

Deep packet inspection has been subject to controversial debates about network neutrality and online privacy for the last few years. In this white paper we will argue that DPI as such is a neutral, neither good nor bad technology, and that it depends on the application that utilizes DPI if and how it will affect the Internet and our society. This paper will focus on Internet bandwidth management based on DPI. Interestingly, the technology has been around in other applications such as firewalls and virus scanners for much longer without sparking similar controversy. After a simple technical explanation of what DPI is – and what it is not –, we will straighten some myths and untruths. Future discussions, particularly in the area of bandwidth management, should not focus on DPI as a technology, but on its specific applications. To facilitate these discussions, we will propose a simple system of categories that classify different Internet traffic management schemes according to their impact on net neutrality, market competition and online privacy.

Introduction

New technologies often spark controversy, particularly if their use has a potential impact on our daily lives. The ability – and necessity – to embark on an open discussion before a wider adoption is an important pillar of modern society. One such technology that recently made rather controversial headlines is deep packet inspection (DPI). A number of quite vocal adversaries has presented a host of concerns, some of them reasonable and worth a discussion, but many also polemic and based on false statements or poor technical understanding. DPI has been branded by some as evil technology that could end the Internet as we know it.

This white paper aims to contribute to this debate by first clarifying the technological background from the perspective of a vendor of networking products based on DPI technology, and by second discussing the potential impact the widespread deployment of DPI applications may have on the Internet and society.

Critics often mix up DPI with a network service or function using DPI as its base technology. Examples of network functions using DPI include spam and virus filters, intrusion detection and prevention systems (IDS/IPS), and firewalls, all of which have been around for many years. And there has hardly been a debate about the perils of any of these. So what is happening in the DPI discussion?

The target of it was not so much DPI, but Internet traffic management based on DPI as a new network function – yet another application using DPI. The core claims of its opponents is the alleged violation of privacy and net neutrality. In fact there are other DPI-based network functions that could be seen even more critical than traffic management.

Examples are network monitoring for lawful interception, which can include mass interception and target profiling, and in-line content injection used for targeted advertisement. So it is all about the application DPI is used for, and not the technology itself. Thus it is important to discuss all these applications separately.

This white paper will focus on DPI-based traffic or bandwidth management. After a technical introduction of DPI and DPI-based Internet traffic management, this paper will extensively discuss the benefits and potential dangers of this technology, including the weakening of net neutrality and freedom of speech in the Internet.

Technical Background: What Is DPI?

At first glance, a technical definition of deep packet inspection is straightforward to write down and in fact very simple. DPI systems inspect entire packets traveling the network as part of a communication, looking not only at packet headers like legacy systems, but also at the packet's payload.

The central point of this definition is the inspection of packet payload. While this seems to be quite clear, both terms require a closer look, not least because this payload inspection constitutes the main draw for criticism of DPI technology. The key problem is that Internet packets do not have only a single header plus payload. Instead, there is a packet header and payload at each layer of the multi-layer Internet architecture that can be found in each network-connected host. A detailed discussion of this header-payload dilemma can be found in the boxed text on the following page.

The most useful definition is based on the demarcation line between IP header and IP payload. It is also used in Wikipedia's definition of DPI¹:

"Deep Packet Inspection (DPI) is the act of any IP network equipment which is not an endpoint of a communication using any field other than the layer 3 destination IP [...]. [...] This is in contrast to shallow packet inspection (usually called Stateful Packet Inspection) which just checks the header portion of a packet."

This many-headers dilemma can be confusing, which becomes apparent in the contradiction in the above definition. Its second sentence implies that stateful packet inspection is the same as shallow inspection. However, stateful inspection – or filtering –, as it is commonly deployed in pretty much all of today's firewalls, keeps track of network connections or flows by grouping all packets with the same 5-tuple {source IP, destination IP, source port, destination port, layer-4 protocol}. Port numbers are encoded in the TCP and UDP headers, which are part of the IP payload. This would be a clear violation of the first statement.

Nevertheless, the IP header boundary is the most commonly used limit for packet inspection and is frequently cited by DPI opponents. It is a sensible and understandable position, and even if one chooses to deviate from it, it is still a useful baseline and starting point for any DPI discussion. We will indeed go beyond this very restricted understanding of what inspection operations should be allowed for a packet in transit. Exactly how deep a DPI system has to look, and what data it needs to gather, strongly depends on the application or network function that it is used for.

Myths and Wrong Analogies

Analogies are a favorite instrument to illustrate technical matters for a non-technical audience. It is often a challenge to get them right. And sometimes a poor analogy is intentionally used to convey a personal opinion instead of technical facts – or worse, to evoke fear, uncertainty and doubt.

