

Programmable Multifunctional Line Rate Analyzer for 10 Gbps Networks¹

Livio Ricciulli and Ajoy Aswadhati

Abstract— This paper examines (1) the network trends driving the needs for 10 Gigabit, line-rate deep packet inspection, and (2) the architecture of Force10's P-Series platform to deliver a very comprehensive, low level and flexible inspection system for high speed networks. The P-Series has two 10 Gbps Ethernet ports and acts as a Layer 1 transceiver with < 2 microsecond latency. The device can be programmed with several thousand stateful signatures that identify which packets are to be captured and/or blocked at line speed. Blocking/monitoring rules (specified as Snort 2.0 rules) can be modified in real-time without interrupting the packet flow. The hardware has been designed to easily integrate with existing open source monitoring software. Using our approach, all existing libpcap-based applications, such as tcpdump, Snort, etc., can transparently benefit from the hardware line-speed acceleration without modification. The programmable nature of this hardware technology can easily be adapted, modified and enhanced to accommodate new user-defined function through a Verilog API.

Index Terms—10 Gbps Network Processing, Intrusion Detection, Deep-packet Inspection, Network Security, Firewall, 10 Gbps security probe.

I. INTRODUCTION

Emerging network security challenges include combating insurgent, self-replicating and self-modifying network attack technologies. These technologies are hard to detect and their invasion is compounded by the increasing mobility of our computing resources (such as laptops and PDAs) which constantly redefine the concept of a security perimeter.

As network speed and complexity grows, network security devices are faced with the short-term technological challenges of (1) meeting the computational requirements for finding the vulnerabilities and (2) the ability to analyze complex event patterns across distributed systems.

Ethernet networks are being called upon to deliver unprecedented volumes of increasingly diverse traffic. In fact, IDC believes that as customers fully exploit the benefits of multicore technology in the 2008-2009 timeframe, they will begin to look seriously at 10 GbE server connections [1].

Following this inexorable rise in the demand for bandwidth, the adoption of 10 Gigabit Ethernet is accelerating. An estimated 60,000 ports of 10 GbE were sold in 2005 and 3

million ports are expected to be sold annually by 2010. 10 GbE builds on the success of previous generations of Ethernet – extending Ethernet simplicity and scalability across the LAN, MAN, and WAN. Whether in a massive data center, in the core of a large campus, or as a transport technology in a service provider or high speed research network, 10 GbE represents a breakthrough technology that will give users a high performance, low-cost alternative to today's backbone technologies – and a considerable competitive advantage.

The speed and growth of high performance networks has always presented a challenge to the IT staff. Today, that challenge is compounded by the very fluid, dynamic nature of traffic and applications being used on the network. As networks speed up, the tools and systems typically deployed for packet inspection have not kept pace.

The single biggest challenge for any surveillance or inspection device is delivering comprehensive, layer 2 through layer 7 inspections of payloads at the performance levels required by today's networks. Put simply, the speeds of networking technologies have far surpassed the capabilities of today's appliances to monitor, analyze, filter, and capture relevant traffic from the network. There is no easy or comprehensive way to see and track 10 Gigabit networks, especially with systems that can become congested and drop packets at 3-4 Gbps throughputs [2].

The result is a Hobbesian dilemma for organizations – slow the network and ensure full monitoring and capture capabilities or provide a high speed network with reduced scope of what to look at.

Resolving this tradeoff creates a parallel and daunting challenge for device and probe vendors. Delivering inspection and full capture services at 10 Gigabit speed is enormously challenging – but adding the requirement for flexibility so that the probe can adapt and automatically change inspection scopes on-the-fly has never been accomplished until now. The heart of the problem lies in the traditional divide between software and hardware. Software provides flexibility but is slow because it runs on general purpose microprocessors.

