seaway NETWORKS INC

Hifn Intelligent Secure Networking

Multi-Gbps Multi-Function Security Gateway System Design

High-Level Application Note

Abstract

The growing list of security devices deployed in enterprises and data centers has created the need for a more integrated network security solution. Network administrators are currently grappling with managing security policies across a number of network devices including firewalls, IPSec and SSL Virtual Private Networks (VPNs), Network Intrusion Detection Systems (NIDS), along with an emerging set of point equipment providing such functions as anti-virus and content filtering.

Combining all the security gateway functions into one device ensures a consistent application of policy and reduces the likelihood of provisioning errors creating the very security gaps that the security gateway is supposed to remove.

Designing a multi-function security gateway is not only about how best to support each individual function, but also how well the dfifferent subsystems interact to create a coherent and unified solution.

A security gateway based on Seaway and Hifn technology can support a multitude of security functions at multi-gigabit rates.

Document SDN 00025
Version 1.0
28 March 2003

For complete Hifn disclaimer information, including but not limited to, licensing, patents, trademarks, and exporting, see HIPP / 8xxx documentation. Hifn reserves the right to make changes to its products, including the contents of this document, or to discontinue any product or service without notice. Hifn advises its customers to obtain the latest version of relevant information to verify, before placing orders, that information being relied upon is current. Every effort has been made to keep the information in this document current and accurate as of the date of this document's publication or revision.

Hifn warrants performance of its products to the specifications applicable at the time of sale in accordance with Hifn's standard warranty or the warranty provisions specified in any applicable license. Testing and other quality control techniques are utilized to the extent Hifn deems necessary to support such warranty. Specific testing of all parameters, with the exception of those mandated by government requirements, of each product is not necessarily performed.

Certain applications using Hifn products may involve potential risks of death, personal injury, or severe property or environmental damage ("Critical Applications"). Hifn products are not designed, intended, authorized, or warranted to be suitable for use in life saving, or life support applications, devices or systems or other critical applications. Inclusion of Hifn products in such critical applications is understood to be fully at the risk of the customer. Questions concerning potential risk applications should be directed to Hifn through a local sales office.

Hifn does not warrant that its products are free from infringement of any patents, copyrights or other proprietary rights of third parties. In no event shall Hifn be liable for any special, incidental or consequential damages arising from infringement or alleged infringement of any patents, copyrights or other third party intellectual property rights.

1. Introduction

Just as firewalls and VPN gateways have combined to form security gateways, the emergence of network intrusion detection systems and other network security point products has created the need for a combined multi-function security gateway appliance as a single managed security entity. According to a recent study by Gartner, "an integrated network security platform approach will increase network security and reduce the cost of ownership for perimeter security, while preserving best-of-breed options."[1]

In order to do so, however, newer and more advanced technologies are required to enable wire-rate processing across a wide range of functions and layers. While it is generally accepted that a software solution on general-purpose processors cannot scale to process these functions at wire-rate, the conundrum is that a software solution on general-purpose processors is the most practical way by which these security applications can be developed. Many of these security functions require application-layer processing on the content of the packets, and the very nature of application-layer software development is characterized by relatively large code size with a high need for portability and flexibility. A solution that combines the flexibility of software on general-purpose processors with the performance of an ASIC is therefore desired.

This paper presents a design of a combined hardware/software multi-gigabit, multi-function security gateway using Seaway Network's SW5000 Network Content Processor and Hifn's 8154 crypto-processor, together with general-purpose processors.

2. System Design Considerations

Designing a multi-function device is challenging, not only because of the inherent issues related to system interactions, but also because the multi-function device is expected to be at least equivalent in function and performance to each individual best-in-class single-function device. The multi-function security gateway, for example, requires performance comparable to the performance of a collection of distinct security devices, while handling such disparate tasks ranging from IP forwarding to software-intensive content layer processing. In reality, however, the benefit due to increased system efficiency and tighter integration can often help mitigate what can truly be overwhelming performance requirements. With separate devices, the protocol stack must be traversed for each device separately, while a highly integrated multi-function device, when designed properly, needs to traverse the protocol stack exactly once.