One such analogy has enjoyed particular popularity in the recent debate about DPI. In its "Memorandum Opinion and Order" from 1 August 2008², the FCC states:

The Header-Payload Confusion

The architecture of any Internet node follows a standardized layering structure. Each layer implements a subset of functions necessary for end-to-end data transmission. Defined interfaces between these layers provide a data hand-over point. In a sending system, each layer receives data via this interface from its upper layer. These data constitute the payload for the current layer. Data are processed and a header is added at the head of the packet. (Sometimes, trailing information is also added to the tail of the packet, usually padding or checksums, but that is irrelevant for our discussion.) This process repeats at each of the layers.

As an example, let us look at the sending of an e-mail. After composing a message and pressing the send button of the e-mail client (e.g. Microsoft Outlook, Mozilla Thunderbird), this is what happens:

1. The message including e-mail-specific header fields (e.g. subject, from, to, cc, bcc, attachments) is encoded in the Internet Message Format (IMF).
2. The IMF-encoded message is sent to the SMTP handler, which in turn encapsulates the IMF payload by adding its header.
3. The SMTP packet is then handed to the sending host's TCP instance, that again adds its header (with port numbers identifying the communicating application, plus other, connection-state and flow control information) to the SMTP payload data.
4. The TCP segment is passed on to the IP instance, that adds an IP header with IP source and destination addresses.
5. The data link layer (Ethernet in most cases) takes the IP packet and encapsulates it in an Ethernet frame, again adding a header with Ethernet addressing information (source and destination MAC addresses).
6. Only now this Ethernet frame is put as an electromagnetic signal, representing the '0' and '1' bit values that comprise the frame, onto the copper or fiber-optical cable.

The same process, only in reverse order, happens again at the receiver.

This description shows that there is no sharp distinction between header and payload in the Internet. An IP packet is an Ethernet frame's payload, a TCP segment (i.e. the TCP packet) is the payload of an IP packet, and so on. So where exactly does 'deep' packet inspection start? While there is no definite answer to this, a demarcation line can be established by looking at the Internet's packet delivery process.

Packets are exchanged between user applications (e.g. e-mail client and server, Web browser and server, or peers in a peer-to-peer network). For packet forwarding in the Internet, however, the applications sending and receiving the packets are irrelevant. Packets are sent from one host (represented by the sender IP address) to another (represented by the receiver IP address). These two addresses are the only information required by Internet nodes (the sending and receiving hosts and the intermediate routers) to deliver a packet.

So one could assume an understanding where only the communicating end systems should look beyond the IP header at TCP/UDP port numbers. That is necessary to deliver data to the correct application, of which several may run on any given host. At any other point along the network path between the communicating hosts, the look-up needs only go as deep as the IP header, as this is all what is necessary to route the packet.

¹ http://en.wikipedia.org/wiki/Deep_packet_inspection, retrieved 1 September 2009

² Document: http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-08-183A1.pdf, linked from http://www.fcc.gov/Document_Indexes/WCB/2008_index_WCB_Order.html, row FCC-08-183, retrieved 1 September 2009

"[...] Comcast opens its customers' mail because it wants to deliver mail not based on the address or type of stamp on the envelope but on the type of letter contained therein."

Later in the same document, a statement by then FCC chairman Kevin J. Martin begins with:

"Would it be OK if the post office opened your mail, decided they didn't want to bother delivering it, and hid that fact by sending it back to you stamped 'address unknown -return to sender'? Or would it be OK, when someone sends you a first class-stamped letter, if the post office opened it, decided that because the mail truck is full sometimes, letters to you could wait, and then hid both that they read your letters and delayed them?"

Unfortunately, that is exactly what Comcast was doing with their subscribers' Internet traffic."

So DPI is like 'opening' a letter and 'reading' its content, right? One could argue about opening because a sealed letter is clearly marked 'private content - do not open', a network packet, if it is unencrypted, is not. But this is debatable. DPI systems, at least those used for bandwidth management as in the case of Comcast, by no means 'read' or even 'understand' the communication content. Instead, they scan for certain markers - or patterns - to classify the protocol or application that generated the packets used to transmit the content. Such systems only find what they are looking for, i.e. if they do not scan for the word 'bomb', they will not know if it is there or not. Or in other words, DPI does not index the content of network packets as search engines like Google do for Web pages.

DPI in bandwidth management systems does not read all packets. Instead, it only scans for patterns in the first few packets of each network flow - about 1-3 packets for unencrypted and 3-20 packets for encrypted communication protocols. The rest is done by a flow tracking - or stateful filtering as known from firewalls. Scanning all packets of a flow would be both unnecessary and rather expensive.

If one really wants to use the 'reading of letters' analogy, it should be postcards instead of letters, and the 'reader' should be one who does not understand the language of the letter and who only scans certain parts of the contents for matching symbols from a list of symbols - or patterns. It is important for our discussion to understand that DPI is not automatically a privacy violation.

