

Scaling Network Appliance Performance

Netronome Flow Manager and Netronome Flow Engine

Solution Overview

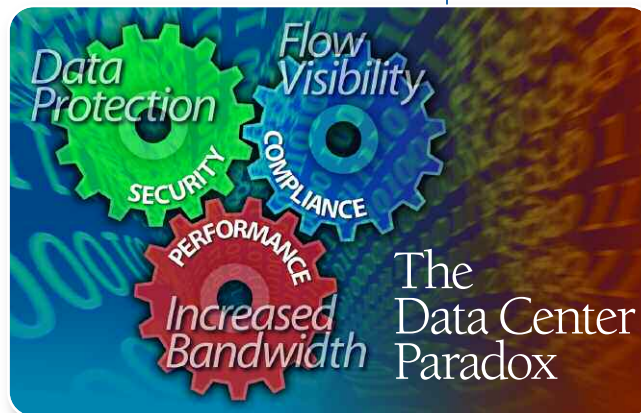
As IT professionals realize the vision of high-speed communications, they are tasked with building high-speed networks and data centers—based on standard packet technologies, such as Ethernet and IP—to provide a universal infrastructure for the rapid deployment of countless software-based applications and services. In a relatively short time, these enterprise networks have moved from 100Mbps to 1Gbps backbones. Just as quickly, many are beginning the move to 10Gbps, with designs for the next generation of Ethernet promising between 40Gbps to 100Gbps. This exponential increase in network performance is done with the expectation of providing the bandwidth required to support a rapidly expanding list of IP applications and services with stringent performance requirements.

These Gigabit and 10 Gigabit Ethernet networks and data centers expose the enterprise to millions of packets per second, comprised of numerous protocols, applications and services. These data rates make it difficult for enterprise IT organizations to guarantee network and application performance, as well as manage, secure and control the overall network usage to protect against threats and unacceptable practices.

COMPLIANCE, VISIBILITY AND CONTROL

Enterprise IT organizations require visibility into the network to maintain sensible levels of control. From outside of the enterprise, attackers can intercept data to steal or compromise information, or attack the enterprise with an arsenal of worms, viruses, spam and other malware. Within the enterprise, the IT staff must ensure that the network is being used for the applications critical to the business' success while simultaneously addressing concerns of accidental or intentional leakage of confidential information from within. Whether driven by corporate acceptable usage policies or government regulation, the control of network access and resources has become one of the most important aspects of IP network communications. IT organizations rely upon a host of network appliances to provide control and visibility into the network communication flows, as well as protection from internal and external threats. These network appliances are often deployed in enterprise data centers and at the LAN-WAN boundary, performing many functions like network firewall, intrusion detection and prevention (IDPS), unified threat management (UTM), lawful intercept, network access control (NAC) and a host of antigen functions to detect spam, viruses and other malware. This enables network administrators to provide a new level of compliance by verifying user access rights, scanning all data for known threats and protecting against confidential data leakage.

Even though the network security appliance market is complex and constantly changing, these devices share several common deployment models. They are either installed in-line, with respect to traffic, actively filtering flows to block attacks; or they



NETRONOME WHITE PAPER

EXECUTIVE OVERVIEW

As networks evolve to multi-Gigabit and 10-Gigabit Ethernet speeds, network appliances can create a significant performance bottleneck because they have failed to keep pace with ever-increasing bandwidth. Traditional server platforms provide poor packet capture performance, due to fundamental architectural issues in hardware and the operating system, making these platforms unsuitable when Gigabit levels of packet capture and application processing are required. There have been many attempts to scale the packet capture, classification and filtering performance of network appliances built on standard server platforms with some performance improvement. Even considering these techniques, these platforms provide actual throughput that is far below what is required to support high-throughput packet capture, classification and filtering applications. To overcome these network appliance performance

issues, Netronome has developed hardware and software components that allow both ISVs and end users to accelerate network packet capture and filtering applications. Netronome's solutions provide a combination of hardware acceleration with an open application programming interface (API)