The multi-function combination of a firewall, IPSec and SSL VPN, content filtering, gateway anti-virus, and intrusion detection and prevention requires the following in a single device:

- Basic layer 2 and layer 3 such as VLANs, IP forwarding, packet filtering
- Bulk crypto-processing for IPSec and SSL VPNs
- Public key processing for IPSec and SSL VPNs
- TCP byte stream reconstruction for intrusion detection and prevention, content filtering, and anti-virus, and full termination for SSL VPNs
- Content searching for intrusion detection and prevention, content filtering, and anti-virus
- Content processing for intrusion detection and prevention, content filtering, and anti-virus

[1] Network Security Platforms Will Transform Security Markets, Research Note, November 7 2002

- Control plane for all of the above

This range of tasks can be offloaded by hardware in varying degrees, from lower-level processing that lends itself well to hardware assists, to content processing that requires significant software involvement. What is not always obvious is the implicit requirement of how these different blocks communicate and pass data to each other.

Figure 1 shows these blocks and the entangled array of communication paths that they require.

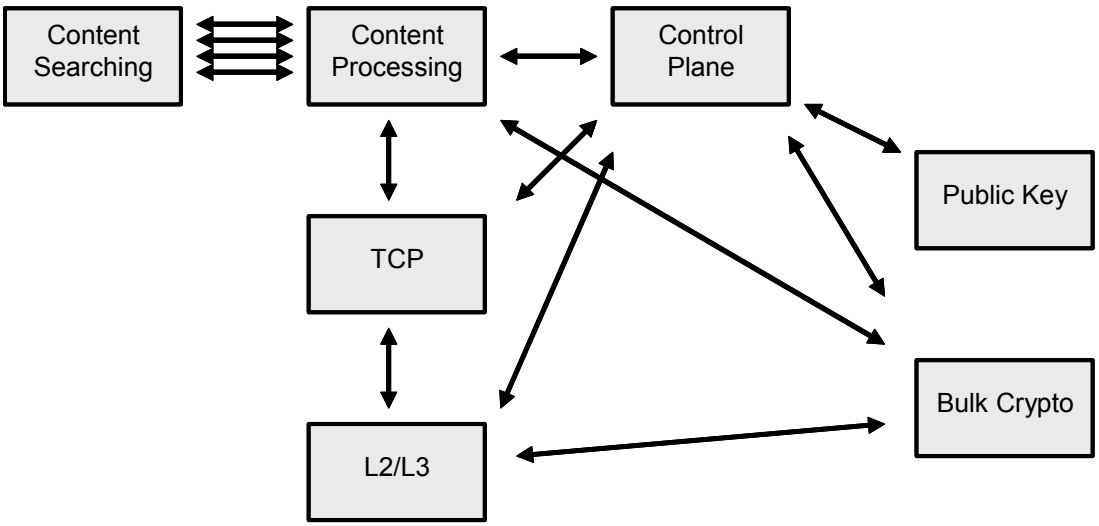

Figure 1: Processing Interactions in a Multi-Function Security Gateway

Unlike single-function devices that may have a single datapath, the multi-function device must handle multiple datapaths. The challenge then becomes how to create an architecture that moves data between these blocks efficiently and at multi-gigabit rates.

2.1 Data Movement and Switching

Frequently, the bottleneck in traditional systems turns out to be caused not by processing limitations, but by memory bus and I/O bus bandwidth constraints. In relatively simpler systems where a limited amount of processing is required on each packet, a single high-speed bus interconnecting processing elements and memory may be sufficient. As the processing and memory requirements increase, however, such as when multiple processing elements need to act on a single packet, changing over to a switched architecture can help reduce bus contention and improve efficiency.

