The debate on DPI technology

Luís Reis Mata
e-Planning PhD student
e-Planning Consortium
luisreismata@gmail.com

***Abstract* - This paper presents some reflections on Deep Packet Inspection technology. The methodology used was the revision of some references about technology and discussion of key points in their use, such as technological environment, security and network management, bandwidth management, benefits and dangers in the use of DPI.**

Keywords – Technology; Internet; Inspection; Packets; DPI

I. INTRODUCTION

Using Deep Packet Inspection (DPI) technology enables Internet service providers (ISPs) to block, alter or prioritize certain data packets traded on the Internet, enabling the creation of policies for traffic management and Quality of Service (QoS).

The debate on DPI technology, its uses, benefits and dangers is extensive.

On one hand, advocates of net neutrality oppose to all technology that distorts the founding ideal of the internet, ie, a network where information flows between its ends without any interference, on the other hand, the benefits brought by this innovation raises questions about the legitimacy of net neutrality discourse.

Thus, in this paper we discuss the key technical concepts of DPI technology and analyze some of its applications.

II. BACKGROUND OF TECHNOLOGY DPI

The internet operates based on a protocols layered architecture such as the Transmission Control Protocol (TCP) and the Internet Protocol (IP), designed to ensure the interconnection between computers. Currently, the most used version is still version 4 - IPv4, in spite the growing migration to version 6 - IPv6. Fractional information in data packets (DP) is sent over the network between sender and receiver, via the IP address assigned to each terminal equipment, ensuring communication on the Internet.

Inspired in the OSI[1]Model, TCP/IP is a layered model by different abstraction levels composed of four layers as shown in figure 1:

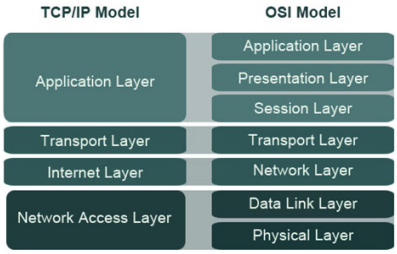

Figure 1 - Relationship between TCP/IPv4 architecture and the OSI model [1].

The fourth layer of the TCP/IP protocol, «Application Layer», is responsible for TCP/IP "applications", which perform the different communication protocols (HTTP, SMTP, etc). The layer of «Transport» (third layer) is responsible for ensuring that DP are delivered error free, in sequence and without loss or duplication of data. The «Internet Layer» (second layer) is responsible for addressing and routing network also by carrying out the fragmentation of DP, assigning them unique identifiers (IDs). The first layer, «Network» Interface, is responsible for inserting a header and footer on the transaction data. All layers introduce information on header, but the footer is filled only by the first layer. The footer is created to perform validation error through cyclic redundancy test (CRC). When a client receives a DP, a CRC is generated and compared with the CRC residing in the footer of DP. If the CRCs match up each other, the DP is considered validated and sent to the next layer. If they don't match up, the data will be considered invalid, and therefore, discarded. The first layer is also responsible for establishing and maintaining connections between sender and receiver [2].

As stated, the DP is composed of three distinct parts [3]:

- Header - contains instructions about the data contained in DP, such as packet length, synchronization information, number of DP transmission protocol (HTTP, SMTP, P2P, etc.), source IP address and destination IP address;
- Body or data - data being transported;
- Footer - contains information about the termination of DP and information for error validation, CRC.

The assessment of DP over the network has promoted the development of solutions for inspection of its content, with the

[1] Open Systems Interconnection (OSI) – Model implemented by the International Organization for Standardization (ISO), with the aim of creating communication standards.

aim of, for example, monitor and control malicious content, or assist in the prioritization of DP traffic.

DPI technology enables greater inspection of DP circulating on the network that achieved by traditional inspection processes, also known as *Shallow Packet Inspection* (SPI). Combining the features of technologies *Intrusion Detection System* (IDS) and *Intrusion Prevention Systems* (IPS), the DPI can inspect the DP in all abstraction layers of TCP/IP protocol.

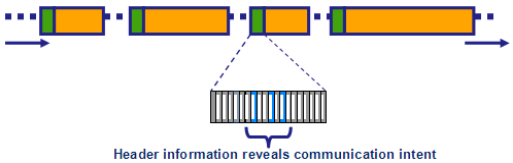

Figure 2 - SPI Inspection - Profile of DP header [4]

Thus, the SPI can only extract the ordinarily information resident in the protocol header (Figure 2), such as the sender and receiver IP addresses and other low-level connection states [4].