Applications and Systems Using DPI

Interestingly, there have been DPI-based applications and network functions deployed in many places across the

Internet for many years without drawing much criticism. Here is a short list of examples:

- E-mail spam filtering
- Anti-virus filtering for e-mail and other Internet content
- Intrusion detection and prevention systems (IDS/IPS)
- Firewalls
- Content caching systems (e.g. for Web pages)
- Network probes for network monitoring and troubleshooting

All these technologies have a certain misuse potential as they all have access to user content data. DPI-based traffic management is just another DPI application on this list and should be treated as such.

Bandwidth Management - A DPI Application



The rest of this white paper will focus on one particular application of DPI: Internet traffic or bandwidth management.

The Big QoS Failure

It is common knowledge that different network applications have varying quality of service (QoS) requirements. For instance, Internet telephony and online games work best under low-latency, low-jitter conditions, but consume little bandwidth. Large downloads are nearly unaffected by latency and jitter and only need as much bandwidth as possible.

Unfortunately, the Internet has so far failed to bring about QoS support. Not that there have been no attempts. ATM³ tried to solve this issue by overturning the entire architecture of the Internet - and failed due to its overwhelming complexity. Then, extensions to TCP/IP were proposed, most prominently Integrated Services (IntServ) and Differentiated Services (DiffServ). The more comprehensive IntServ failed for its poor scalability. The simpler DiffServ failed because it would have required the support by all router hops of an end-to-end communication path, hence the cooperation of

³ ATM: Asynchronous Transfer Mode is a standards suite developed by the International Telecommunications Union (ITU) and the ATM Forum (an industry consortium) with the ultimate goal to replace the entire Internet infrastructure from the core to the end system, including the TCP/IP protocol suite that is the very basis of the Internet.

all network operators along the path – something which is and will be difficult to achieve.

All these legacy mechanisms for providing QoS guarantees to certain types of network traffic relied on marking of specific requirements by the end points. Basically, the sending application would mark a certain packet as, for instance, real-time traffic (i.e. requiring low latency and jitter) with a certain data rate profile. All intermediate routers would then assign the marked packets to a separate queue that gets special treatment according to the required QoS profile. As mentioned above, one challenge that prevented that approach from being widely implemented is the missing cooperation among all network operators. A second problem for such an implementation would be the proper marking of network traffic. The routers along the path would have to rely on the marking as set by the sending application. Without any additional access policing at the network edge, mismarkings cannot be avoided and can potentially mess up the QoS guarantees for an entire traffic class (e.g. if a large download would be marked as time-sensitive voice traffic).

DPI Bandwidth Management

DPI can effectively solve the traffic marking problem by its capability to classify network flows according to the communicating application. This marking would be a network function – as opposed to an end system function – and thus be immune against intentional mismarkings. For this classification to work reliably, DPI is a necessity. Inspecting headers only – looking at port numbers in TCP and UDP headers – does not yield a reliable protocol or application classification anymore, as many modern applications use dynamic ports or even ports that have traditionally been used by other applications (e.g. Skype and other peer-to-peer systems often used port 80 that is historically reserved for Web traffic). Note that even this type of header inspection would already violate the above mentioned DPI demarcation line between IP header and payload. Interestingly, this has never drawn any complaints from privacy or net neutrality activists.

The DPI-based protocol and application classification is achieved using a number of different techniques:

- Pattern matching:
the scanning for strings or generic bit and byte patterns anywhere in the packet, including the payload portion, usually at fixed locations.
- Behavioral analysis:
the scanning for patterns in the communication behavior of an application, including absolute and relative packet

sizes, per-flow data and packet rates, number of flows and new flow rate per application.

- Statistical analysis:
the calculation of statistical indicators that can be used to identify transmission types (e.g. real-time audio and video, chat, or file transfer), including mean, median and variation of values collected as part of the behavioral analysis, and the entropy of a flow.

Once the classification has been done, bandwidth management systems installed at those points in the network that will most likely become a bottleneck during periods of high network load can at least provide ‘soft’ QoS guarantees similar to what DiffServ could have offered. ‘Soft’ means that the guarantees would only be network-local to where the bandwidth management system is deployed. While this falls short of providing full, end-to-end QoS support, it can solve the most common problems in today’s Internet infrastructure: congestions in network access links at the edge (i.e. at the DSLAM or CMTS level), and at transit and peering points.

The following list gives some examples of potential uses for DPI-based bandwidth management. More can be found in the last part of this paper.

- Prioritize interactive real-time applications such as Internet telephony, online gaming and remote access
- Rate-limit bandwidth-intensive applications such as large downloads from peer-to-peer (P2P) networks and Web-based file hosting services during periods of congestion
- Block access to undesired applications such as P2P file sharing in an enterprise environment

DPI bandwidth management is a rather new technology. A commonly cited reason is that powerful-enough hardware to inspect the content of network packets has become available only in recent years. This is wrong. In fact, Internet data rates have been growing faster than the performance of any other information technology, including CPU, memory and bus systems. DPI would have been easier ten years ago using the then available technology than it is today. No, the right reason for the advent of DPI bandwidth management is the continuing increase of data rates along with proprietary, non-standard applications that use arbitrary TCP and UDP ports for their data exchange, defying legacy methods of application classification necessary for network management, troubleshooting and capacity planning.