Key to the efficacy of handling multiple functions is the ability to switch data amongst the different processing elements. Sometimes a packet requires IPSec processing, another packet may require TCP termination, yet others may require gateway anti-virus scanning, and others still may require all these functions. Because each network administrator may choose to enforce different policies, the behaviour of a packet as it enters the system cannot be hardcoded in the design of the appliance, nor can one function be optimized at the expense of others. It is important for the system architecture to have the flexibility to handle different situations, where a single packet may need to be processed once, or numerous times by different elements, as dictated by packet contents and policy.

Although a switched architecture may be unjustified in simpler devices, it is highly desirable in more complex systems. In contrast, a shared bus architecture is usually adequate in simpler devices, but has well-known drawbacks in multi-processing systems:

- 'Real-life' performance is difficult to predict, as a single packet may traverse the shared bus for an unknown number of times, potentially affecting other flows on the shared bus.
- The onus of traffic management is shifted to each element connected to the shared bus, making each element more complex, and the overall system behaviour possibly less consistent.

A switched architecture solves these problems by enabling a point-to-point connection to each element, and by providing traffic management functionality within the switch. **Figure 2** shows a conceptual sequence through a switched system for an IPSec packet that requires subsequent processing of the content.

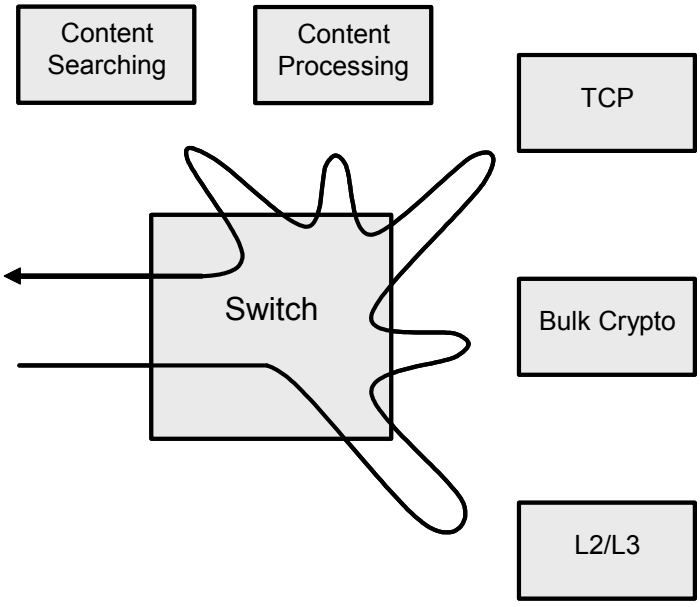

Figure 2: Conceptual Processing Sequence using a Switch

First, Ethernet, IP, and IPSec header processing are required to properly classify the packet. Next, the packet is decrypted, and the TCP byte stream recovered. Processing of the content can now begin and usually entails searching the byte stream for specific signatures or strings.

Although this is a conceptual diagram, it nevertheless illustrates how a switch can help coordinate the flow of data in a multi-function and multi-layer processing system.

2.2 Dynamic Pipelining

The multi-layered functionality of a security appliance also lends itself well to a pipelined architecture where data is passed through different hardware elements as the stack is traversed. The pipeline cannot be static, however, due to the diverse set of functions and layers required. For example, as described previously, content processing of data within an IPSec tunnel first requires decryption of the IP packet, then TCP byte stream recovery. An SSL VPN, however, would require TCP termination first before decryption, while a plain firewall application may not require TCP termination or crypto-processing at all.

While this sounds simple enough, a pipeline that has a multitude of paths that depend on a multitude of external factors is difficult to design without a switch. A well-designed switched architecture can readily accommodate the dynamic pipelining required for multi-function processing.

2.3 Efficient Inter-Layer Handoff

Sometimes, inefficiencies can arise when handing off data from layer to layer. For example, SSL VPN processing requires the full SSL record to be received before decryption. If spread across multiple TCP segments, this involves accumulating enough segments to complete the record before proceeding with SSL processing. Conceptually, an intermediate stage is required that recognizes situations where TCP processing is complete but SSL processing cannot yet begin. The ability to delay scheduling the recovered TCP byte stream and allow the data to accumulate automatically to the full SSL record before handing off to the SSL layer is desirable.