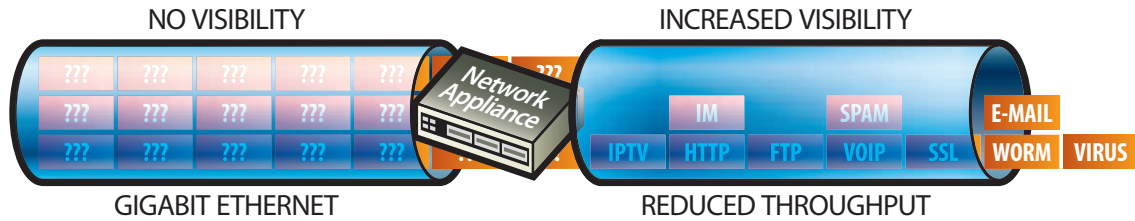
for packet classification, zero-copy drivers and libpcap load balancing.

NETRONOME™

Intelligent to the Core.™

netronome.com

Network appliances have failed to keep pace with improvements in network performance.



are deployed on span ports, taps or mirrored interfaces, only monitoring network activity “offline” to identify attacks or record traffic. Regardless of the deployment model, these devices typically experience little performance impact when used in 10/100Mb and underutilized Gigabit Ethernet networks. However, as networks have evolved to line-rate, multi-Gigabit and 10-Gigabit Ethernet speeds, network appliances can create a significant performance bottleneck because they have failed to keep pace with ever-increasing bandwidth. The network I/O, memory and CPU utilization of typical security network appliances fundamentally cannot support Gigabit performance levels. As a result, these security appliances are typically rated for use by some amount of aggregate bandwidth or a number of users, sessions and flows. When any of these are exceeded, the appliance becomes a bottleneck that can only be relieved by the addition of another network appliance(s).

NETWORK APPLIANCE ARCHITECTURE

Network security appliances are usually software applications running on top of standard Intel® Architecture (IA)/x86 servers. This allows Independent Software Vendors (ISVs) to develop and host their applications on PC server platforms, providing a low cost of entry and rapid development cycle with reasonable performance and a visible road map. It also allows vendors to standardize on open source operating systems, like one of the many Linux® variants available, and leverage open source packet capture (pcap) applications (like SNORT®, wireshark, tcpdump, and others) where possible. While the computing horsepower of these standard PC platforms will continue to increase dramatically according to Moore’s law, the network speeds feeding these devices are growing at similar rates. To further exacerbate this throughput scaling problem, the nature of the applications in question—such as IDPS, test and measurement, network forensics, billing and virus scanning—require extremely complex packet processing at Layers 4-7 of the data stream. These applications may require deep packet inspection and classification, packet filtering, forwarding, statistics gathering and recording to disk. This is in contrast to less-complex network functions, like traditional switching and routing applications that typically only look at L2-3 packet headers. Dealing with increasing network speeds for L4-7 applications, needs specialist, optimized, processing elements that are architected for network flow processing, supplemented by general-purpose IA.

The need to passively capture network traffic, as well as provide packet classification and filtering in-line, are fundamental requirements for any network security appliance. Traditional server platforms (even high-performance configurations with 4-8 CPU cores operating at ~3Ghz with 16-32GB RAM and standard Gigabit Ethernet network interface cards [NICs]) provide poor packet capture performance due to fundamental

architectural issues in hardware and the operating system, making these platforms unsuitable when Gigabit levels of packet capture and application processing are required. Consider a typical PC with a basic (Linux) networking stack and standard NICs functioning as a network appliance packet capture system. As bandwidth to the device increases, the appliance is increasingly unable to receive packets and pass them to the application in Linux user space. As throughput increases to Gigabit levels, the host system CPU utilization can spike at or near 100% because the system kernel is consumed by answering interrupts to move packets from the network interface(s) to the host. This creates a situation where the entire CPU (or CPUs) is busy answering interrupts and has no spare cycles to actually empty packets from the user space buffer and process them at the application layer. The inevitable consequence is packet loss, poor network performance and application ineffectiveness. With the critical nature of the network monitoring and filtering applications traditionally built on these platforms, the application layer needs to see every packet in the data stream. Each packet lost before being analyzed at the application layer may represent a missed opportunity to find threats, identify vulnerabilities, filter attacks, recognize confidential data leakage or record suspicious activity, making the utility of the packet capture application questionable.