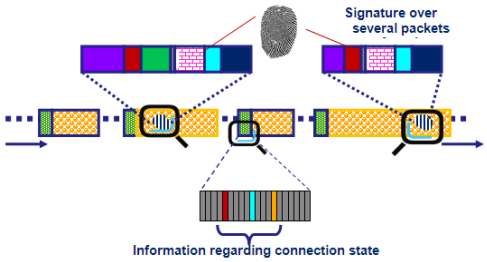

Figure 3 - DPI – Analysis and encapsulation of DP contents [4]

The DPI technology, by content inspection of header and body can determine the application that generated the DP (http, SMTP, P2P, etc.), enabling the differentiation and classification of traded DP. The DPI uses a signatures system (Figure 3) to identify a particular data stream and encapsulate the DP per application protocol levels [5], allowing prioritization.

The increase of technological solutions to specific needs, whether in terms of latency, or in terms of bandwidth consumption levels, motivated the implementation of prioritization policies and traffic management, considered vital by ISPs.

III. APPLICATIONS OF DPI TECHNOLOGY

Initially developed to ensure the safety of local networks (eg, companies or universities) in order to block unwanted traffic, the DPI is, for many years used in various internet applications: filter and preventing spam and viruses, firewalls, prevention and intrusion detection systems or cookies.

Following, we explore two main motivations for the use and application of DPI technology: spam (detection and blocking) and peer-to-peer (prioritization of content).

A. Spam

The term spam is used to classify the messages sent to a large group of recipients in an attempt to force unsolicited content to people who otherwise would choose not to receive them.

Spamming has several purposes, from advertising to sending malicious programs such as viruses, phishing, malware or spyware.

The e-mail spam volume has grown over time at an accelerated rate, since its mass mailing is extremely easy to implement at practically negligible costs.

As it causes the loss of resources (time, storage and bandwidth), spam enforcement is considered a priority with regard to the security policies of work environments, personal or business.

B. Peer-to-Peer (P2P)

In traditional forms of content distribution, based on the client-server model, a reduced number of servers satisfy the service requests made by customers (centralized architecture).

The P2P architecture is also a model of distributed computing, but with one significant difference. The P2P architecture is a decentralized architecture (Figure 4), where exist equal status between participants of transactions, ie, an entity can simultaneously request a service, running as a client, or providing a service, acting as a server.

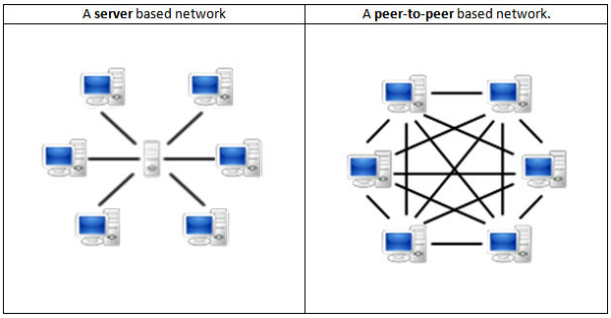

Figure 4 - Comparison between architectures client-server and P2P [6]

Given the P2P properties, the ends strive to maintain the highest number of active connections and bandwidth, significantly degrading the quality of the internet service. Then becomes necessary to monitor and control data streams originated by P2P networks, in order to relieve pressure on the network load [7] [8].

Therefore, ISPs and technology service suppliers (network) need to know the different types of clients they have and what their impact on network characteristics. Many business decisions, such as setting different tariff models, scalability determination in network design or their dimensioning are heavily dependent on the demands of its customers [9].

Mochalski and Schulze [10] refer that the need to manage bandwidth is related to the different levels of Quality of Service (QoS) required by the various types of applications. So, while the internet telephony (VoIP) and online games work best at low latencies, but consume little bandwidth, downloads are practically not affected by latency and use the maximum possible bandwidth.

These authors advocate the usefulness and importance of DPI[2] in: a) the prioritization of real-time applications such as VoIP, online games or remote accesses; b) limiting the available width during periods of congestion to the intensive use of wide bandwidth applications, such as, large downloads originated by peer-to-peer (P2P) connections or hosting file servers; c) blocking access to unwanted applications in enterprise environments, e.g., file sharing by P2P.

As benefits in bandwidth management, these authors also consider:

- Can improve the average performance of the internet, at the expense of a small fraction of users [3];
- Can provide users with a personalized service, including some QoS guarantees at a price higher or lower, depending on the level of service required. Users that only use email and web pay a lower price, implementing the principle of the pay per use.

The use of DPI in network management, discriminating the DP through its contents, considering its use for legal purposes and not distorting the competition mechanisms, promotes the paradigm shift of network neutrality.