2.4 Offload Engines

Only after the data movement problem is resolved can a meaningful examination of offload engines be made. Offload engines can perform in hardware, the repetitive, processing-intensive tasks that would otherwise burden the general-purpose processor. Some of the more obvious offload functions required include TCP offload and content searching. In both these cases, a total system perspective should be taken since performance is dictated not only by how well each offload engine performs, but also by how well it interacts with other subsystems. For example, a large part of TCP processing 'overhead' is in the movement and copying of data after the TCP byte stream has been recovered, rather than the actual termination of the protocol itself. By examining TCP offload engines in isolation and not considering the underlying system, potential inefficiencies can be overlooked.

2.5 Crypto-Processors

Crypto-processors come with varying degrees of functionality, from pure algorithm accelerators that apply raw transforms individually, to protocol-aware crypto-engines that can interpret and manipulate (modify, insert, delete) packet headers and perform multiple transforms in a single pass. Systems relying on pure algorithm accelerators often require the external processor to perform all packet header manipulations, maintain all session contexts, and submit each packet to the accelerator multiple times, once for each transform.

2.6 New Protocols and Attacks

In the past, the ability for a system to accommodate future changes in the protocol represented an unnecessary luxury, but with emerging applications such as intrusion detection where new attacks must be continuously detected, the flexibility that software provides to detect and thwart future attacks is of paramount importance. Intrusion detection is truly an example where embedding protocol handling completely in hardware would be a poor design choice.

2.7 Software Portability

Lastly, the best hardware architecture is inadequate if existing software cannot be reused. The ability to allow software to be ported easily from other systems is beneficial, especially in an environment where vendors need to leverage their software investment in their other products. The simplest porting activity is from one general-purpose processor with a standard ISA and toolchain to another general-purpose processor with a standard ISA and toolchain. Using standard rather than proprietary development environments and tools reduces the R&D investment required to bring the system to market.

An added benefit is the ability to port software easily between different members of the same processor family, and take advantage of the subsequent rollout of newer and faster processors. By swapping out older processors with newer ones, performance can instantly be improved with relatively minor development effort.

3. Multi-Gigabit Multi-Function Security Gateway

Together with general-purpose processors, a Seaway and Hifn solution combines the SW5000 Network Content Processor (NCP) and Hifn's 8xxx series crypto-processors for a complete security solution. The SW5000 NCP contains an embedded switch that allows TCP byte streams and content to be efficiently maneuvered while seamlessly connecting to external processors and Hifn's 8xxx series crypto-processors.

Figure 3 shows a block diagram for a two-port appliance (extensible to multiple ports).

