White paper on

Deep Packet Inspection

Deep Packet Inspection

1.0 Introduction

Deep Packet Inspection (DPI) is a technology that enables the network owner to analyse internet traffic, through the network, in real-time and to differentiate them according to their payload. Since, this has to be done on real time basis at the high speeds it cannot be implemented by software running on normal processors or switches. It has only become possible in the last few years through advances in computer engineering and in pattern matching algorithms.

Originally the Internet protocols required the networkrouters to scan only the header of an Internet Protocol (IP) packet. The packet header contains the origin and destination address and other information relevant to moving the packet across the network. The “payload” or content of the packet, which contains (all or part of) the text, images, files or applications transmitted by the user, was not considered to be a concern of the network operator. DPI allows network operators to scan the payload of IP packets as well as the header. Figure 1.shows the domain of packet inspection required in internet protocols and in DPI.

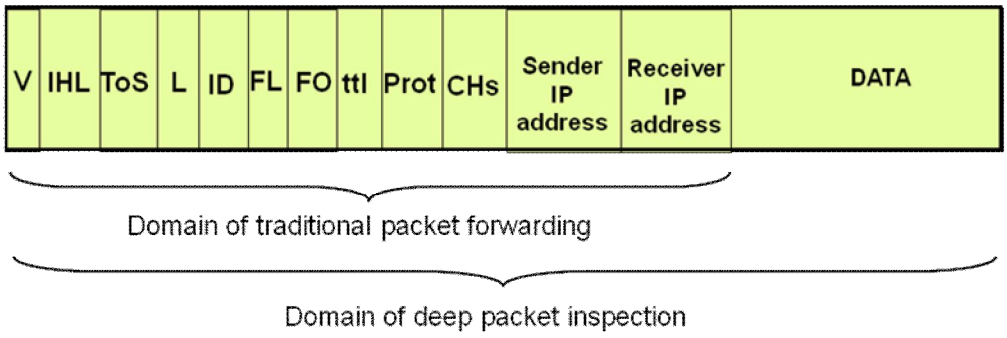

Figure 1. Domain of Deep Packet Inspection

DPI systems use expressions to define patterns of interest in network data streams. The equipment is programmed to make decisions about how to handle the packet or a stream of packets based on the recognition of a regular expression or pattern in the payload. This allows networks to classify and control traffic based on the content, applications, and subscribers.

2.0 Levels of Packet Inspections

Many of the functions provided by DPI technology have been available before to limited extent depending on the level of packet analysis. Packet inspection

technologies that have been in use in networking environments can be classified in three classes. These three classes are ‘shallow’, ‘medium’, and ‘deep’ packet inspection. Figure 2 provides a visual representation of the depth of inspection each of these technologies allows for.

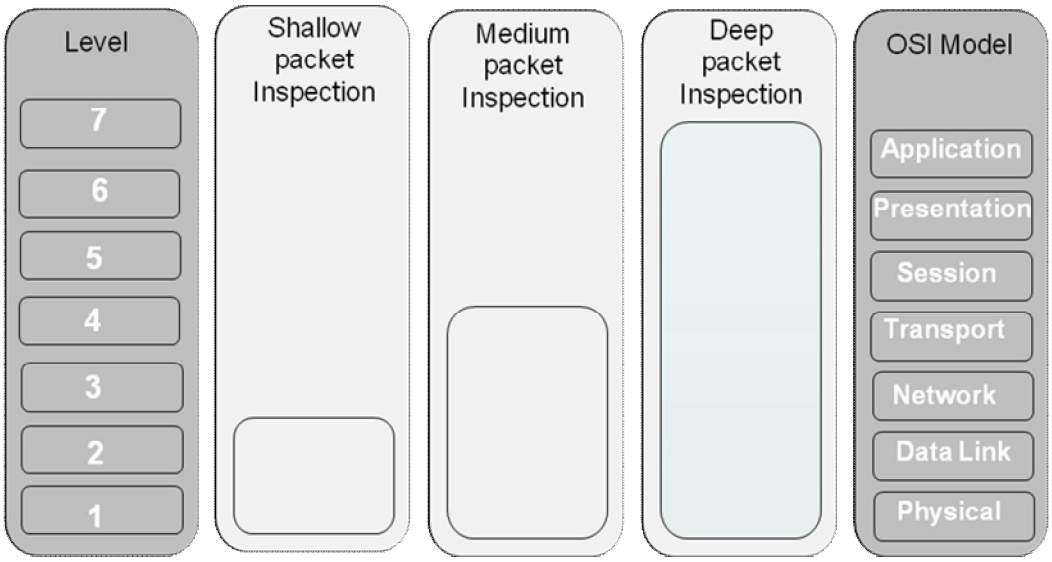

Figure 2. Packet Inspection Depth

2.1 Shallow Packet Inspection

Shallow packet inspection (SPI) examines the headers of the packets (which is the information placed at the beginning of a block of data, such as the sender and recipient's IP addresses), as opposed to the body or “payload” of the packet. This kind of packet inspection allows the communications to remain 'virtually anonymous' since the content of the packets is not observed, and the information in the header is used only to route the packet.