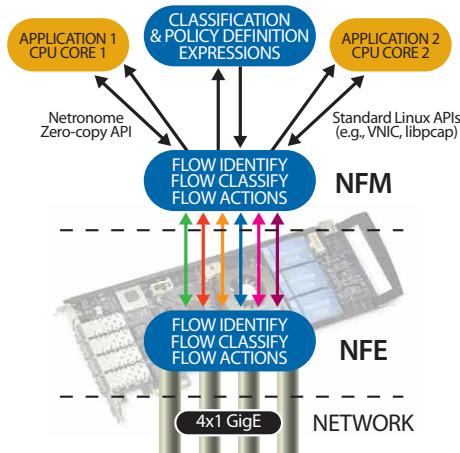
There have been many attempts to scale the packet capture, classification and filtering performance of network appliances built on standard server platforms. The simplest solution is “throwing” CPU cycles at the problem. CPU vendors are now shipping 4-8 core processors, with 16 core devices on the horizon operating at over 3Ghz. While additional CPU cycle availability can scale network performance to some extent, this solution does not fundamentally change the architecture of the underlying Ethernet NIC hardware or the standard operating system interfaces between the network, operating system kernel and user space.

Other approaches have been tried that attempt to improve the operating system implementation for packet capture, called libpcap. Typically, device drivers use libpcap to move packets from network interfaces to host processes by generating interrupts. In high-data-rate situations, though, the CPU can get in a livelock state where the processor cycles are fully consumed for interrupt handling rather than actually passing packets from the kernel into user space for processing the data stream. To avoid interrupt-livelock, device polling has also been implemented where, instead of servicing every interrupt, the interface card is periodically polled to see if data is present. This gives the operating system control over the NIC to assure that the CPU is not held hostage to servicing interrupts at high-input data rates. Other approaches to improve the packet capture performance for off-the-shelf server platforms, such as interrupt coalescence, zero-copy drivers and ring buffers, have been used with some

performance improvement. Even considering these techniques, standard server platforms still provide actual throughput that is far below what is required to support high-throughput packet capture, classification and filtering applications.

DEEP PACKET INSPECTION AND FLOW ANALYSIS AT 10GBPS

To overcome network appliance performance issues, Netronome has developed hardware and software components that allow both ISVs and end users to accelerate network packet capture and filtering applications. Netronome's solutions provide a combination of hardware acceleration for packet processing with an open application programming interface (API) for packet classification, zero-copy drivers and libpcap load balancing. This solution provides hardware acceleration, based on the Netronome Flow Engine (NFE) family of PCIe network interface cards, and software acceleration, in the form of the Netronome Flow Manager (NFM) API, which includes libpcap load balancing capabilities.



SIMPLIFIED ACCELERATION FOR NETWORK AND SECURITY APPLICATIONS

Supported across a wide range of Linux environments—including CentOS, Fedora Core, Ubuntu, Mandrake and Gentoo, in both 32- and 64-bit variants—the NFM software provides an open API for network and security appliances and applications that dramatically improves server/appliance performance. The combination of NFM with the NFE acceleration hardware significantly reduces appliance CPU utilization and packet delay/jitter by offloading complex flow analysis and deep packet inspection to the Netronome accelerator coprocessor. In addition to line-rate, 4x1 gigabit Ethernet flow processing and cut-through (per NFE), the solution also provides the ability to load balance flows across CPU cores so that single-threaded applications can take advantage of IA/x86 multi-core architectures.

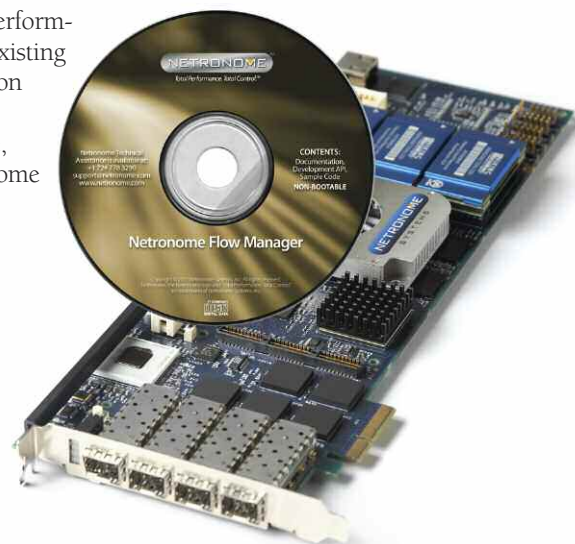
These capabilities provide appliance manufacturers and end-users with the ability to quickly improve the performance of existing network security applications and accelerate the development of their next-generation products, while reducing overall development costs. By abstracting complex NFE microengine programming with standard API calls, the NFM programming interface allows users to focus development on their software applications, while simultaneously benefiting from NFE-based hardware acceleration that significantly increases application performance. With the use of the NFE and NFM, network appliances can increase their performance and I/O throughput without requiring modification to their existing applications or complicated programming of underlying acceleration hardware in a standards-based Linux and IA/x86 environment.