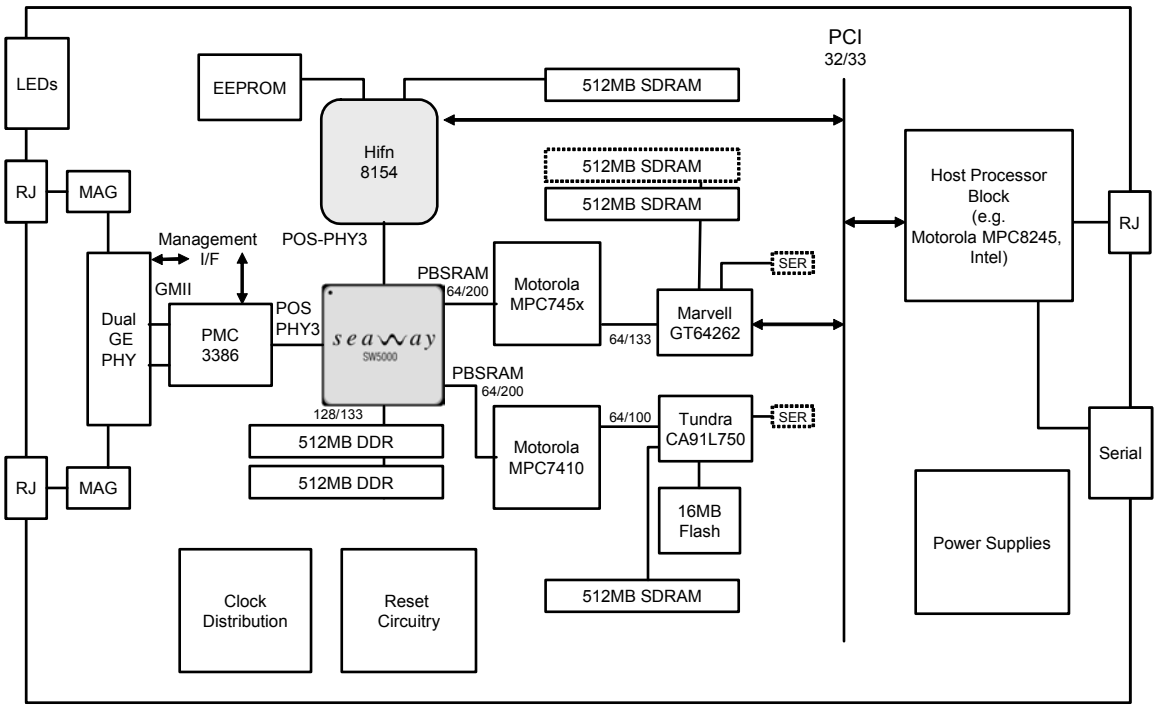

Figure 3: Dual-Port Multi-Function Security Gateway Block Diagram
(Firewall, IPSec and SSL VPN, Intrusion Detection and Prevention, Anti-Virus, Content Filtering)

The Hifn 8154, Motorola MPC745x and MPC7410 are adequate for a 2 Gbps full duplex appliance. All firmware on the MPC7410 is supplied by Seaway Networks. Application software resides on the MPC745x, with APIs to the SW5000 supplied by Seaway Networks, and APIs to the Hifn 8154 supplied by Hifn. The PMC3386 is a POS-PHY3 Dual GE MAC device. The Dual GE PHY can be the Marvell Alaska II 88E1020 or equivalent. The SW5000 is shown with 1GB of DDR SDRAM, which is sufficient to buffer close to one second of data at wire-rate. The SW5000 can support up to 4GB of memory. The MPC745x is shown with 512MB to 1GB of memory. The actual memory required is dependent on the application software running on the processor. The system controller shown for the MPC7410 is the Tundra PowerPro, and for the MPC745x, the Marvell GT series. All memory shown has ECC protection. The host processor block performs system housekeeping functions, and can be based on a basic integrated processor, such as the Motorola 8245 or an Intel equivalent.

This appliance can perform all security functions at multi-gigabit rates, including full TCP byte stream recovery and termination, IPSec and SSL crypto-processing and content searching.

Key components of this architecture are the SW5000, the Hifn 8154 crypto-processor, the Motorola PowerPC 7410 which is programmed for the packet processing functions, and the Motorola PowerPC 745x which is programmed for the content processing functions.

3.1 SW5000

The SW5000 provides a set of autonomous hardware assists that handle the parsing, validation, classification and movement of data. These are broken down into three distinct entities:

- Packet engines that parse, validate, and classify packet headers and flows; transform received packets into ordered byte streams; and create and format packets from byte streams for transmission
- A content engine that searches, modifies and replicates content
- An embedded fabric that switches data streams to and between the different externally-attached components

The SW5000 does not contain nor does it execute software. Instead, it works in conjunction with software on external processors to provide the system functions. Specifically, the packet engine works in conjunction with a packet control processor (shown as the MPC7410) and the content engine works in conjunction with a content control processor (shown as the MPC745x).