DPI & Encryption

It is a common claim that encryption and obfuscation prevent DPI systems from being able to classify the encrypted

The P2P Encryption Lie

The common claim among P2P users and DPI opponents that the use of encryption and obfuscation in P2P networks like eDonkey and BitTorrent is a measure to ensure the users' privacy is plain dishonest. Even if encryption is enabled, files are still shared with the general public, so for everybody to download, store and read – of course in unencrypted format. This is also why encryption does not provide any protection against copyright investigations in P2P networks, where investigators use normal P2P clients to participate in the network and download files from potential infringers. The only sensible reason for encryption is the attempt to circumvent bandwidth limitations imposed for P2P transfers by the ISP. However, with modern traffic management systems, which are able to reliably detect obfuscated and encrypted P2P traffic, this measure is totally ineffective.

network flow. While it is true that plain pattern matching does not work with encrypted communication, modern DPI systems go beyond this simple method. They use behavioral and statistical analysis as described above. In fact, encryption has very little effect on the classification ability and accuracy of advanced DPI equipment.⁴

Of course encryption prevents inline systems to 'read' packet content thus protecting the privacy of a communication, at least in cases where information is not shared with the general public (see box "The P2P Encryption Lie").

Benefits of Bandwidth Management

- Can improve the economics of providing access to remote and rural geographic areas
- can improve the average performance for Internet users (at the cost of limiting the resources for a few excessive users)
- Can provide users with a tailored service, including 'soft' QoS guarantees, at a higher or lower price, depending on the required service level; users that only use Web and e-mail would get a lower price; everyone pays only for what they use

Dangers of Bandwidth Management

- Can limit access to certain Internet services (e.g. P2P file sharing)
- Misuse potential for protocol/application censorship
- Can stifle innovation by slowing down capacity extension of the Internet

The Potential Impact of DPI Applications on Society

DPI opponents regularly point out the potentially devastating effect an extensive deployment of such technology would have on the Internet and society in general. Here it is important to differentiate between two main uses of DPI:

- DPI used by ISPs for commercial reasons
- DPI as a technological basis for new regulatory systems

DPI and Network Operator Business Models

Internet service providers have deployed or plan to deploy DPI-based traffic management systems. Their interests are almost always commercial – maintaining or, better, increasing their revenues. Foremost, that means attracting new customers and reducing subscriber churn – the loss of customers to competitors. At the same time, they need to protect their infrastructure investments from congestion by a few users or applications. These two goals are somewhat contradictory, so it is important to find the delicate balance between them. At the end, the use of this kind of DPI traffic management will be regulated by the market. Governments only need to ensure a properly working market with sufficient competition.

Internet Regulation

So far in human history, every new technology, such as the automobile, the television and even the printing press and with it freedom of speech, eventually got regulated. Such regulation always has to take into account the specifics of a new technology. For printing press and television it is sufficient to make publishers and senders liable for the distributed content because it is easy for law enforcement to identify the source. For cars, there is a speed limit (at least in most countries) and tight registration and liability laws.

The same legislative process will happen for the Internet as well. In theory, existing offline laws covering issues like libel, organized crime, terrorism and child abuse already apply. Unfortunately, due to the distributed, multinational nature of the Internet, current national legislation rarely provides a satisfying solution. In many cases, suspects and criminals are beyond the reach of the executive of a country. Solutions can be investments in online law enforcement along with international cooperation agreements, but also new Internet-specific regulation relying on technological advances.

DPI and Regulation

Traffic filtering systems that use DPI have been proposed as a technical regulatory measure in various countries to en-

⁴ In January 2009, the European Advanced Networking Test Center (EANTC) conducted an independent test of DPI system with special focus on the detection capabilities with respect to encrypted P2P protocols. The results clearly showed that the three participating vendors had no difficulties with encryption. The test results are available at Internet Evolution: http://www.internetevolution.com/document.asp?doc_id=178633.



force existing or new legislation. In most instances, DPI is supposed to enable a fine-grained content classification. This goes far beyond what most ISPs use – or plan to use – DPI for, who are usually only interested in protocol or application classification for commercially driven bandwidth management purposes. Examples of potential regulatory uses of DPI filtering systems are:

- Blocking of illegal (i.e., in accordance with local laws) contents such as child pornography
- Blocking of encryption and tunneling systems that render lawful interception systems (as required by many legislations) ineffective
- Blocking of unregulated Internet telephony

Net Neutrality

One of the most common allegations against DPI and DPI-based traffic management is that it violates net neutrality. Yet, there is no clear and generally applicable definition of net neutrality. It is a principle that is described differently by different people coming from different regional, social and scientific backgrounds. Wikipedia provides a short definition⁵:

“At its simplest network neutrality is the principle that all Internet traffic should be treated equally.”