No longer should be define neutrality as the imposition of network operation modeled pre-DPI, since, as described above, the prioritization of content for lawful purposes is a necessity, and cannot claim to non-discrimination of DP, regardless the resources on bandwidth consumed. Properly, it should be stated that the net neutrality as a principle of equality never occurred because not all users have the same access to the internet. Cases from countries like China or North Korea that control and manipulate contents that will be allowed to be consumed by its inhabitants or the different conditions of access for users through the conditions of local market competition.

However, the concept of net neutrality continues to have relevance, understood as a sponsor principle for biggest and best conditions of transparency and equal conditions of access to provided technology and services.

In addition to the above reasons, such as protecting privacy by blocking spam, phishing, viruses, defense or preservation of copyright, there are many other situations where their use raises questions or promotes acts of iniquity and lawlessness.

One example where the use of DPI can promote such practices relates to commercial reasons. Discrimination of pages of direct competitors or trading partners of competitors or, the implementation of new internet plans based on the consumption made. This type of activity can promote operator changes by consumers, getting the National Authorities with responsibility for promoting competitive markets.

One of the main criticisms in the use of DPI management and network traffic prioritization indicates a possible reduction in the generation of innovation by the service providers regarding to the increase of network capacity.

CONCLUSIONS AND FUTURE WORK

The DPI use, as the vast majority of innovations, is subject to the purposes of those who use it. There are good and bad purposes. So, exists the need to promote good uses and seek to create mechanisms of control over inappropriate or abusive use.

In this paper, we analyzed the generic operation of DPI technology and some of its uses.

As future work, we consider relevant expand this study by introducing and exploring other technologies that promote the need for network management, such as VoIP, IPTV and relate them to the concepts of New Generation Networks (NGR) and with the IPv6 protocol. Further work on the benefits and dangers may also come to be realized

REFERENCES

[1] Yanuar, Muhamad. tempayan.com. tempayan.com. [Online] 2010. [Citação: 2010-07-20] http://tempayan.com/page/2?topic=14.0.

[2] SCRIMGER, ROB, LASALLE, PAUL e PARIHAR, MRIDULA. TCP/IP - A BIBLIA. Rio de Janeiro : Campus, 2002. 13:978-85-352-0922-8.

[3] HSW, How Stuff Works. O que é um pacote de dados? How Stuff Works. [Online] 2000. [Quote: 2010-07-19] http://informatica.hsw.uol.com.br/questao525.htm.

[4] Shess. Digging Deeper Into Deep Packet Inspection (DPI). [d]packet.org. [Online] 2007. [Quote: 2010-07-19] https://www.dpacket.org/articles/digging-deeper-deep-packet-inspection-dpi.

[5] Cascarano, Niccolò, Ciminiera, Luigi e Risso, Fulvio. Accelerating DPI Traffic Classifiers. Netgroup - Politecnico de Torino. [Online] 2008. [Quote: 2010-07-19] http://netgroup.polito.it/Members/niccolo_cascarano/pub/lightweight-DPI.pdf.

[6] GigaTribe. Need help with GigaTribe? GigaTribe - Share with People You Trust. [Online] 2010. [Quote: 2010-07-20] http://www.gigatribe.com/en/help-p2p-intro.

[7] Wang, Chunzhi, et al. Design of P2P Traffic Identification Based on DPI and DFI. IEEEXplore. [Online] 2010. [Quote: 2010-07-19] http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5374577.

[8] Tomoya, Kato e Shigeki, Yokoi. Application of P2P (Peer-to-Peer) Technology to Marketing. IEEXplore. [Online] 2004. [Quote: 2010-07-20] http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1253478.

[9] Kolbe, H.J., Kettig, O. e Golic, E. Monitoring the Impact of P2P Users on a Broadband Operator's Network. IEEEXplorer. [Online] 2009. [Quote: 2010-07-20] http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5188835.

[10] Mochalski, Klaus , Schulze, Hendrik. Deep Packet Inspection: Technology, Applications & Net Neutrality. Ipoque White Paper. [Online] [Quote: 2010-07-19] http://www.ipoque.com/userfiles/file/DPI-Whitepaper.pdf.

[s.r.] Cruz, Ademar. A versão actual (4) do protocolo IP. IPV4. [Online] 1999. [Citação: 2010-07-19] http://civil.fe.up.pt/acruz/Mi99/asr/IPv4.htm.

[2] Usually the DP detection and identification is made by concerted effort between DPI technology and DFI technology (*Deep Flow Inspection*).

[3] Many studies indicate that the use of Internet resources follows the Pareto Law, ie 80% of bandwidth consumption is made by only 20% of users.