network flow. Other trends such as IoT with billions of connected devices, each pushing packets onto the network infrastructure, will lead to a further increase of networks flows. A 100GbE network pipe can easily carry millions of flows in a real world scenario.

These two effects, increasing network throughput and a vastly increased traffic complexity, represented by the growing number of flows, impose major challenges to any network infrastructure.

Network operators have been deploying dedicated platforms from different vendors to address these requirements. Telecom and enterprise network operators now look for higher-performance, scale-out architectures to optimize and consolidate added value services such as policy enforcement, network analytics, intrusion prevention, firewalls or WAN optimization where complexity rapidly increases in terms of intelligence and throughput.

Platform Evolution

Performance Bottlenecks

The network transformation drives the evolution of the infrastructure to support next-generation services and speeds. Network computing is mainly performed by Intel® processors thanks to their high performance within a standard architecture enhanced by Intel's heavy investment in network specific features. Technologies such as Intel® QuickAssist or Data Plane Development Kit (DPDK) offload cryptographic and data compression workloads, accelerate packet processing, and provide advanced support to virtualized

environments. However, when this CPU horsepower is integrated into multi-processor platforms targeting hundreds of gigabits per second of I/O bandwidth several performance bottlenecks appear:

1. **Delivery of packets from the physical network interface card to the designated virtual function:** once a packet arrives at the network port it has to be processed and analysed so as to deliver its payload to the right core or virtual machine. In traditional architectures, sets of I/O ports are connected to a specific CPU socket through PCIe. The process of classifying packets and delivering them to the right destination not only involves protocol computation and packet movement that consume CPU cycles but also intensifies out-of-cache memory accesses and QPI traffic between sockets when traffic needs to be processed on other than the local CPU socket. This results in reduced platform throughput, greater latency and reduced scalability.
2. **Moving traffic through different virtual machines:** the result of an operation performed on a packet usually requires further packet switching between different threads, cores or sockets. This switching grows exponentially in a virtualized environment. The widely adopted Open vSwitch helps perform L2-L4 switching by computing resources within the platform but its computational cost in terms of CPU core cycles turns it into a bottleneck itself.
3. **System overhead:** a standard operating system stack running in a virtual machine (or in a physical appliance) typically limits packet processing performance due to poor scalability over multiple cores, significant overhead and interrupt latency.

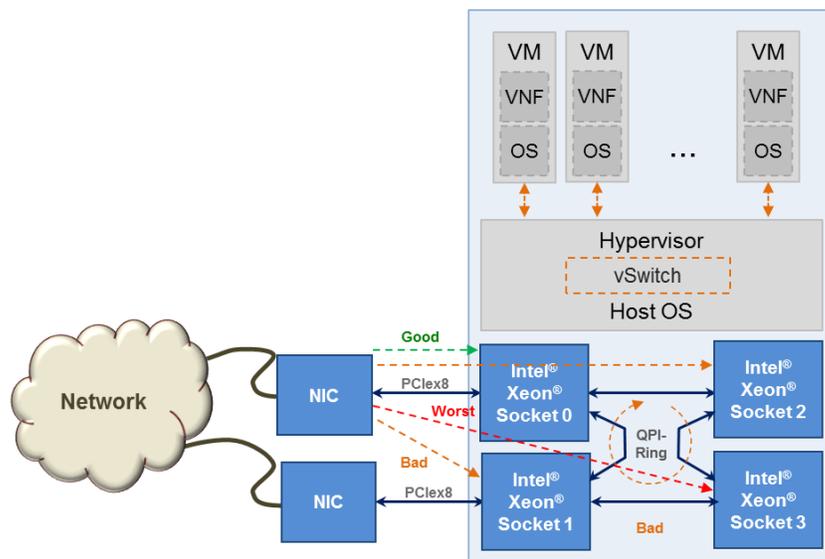


Figure 4. Whitebox System Architecture Limitations

100 Gigabit Ethernet Traffic Handling

In standard server hardware architecture, CPU cycles are used to perform networking as well as heavy lifting packet processing and classification tasks. The architectural bottlenecks described above can reduce application throughput and also limit overall platform throughput. To compensate these inefficiencies when dealing with millions of flows embedded in 100GbE pipes, additional platforms and extra load balancers need to be provisioned to deliver the same service, which limits scalability and has a greater impact on the total cost of ownership.