A more precise technical definition would be that all IP packets should be treated equally on a best-effort basis. But net neutrality is not a technical principle but a social paradigm that strives to preserve the Internet in a perceived state of maximum freedom and equality among its participants. Hence, ultimately, society will have to decide about how much freedom and equality there will be in the Internet of the future. Here we will focus on the more technical aspects of net neutrality.

If net neutrality is to be understood in a way that it guarantees equal access to all its users, then certainly the Internet of today is by no means neutral. Usage statistics across many regions gathered over many years agree that less than 20 percent of network users generate over 80 percent of the traffic. This phenomenon can not only be explained by mere differences in demand. Instead, tech-savvy Internet users can get significantly more than their fair share of the available bandwidth, which is particularly critical in times of network congestion when this inequality adversely affects the performance of other users. An example are P2P file sharing applications such as BitTorrent and eDonkey. To maximize download speeds, they open many, sometimes hundreds of parallel connections to many different peers. Legacy client-server applications such as Web browsing only open one connection to a server, or in some cases, a few connections to a small number of servers. The multi-connection applications will also win in the competition for bandwidth in a network without bandwidth management. Worse, they can completely displace low-data rate, real-

TCP and Net Neutrality

The Transport Control Protocol (TCP) is the workhorse of the Internet providing reliable data transport from host to host for the large majority of all traffic. One of its important properties is a mechanism called sender-based flow control which ensures that each sending TCP instance does not overload the network path to the receiver and, on average, gets its fair share of the available bandwidth on that path. TCP provides fairness among data transport connections.

As long as each network application utilizes one TCP connection to send data, this transport fairness also ensures fair resource sharing among applications. TCP has no enforcement mechanism that limits the number of connections per host. This is important because server applications usually serve many clients in parallel, each over a separate connection. But this also means that two hosts can open parallel connections between them to achieve higher transfer rates. The math is simple: opening ten parallel connections provides this aggregate connection with a transfer speed ten times higher than its fair share. And P2P file sharing applications, for example, often open hundreds of parallel transport connections. This is particularly critical in congested networks where this behavior will deprive single-connection applications (which are most other applications) of their fair bandwidth share.

time applications like Internet telephony and online games, effectively rendering them unusable. Hence, an unregulated network cannot be called neutral because it does not guarantee fairness among users.

Bandwidth management can correct this shortcoming of today's Internet. It can enforce a fair bandwidth distribution among network users particularly in times of network congestion. The simplest form would be a fair distribution of available bandwidth among all network users. This protocol- or application-agnostic bandwidth management would indeed not require DPI technology.

However, this may not always be the smartest approach since bandwidth or, more generically, quality of service requirements differ widely between applications. For example, an Internet telephony application requires very little bandwidth but with a guaranteed minimum at all times. A file sharing application requires as much bandwidth as possible, but can sustain periods of very low data rates without noticeable service degradation. As mentioned above, the Internet has failed to provide a QoS reservation and guarantee mechanism. Application-aware traffic management can improve this situation by providing bandwidth guarantees, priorities, or a combination of both. This kind of bandwidth management requires DPI-based classification of application traffic because particularly those file sharing applications that utilize multi-connection mechanisms tend to obfuscate or hide their activities. Generally, DPI and derivative technologies such as behavioral and statistical analysis provide the only way in today's Internet to reliably classify applications and application types.

⁵ http://en.wikipedia.org/wiki/Net_neutrality, retrieved 1 September 2009

The simplest form of such an application-specific traffic management would be the assignment of priorities to different application classes. A ruleset could for instance be:

- Internet telephony (e.g. SIP, H.323, Skype) gets the highest priority
- Interactive applications (Web, instant messaging) get high priority
- Non-interactive applications (FTP, e-mail) get normal priority
- High-bandwidth downloads (P2P file sharing, file hosting⁶) get low priority

It is important to understand that providing priorities to selected applications does not necessarily cause a service degradation. For instance, giving voice traffic a higher priority than P2P will not at all affect the bandwidth available to P2P. This is because only less than 1 percent of all Internet traffic is voice versus at least 50 percent P2P traffic. The voice traffic increase will be unnoticeable and insignificant relative to the P2P traffic volume. The fear that low priority does automatically mean a slower application is unfounded.

However, if there are two types of high-volume applications, for instance P2P and Internet TV, then priorities can indeed have an adverse effect on the lower-priority application. In the specific case of Internet TV, which requires a lot of network resources, this is why most service providers who offer such a service have chosen to build a separate network dedicated to this service only.