This unique combination of accelerated network performance, granular flow visibility and reduced time-to-market makes Netronome technology an ideal solution for network appliances for:

- Security policy enforcement and compliance monitoring;
- Intrusion detection and prevention;
- Unified threat management;
- Deep packet inspection;
- URL filtering and reverse proxying;
- LAN/WAN bandwidth optimization;
- Application load balancing;
- Test, measurement and service assurance; and
- Lawful intercept (CALEA).

libpcap

libpcap is a system-independent library for user-level network packet capture typically found on UNIX-like systems and Linux. All packets on the network segment—even those destined for other hosts—are accessible through this mechanism. libpcap provides a portable framework for low-level network monitoring, as well as for saving captured packets to a file and reading files containing saved packets. libpcap is the packet-capture and filtering engine of many open-source and commercial network tools, including protocol analyzers, network forensics devices, network intrusion detection and prevention systems, traffic-generators and network test and measurement appliances.



COMPLEX FLOW ANALYSIS AND DEEP PACKET INSPECTION

The NFM provides support for a full suite of L2-7 flow analysis and deep packet inspection capabilities, including classification of flows based on standard packet header fields such as IP and Ethernet (including 802.1p/q) and identification of applications and protocols with fixed or well-known TCP/UDP ports or IP types. Beyond classification at Layers 2-4, the NFM also supports identification of applications and protocols through deep packet inspection at Layers 4-7. Application protocols supported include HTTP 1.0/1.1 (with support for embedded transactions and chunked encoding), as well as classification of common e-mail protocols such as POP3, SMTP and IMAP. Additionally, the NFM also provides classification of the following:

- Protocols embedded in HTTP, such as SOAP and web conferencing;
- Associated media and data flows, including FTP and SIP;
- VoIP, IPTV and other streaming media, such as SIP, RTP and VLC;
- IP Tunnels, including GRE, L2TP, PPTP, IPsec, IP in UDP or TCP; and
- Common peer-to-peer applications, such as BitTorrent, Gnutella, FastTrack, Jabber and WinMX.

Additional features and benefits include:

- A table-driven classifier (list of rules) simplifies the porting of applications that already use a list-of-rules mechanism, such as Linux iptables.
- Custom classification criteria can be implemented via arbitrary expressions that examine protocol-specific fields (e.g., extract HTTP URLs and pattern-match them), providing a level of expressiveness that is not possible with packet field (“n-tuple”) rules.

GRANULAR FLOW PROCESSING

In addition to packet classification, this unique combination of NFE-based hardware acceleration and the NFM API allows enterprise and networking applications to perform one or more unique actions on flows, once they have been identified. Whether deployed in flow-forwarding mode (where the appliance is deployed in-line as a “bump-in-the-wire”) or in passive mode (where the appliance is deployed off a span port, tap or mirrored interface), classified flows can be:

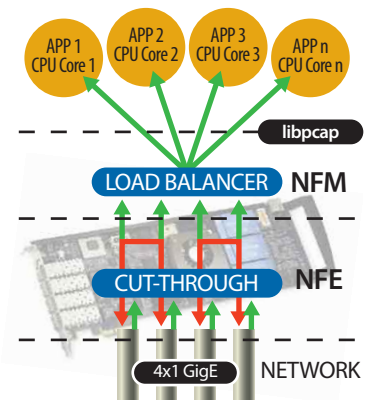
- “Cut-through”: Cut-through provides increased application CPU utilization by allowing flows to bypass the host CPU application processing. Classified flows are cut-through the network appliance in hardware (TCAM) between gigabit Ethernet ports on the NFE.
- “Load balanced”: Flows can be load-balanced across CPU cores for added application performance.
- “Redirected”: Classified flows can be diverted to the CPU for processing by the host application.

