Spring March 12, 2009

DPI Considered Not Harmful

Matthias Bärwolff

DPI Considered Not Harmful

Matthias Bärwolff*

March 12, 2009

Abstract

There is a strong sentiment both in technology and in policy circles against deep packet inspection by ISPs as a means to augment their services or control and shape the uses of their network. We argue that the premise of full transparency of the network is no more viable, and should give way to a notion explicitly recognising the role of networks in assuming functions above the network layer on behalf of the end-systems, and the end-systems controlling, endorsing, and verifying the process of having functions delegated towards the network. Such a paradigm may be a step towards conceptualising an architecture which allows for a broader division of functions between end-points and networks.

1 Introduction

This brief note deals with the internet tussle space opened up by the interaction of end-users, network operators, and application developers. Our motivating question is that of deep packet inspection (DPI) and the issues that it creates. We conclude that DPI is not a problem that need to be fought, but much rather a sensible means for network operators to control traffic on their networks and implement contract carrier (as opposed to common carrier) functions. It is on the part of end-users and

*PhD student at Technische Universität Berlin; currently as a visiting PhD student at MIT's CSAIL

application developers serving them, that we see problems that have to be addressed by research and policy efforts. While the network neutrality debate had been intended to help solving some of the issues ahead, we feel that it provides a wrong and misleading frame for even discussing the problems.

The paper proceed as follows. First, we will point to DPI as the primary mechanism for ISPs to transparently augment their service to end users. Then, we will discuss the robustness and resilience of the core internet protocols in spite of DPI. And, last, we will briefly touch on the broader implications of the tussle between end-users, network operators, and application developers.

2 DPI Everywhere

It is intuitively obvious that for a router to peek into IP packets is typically so easy as to be "irresistable" (Clark et al. 2002). Strong end-to-end encryption is no intrinsic part of the internet architecture, routes are typically very stable, and unwrapping an IP packet is not much different an operation for a legitimate recipients than it is for an illegitimate intermediary (assuming such dichotomy in the first place). The requisite string matching capabilities are often straightforward, and current generation router equipment is capable of inferring — for all intents and purposes — the type of application and content of IP traffic at rates of multiple million packets per second by inspecting the IP payloads. On that basis, routers can then implement admission and traffic shaping policies.[1]

Although we must be very careful in comparing today's situation with those back in the 1970s and 1980s, it is important to note that discriminat-

[1] See, e. g., the advertised features of one such box at http://www.ipoque.com/userfiles/file/modules_rev2008-07-15_web.pdf.

It should be noted that technically speaking the proper termination of an application protocol is quite different from DPI, and generally more "complete". A middlebox may not get to see all packets of a given data transfer; after all, IP packets belonging to one transfer may take different paths through the internet. Also, there might be considerable state at the application layer that may be hard to maintain at middleboxes not conceptually part of the application layer transaction. However, closer towards the edges the variation in routes typically collapses towards one.

ing packets based on inspecting their payload has not been "invented" by commercial ISPs. Much rather it has been an essential part of the internet ever since its very early days. In the mid-1980s the Fuzzball routers used in the NSFNET backbone actively prioritised TELNET applications at the expense of file transfer and email, not by looking at any priority bits but by inferring the type of application from looking at the TCP port numbers (Mills 1988). While there were some minor side-effects, the result of that scheme was largely positive:

> Customers of the NSFNET Backbone were thrilled when TELNET response dramatically improved after the new scheme was installed. (p. 119)

In fact, it was precisely such "tacit" packet inspection and traffic prioritisation that made the internet appeal to many in the research community who had been used to the then widespread DECNET equipment featuring such prioritisation mechanisms (Clark 2009).

The crucial difference between today and back then is that today it is the extraordinary capabilities in terms of memory and processing power that *allow for* most elaborate traffic policy schemes, whereas back then it was the *limitations* of hardware that *necessitated* basic prioritisation schemes. Today, few if any of the traffic on the internet at large ever encounters congestion.[2] Back then, on the other hand, there were often not enough buffers in routers to allow anything near a laissez faire approach (Rosen 1980).[3]

[2]Note that unless buffer space was truly unlimited, without buffer management there would always be the theoretical possibility of deadlocks stemming from processes taking up all the available buffer space. However, today, buffer space in routers generally accommodates all input and output processes without having to resort to any packet discard schemes. Davie (2003) notes that "[m]any ISPs today continue to run networks at such low utilization that queuing delay is negligible and packet loss due to buffer overflow is almost unheard of" (p. 134).