Now even if everybody agreed that priorities are a good idea, one open problem remains: who decides what application gets what priority? One option would be to let users pick their priorities themselves. This option has two problems. First, it requires knowledge about the quality of service requirements of applications and network protocols, and second, users would most likely tend to over-prioritize their own traffic. So the other option would be to have the Internet service provider assign priorities. Here it is important that assignments are not driven by the interests of a certain ISP, but only by the QoS requirements of an application or application class. Even an international standardization process is conceivable. The same applies to bandwidth management that goes beyond simple priority management by assigning application-specific bandwidth guarantees.

A totally different solution to this fairness problem among users would be going back from flat rate Internet access fees to volume-based billing. While this would provide for maximum fairness among users – yes, there is a cost per transmitted byte! – and indeed most users, probably over 80 percent, would financially benefit by paying less for their Internet access, it would also severely limit the Internet's potential to foster innovation.

Ultimately, the long-term solution to the net neutrality dispute is rather simple. Governments have to ensure a competitive environment among service providers. And then, the market – including ISP subscribers – can and will decide.

Privacy

DPI as such has no negative impact on online privacy. It is, again, only the applications that may have this impact. Prohibiting DPI as a technology would be just as naive as prohibiting automatic speech recognition because it can be used to eavesdrop on conversations based on content. Although DPI can be used as a base technology to look at and evaluate the actual content of a network communication, this goes beyond what we understand as DPI as it is used by Internet bandwidth management – the classification of network protocols and applications. Other applications of DPI, for instance lawful interception and targeted injection of advertisements, do indeed go further, but they are beyond the scope of this paper.

Ultimately, it is again a matter of regulation and social discourse to decide what levels of DPI and what applications are considered acceptable. But it is also naive to believe that intelligence services will refrain from using the latest available technology for wiretapping. This, too, is a matter of regulation. *Quis custodiet ipsos custodes?*

Content-Specific Filtering

Filtering of data transfers based on their content is one application where DPI goes beyond a simple protocol or application classification. Here, not only the application or communication protocol get classified, but the content that is exchanged. After this classification, certain content types may be blocked. Today, this type of content filtering is usually limited to Web traffic and is only deployed in certain countries.

This DPI application does indeed have a potential impact on net neutrality and freedom of speech and thus becomes a matter of national – and maybe also international – legislation. Every country has its own rules on what is legal and what is not. Freedom of speech is not unconditional even in the USA, meaning there are limits to what kind of content can legally be made publicly available. This kind of regulation of course exists in most countries for non-Internet content. There are age ratings for movies, and one country would certainly not accept the categorization of another country. Access to movies is controlled based on these ratings. There is no similar classification scheme along with access control for Internet content. This is something we could see in the Internet of the future, and whether this is desirable or not needs to be decided by society.

⁶ "File hosting" refers to Web-based services that allow to upload files, including very large ones, and then provide a URL, or link, to that file which can be shared with other users who can then simply download the file by following that link. These services are also known as "direct download links" (DDL). The largest operators of such services currently are RapidShare and MegaUpload.

This debate of content filtering is already happening in some countries. Currently, there are discussions in Germany, France and other countries about technical solutions for filtering of child pornography. This is a good prototype for any Internet filtering discussion because the distribution of such material is clearly outside any freedom-of-speech legislation. No one seriously challenges the illegality of pedophilic material. This allows to focus on technical challenges, their possible solutions, and what advantages, disadvantages and dangers each proposed solution implies.

The current proposals for pedophilia filtering solutions in Germany are not based on DPI. Instead they will use block lists of DNS host names and IP addresses. The proposed law would obligate ISPs to deploy an encrypted black list of host names and URLs provided by the Federal Criminal Police Office (Bundeskriminalamt, BKA) to their DNS servers to send users trying to access these addresses to a block page with a stop sign. This DNS-based access blocking is the minimum requirement of the law. It can be easily circumvented by using external DNS servers. The optional IP address blocking also has its problems. It potentially blocks access to legitimate content that is hosted on the same server as the illegal material.



A danger that is generally seen with such filtering legislation is the future emergence of a wider censorship in the Internet. And indeed, some politicians already call for filtering of other illegal material such as extreme-right propaganda or copyright-protected media files.

A different, to a certain degree already implemented measure is the threat of prosecution. This is similar to traffic laws. Cars, for instance, can also drive faster than the speed limit, and it is also possible to drive the wrong direction into a one-way road. The law is not directly enforced. It could well be that this is a sufficient deterrence also for the Internet if it was implemented with a similar coverage as in the offline world. This would require a huge additional effort in most countries for law enforcement activities in the Internet.