- “Tee’d”: All approved flows can be cut-through the appliance in hardware and copied to the CPU where further processing can be performed by the host application.
- “Tunneled”: Select classified flows can be tunneled through virtual overlay networks by terminating and re-originated IP tunnels (e.g., GRE).

Beyond the classification actions previously described, the Netronome solution supports a hardware-based statistics and monitoring mode, where the NFE/NFM are used to gather detailed packet and flow level statistics. This allows an application to capture statistics in hardware rather than in the operating system, providing visibility and flow monitoring necessary for any high-performance packet capture or filtering application.

PCAP LOAD BALANCING

The nature of packet capture and filtering applications, such as network forensics and intrusion prevention and detection systems, is such that they often rely on seeing all network traffic in order to be effective. Lost packets means threats are missed, limiting the effectiveness of these applications. Unfortunately, many of these applications are not multi-threaded and, therefore, cannot take advantage of multi-core CPU systems to increase performance. A typical solution to scale application performance in these cases is to run multiple application instances on a single network appliance (typically one process per core).



Single-threaded applications (like SNORT) can benefit from multiple application instances; but even with this workaround, in traditional server environments, flows cannot be balanced across these application processes evenly. Physical interfaces are tied directly to CPUs where all traffic received on a particular Ethernet segment is forwarded to a specific CPU core and application. Therefore, network segments with more traffic will experience proportionally worse performance than lightly loaded segments.

As a complement to the broad set of deep packet inspection, cut-through and flow analysis capabilities offered by the NFM, users can improve their application performance by load-balancing flows across x86 CPU cores using the zero-copy technique to deliver packet or flow data directly to Linux user mode applications (bypassing the Linux kernel and networking stack, thus avoiding copying from kernel memory to user memory). To improve application performance in multi-core environments, Netronome has modified the libpcap packet-capture library to utilize the NFM capabilities and provide an interface to pcap-based applications to take advantage of the underlying acceleration hardware. This approach load-balances traffic equally across all CPU cores (or a subset) smoothing out uneven network segment utilization and taking full advantage of all available computing resources. This zero-copy mechanism frees the network appliance CPU(s) from the cycle-intensive task

of getting packets from the Ethernet port(s) to the application(s) by minimizing kernel mode to user mode transitions and data copying, thereby improving performance. Flows are balanced across CPU cores by computing a hash key for TCP and UDP traffic from the following packet header fields:

- IP Destination Address
- IP Source Address
- Protocol Id
- TCP/UDP Source Port
- TCP/UDP Destination Port

This zero-copy load-balancing extends the NFM's capabilities to existing libpcap-based applications such as:

- tcpdump, a tool for capturing packets for further analysis;
- Wireshark (formerly Ethereal), a graphical packet-capture and protocol-analysis tool;
- SNORT, a network intrusion detection system;
- Nmap, a port-scanning and fingerprinting network utility;
- Bro IDS, a network-monitoring platform; and
- Clam AntiVirus, an anti-virus toolkit designed especially for e-mail scanning on mail gateways.

PCAP BLOCKING MODES

There are three different blocking modes that can be used with the libpcap available in NFM:

- Blocking
- Non-Blocking
- Adaptive Blocking

When a pcap application is running in blocking mode, the library goes to sleep until there are packets available to handle. Blocking mode will typically show lower performance than its counterparts; however CPU utilization is more efficient at low to medium data rates.