Hardware, on the other hand, has traditionally been high performing, but has not been flexible and therefore not ideal for environments in which the requirements are constantly changing. First generation inspection and probe devices attempted to resolve this challenge by running software on faster and faster microprocessors. While this improved speed somewhat, the maximum performance remained well below

¹ This work was supported by the Division of Design Manufacturing and Industrial Innovation of the National Science Foundation (Awards 0339343, 0443534, 0521902) and the Air Force Rome Laboratories.

Gigabit speeds. The second generation of devices used hardware to assist the software in an effort to offload some of the packet analysis. This did improve performance but again topped out near 1 Gigabit per second. The second generation also suffered from being exceptionally expensive which limited its adoption in the market.

In sum, inspection and monitoring appliance systems have existed on a hard divide – deliver the software flexibility necessary to deliver adaptable scope rules or provide high performance by utilizing dedicated, permanent silicon-based computing systems such as those delivered from ASICs but give up flexibility in trade.

The Force10 P-Series is the first of a new generation of inspection platform. It is based on a fundamentally different technology called DPI (Dynamic Parallel Inspection), which provides the performance of custom hardware but the flexibility of software. DPI overcomes the software-hardware conundrum through the patented use of a massively parallel inspection process deployed entirely in field programmable gate array technology (FPGA).

The P-Series 10 Gigabit system takes an entirely different, deterministic approach to deep packet inspection. In fact, the design of the system ensures that all traffic continues to pass through the system, regardless of load, or packet size, IPv4 or IPv6, all the time.

In this paper Force10 Networks describes the P-Series appliance, (the world's first line-rate, programmable 10 Gbps deep packet inspection system). The P-Series is a programmable multifunctional line rate analyzer designed for packet capture, surveillance, network monitoring and packet filtering applications.. This technology is now available as a 1U appliance capable of full-duplex 10 gigabit line-rate processing.

The P-Series is based on patented Dynamic Parallel Inspection (DPI) technology, derived from an innovative Multiple Instruction Single Data (MISD) processing architecture. Using this architecture the P-Series processor simultaneously applies thousands of rules to each packet at the same time. In fact, inspection and analysis rules are dynamically embedded into the fabric of the hardware allowing the P-Series to deliver predictable performance and signature scalability under all traffic conditions.

II. P-SERIES KEY APPLICATIONS

The combination of high performance line-rate 10 GbE densities, ultra low latency, total open flexibility, and reliable and resilient operation are driving five P-Series solution architectures:

A. *Lawful Intercept:*

In surveillance, lawful intercept and Communications Assistance for Law Enforcement Act (CALEA) compliance applications, new aggregate network speeds and recently increased illegal activity on the Internet have challenged law enforcement's ability to conduct authorized electronic surveillance online. The P-Series, in its ability to stay fully

deterministic under full traffic conditions, to deliver inspection reliability through a 100% hardware-based inspection engine, and the full customization to search anywhere in the data stream, including IPv6 header information, provides the ideal solution to this challenge. The P-Series dynamic packet and traffic capture capabilities are a strong compliment to Force10's switching solutions already gaining traction in this application.

B. *Custom Open-Ended Wire-Speed Applications:*

Due to the unique properties of the DPI technology – wire speed, predictable, programmable packet analysis – customers are also pursuing new and unique applications for the P-Series using the application programming interface (API). Examples of customer-driven P-Series applications include a leading European service provider's network monitoring and diagnostic application, and a U.S. financial services customer writing a latency validation application for Tibco message data.

C. *Regulatory Compliance Applications:*

The ability to continuously guard, monitor, and capture key data are also key elements of both the Sarbanes-Oxley Act (SOX) of 2002 and the Health Insurance Portability and Accountability Act (HIPAA) of 1996. Both regulations were driven by the need to improve how we report, govern, and disclose public information and manage confidential record keeping. Because the P-Series can identify traffic by any nature – machine, address, traffic type- and provides line-rate monitoring and packet capture of potentially 1000s of search strings, customers have been able to insert the P-Series at key network intersections, and provide compliance visibility to these regulations.