SPI technologies drive the (relatively) simplistic firewalls found in the recent generations of operating systems, such as Windows XP, Windows Vista, and OS X. These firewalls stand between a particular client computer and the network that it is attached to. They limit user-specified content from either leaving, or being received by, the client computer. When a server sends a packet to a client computer, SPI technologies examine the packet’s header information and evaluate it against a blacklist. These firewalls, specifically, focus on the source and destination IP address that the packet is trying to access. If the packet’s header information is on the blacklist, the packet is not delivered. When SPI technology refuses to deliver a packet, the technology simply refuses to pass it along without notifying the source that the packet has been rejected.

SPI cannot read beyond the information contained in a header and focuses on the second and third layers in the OSI model. SPI examines the sender's and receiver's IP address, the number of packets that a message is broken into, the number of hops a packet can make before routers stop forwarding it, and the synchronization data that allows for reassembling the packets into a format that the receiving application can understand. SPI cannot read the session, presentation, or applications layers of a packet; it is unable to peer inside a packet's payload to survey the packet's contents.

2.2 Medium Packet Inspection

Medium Packet Inspection (MPI) is typically used to refer to 'application proxies', or devices that stand between end-users' computers and ISP/Internet gateways. These proxies can examine packet header information against their loaded parse-list.When a packet enters the proxy, it is analyzed against a parse-list that system administrators can easily update. A parse-list allows specific packet-types to be allowed or disallowed based on their data format types and associated location on the Internet, rather than on their IP address alone.

MPI devices can read the presentation layer of the packet's payload and identify facets of the application layer. Using MPI devices, administrators could prevent client computers from receiving flash files from YouTube, or image files from social networking sites. MPI technologies can prioritize some packets over others by examining the application commands that are located within the application layerand the file formats in the presentation layer.

MPI devices suffer from poor scalability which limits their usefulness for ISPs, where tens of thousands of applications can be transmitting packets at any given moment.

2.3 Deep Packet Inspection

Deep Packet Inspection (DPI) technologies are intended to allow network operators precisely to identify the origin and content of each packet of data that passes through the networking hubs.

Whereas MPI devices have very limited application awareness, DPI devices have the potential to look inside all traffic from a specific IP address, pick out the HTTP traffic, then drill even further down to capture traffic headed to and from a specific mail server, and can then reassemble e-mails as they are typed out by the user.

DPI devices are designed to determine what programs generate packets, in real-time, for hundreds of thousands of transactions each second.

3.0 Applications of DPI

DPI can be used by public and private entities to view the contents of packets of information being sent over the Internet, and act in various ways on this information. Though, it was originally intended to be a mean of managing the network to safeguard Internet users from malicious programmes, being sent over the Internet, by intercepting them before they reached the end-users. Now the technology is considered for other uses or applications such as network management, government surveillance, targeting advertising and dealing with copyright infringements.

A brief overview of some of the DPI applications is given in the following paragraphs:

3.1 Network security

DPI was originally developed to secure local area networks (LANs), which are used to cover small geographical areas such as a company or university, in order to ensure there is no unwanted traffic coming in from outside the network. This task used to be accomplished by firewalls, but due to developments in Web applications the delimitation between the internal LAN and the external Internet is not so well-defined, and so network administrators must now fully inspect the data coming in and out of the LAN to achieve this.

DPI equipment allows network operators to detect and intercept recognized forms of mal-ware (viruses, Trojans, worms, and other dangerous code) before it reached their customers or employees.