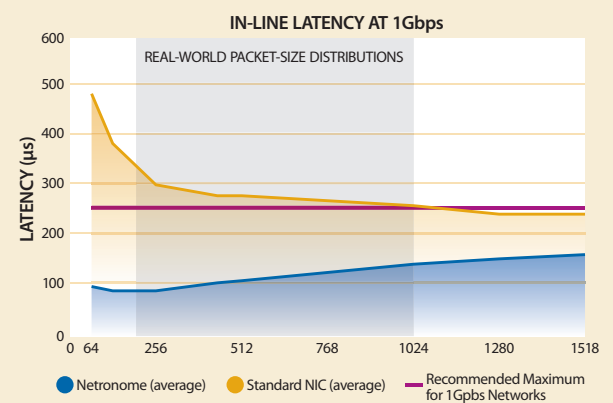
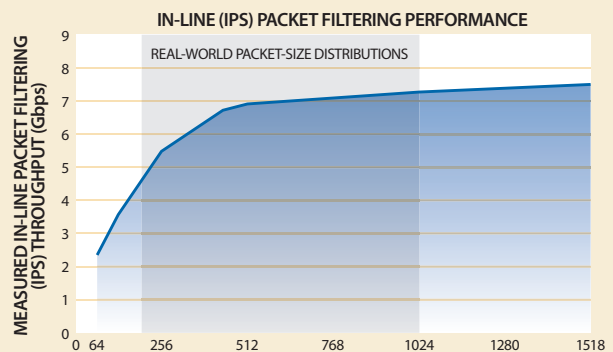
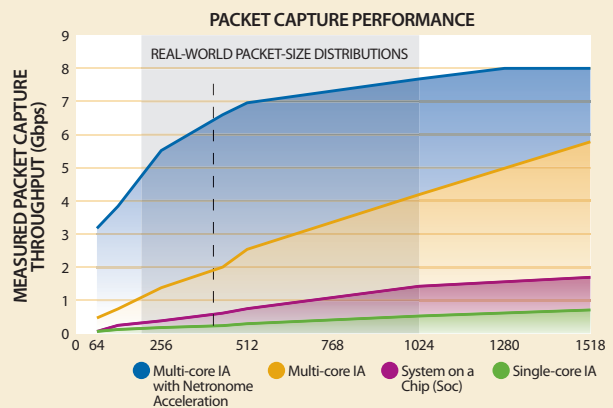
In non-blocking mode, the library never goes to sleep and constantly attempts to read packets from the capture descriptor. This mode will typically show higher performance over the other modes. However the CPUs report full utilization, even at low data rates.

Adaptive blocking mode is a mode introduced by Netronome with the integration of libpcap into NFM. Adaptive mode is a hybrid of blocking and non-blocking. This mode will behave as non-blocking does for a certain number of attempts before switching to blocking. In NFM, this number is usually set to 1000. This means the application will make 1000 attempts to read packets from the packet descriptor. If, after 1000 attempts, there has not been one packet returned, the application switches to a blocking mode and sleeps until a packet arrives. In units of time, this would be roughly 600µs without a packet. While packets are arriving on the interfaces, the CPUs are near full utilization. At times of low throughput rates, CPU utilization is similar to the utilization observed in blocking mode. The non-blocking mode is preferred in high-throughput environments where CPU utilization will always be high and optimal performance is required. In low-throughput environments, this mode may still be desired if the manufacturer is not concerned with reported CPU utilization.

APPLICATION ACCELERATION AT ALL PACKET SIZES

The Netronome solution, combining the Netronome Flow Engine and Flow Management software, can markedly increase server application packet capture and filtering performance in IA/x86 server platforms over non-accelerated solutions. Off-the-shelf solutions are adequate for low-throughput applications, but struggle to maintain line-rate performance for small packet sizes and at high data rates. As shown below, the Netronome solution can provide 2-4 times the measured performance when compared to an off-the-shelf solution. Offering up to 8 Gbps of throughput, the Netronome solution is the ideal hosting platform for custom and open source applications.

The Netronome solution, containing the NFE-i8000 PCIe accelerator cards, can provide 2-4 times the measured performance when compared to an off-the-shelf solution with standard Gigabit Ethernet Network Interface Cards (NICs).



Testing was performed on an Intel NSC2U server platform containing two Quad-Core Intel Xeon® processors at 2.83Ghz and 16 Gigabytes of RAM.