Packet Engine

At the packet level (TCP layer and below), the SW5000 packet engine performs full Ethernet/VLAN/PPP/IP/IPSec/TCP/UDP protocol header parsing and validation, full layer 2 classification including VLANs and Multi-Link Trunks (MLTs), IP fragment reassembly if necessary, IPSec classification, ACL lookups, 3-tuple and 5-tuple flow classification, and Ethernet/IP/TCP/UDP CRC and checksum validation/calculation, along with a number of packet-formatting assists.

On ingress, packets for each flow are ‘converted’ into individual byte streams, while on egress, byte streams are ‘converted’ back into packets for transmission. Conversion on ingress entails the processing of and removal of packet headers, and the subsequent enqueuing of the contents of the packet onto a stream queue. Conversion on egress entails the dequeuing of the contents from a stream queue and the addition of packet headers. By decoupling the application contents from artificial packet delimiters, the content processing layer is no longer encumbered by packet-by-packet processing, and can process data in the format used by the application itself.

The packet engine is not limited to data plane processing, and offloads the control plane by examining and mapping the TCP control bits to different codepoints so that the correct software can be executed. The codepoint, context, and parsed data are placed into a memory-mapped digest along with a copy of the first block of the packet. This digest contains enough information for software on the MPC7410 to subsequently execute on the correct thread with the correct context at the correct code entry point. The actual packet is stored in the SW5000 local DDR memory, and does not need to be moved across to the MPC7410 local memory.

Additionally, the packet engine has the capability of recognizing new connection requests and creating hardware table entries autonomously. For example the SW5000 can be instructed by software on the MPC7410 to autonomously create a table entry for any new TCP connection request to a specified IP address (exact, mask or wildcard) over a specific port (e.g. port 80). When a connection request is received that matches the set criteria, the SW5000 autonomously creates a table entry for this connection and composes a digest of relevant information so that the correct thread and context can be executed on the MPC7410.

Content Engine

At the content level (layers 5 through 7), the SW5000 content engine provides assists for searching, modifying and replicating data streams. By operating with byte streams rather than packets, the SW5000 can be used to detect security attack signatures that cross packet boundaries just as easily as if the signature is contained within a single packet. The content engine can also be programmed by the software on the MPC745x to accumulate a specified minimum amount of data before passing the data on for content processing. This can be done on a per-stream basis.

The Content Engine contains an internal Dual-Ported RAM (DPRAM) that holds data currently being processed by the MPC745x. The MPC745x can make use of hardware assists to search, modify, or replicate content (for lawful intercept applications), all without having to copy the data across to its front-side memory.

Streamwise™ Stream Switch

At the core of the SW5000 is the fabric, or Stream Switch, that controls memory access, manages streams, and provides interconnection among the major functional blocks and external interfaces. It functions as a 20 Gbps shared-memory switch with a 128-bit wide, 133 MHz DDR SDRAM that provides 34 Gbps of raw bandwidth.

The Stream Switch has been specifically designed for content applications. Unlike traditional L2/L3 switch fabrics where the switching decision is made before data is enqueued to the destination, the Stream Switch allows the switching decision to be made before, during, or after the arrival of data. This is possible because the Stream Switch operates on stream queues rather than packets. Stream queues are switched rather than individual packets themselves.

This design is tailored to situations where the destination or consumer of the switched data changes frequently, as in the case of security processing where some upper layer content of the data stream must be examined before a switching decision can be made. By switching stream queues rather than packets, all existing data enqueued at the time the switching decision is made can be sent to the new destination immediately, instead of waiting for the enqueued data to be processed first.

This is useful, for example, in a security appliance performing URL filtering, where the URL in the data must be located and compared to a list of undesirable URLs before a forwarding decision can be made. Once the decision is made, however, all subsequent data for this flow can be sent to the new destination without having to examine the content further.

3.2 MPC7410 as the Packet Control Processor

The MPC7410 in this architecture provides protocol intelligence and flexibility at the packet layers (TCP and below) while using packet and session assists on the SW5000. All firmware on the MPC7410 is supplied by Seaway Networks.