To achieve long term benefits of an efficient 100GbE infrastructure, a new breed of platform is required. The concept of an ultra high-end network appliance is to put an additional layer of intelligence in front of the Intel® Xeon® processors. This smart component balances the load between the network ports and the

applications by performing tasks such as packet classification, filtering and switching, freeing up Intel® Xeon® processor cycles to do the real application processing. Most importantly, it allows packets to be delivered to the target CPU socket, cores, virtual cores and even threads while avoiding cumbersome memory and QPI bottlenecks to make best use of the processing power of the Intel® Xeon® processors.

By presenting what appears to be a Network Interface Controller (NIC) to each CPU socket this approach boosts application performance without breaking the current software model. This consistency in programming model presents a unique approach that allows for a seamless migration from existing platforms with reduced R&D investment, faster time to market and lower software costs. The new ultra high-end appliance improves efficiency and solves the performance scalability problem of data plane intensive applications migrating to 100GbE interfaces at lower total cost of ownership.

The Ultra High-End Network Application Platform



Figure 5. Advantech FWA-6522C

The FWA-6522C is Advantech's flagship network appliance to boost performance of high demanding packet processing applications looking to scale to hundreds of gigabits per second of throughput in a reduced physical footprint. The system integrates four CPUs based on the Intel® Xeon® Processor E5-4600 v3 series with 32 DIMMs and up to 32 PCIe lines per socket. It provides up to 400 Gbps of network I/O which is configurable through four PHY Mezzanine Modules. Functions like packet classification, load balancing and soft switching

can be offloaded to a Netronome FlowNIC that is connected to each CPU socket via PCIe x8 gen.3 ports. While offering performance levels that previously required more complex and costly architectures, Advantech's FWA-6522C seamlessly integrates with standard software frameworks such as Linux, DPDK and Open vSwitch and is compatible with industry standard OpenFlow and Open Daylight controllers and OpenStack orchestrators.

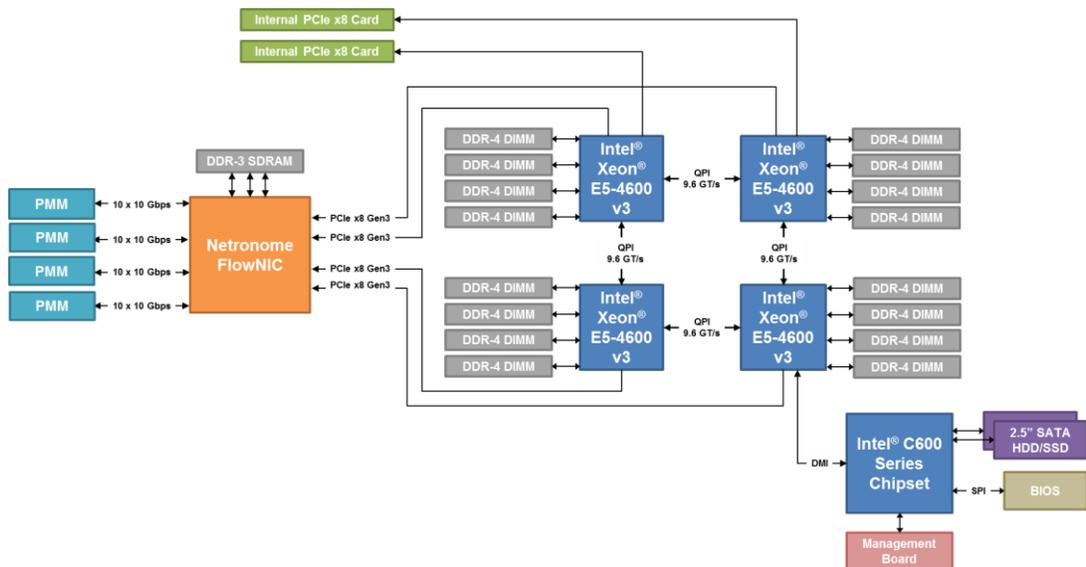


Figure 6. FWA-6522C Block Diagram

Innovative Two Piece Motherboard Design

The FWA-6522C integrates four sockets for the Intel® Xeon® Processor E5-4600 v3 series into a compact 2RU footprint using an exclusive internal architecture based on a stacked motherboard configuration. The Intel® Xeon® Processor E5-4600 v3 family is the next generation of server processor built on 22-nm process technology. It features two Intel® QuickPath Interconnect point-to-point links capable of up to 9.6 GT/s.