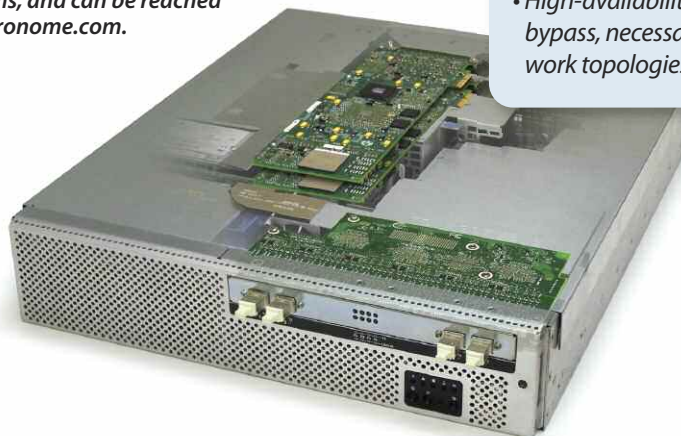
IT Paradox to IT Paradise

A new class of network and security appliances has entered the market. These devices provide solutions for performance, visibility and control, and security in a single system. They are the bridge between the high-speed network and the multi-core, multi-threaded, virtualized appliances that host the network and security applications. These devices offer line-rate network throughput, application and protocol acceleration, deep packet inspection and flow analysis for both plaintext and encrypted communication streams.

These next-generation appliances also serve as host systems for the many network and security applications that enterprises consider staples of their Ethernet and IP infrastructures, such as firewalls, IDS, IPS, UTM, NAC, anti-spam, anti-virus, etc. In addition to serving as application hosts, they have the ability to transparently intercept encrypted communications and provide the hosted applications with all requisite flows for analysis. The ability to provide the network and security applications with both non-encrypted traffic and the plaintext of encrypted flows extends the usability of those appliances, ensuring that the applications will be able to reliably perform as intended.

The Netronome solution provides line-rate throughput, deep packet inspection, flow analysis and application acceleration. These platforms allow a single network and security appliance to scale to line-rate network performance for a larger number of users, sessions and flows. IT organizations are now able to vertically consolidate the data center by not requiring stacks of redundant appliances that exist due to their inability to scale. These appliances also offer virtualization of both the network I/O and host system, enabling the integration of additional network and security appliances into an even smaller number of systems. The combination of performance, control and security allows enterprise IT managers to reinvent their data center by solving long standing problems, while simultaneously reducing capital and operational expenses.

Daniel Proch is the director of product management at Netronome Systems, and can be reached at daniel.proch@netronome.com.



10Gbps Network Flow Processing

Netronome's PCI-Express based acceleration solution offers dense front-facing Ethernet I/O options specifically designed for the Kontron IP Network Server, NCS2U. The 4-port SFP+ 10GigE short-reach optical, and 12-port 10/100/1000 BaseT deliver high-performance and low-latency with integrated high-availability via software configurable fail-to-open/close support. These load balancing interface modules (LBIM) are tightly coupled with the NFE-i8000 acceleration hardware and NFM software to scale server and appliance performance s to 10Gbps by delivering unmatched visibility and control of traffic at L2-L7 for over two million simultaneous flows.

Netronome's Network Flow Processing solutions deliver high-performance packet processing with intelligence, security and virtualization for millions of simultaneous flows through several unique features, including:

- *Hardware-based flow processing with line-rate packet capture of network traffic, enabling lossless deep-packet inspection applications with granular application- and content-specific policy control.*
- *Flow-aware adaptive load balancing to exploit multicore CPU parallelism and an efficient zero-copy driver, combine to quickly and evenly distribute 10Gbps of network traffic, maximizing host CPU performance.*
- *Up to 40Gbps of total throughput for real-world traffic patterns, with variable packet sizes and protocol mixes, to support the increasing bandwidth requirements of enterprise and service provider networks.*
- *Less than 80 microseconds of delay, far exceeding industry standard requirements for latency in 1Gbps and 10Gbps in-line network and security appliances.*
- *High-availability network interfaces, including integrated bypass, necessary for resilient deployment in in-line network topologies.*

Netronome's 10Gbps Network Flow Processing Solution

TM Netronome, the Netronome Logo, "Intelligent to the Core." and Open Appliance Architecture are trademarks of Netronome Systems, Inc.

All other trademarks are the property of their respective owners.

© 2008 Netronome Systems, Inc. All rights reserved. Specifications are subject to change without notice. (10-08)



Intelligent to the Core.™

Netronome has operations in:
USA (Pittsburgh [HQ], Santa Clara & Boston),
UK (Cambridge),
Malaysia (Penang),
South Africa (Centurion) and
China (Shenzhen, Hong Kong)
info@netronome.com
+1 877 638 7629
netronome.com