D. *Dedicated IDS/IPS:*

For intrusion detection and protection, the P-Series supports existing, open source network security and monitoring applications by specifying capture and filtering policies using public domain IDS signatures such as Snort and Bro or standard network packet capture (Libpcap) monitoring expressions.

E. *Firewall and IDS/IPS Pre-Processing:*

Because of its performance characteristics, the P-Series is also being deployed in front of other IDS/IPS and firewall systems as a security accelerator. P-Series technology provides developers an API for creating custom network security and monitoring applications such as stateful firewalling, DOS and DDOS, and packet and flow analysis applications.

III. THE P-SERIES DYNAMIC PARALLEL INSPECTION (DPI):

A. *Architecture*

The foundation of our architecture is a network processing module with two network physical layer interfaces (PHYs) and a host connection. We use a digital circuit to pass signals

between two PHYs. This circuit simply receives digital signals from one PHY and directly transmits them to the other PHY unchanged, creating a Layer 1 transceiver. This forwarding can be halted by gating signals to the Layer 1 transceiver. The data forwarded between the PHYs is also transmitted to another system through an additional bus for analysis.

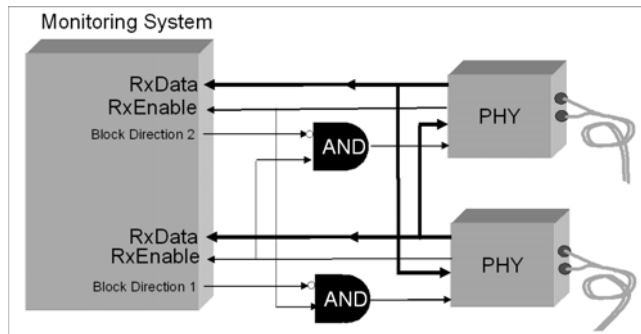


Figure 1: Network Interface Architecture

The P-Series P10 provides two, full duplex 10 Gigabit pluggable XFP “sensing” interfaces. These interfaces can be used to tap two separate 10 gigabit traffic streams, or to sit inline on one 10 gigabit link.

Figure 1 diagrams this architecture. An FPGA serves as “forwarding engine” that copies traffic to the monitoring system as well as mirror, block or forward traffic on the wire. The P-Series is unique in its architecture in that traffic forwarding is independent of the traffic monitoring process, ensuring line-rate traffic with the industry’s lowest latencies for all traffic types.

In addition to providing predictable performance on the network, this architecture gives the P-Series unprecedented network transparency. To the network, this is a layer 1 extension of the fiber. There is no MAC address, or any layer 2 or layer 3 connection occurring when the P-Series is placed into the network. In fact, given the sub-one microsecond latencies of the P-Series, it would be extremely difficult to even detect the presence of the P-Series from the service interfaces on the network in either tapped or inline deployments.

Lastly, by using the forwarding architecture, the system creates a very high degree of service availability during various failure scenarios. As packets are forwarded back onto the wire as a first operation, the P-Series can stay up and operational in numerous fault scenarios. Specific failure scenarios that will NOT affect the line-rate traffic forwarding of the P-Series are (1) hard drive failure, or (2) host interface, or operating system suspension or crash. Technically, the P-Series will continue to forward traffic even when the monitoring system itself is suspended.

B. Monitoring System

Currently network processing performance for security applications falls short of its operational requirements. In part this is because current network packet inspection processing hardware support is based on retrofitted legacy processors.

Using the Multiple Instruction Multiple Data (MIMD) processing model, current network processors parallelize across packets, but each packet is processed serially. This is well suited when the computation to be performed on each packet is sequential in nature (like routing) but does not parallelize efficiently for network security workloads.

This computational workload is extremely parallel in nature and is performed much more efficiently using a *Multiple Instruction Single Data* (MISD) computational model. Using this paradigm a single data stream is broadcast to multiple computational units, each executing different instructions. More specifically, the data stream (packets) produced by the network physical layer interfaces (PHYs) are first de-serialized and then adapted from 8 to 1024 bits for broadcast inside one or more separate processors. Thus a large number (thousands) of simple execution units share the data and concurrently implement different packet matching operations. The execution units either reside within the same processor or are distributed across several processors.