[3]It is worthwhile to digress briefly and quote Haverty (2009) to give the reader an idea of just how severely limited resources were:

> I remember one router, connecting ArpaNet and Satnet, which at one point had only enough memory to buffer exactly ONE packet. So there wasn't much question of how to handle priority in the "queue". [...] So the queue-management approaches were severely affected by the realities of

However, today as yesterday, packet inspection is a means to other ends, and while many of the policies that can be implemented upon such inspection can have a downside, most have considerable benefits, too (prioritising gaming, voice and other traffic requiring low latencies). It is up to the router operators to resolve the ensuing trade-off, and for the customer (and the application developers) to decide whether to accept that result and what to do about it. E. g., while it is possible to block or slow down P2P traffic in favour of other more time sensitive traffic, it is also possible for a customer to circumvent P2P filters, rendering the ISP's filtering efforts futile altogether.[4]

3 Is There a Problem?

The policy question that follows from the above is whether there is a case for regulatory intervention, an issue very much debated but largely undecided at the current moment in time. As for the positive and normative issues in packet inspection and policy implementations by routers, a breadth of arguments can plausibly be pursued. On the one hand, one may argue that the IP protocol is indifferent about discriminations on the basis of packet inspections. The relevant technical standards say little to nothing about values such as equality, neutrality, and innovation. In fact, at the lowest common denominator IP's "best effort" service offers no service whatsoever. DPI should thus be no cause for concern as long as it does not

---

> the hardware, modem speeds, etc. at the time.
>
> [. . .]
>
> [Details about TOS, TTL, and the underlying networks] were all fodder for the thinking about how to manage the queue. Sometimes, if TOS indicated that the packet needed fast service (real-time voice), but the appropriate output queue was clogged, the "high-priority" packet was discarded — the theory being that it wouldn't get there in time to be useful anyway, so it was better to drop it early and avoid wasting bandwidth downstream. The same kind of argument applied to TTL — "if this packet is so old it's likely not going to be useful when it finally gets there, so let's drop it now."

[4]See the latest developments in the tussle between P2P users and ISPs as exemplified by the introduction of the uTorrent client software that explicitly aims at circumventing traffic restrictions by ISPs.

"break" IP, which is highly unlikely given the dependence on a common network protocol transcending adminstrative and operative domains. On the other hand, some argue a case for "network neutrality", according to which a router should not look at, let alone change or discriminate upon the payload or source of the IP packets. This, too, is compatible with the notion of best effort; however, it imposes constraints upon router operators that do not follow from the common standards and interconnection at the heart of the internet.

In the following section we shall briefly discuss those two perspectives as well as the space in between. We feel that they are best framed by positioning them in the layered architecture of the internet protocol stack, the case for IP resilience at the network layer, and that for network neutrality at above the network layer.

4 The Resiliant Waist of the Internet's Hourglass

The internet is very much a network of privately owned and managed networks. Granted, some networks are heavily subsidised or ran by government conferred monopolies, yet the bulk of today's internet is made of commercial entities. There are no laws that stipulate technical standards or terms of interconnection. If there is one common ground the internet is built on, it is the IP protocol (Postel 1981). Any party connecting to the internet will have to abide by that protocol, otherwise it will not be able to interconnect with others on the internet.

The IP protocol, in turn, is remarkably resilient to change. In fact, it is fair to say that the distributedness of and lack of central control over the internet has led IP to ossify completely (Handley 2006). Even the option fields of IPv4, intended to provide some degree of flexibility to IP, are rapidly losing their meaning in a world where using IP's TOS facilities deems packets fit to be dropped by some routers, and the fast hardware switched lane through routers giving substantial preference to packets using but the most basic header fields. And, not only is the internet protocol heavily entrenched, so are widespread implicit assumptions about transparency, addresses, identity and location, etc., making it harder still to introduce any innovation to the core internet protocols (Thaler 2008).

It is thus the very distributedness and commercialisation that keeps the narrow waist of the hourglass of the IP architecture safe, maybe a little too safe.

Importantly, the impossibility of changing the IP protocol also limits what can be done by means of packet inspection and policy implementations. No matter how complete a router's knowledge about an IP packet's content, purpose, and context (and malicious the intent), at the IP layer all that can be done transparently and in accordance with the IP standard is dropping the packet, delaying or prioritising it, and choosing a next hop to which to forward it. Beyond that, it is also possible for intermediary nodes to change header and/or payload of IP packets. However, such interferences generally violate basic assumptions of the end-points about the functions of the network and are readily detected by trivial end-to-end measures. Thus they are generally only feasible if the end-points agree with such interference, e. g., that by a NAT box.

