Introduction to Deep Packet Inspection

by Grant Kirkwood, CTO, PacketExchange

> *Grant Kirkwood is the CTO at PacketExchange, the leading provider of customized Ethernet network solutions for business. With a private, secure global fiber optic network powering a comprehensive suite of Internet and private Ethernet network services, PacketExchange allows companies to purchase their entire network infrastructure as a fully managed service – making the need to maintain expensive network assets in-house obsolete. For more information about Network Infrastructure as a Service (NIaaS) please visit www.packetexchange.net.*

Deep Packet Inspection – the very words can cause a wide range of responses from IT administrators, network engineers, politicians, policy makers and savvy consumers alike. Advances in computing technology have put a level of traffic analysis previously unattainable within reach of the IT department, but the implications of DPI go far beyond enterprise security.

What is it?

Data communications, whether on the Internet or on private networks, is almost universally transmitted using the Internet Protocol – IP. One of the inherent features of IP is that streams of data are broken into packets for transmission. The reasons for this are numerous and have to do with the architecture of the Internet, which we'll not discuss here. But to understand Deep Packet Inspection ("DPI"), an understanding of the composition of these packets is important.

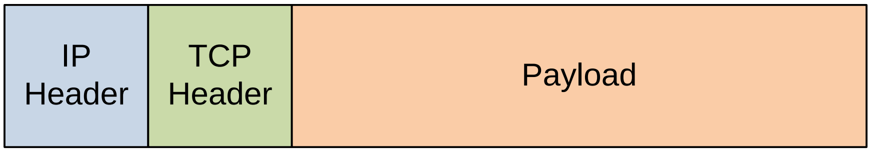

Figure 1: simplified IP packet construction

An IP packet contains several component parts. The first two parts in the diagram above are called "headers" – that is, the delivery information that tells network equipment how to process the packet. Basic routing information such as the source and destination addresses of the packet lives in the header. This information tells routers and switches how to forward the packet along to its destination. Every IP-aware device that the packet traverses must "inspect" this part of the packet to know how to handle it, but these devices do not need to look any further at the content of the packet to perform their routing function.

Further into the packet header is protocol-specific header information. Transmission Control Protocol (TCP) and others have information about the state of sessions between two hosts in this part of the header. TCP is designed to ensure that packets are properly reassembled back into a stream of data when received, and that all packets are verified as having been received. Two computers exchanging data have a constant stream of two-way messages checking to make sure packets are being received. This data lives in this deeper part of the packet header. Routers and switches do not need to know the contents of this information, but firewalls often do. These TCP messages control the flow of data between hosts, and each stream of data forms a "session." Use of this information is considered to be “shallow” packet inspection, although it’s generally referred to Stateful Packet Inspection (hence the term Stateful firewall). By looking at this level of the packet, firewalls and other security appliances can create a profile or “signature” of the traffic, allowing more intelligent decisions about data flows between two hosts.

Beyond the packet headers exists the actual user data being transmitted, referred to as the payload. The payload contains the application-layer information. Images, music, email – all the content is contained within the payload. Inspection of this data is referred to as Deep Packet Inspection.

DPI can mean several things – the “depth” of the inspection can vary depending on how the data will be used. Service providers might want to collect statistics on the types of traffic on their network in order to build quality of service policies. On the other hand, governments or law enforcement entities might want to inspect the actual content of the data for political or public safety reasons.

Who’s using Deep Packet Inspection and why?

Enterprise security has traditionally taken a perimeter-based approach. Firewalls and policies are put in place to control what information can pass in and out of the perimeter of the organization. However new methods of accessing enterprise networks have created the potential for vulnerabilities to be exploited and private information to be obtained without the organization’s consent. As laptops are taken offsite and outside the enterprise perimeter, often to insecure locations, the possibility of becoming infected by viruses or worms increases. When that laptop is carried back into the enterprise and reconnected to the network, effectively bypassing the perimeter firewall or intrusion detection system, whatever it became infected with can spread freely on the enterprise network. DPI provides the enterprise the capability to monitor all traffic on the network for potential threats, by monitoring the actual content of traffic between hosts. DPI can also be placed at the edge of the enterprise to monitor or intercept content arriving or leaving the organization.

Legislation in the U.S. and most other countries has required service providers to provide the ability for law enforcement to perform lawful intercept on their networks for some time. Recently law enforcement organizations have been turning to DPI to provide more intelligent collection and analysis of content in order to better carry out their duties. Lawful intercept has long been used as a surveillance technique, but DPI can separate traffic of interest on the fly, greatly reducing the requirement for large amounts of data storage.