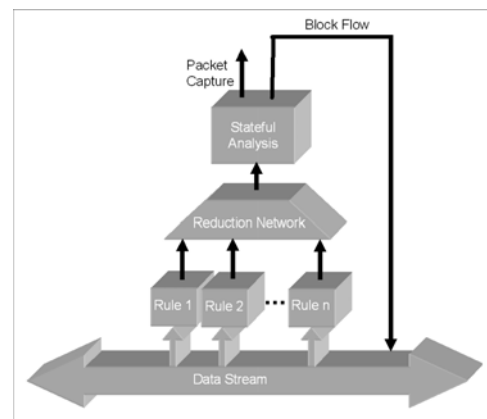


Figure 2: Data Stream Processing

In figure 2, the data stream is concurrently presented to a number of execution units (Rule 1, Rule 2, through Rule n). Each unit is responsible for independently performing wire-speed packet processing and outputting a number of signals. Each of the inspection rules which are embedded in the execution units can be changed dynamically. As new needs emerge, new rules can be written and pushed into the units. This can be done online, on the fly, or offline. In fact, these rules can be changed in a production system and are applied in less than 1/1000th of a second. During the application of new rules, the system will maintain all state and continue to apply all existing rules without interruption.

By splitting analysis rules into many discrete engines that can run on the same data in parallel, and by embedding these rules in the gates of an FPGA, DPI can achieve both record-breaking inspection throughputs of 14.88 million packets per second, while doing so in approximately 1 microsecond.

This processing technology can scale in both speed by employing larger de-serialized words to balance faster serial links and in the number of signatures by adding more matching logic to store the additional signatures. The

additional matching logic may be gained through adding FPGAs or ASICs or by using a larger chip.

The DPI also tracks state for each flow through the use of an external memory table. This memory table provides very high performance state memory management to handle up to 300,000 new flows per second (10x better than traditional firewalls), and up to 8 million concurrent stateful flows in the aggregate.

Mirroring ports enable advanced surveillance applications that require to immediately move matched traffic to separate collection systems. In essence, the P-Series can provide real-time matched traffic forwarding, with no store or caching of sensitive network data.

By using true hardware separation of the blocking, mirroring and capturing sub-systems, DPI ensures there is no impact to signature inspection logic or its wire-speed operation under load, under full traffic capture, or under any traffic mirroring or blocking scenarios. This leads to identical performance, identical throughput, and identical latency with any traffic load, and under full use of the system's analysis policies.

In addition, the DPI's programmable features (built entirely in FPGAs), can be totally customized to various capabilities in pattern matching, network monitoring, surveillance, or network capture. The entire system can be reprogrammed at the level of the silicon "assignments" of each gate on the FPGA. This flexibility is also driven by the open-ended nature of the DPI inspection logic. There are no protocol, packet, framing requirements, and as such, customizing new inspection capabilities (IPv6 traffic inspection for example) can be quickly facilitated.

In sum, DPI achieves both the dynamic, real-time flexibility required for today's fast-changing surveillance, monitoring and network analysis requirements, while introducing the industry's first line-rate 10 Gigabit inspection and packet capture system.

IV. THE IMPLEMENTATION: OPEN-SOURCE UNIX APPLIANCE WITH DPI PROCESSORS

A. Processor module

Our hardware network processing architecture is implemented as 10 Gbps DPI cards, which are capable of line-rate processing several thousands of stateful IDPS signatures.