While there may be (and have occasionally been) instances of discrimination against users and applications (Windhausen 2006), the empirical evidence for material and sustained interferences that go against the interests of the end-points is rather slim. This is not surprising, for deviating in a material way from the behaviour of a "normal" router *necessarily* makes those effect felt by the end-points, applications, or, ultimately, their users. Policies that go against the interests of an ISPs subscribers have thus far been unsustainable, see the Comcast BitTorrent blocking incident of late 2007 and its repercussions.

It is for this very reason that ISPs prefer to increase the value of their subscribers' internet connections by means of discrimination rather than lower it. Examples for such policies are temporary download bandwidth boosts in order to help loading websites faster,[5] and prioritisation of gaming and voice traffic over latency insensitive traffic such as file transfers.[6] Also, ISPs will often want to avoid delivering unwanted traffic such as that created by malicious attack from reaching its subscribers by means of

[5]See, e. g., Comcast's PowerBoost feature discussed at http://ask.slashdot.org/article.pl?sid=08/02/19/0434234.

[6]See, e. g., the policies of UK's PlusNet at http://www.plus.net/support/broadband/quality_broadband/traffic_prioritisation.shtml.

traffic management. This is sensible because ISPs are often in a superior position to do so: any successful measures against such attacks scale much better when placed inside the network (Savage 2005). While this effectively means that ISPs come to act more like contract carriers rather than common carriers, this is precisely what economic efficiency in the absence of a government controlled monopoly would dictate. It is thus plausible to argue that DPI is actually benefiting users and applications to a greater degree than it is imposing harm upon them.

Last, as for speculative fears about the internet losing its innovation potential, we may safely state along the lines of Gillett et al. (2001) that it is actually not interference by ISPs that threaten internet innovation, but rather the pace of internet innovation that threatens any such policies which are not adaptable and controllable enough through the end-users.

5 Outlook: The Case for Regulation

Generally, it is questionable whether ISPs have the power to turn the internet into an integrated monopolised innovation-adverse network, when they cannot even implement modest changes to IP. The core protocols of the internet are thus safe; safer, in fact, than they should be.

And, while ISPs may discriminate IP traffic on the basis of DPI, and even change the IP header and IP payload, there is little point for them in doing so against the interests of the end-points. Any material interference with or alteration of IP traffic or its payload can readily be observed by trivial end-to-end checks, and often be minimised using encryption. Interference with IP traffic will thus generally be aimed at transparently improving the overall service of the network or be done on explicit (often contractual) behalf of the end-points.

Also, there is little reason to worry about the innovation potential of the internet when the network assumes more functions (van Schewick 2007) and the end-points readily hand those functions over to the network (Zittrain 2008). As long as control over the internet is as distributed as it is today, and as long as there are people who build on the flexibility of IP, there is no reason to believe that internet innovation will come to a halt.

As for the need for network neutrality regulation, we feel that it is

time to realise that the notion of a transparent network is untenable. The internet has forever been more than a neutral network; after all, it does *something*, being entirely neutral is thus an elusive notion. Any reasonable notion of network neutrality — one that acknowledges technical, operational, and business grounds for discrimination — will, in fact, become so "differentiated" that a metaphor other than neutrality may well be more suitable to guide the intellectual effort towards understanding and framing the due policy implications of the tussles between end-users and network operators.

The dismissive discussion of the merits of network neutrality above notwithstanding, there are issues regarding that tussle that deserve our serious attention. An important ground for concern is the ignorance of end-users about controlling that the applications they use act truly on their and not others' behalf. For one, many users simply lack the expertise in implementing efficient end-to-end checks that ensure the integrity of the applications' operations. And two, many applications simply do not bother to implement end-to-end checks aimed at giving users control over the operation of applications transcending potentially adverse networks on the internet. With all due respect to Stallman (2008), the principal problem regarding end-users is that they have very little economic incentives to take control over the software they are using. The cost of doing so are often prohibitive, and the benefits of many applications outweigh the costs of giving up control on them. Consequently, much of the control over the functions of an application is not with the users but with application developers, application providers, and network operators — a result, which may be considered unfortunate, but which, as a second best solution, is economically sensible (Clark and Blumenthal 2007).