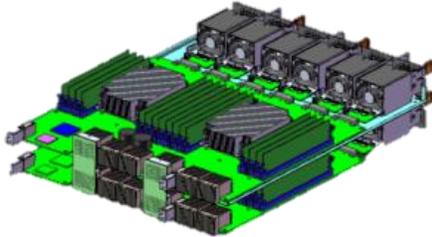


Figure 7. FWA-6522C Dual Motherboard

The unique FWA-6522C dual motherboard design considerably improves density, cooling and reliability when compared to other quad socket platforms. Traditional one piece motherboards are limited by both mechanical and thermal constraints. The FWA-6522C motherboard allows for richer PCIe connectivity without sacrificing signal integrity and thermal performance. Its 32 PCIe 3.0 lanes per CPU socket and full 8Gbps per lane support a higher throughput I/O subsystem in a compact 2U mechanical design.

The FWA-6522C's innovative motherboard design also improves thermal performance by uncoupling the thermal influence between CPUs. In most server designs, the second CPU and DIMMs are shaded in airflow direction and are therefore exposed to pre-heated air. The FWA-6522C avoids the "shaded core" problem by placing its four CPUs side by side and at different levels. This advanced thermal design supports higher dense subsystems which are less difficult to cool, allowing for lower fan speeds that provide best in class robustness at minimum acoustic noise.

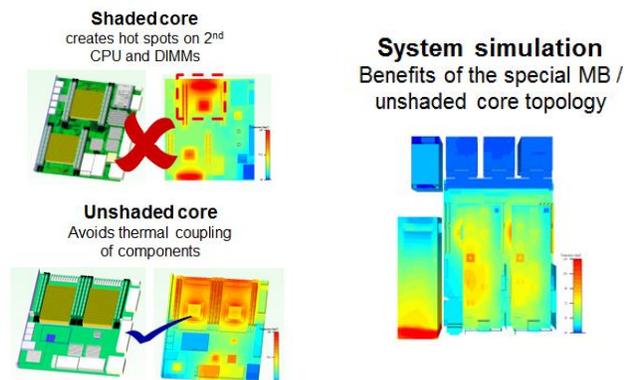
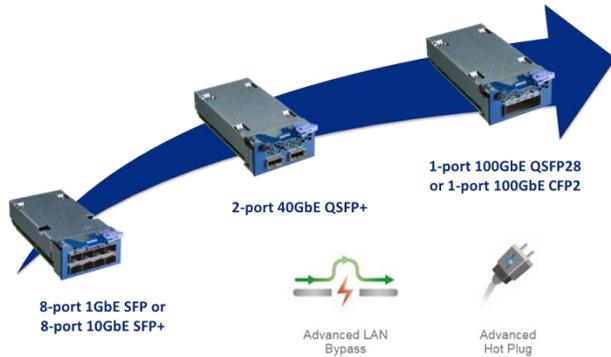


Figure 8. FWA-6522C Advanced Thermal Design

The FWA-6522C supports up to 32 DIMMs, 8 per CPU socket, for increased memory density which become critical when working with large number of virtual machines. Configurations up to 1TB RAM using 32GB LRDIMMs offer greater flexibility and granularity to network equipment providers looking to virtualize their network infrastructure. Extra offload capabilities can be added through two internal PCIe x8 slots that can host accelerators such as Advantech PCIE-3215 dual Intel® QuickAssist card for common security and crypto offload. Two hot-swappable SATA disks or SSDs (2.5") can be mounted at the front of the system.

Modular I/O for Various Deployment Scenarios

FWA-6522C I/O is implemented through four front-loading, field replaceable PHY Mezzanine Modules (PMMs) that provide greater modularity and flexibility with various network interface options ranging from Gigabit Ethernet to 100GbE. Leveraging Advantech PMMs, network application developers can scale to address different deployment scenarios within just one platform.