3.2 Network management

DPI can be used for the purpose of network management, which involves various functions such as ensuring a basic quality of service (QoS) for the end-users, preventing congestion on the network, and facilitating the creation of different packages of Internet access for consumers. It enables ISPs to discriminate among different types of traffic streams to try to maintain quality of service standards, or to throttle down “excessive” traffic, such as peer-to-peer file-sharing or voice over IP calls on mobile networks.

a. **Content optimization:** DPI can be used for optimization of contents by way of acting as proxy and modifying contents such as by reducing still and video image quality, reformatting web pages for mobile devices, and other techniques as per the bandwidth available and device constraints so that users can enjoy content with acceptable performance than otherwise.

b. **Billing and Metering applications**: DPI can be used to count the volumes or rates of traffic, but with more complex schemes to account for a mixture of free, partner, and paid traffic to support schemes where a subscriber's traffic may be capped to a certain volume, may be paying by the byte, or other schemes (such as payments between a service provider and content provider for certain types of traffic and/or application usage).

c. **Application Distribution and Load Balancing:** DPI technology can be deployed to examine packet content and re-write the packet to direct the packet to a different destination for purposes of load balancing, fault tolerance, etc.

d. **Network and subscriber analysis** : DPI based applications can help operators in gauging the overall health of the network by pinpointing performance and capacity as well as provide the service provider with a greater understanding of their subscriber's behaviour that may be used to enhance marketing revenues.

3.3 Monitoring and interception:

DPI equipment can be used for real-time monitoring and interception purposes to comply with the surveillance requirements of Internet communications specified by LEAs.

3.4 Targeted advertising

DPI can enable ISPs to inject advertisements into websites that match the assumed interests of the users. Organisations can target advertisements to Internet users depending on their interests, which are defined according to the users' web-browsing habits. Data on what users are looking at on the Internet is gathered by ISPs, which is analysed and used to show individualised advertisements which follow users across the Internet and directly pertain to his or her demonstrated interests. ISPs are particularly well-placed to facilitate this kind of advertising as they have access to all their subscribers' web surfing data being transported by their network.

3.5 Copyright enforcement

A number of big players in the content industry are pushing for mandatory filtering of copyrighted material that is shared on peer-to-peer platforms. ISPs can use DPI filtering equipment that would automatically detect and block unauthorized sharing of music or video files.

3.6 Content regulation

Content regulation: DPI can be used to recognize and block access to content deemed illegal or harmful. e.g., filtering child-abuse websites, censoring anything that is considered a threat to the government and public stability

4.0 Issues related to DPI

In addition to the various benefits enumerated above, the DPI usage has been claimed to be violation of privacy, transparency, free expression, competition, etc. The discussion on these issues generally leads to the conclusion that DPI is neither per se legal nor illegal – the legality of DPI depends on who is using it,how it is being used and the purpose for which it is being used.

5.0 Standardization activities at ITU

At present in ITU the work is in progress on following two Recommendations:

i. ITU-T Recommendation Y.dpireq "Requirements for Deep Packet Inspection in NGN"

The scope of this Recommendation covers the specification of requirements for Deep Packet Inspection (DPI) in NGN, addressing in particular aspects such as application identification, optional flow identification, monitored traffic granularity, signature management, reporting to the network management system (NMS) and interaction with the policy decision functional entity. It also identifies the requirements associated with traffic in non-native encoding formats (e.g., encrypted traffic, compressed data, transcoded information, etc.).

DPI application scenarios and complementary information such as example rules for packet identification, policy enforcement process, policy specification languages, DPI in layered protocol architectures, formal specification of major terminology and illustration of terminology is also included in the document.

ii. ITU-T Recommendation Y.dpifr "Framework for Deep Packet Inspection"

This document is in initial stages and it intends to provide a framework outlining the major boundary conditions of a DPI deployment in a network infrastructure. It also covers Protocol architectural framework and provides an overview of typical network applications with respect to required involvement of "packet inspection levels".

6.0 Conclusion

DPI introduces intelligence into the internet network, which used to simply transport the packets irrespective to its contents, facilitating comprehensive surveillance and discrimination of data packets moving through the network. Internet Service Providers who use DPI can effectively monitor, speed up, slow down, block, filter, or otherwise make decisions about the traffic of their users, based on knowledge of what kind of information they are transmitting. This could potentially have a major impact on privacy, the free flow of information, intellectual property protection online, network security and virtually all other internet governance issues.

Glossary

DPI	Deep Packet Inspection
MPI	Medium Packet Inspection
HTTP	Hyper Text Transfer Protocol
IP	Internet Protocol
ISP	Internet Service Provider
ITU	International Telecommunication Union
LAN	Local Area Network
LEA	Law Enforcement Agency
NMS	Network Management System
OSI	Open Systems Interconnection
QoS	Quality Of Service
SPI	Shallow packet inspection

References

1. http://www.itu.int/ITU-T/studygroups/com13/
2. Deep Packet Inspection:Technology, Applications & Net Neutrality by Klaus Mochalski, Hendrik Schulze
3. Deep Packet Inspection in Perspective:Tracing its lineage and surveillance potentials by Christopher Parsons