On reception of a packet, the MPC7410 is notified after the SW5000 has validated and classified the packet and flow. The MPC7410 reads the digest provided by the SW5000 and invokes the appropriate software function based on the supplied function pointer using the appropriate software context. For transmission of a packet onto the link, software on the MPC7410 can make use of a number of packet formatting assists and checksum calculations on the SW5000.

With the SW5000 handling all the processing-intensive tasks, the MPC7410 can process TCP data and connection setups at multi-gigabit rates.

If crypto or content processing is required on the recovered byte stream, the MPC7410 invokes the SW5000 embedded switch to enqueue the byte stream onto the appropriate stream queue and schedule it for the content engine or the Hifn 8154 co-processor.

3.3 Hifn 8154

The Hifn 8154 network security processor handles all the cryptographic processing of IPSec and SSL traffic in this system. The 8154 receives entire IP packets (or SSL records) from the SW5000, applies one or more cryptographic transforms (compression, encryption, authentication) on each packet, and returns a properly formatted IP packet (or SSL record) to the SW5000. Its packet processing architecture goes beyond supporting simple algorithm acceleration, and enables payload extraction, compression, encryption, authentication, IP header manipulation, and packet assembly, all in a single pass at wire-speed. The capability to manipulate headers is useful, for example, in a VPN tunnel where an outer IP header needs to be inserted in front of the encrypted and authenticated original IP packet.

The 8154 connects to the SW5000 across a POS-PHY L3 packet data interface. With security associations stored locally, and with a separate PCI interface for configuration and non-wire speed functions such as session management, the full capacity of the POS-PHY L3 interface can be used for packet data. The 8154 can deliver a gigabit full-duplex performance on large packets.

The 8154 also supports packet compression using Hifn-patented LZS compression. Compression significantly reduces packet fragmentation across the network. Without compression, packets can grow beyond MTU limits when encryption is applied, causing downstream routers to split the packets in two, increasing the amount of traffic in the network. When compression is enabled in conjunction with encryption, there is no impact in network bandwidth consumption.

The 8154 contains a highly integrated public key (PK) sub-system with parallel PK processors. A dispatcher performs automatic load balancing between PK processors making the PK cores appear as a monolithic unit. Public key processing is required for IKE and SSL handshaking.

3.4 MPC745x as the Content Control Processor

The MPC745x is responsible for the actual processing of the content for the application. The MPC745x operates on data within the SW5000 DPRAM workspace directly, and uses the content searching, modification, and replication assists within the SW5000.

The MPC745x houses application software to perform intrusion detection and prevention, content filtering, and virus scanning. All the processing-intensive functions are taken care of by the SW5000.

3.5 Host Processor Block

The host processor is responsible for the system housekeeping functions and may take on some control plane functions as required. Because, in this architecture, the host processor is not performing any processing-intensive work, selection of this processor should be based on its cost and degree of integration rather than how much processing capability it possesses.

3.6 Instruction Set Architecture (ISA) and Toolchain

Software development on this system takes place on the MPC745x, with the SW5000 and Hifn 8154 appearing as devices. Firmware on the MPC7410 is supplied by Seaway Networks. All Seaway Networks software and APIs are written in ‘C’, and are C/C++ compatible. All Hifn software and APIs are written in ‘C’, and are C/C++ compatible. A standard ISA and toolchain can be used.

4. Key Observations

Multiple offload engines are packaged together. The SW5000 provides TCP offload, content searching, stream switching, as well as a number of assists at the packet layer and content layer. The Hifn 8154 provides both bulk encryption and public key processing for both IPSec and SSL.

Memory I/O is distributed. Packet buffer memory is separate and distinct from content processing memory. Packets are buffered in the local DDR SDRAM, while content processing operates on data within the SW5000 internal DPRAM. More specifically, packet buffers, packet and flow classification tables, ACLs and other hardware tables, are stored in the DDR memory. Software packet processing tables and context are stored in the packet processor's SDRAM. Current content is stored in the DPRAM. Software content processing tables and context are stored in the content processor's SDRAM.