Country Borders & the Internet



It is another common statement that national regulation is impossible to enforce in the Internet. That this is wrong can be seen by the many examples of national access restrictions to certain applications or content. iTunes content availability strongly depends on the country the user is in. Pandora, an Internet streaming radio site, is unavailable outside the United States. So it seems that country borders are existent even in the transnational Internet – if there is sufficient commercial interest.

Admittedly, in all these examples a company stands behind the offerings that can be held responsible for its online activities. DPI technology in combination with traffic management can extend this control to any entity offering online content. If, for instance, a certain kind of content is legal in country A, but illegal in country B, access to such content can be selectively blocked in country B. Whether such a 'nationalization' of Internet access regulation is desirable is again a matter of social discourse and regulation.

Levels of Bandwidth Management



The focus of this paper is Internet bandwidth management based on DPI. The previous sections have explained the technical, legal and social aspects of this technology. In many of the public discussions, the participants are in irreconcilable opposition. Particularly DPI opponents often assume a very extreme position in their arguments. FUD and other scare tactics are no rarity.

We strongly believe that a more differentiated discussion has been long overdue. For this reason we propose a classification scheme with seven levels of bandwidth management – some involving DPI, some not. The following list is in ascending order according to a bandwidth management policy's potential impact on net neutrality. All measures could be deployed separately or in combination.

Independent of the bandwidth management policy implemented by ISPs we strongly believe that this policy should be openly communicated to customers and – more importantly – to prospective customers. This is also where legislation, if deemed necessary, should put its focus on. Instead of trying to define what kind of bandwidth management is acceptable, it should enforce transparency and let the market do the regulation.

Best Effort Service

This has been the status quo in the Internet since its inception. Every packet is treated equally independent of its type or content. In case of congestion at a particular router hop along a network path, packets are randomly dropped depending on their arrival time and router buffer occupancy.

Pros:

- Provides maximum net neutrality according to some definitions
- No additional implementation cost

Cons:

- Prevents the implementation of QoS guarantees
- Unfair to the majority of network users

Per-User Bandwidth Fairness

Currently, the Internet only provides per-connection fairness for the TCP transport protocol as described above. Bandwidth-greedy applications that use UDP for bulk data transfer or open many simultaneous TCP connections can easily circumvent this transport capacity fairness and use more than their fair share of the available bandwidth. A traffic management system can rather easily enforce a per-subscriber bandwidth usage fairness that ensures all users getting on average an about equal share of the available bandwidth, which is particularly important during periods of network congestion.

Pros:

- Heavy users have no negative performance impact on others
- Fair distribution of available resources among all users
- No DPI required

Cons:

- None found

User-Configurable Disabling of Selected Applications

The ISP offers its subscribers the ability to block access to selected protocols, applications or even content as a managed service. Residential customers can use this feature for parental control and enterprise customers for blocking of non-work-related applications. For example, residential subscribers may choose to disable P2P file sharing to avoid prosecution for copyright infringements done by their children. The same could be done in a company network or at a public hotspot to avoid any liability issues for user activities. Also, access to recreational applications (e.g. media streaming, social networking sites, online games) could be blocked for company staff.

Pros:

- Improved security and application control for Internet users
- Protection against copyright liabilities
- Protection against application-specific attacks

Cons:

- Requires DPI equipment

Application-Aware Congestion Management

Based on the fact that certain QoS guarantees (e.g. minimum available bandwidth, maximum delay and jitter, maximum packet loss) are more critical for some applications than for others, an ISP implements a QoS management scheme taking into account the specific requirements for an application or application class. In its simplest form, this could be a tiered priority scheme as in the following example:

- Highest priority: network-critical protocols such as BGP, ICMP, DNS, maybe TCP FIN and ACK packets

- High priority: interactive real-time applications such as VoIP, online games, remote control software
- Default priority: all applications with no specified priority
- Low priority: high-bandwidth applications such P2P file sharing, large Web downloads, NNTP, e-mail

In addition, bandwidth guarantees can be assigned per application – either aggregated for an entire network or even per individual subscriber or subscriber group.

Pros:

- Better congestion protection
- Better QoS for network users with the same available bandwidth
- Better resource utilization at the ISP which can mean lower charges for Internet access service

Cons:

- Low priority applications will get slightly less bandwidth in times of network congestion
- Requires DPI equipment

Tiered Services and Pricing

Internet access fees have seen an evolution from online-time, over data volume charges, to today's prevalent model of flat rates that differ mostly by maximum access data rates. Usage-based charges are still the norm in mobile networks, but even in wireline networks they have reappeared in the discussion due to the huge disparity in data volumes between normal and heavy users. This is a bad idea because – and this is a widely accepted assumption – it would stifle innovation in the Internet.