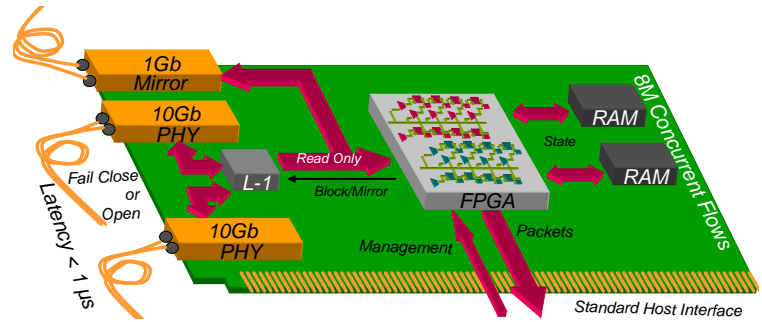


Figure 3: P-Series DPI Card

As shown in Figure 3, the DPI cards feature (1) two 10Gb Layer 1 transceivers that can forward at wire-speed with latency of $< 1\mu s$ and (2) two 1Gb Layer 1 mirroring ports (for viewing clarity, only 1 mirroring port is shown) that can transmit matched packets/flows to external devices. Our FPGAs apply several thousand user-defined stateful policies (up to 8 million concurrent flows) using MISD processing. These policies are directly derived from public-domain Snort rules or are custom defined. The DPI cards concurrently apply all policies in the fast-path. The packets that are matched by the policies can be (1) captured from the fast-path for the host, (2) mirrored to an external device and/or (3) blocked.

V. OPEN-SOURCE SOFTWARE SUPPORT

The P-Series appliance supports existing, open-source network security and monitoring applications by specifying capture and filtering policies using public domain IDS signatures or standard network monitoring libraries. The cards also offer a rich application programming interface for creating custom network wire-speed applications (see Section VII).

A. Standard NIC Emulation

The P-Series DPI cards appear as a standard NIC to the host OS. In Figure 5, *eth0* is a regular 100 Mb NIC (for management), *eth1* is a Gigabit NIC (for testing), and *eth2* a DPI card. Notice that the MAC address of the DPI card is a dummy value of *00:01:02:03:04:05* because the DPI cards are not MAC-addressable devices; the dummy value is there to satisfy the OS.

```

root@localhost:~# ifconfig
[root@localhost root]# ifconfig
eth0      Link encap:Ethernet  HWaddr 00:30:48:2A:33:27
          inet addr:192.168.1.120  Bcast:192.168.1.255  Mask:255.255.255.0
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:137  errors:0  dropped:0  overruns:0  frame:0
          TX packets:75  errors:0  dropped:0  overruns:0  carrier:0
          collisions:0  txqueuelen:100
          RX bytes:13572 (13.2 Kb)  TX bytes:12245 (11.9 Kb)
          Interrupt:22

eth1      Link encap:Ethernet  HWaddr 00:30:48:2A:33:26
          inet addr:10.1.1.120  Bcast:10.255.255.255  Mask:255.0.0.0
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:0  errors:0  dropped:0  overruns:0  frame:0
          TX packets:4  errors:0  dropped:0  overruns:0  carrier:0
          collisions:0  txqueuelen:100
          RX bytes:0 (0.0 b)  TX bytes:256 (256.0 b)
          Interrupt:54  Base address:0x3000  Memory:fc200000-0

eth2      Link encap:Ethernet  HWaddr 00:01:02:03:04:05
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:0  errors:0  dropped:0  overruns:0  frame:0
          TX packets:0  errors:0  dropped:0  overruns:0  carrier:0
          collisions:0  txqueuelen:100
          RX bytes:0 (0.0 b)  TX bytes:0 (0.0 b)
          Interrupt:28  Base address:0x4000

lo        Link encap:Local Loopback
          inet addr:127.0.0.1  Mask:255.0.0.0
          UP LOOPBACK RUNNING  MTU:16436  Metric:1
          RX packets:0  errors:0  dropped:0  overruns:0  frame:0
          TX packets:0  errors:0  dropped:0  overruns:0  carrier:0
          collisions:0  txqueuelen:0
          RX bytes:0 (0.0 b)  TX bytes:0 (0.0 b)

[root@localhost root]#

```

Figure 5: DPI Card Emulating a Standard NIC

B. Performance Boost for Trusted Tools

Because the DPI cards appear as regular NICs, they can seamlessly run a variety of standard application software at much faster speeds. For example, open-source Snort IDS software can monitor a few hundred megabits of traffic with a standard NIC [3]. With the DPI cards, Snort can monitor a full 10 Gbps of traffic without modification. The DPI cards are also compatible with other popular libpcap-based network monitoring applications such as tcpdump.