At any given time, the packet processor can be processing data and looking up tables, the content processor can be processing content and looking up tables, and packets can be arriving from the link and processed by the SW5000 and stored in packet buffers, all using different memory buses.

No processor copying is required from the time the Ethernet frame enters the system, through IP/IPSec and TCP processing, through content processing including scanning the content, and finally out of the system as an Ethernet frame. The packet processor operates on data directly from SW5000 registers while the content processor operates on data within the internal dual-ported RAM directly, with neither processor needing to move data across to its front-side bus.

The SW5000 provides assists for all the processing-intensive tasks. All protocol responses are handled **under software control** by the packet control processor and the content control processor. The flexibility that software provides can allow, for example, new protocol-anomaly attacks to be addressed without invalidating any of the hardware.

This solution can **scale with the processor performance curve** of general-purpose processors. Performance can be improved simply by upgrading the processors or co-processors.

The **crypto-processing function can scale independently** from the rest of the system. For example, if it turns out that a large portion of the incoming traffic requires IPSec processing, then multiple Hifn 8xxx processors can be attached to the host/co-processor interface of the SW5000.

Software development on this system uses an **industry-standard ISA and toolchain**. No learning of proprietary systems is required. Both the SW5000 and the Hifn 8154 act as devices, and do not require a proprietary development environment.

5. Performance

The appliance as described supports the following performance capabilities:

- Firewall throughput of 2 Gbps full-duplex
- 66 million ACL comparisons per second
- IPSec VPN throughput of 1.2 Gbps full-duplex
- More than 500 IKE tunnels/second
- SSL throughput of 1.2 Gbps full-duplex
- More than 900 SSL sessions/second
- Up to 1 million SSL sessions or 2 million IPSec tunnels
- 1 million layer 4 sessions (per-flow queues)
- 200,000 TCP connection setups per second
- Content searching at 17 Gbps
- Total throughput of 2 Gbps full-duplex

Performance can be improved by increasing memory, by adding a second MPC7410 and/or additional Hifn 8xxx series processors, for example, to split out the public key functionality from bulk encryption.

6. Conclusions

A multi-gigabit multi-function security gateway appliance requires the performance of specialized hardware combined with the flexibility of software. A solution based on Seaway and Hifn technology can leverage general-purpose processors to provide multi-gigabit wire-rate processing for a number of functions including firewall, VPN, SSL VPN, intrusion detection and prevention, gateway anti-virus, and content filtering. Other security functions such as lawful intercept and proxy gateways are possible as well.

A number of aspects make the Seaway Networks SW5000 Network Content Processor and the Hifn 8154 crypto-processor well suited for multi-function applications:

- The SW5000 hardware assists offload the most processing-intensive tasks, while being sufficiently generic to be applicable across a multitude of applications and functions.
- The SW5000 contains an embedded switch fabric that enables efficient movement of data across processing elements.
- The SW5000 Streamwise™ Stream Switch supports a dynamic pipelining architecture that is well suited for multi-function, multi-layer processing.
- The Hifn 8xxx family of crypto-processors supports both IPSec and SSL, bulk processing as well as public key cryptography, in a wide range of performance options.

7. Contact

For more information on the Seaway Networks SW5000 Network Content Processor, please contact:

Seaway Networks Inc.
One Chrysalis Way
Ottawa, ON
Canada K2G 6P9

E-mail: info@seawaynetworks.com
Web: www.seawaynetworks.com
Phone: 613.723.9161
Fax: 613.723.8244

For more information on the Hifn family of crypto-processors, please contact:

Hifn
750 University Avenue
Los Gatos, CA
United States 95032

E-mail: info@hifn.com
Web: www.hifn.com
Phone: 408.399.3500
Fax: 408.399.3501