A possible way out of this dilemma for ISPs and their subscribers is an approach that strikes a balance between flat rates and usage-based charging. The basic idea is to offer customers a choice of which services they require from their ISP – and they are happy to pay for – and which they do not. Below is a short list with examples of different services that could be offered by an ISP with such a tiered services and pricing model:

- A very cheap or ad-financed Web-only service
- A cheaper service that excludes certain high-bandwidth applications
- In addition to the previous service, allow customers to enable excluded services for an additional one-time fee on demand via a customers portal
- A more expensive all-inclusive service
- An expensive business service with QoS guarantees for user-selected applications such as VoIP, the corporate VPN, and business-critical SaaS sites like Salesforce.com

Pros:

- Better, more flexible access services
- More fairness among subscribers (normal vs. heavy users)
- Subscribers get more control over access fees

Cons:

- More expensive for heavy users
- More complex tariff models
- Requires DPI equipment

QoS Guarantees for Provider Services

Triple play providers offering Internet, telephone and TV service over a single broadband connection need to ensure that each of these application classes gets its required QoS parameters. Some run entirely separated networks for each service so that there is no QoS interdependency between them and with third-party services. A less expensive way to solve this problem is to simply prioritize VoIP over IPTV over everything else that runs through a shared pipe. The prioritization only has an effect during network congestion, and it would be limited to the ISP's VoIP and IPTV services.

Higher priorities for a provider's own service always has a certain misuse potential. A priority advantage of the ISP's services over competing, over-the-top third-party services limits competition and could in turn drive up prices. A clear regulation that defines how much resources can be used exclusively by the infrastructure provider versus resources that need to be available for third-party use would be desirable.

Pros:

- Guaranteed performance for providers' business-critical applications
- Better resource utilization, which can potentially mean a cheaper Internet access for subscribers

Cons:

- Misuse potential requires regulation
- Depending on the specific infrastructure, DPI equipment may be required

Revenue Protection and Generation

An ISP blocks services that directly compete with its own, revenue-generating product. If the ISP offers a triple-play package including VoIP and IPTV, and there is a usage-based charge for instance for international calls, services like Skype are a clear competitor and decrease the potential revenues of the provider's service. Free voice services have caused headaches particularly for mobile operators. Customers are asking for data flat rates, and the operators want to offer them for a fee to generate additional revenue, but they fear that Skype and free SIP services will bite big chunks out of their normal voice revenues. This fear has so far limited the introduction of data flat rates in the mobile market.

The networks of mobile operators are also more susceptible to congestion due to their limited capacity. A few P2P users can have a devastating effect on the performance of an entire network cell. Thus, providers may choose to exclude high-bandwidth services that have a negative performance

impact on other subscribers as a form of infrastructure investment protection.

In addition, ISPs have the option to monetize their gateway position between the subscriber and the Internet. By monitoring the online behavior of their customers, they can serve targeted advertisements to generate additional revenue. Transparency – or the lack of it – is a big problem for this kind of activity. The DPI equipment required for this ad injection needs to have special capabilities to extract information on content downloaded by subscribers to serve relevant advertisements. This goes far beyond what DPI bandwidth management systems do, at least for a small subset of the entire traffic as this kind of monitoring is usually limited to Web traffic. On the more open side, the ISP could offer this ad injection as a customer-selectable option that reduces the monthly Internet access fee.

Pros:

- Allows the ISP to monetize on advertisements, which has been traditionally limited to content providers
- Can reduce Internet access fees

Cons:

- Privacy and transparency problems
- Requires special, single-purpose DPI equipment

Feedback Welcome!

The presented list is not meant to be complete, but as a contribution to bring more structure into the public debate about DPI Internet traffic management. Feedback and comments are always welcome.

About ipoque

ipoque is the leading European provider of deep packet inspection (DPI) solutions for Internet traffic management and analysis. Designed for Internet service providers, enterprises and educational institutions, ipoque's PRX Traffic Manager allows to effectively monitor, shape and optimize network applications. These include the most critical and hard-to-detect protocols used for peer-to-peer file sharing (P2P), instant messaging (IM), Voice over IP (VoIP), tunneling and media streaming, but also many legacy applications. For further information, please visit www.ipoque.com.

Contact

ipoque
Mozartstr. 3
D-04107 Leipzig
Germany

Tel.: +49 (341) 59 40 30
Fax: +49 (341) 59 40 30 19

E mail: pr@ipoque.com
Web: www.ipoque.com

Distributed by

BRAIN FORCE

BRAIN FORCE Software GmbH supplies intelligent IT solutions based on best practices, effective services, and innovative products in the business solutions and infrastructure optimization areas. Customers profit from tailored service offerings, flexible solutions, and innovative goods. With its product offering for professional IT processes, BRAIN FORCE is reducing costs and thus contributing to the economic success of customers.

Contact:

BRAIN FORCE Software GmbH
Ohmstr. 12
63225 Langen (near Frankfurt)
Germany

Tel.: +49 (0)6103 906-767
Fax: +49 (0)6103 906-789

E-mail: ipoque@brainforce.com
Web: www.brainforce-channel.com