However now service providers are beginning to use DPI for their own internal engineering purposes. As ISPs wrestle with exploding traffic (driven by things like peer-to-peer filesharing and online video) and the need to provide varying qualities of service, prioritizing traffic based on its performance characteristics and requirements becomes critical. Peer-to-peer filesharing can consume massive amounts of bandwidth, potentially saturating connections and making real-time communications via applications like Skype impossible. Using DPI allows the service provider to determine what traffic is time sensitive and what isn’t, and they can create Quality of Service (QoS) policies to ensure time-sensitive traffic is delivered first. Conversely, proponents of net neutrality legislation often cite the ability to perform real-time DPI as one of the methods a service provider could use to de-prioritize competitors' traffic on their networks.

Mobile data providers are increasingly using DPI to prioritize (or block) traffic. According to most reports, mobile data is the fastest growing segment of Internet data usage. Providing sufficient backhaul capacity is a significant challenge, as cell sites are often bandwidth-constrained. A recent report confirmed that T-Mobile uses DPI at cell sites to give higher priority to low-latency traffic such as video while de-prioritizing background file transfers. This ensures that mobile video users have the best experience possible, but conversely mobile data DPI can be used for anti-competitive reasons. U.S. wireless carriers have until recently blocked Skype from being used on their networks as it competes with their core voice service.

Perhaps the most controversial use of Deep Packet Inspection is by government, both foreign and domestic. China is the oft-cited example of aggressive content filtering and intercept. The Chinese government uses DPI to monitor and censor network traffic it deems harmful to government interests, and is rumored to have an “Internet police” force exceeding 50,000. Content commonly blocked in China includes pornography and religious material, but also material of political dissent or historical accounts of events such as the Tiananmem Square protests. In the modern information age, the spread of information can prove dangerous to a controlling, suppressive regime, as we've recently witnessed in Egypt.

I had the good fortune of visiting friends in Egypt a few weeks before the so-called “Facebook revolution” began. Egypt is not known to have any content inspection systems in place – despite the well-publicized “Internet shutdown” that took place there (which was accomplished by essentially disabling links to the outside world). As information could essentially flow freely in Egypt, a multi-year campaign started on Facebook resulted in the successful ouster of the country's suppressive regime. The events in Egypt will likely reinforce the DPI activities of governments interested in limiting political dissent.

Following the terrorist attacks of September 11th, the U.S. Congress passed legislation giving the President broad authority to fight a war on terror. This arguably included the ability to conduct warrantless electronic surveillance on suspected terrorists. In the years that followed, it was discovered that the largest telecommunications providers in the country were cooperating with the NSA, allowing them unrestricted access to all traffic, regardless of its source or destination. The government's authority to conduct DPI on U.S. citizens is under intense scrutiny, and in fact since 2007 is the subject of over a dozen ongoing lawsuits against various elements of the U.S. government. The exchange of information within the U.S. is generally considered to be among the least restricted in the world, but despite this DPI is in use by the government on a daily basis and in ever-increasing capacities.

What are the implications of DPI?

The technology to perform DPI at commercially-available prices and levels has only been made commonly available in recent years. Previously, capturing data required a large amount of storage capacity and offline processing to sift through data. Computing systems have now become fast enough to allow for intelligent data analysis in real-time, negating the need for massive storage and enabling parallel response to vulnerabilities or real-time content filtering and censorship.

As the technology progresses, new uses will be discovered for the data made available by DPI. Although plans have been recently suspended by several large consumer ISPs in the U.S. and the U.K., one such proposed use has centered around topical advertising to consumers based on their Internet browsing habits. As consumer ISPs have the ability to perform DPI on all unencrypted Internet traffic, a highly detailed profile and history of a user's Internet usage can be created. This information could be used to target advertising, or more perhaps more significantly, be sold directly to advertisers in a consumer database format. Considering that so much of modern life revolves around Internet-based transactions, the amount of personal data potentially available to broadband ISPs is unprecedented.

DPI has significant implications for the Net Neutrality debate. As Internet traffic continues to grow at an increasing rate, network infrastructure (particularly "last mile" access) will sometimes be unable to keep up. Proponents of DPI as a traffic management tool advocate utilizing it to prioritize traffic based on performance considerations. Real-time video and two-way communications require low latency to function, but file transfers and peer-to-peer filesharing can be done in the background as bandwidth becomes available. DPI allows for the on-demand video to be prioritized over the file transfers, ensuring a quality user experience. However this same functionality gives the service provider the opportunity to deprioritize their competition's traffic, potentially creating an Internet that no longer adheres to the tradition of openness. Net Neutrality legislation attempts to prevent this, but naturally legislation rarely manages to keep up with the onward march of technological innovation.

Whatever the reason for Deep Packet Inspection—enterprise data security, law enforcement, traffic prioritization, information contract, anti-terrorism—the debate about the merits and potential for abuse of this technology is only just beginning.