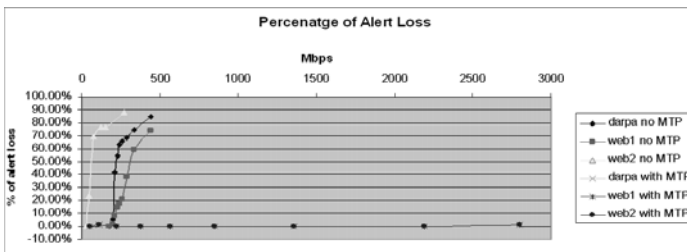


Figure 7: Snort Benefiting from Our Technology

Figure 7 shows attack detection by Snort under increasingly higher loads. Notice that without the P-Series (also called MTP), as the packet rate per second increases beyond a few hundred Mbps, Snort loses more and more attacks, quickly becoming ineffective. The P-Series insulates Snort's performance from extremely high traffic loads (in this case up to 2.8 Gbps).

VI. MANAGEMENT INTERFACE

The DPI processing is managed with a simple open-source, rule-management screen or CLI interface that enables users to:

- Start and stop the DPI processor
- Manage simple runtime parameters such as packet truncation, flow length, and timeout
- Set capture and forwarding policies for each rule

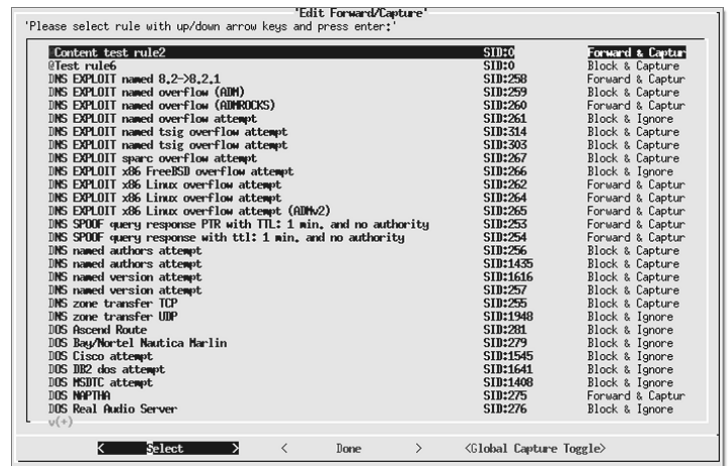


Figure 6: DPI Processor User Interface

Figure 6 shows our nurses interface through which the user can direct the DPI to:

- Capture packets for the host
- Forward packets (with negligible delay)
- Block packets
- Mirror packets to the mirroring interface

A. Rule Writing

Two types of rules can be uploaded to the DPI processor.

- Static rules are compiled by the user to become part of the firmware and are mapped directly into logic gates. Static rules can be disabled/enabled individually, but they cannot be changed once they have been loaded into the FPGA (unless the firmware is recompiled and reloaded by the user).

- Dynamic rules are programmed at runtime in the DPI hardware registers and can be configured without changing the firmware. These rules (like static rules) can be disabled/enabled individually.

Each signature can be pre-programmed to match network traffic by:

- TCP/IP address
- MAC address
- Port, protocol and application data
- Complete “5 tuples” flow tracking
- Payload content at a specific offset from the start of the packet or unanchored

B. Stateful Matching

Stateful matching enforces a time dependency between the matching events. With stateful pattern matching, it is possible to specify which matching events needs to occur at which time with respect to other matching events. In order to achieve this, it is necessary to store information (state) on which events occurred and use such information each time a new event occurs.