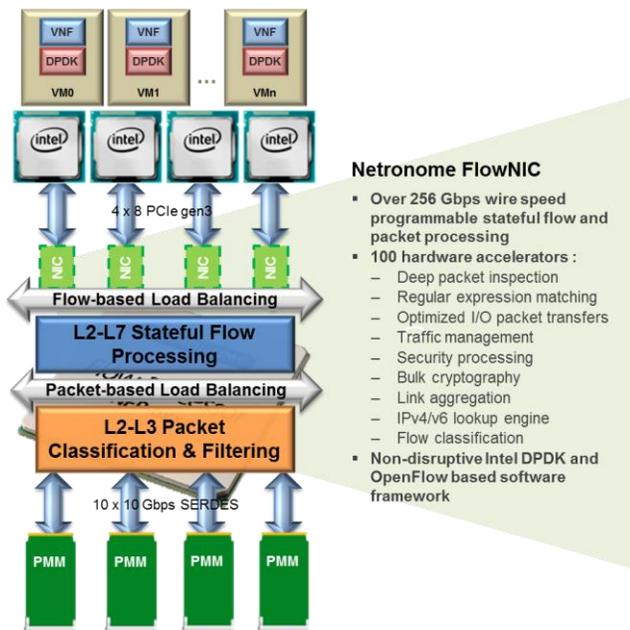
The PMM is a future-proof standard Advantech product, re-usable across different product lines, that provides enhanced features for high-end networking applications and can be customized to meet specific customer requirements. PMMs are 25 Gbps SERDES ready and support the latest I/O standards such as Intel® Silicon Photonics (SiPh).

PHY Mezzanine Modules integrate Advantech's advanced LAN bypass and hot-plug technologies. Advantech's

advanced LAN bypass provides a userspace API that makes it accessible to multiple applications running as different instances on the same OS or on distinct bare metal as well as virtualized environments. Advanced LAN bypass segments can be controlled independently over a PCIe interface avoiding software dependencies and reducing latency. Advantech's advanced hot-plug feature is controlled through an application-aware extension that manages the replacement process of failing IO modules. PMMs can be removed without shutting down the OS or the application. Unlike kernel based hot plug support, Advanced Hot Plug notifies the application layer before network ports disappear and allows healthy application software to perform its housekeeping such as cleaning state tables or re-routing traffic to other interfaces.

State-Aware Flow Processing at Over 200 Gbps

The advanced integration of compute and FlowNIC based I/O, able to handle millions of flows, is the cornerstone of the FWA-6522C's innovative architecture. This is achieved by offloading load balancing and packet processing tasks to a Netronome FlowNIC that is placed between the PMMs and the CPU sockets.



The FlowNIC circumvents bottlenecks by delivering flows to the targeted Intel® Xeon® processor sockets, cores and threads, which becomes an essential requirement in high-end 100GbE based platforms. It thereby also reduces Intel® Xeon® processors utilization and relieves them to run application software.

The FlowNIC features more than 200 individual processing cores and dedicated hardware accelerators and it is produced using Intel® 22nm process technology. The FlowNIC maps the Intel® Xeon® processors to the PMMs over a high-speed PCIe Gen 3.0 data path based on a standard NIC programming model that supports an Intel® DPDK API. Stateful packet and flow processing is transparent to the application software developer and also accessible through APIs such as an OVS-DB plugin that control wildcard match and exact-match flow state tables that support for over 100 million rules and flows.

Figure 9. FWA-6522C Flow Processing

This granular flow processing also enables application-aware hash-based and dynamic load balancing capabilities that can forward flows to specific sockets, cores or threads, eliminating performance bottlenecks caused by extra packet moves between CPU sockets. Beyond internal flow classification and load balancing, the FlowNIC can also perform traffic forwarding, filtering and balance traffic to external compute nodes. These functions can be implemented in scenarios such as WAN optimization to balance qualified traffic to specific processing appliances for greater scalability of the solution.

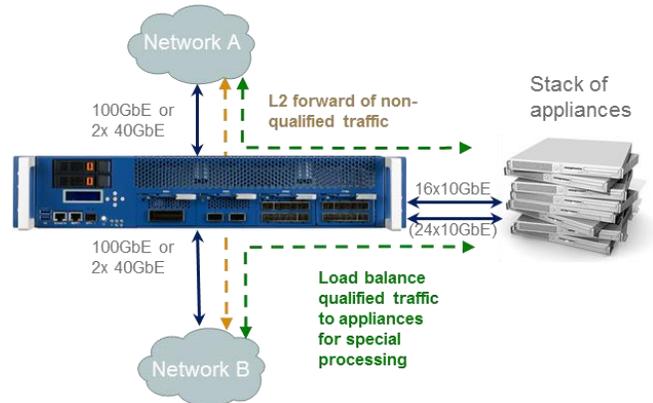


Figure 10. Intelligent Traffic Handling Application

Accelerated vSwitch

Virtual environments have widely adopted Open vSwitch (OVS) to complete the built-in L2/L3 switch on Linux-based hypervisors and provide the advanced multilayer functionality required in multi-server virtualization environments. The OVS switches traffic between virtual machines (VMs), making use of compute resources previously reserved for host applications which degrades overall system performance. The FWA-6522C overcomes this limitation by offloading the OVS datapath forwarding tasks to the integrated FlowNIC, enabling the new virtualized architecture at minimum impact on the Intel® Xeon® processors utilization.

The accelerated OVS datapath supports all match and action fields used in OVS 2.3 and leverages FlowNIC exact match flow classifier capabilities for stateful packet processing. Control tasks are handled by the Intel® CPUs and remain the same to that of the non-accelerated, open-source version making the solution fully compatible with OpenFlow. With over 200 cores and 960 threads, the highly parallel nature of the FlowNIC scales OVS flow forwarding throughput to 200 Gbps and over 15M connections per second without performance degradation.

Carrier-Grade Platform Management

All design aspects of the FWA-6522C platform has been tailored to provide advanced reliability and serviceability to high-end networking applications. The FWA-6522C integrates Advantech's Advanced Platform Management which provides all required IPMI v2.0 Baseboard Management Controller (BMC) functionality and also additional features that allow local and remote users to early detect system degradation, avoid system downtime and shorten mean time to repair.

Advanced IPMI and BIOS features include:

- Fully redundant firmware images not only for BMC firmware but also for BIOS.
- Out-of-band management interface over LAN via shared management Ethernet port that provides access to the BMC and enables remote BIOS and firmware updates and configuration. The system's console can also be remotely accessed using Serial-over-LAN. Both LAN interfaces are encrypted using IPMI defined RMCP+ protocol for enhanced security.
- Fail-safe HPM.1 updates of all components that support automatic failover in case of startup problems, with BIOS POST sensors logging the POST code where the board hangs, and automatic rollback in case of update failures or integrity issues.



The FWA-6522C provides a comprehensive set of front panel features including 2 x RJ45 GbE management ports, 1 x RJ45 RS232 console, 1 x USB, 1 x power button, 3 x status LED (power LED, status LED, alert LED) and a graphic LCM with 5 buttons. The six hot-swappable cooling fan modules in the rear are distributed between the top and bottom motherboards and can be extracted independently from the motherboard. The platform is available with a choice of redundant 1400W AC or DC Power Supply Units (PSU). Field Replaceable Units include the PHY Mezzanine Modules, fan modules and PSUs.

Use Cases

Enterprise Security

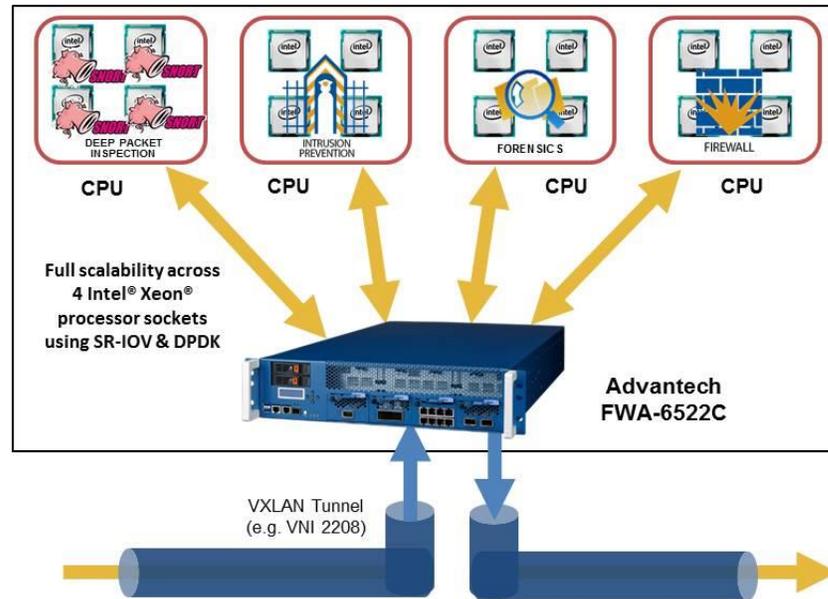
Network security evolves as rapidly as new threats spread. Security applications protect services and users in a variety of network deployments with different architectures. However, they all share a basic requirement to be effective: complete visibility and control over the traffic crossing the network. Applications such as intrusion detection and prevention, SSL inspection, Unified Threat Management or next-generation firewalls need to capture 100% of the traffic across all packet sizes and cannot risk any portion of data.