Stateful matching improves the accuracy of detection

because it adds ordering when specifying behaviors across multiple matching events. As in many designs, one of the great challenges of stateful matching is to efficiently manage the matching information as it unfolds. It is often the case that there are limited resources to record the information and thus techniques are needed to reclaim stale resources for new matching events. Because wire-speed hardware-based matching systems work synchronously without the aid of operating systems, they need to manage state in a simple and deterministic way.

The DPI incorporates a lossless state management algorithm that allows a deterministic use of memory resources by matching state to follow a non-cyclic pattern.

VII. DPI VERILOG API

The P-Series product provides a low-level API for customizing the traffic processing features of the DPI cards.

P-Series supports a hardware-based API mechanism where customers can write their own 10 Gbps, line-rate IDS/IPS policies using verilog that can be mapped directly to the FPGA. This offers unmatched flexibility in writing custom rules and actions to be taken. The user can write Verilog modules using a standard text editor and subsequently invoke the Unix command “make”. This will synthesize the module and link it with the rest of the DPI firmware. The resulting FPGA “bit” files can then be uploaded and managed through the existing DPI management applications through the host interface. The DPI API is implemented on some of the largest FPGAs available today. Our current architecture assumes that the P-Series would be used for IDS/IPS or statistical flow-based analysis and the API exposes the on-board memory to store flow state information. Future extensions to the API could enable different uses of this memory.

VIII. SUMMARY

Force10’s P-Series product, and its DPI technology, represent the next generation inspection, packet monitoring, and packet capture platform. Based on DPI integrated hardware packet processing technology, the P-Series is capable of deep packet inspection at line-rate 10 Gigabit speeds, enabling the P-Series to monitor, capture and block malicious traffic without impacting performance. The P-Series is adaptable and extensible to provide open flexibility – especially important in surveillance, and security filtering applications.

Force10 is the performance and resiliency leader in Gigabit and 10 Gigabit Ethernet switching and routing. With the launch of the P-Series, we extend our technology leadership into the inspection and monitoring market, while introducing industry-leading innovations in high speed DPI designs. We believe the P-Series, based on the advanced DPI technology brings new performance scaling, reliability, flexibility and predictable performance to stateful packet inspection, filtering and capture and blocking applications

ACKNOWLEDGMENT

We thank Jagjit Choudary, David Voskian, and Jeremy Stieglitz for help in reviewing this paper.

REFERENCES

- [1] Lucinda Borovick Worldwide Enterprise Datacenter Network Forecast, Ethernet Switch Report, Five Year Forecast, 2006-2010, IDC., March 2007 2 Dell Oro Group,.
- [2] David Newman. (2006, November 09). “IPS performance tests show products must slow down for safety,” Network World. Available: <http://www.networkworld.com/reviews/2006/091106-ips-test.html>
- [3] Kerry Cox and Christopher Gerg, Managing Security with Snort and IDS Tools. O’Reilly, 2004, pp. 226-227

Livio Ricciulli is the Chief Security Scientist at Force10 Networks where he leads research efforts in high speed network security. Livio is the inventor and architect of the P-Series and was previously the Chief Scientist at MetaNetworks Inc and a Senior Network Scientist at SRI International. Livio has led several award-winning and commercially successful research projects for both Government and Industry for the past 15 years. Livio has a M.S. from the University of Southern California and a B.S. from the University of Arizona both in Computer Engineering.

Ajoy Aswadhati is a Senior Director of Engineering, at Force10 Networks Inc. He manages the Edge Development Unit at Force10 Networks, which is responsible for P-Series and the C-series product lines. He is interested in designing and bringing products to market that target high speed networking and security.

He has an M.S in Computer Science from the University of Texas at Dallas, U.S.A and an M.S in Electrical Communication Engineering from the Indian Institute of Science, Bangalore, India.