Cybersecurity applications perform forensic analysis and extensive pattern matching functions on all traffic requiring a high amount of compute, extensive memory capacity and a high performance IO subsystem.

Security applications strongly rely on Deep Packet Inspection (DPI) techniques to accurately classify network traffic. Traditional DPI stopped at the application identification but latest application-aware solutions can classify both enterprise and consumer applications and protocols, and extract valuable insights up to Layer 7. The amount of processing power required by these applications grows exponentially with increasing line rates.

The FWA-6522C is ideal for network equipment manufacturers looking to scale their x86-based security applications throughput to over 100 Gbps without breaking the NIC based programming model. The unique architecture of the FWA-6522C unleashes the full performance of its four Intel® Xeon® processors by

avoiding performance bottlenecks and by providing state aware flow pinning via a Netronome FlowNIC. The integration of high computing capabilities, support for up to 1TB of memory and its high performance, smart networking interfaces makes the FWA-6522C the first choice for the most demanding security applications.



Enhanced Subscriber Experience

With the high market penetration of internet-capable mobile devices, improving subscriber experience is viewed by mobile service providers as an opportunity to differentiate their mobile offering and generate alternative revenue streams. Understanding user behaviour is key to providing tailored added-value services and requires complete application visibility.

L7 network analytics is the basis of the engines that provide these advanced network insights and that work to maintain network health through capabilities such as congestion prevention, bandwidth management and content optimization. These are business critical applications that require real-time network visibility to gather comprehensive application information and early detect network performance threats.

The main requirement around which underlying platforms are built is lane-rate packet capture and delivery to the host. Advantech's FWA-6522C provides deep visibility into network flows at higher throughputs. Advantech's ultra high-end network platform offloads flow-based forwarding and alleviates the Intel® Xeon® processors from compute intensive L2-L7 packet processing tasks such as packet pre-classification and filtering and per-flow action processing. The FWA-6522C flow-aware dynamic load balancer creates a stateful high-speed datapath over the 32 PCIe gen3 lanes that reduces latency of feature-rich network analytics applications whose intelligent nature intensifies I/O to CPU traffic.

Network Functions Virtualization (NFV)

The goal of Network Functions Virtualization is to replace costly purpose-built equipment with more flexible, programmable x86-based systems. The FWA 6522C's quad processors make it an excellent scale-up platform to run CPU hungry applications, and even run multiple Virtual Network Functions (VNFs) in a single platform. Firewall, load balancing, virtual routing, gateway and Evolved Packet Core (EPC) applications are excellent candidates for the FWA-6522C in both service provider and data center environments. NFV makes

extensive use of virtual switching. As mentioned previously, offloading OVS via the FlowNIC yields excellent throughput at all data rates and returns valuable CPU cycles to run virtual network functions.

Conclusion

With the arrival of 100 GbE interfaces, commodity x86 servers and legacy network appliances struggle to run compute and I/O intensive applications that have to deal with millions of flows. Complex packet processing tasks handled by the Intel® Xeon® processors consume valuable CPU resources creating bottlenecks that can degrade overall application performance. To efficiently scale such network applications a new breed of system is required. Advantech FWA-6522C ultra high-end network appliance provides a fast migration path to x86 application developers looking to upgrade their solutions to hundreds of gigabits per second of throughput. The FWA-6522C integrates four of the latest generation Intel® Xeon® processors and 400 Gbps or modular I/O within innovative architecture that provides an extra layer of intelligence between the network interface and the CPU sockets. High-end network applications can make full use of the Intel® Xeon® processors by leveraging the built-in FlowNIC to easily offload stateful packet processing and application-aware load balancing via APIs and without breaking the existing NIC programming model which minimizes development efforts and reduces time to market. The FWA-6522C is ready to be deployed in virtualized environments offering competitive advantages through capabilities such as Open vSwitch acceleration. Mission-critical applications can rely on the carrier-grade features of the FWA-6522C whose superior performance, scalability and efficiency take high-end networking to the next level.

Advantech Contact Information

Hotline Europe: 00-800-248-080 | Hotline USA: 1-800-866-6008

Email: NCG@advantech.com

Regional phone numbers can be found on our website at <http://www.advantech.com/contact/>

www.advantech.com/nc

www.netronome.com