Moving to a better solution may or may not be possible. Here, we would only like to offer some general thoughts on the issue. It is well-known that individuals are notoriously bad at estimating low probability risks. Generally, they will underestimate the costs of low probability risks. Yet they among the stakeholders in internet applications are the ones who assume the residual risks: it is their transactions that may go wrong because of unexpected application behaviour, and it is their private data that may be misappropriated by third parties with malicious intents. Both application developers and network operators can often disclaim

liability, the former because they have successfully fought any attempts to introduce producer liability for software vendors, and the latter because they can often rely on their status as mere conduits which exempts them from liability.

How this would give rise to any regulatory implications is beyond the scope of this paper. Intuitively, shifting liability to those parties that can best minimise the costs of residual risks, one of which are the ISPs, is tempting (Lichtman and Posner 2004). However, because of the complex interrelations between internet infrastructure, applications, and end-points it is not trivial to judge the net effects of such liability regimes. It is probably fair to say that informed and "empowered" end-users would be in the best position to safeguard their privacy, and the integrity of their transactions. How to get there is the hard part. And, it is unlikely that we can have both: a perfect and an open internet.

6 Conclusion

In conclusion, we would like to reiterate that DPI in and by itself is no more than a mechanism that can be used to implement both useful and malicious policies on the part of network operators. The notion of insisting on the network to be perfectly transparent and neutral is not helpful in dealing with the potential issues DPI gives rise to.

Instead of pursuing dead ends under the umbrella metaphor of network neutrality, we should focus on how to resolve the economic incentive and asymmetric information problems in the tussle space of end-users, application developers, and network operators.

References

Clark, D. D. (2009). Private communication on the dynamics between DECNET and internet.

Clark, D. D. and M. S. Blumenthal (2007). The end-to-end argument and application design: The role of trust. In *35th Telecommunications and Communications Policy Research Conference (Papers online)*, Number

ID: 748, pp. 1–24. http://web.si.umich.edu/tprc/papers/2007/748/End%202%20end%20and%20trust%2010%20final%20TPRC.pdf.

Clark, D. D., J. Wroclawski, K. R. Sollins, and R. Braden (2002). Tussle in cyberspace: Defining tomorrow's internet. In *SIGCOMM '02: Proceedings of the 2002 conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*, New York, NY, USA, pp. 347–356. ACM. http://www.sigcomm.org/sigcomm2002/papers/tussle.pdf.

Davie, B. (2003). Deployment experience with differentiated services. In *RIPQoS '03: Proceedings of the ACM SIGCOMM workshop on Revisiting IP QoS*, New York, NY, USA, pp. 131–136. ACM.

Gillett, S. E., W. H. Lehr, J. T. Wroclawski, and D. D. Clark (2001, Oct). Do appliances threaten internet innovation? *Communications Magazine, IEEE 39*(10), 46–51.

Handley, M. (2006). Why the internet only just works. *BT Technology Journal 24*(3), 119–129.

Haverty, J. (2009). Re: [ih] Secret precedence schemes back then. Email to internet history mailing list at 2009-01-27 22:34. Archive at http://mailman.postel.org/pipermail/internet-history/.

Lichtman, D. G. and E. A. Posner (2004). Holding internet service providers accountable. John M. Olin Working Paper 217, University of Chicago. http://www.law.uchicago.edu/Lawecon/WkngPprs_201-25/217-dgl-eap-isp.pdf.

Mills, D. L. (1988). The Fuzzball. In *SIGCOMM '88: Symposium proceedings on Communications architectures and protocols*, New York, NY, USA, pp. 115–122. ACM. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.29.8650 (reprint with different layout).

Postel, J. (1981). Internet Protocol: DARPA internet program protocol specification. RFC 791 (Standard). http://tools.ietf.org/html/rfc791.

Rosen, E. C. (1980). Issues in buffer management. IEN 182. http://www.postel.org/ien/pdf/ien182.pdf.

Savage, S. (2005). Internet outbreaks: Epidemiology and defenses. http://www.cs.ucsd.edu/~savage/papers/InternetOutbreak.

NDSS05.pdf. Presentation at the 12th Annual Network and Distributed System Security Symposium, San Diego, CA (NDSS 2005).

Stallman, R. (2008). The root of this problem is software controlled by its developer. *Boston Review 33*(2). http://www.bostonreview.net/BR33.2/stallman.php (a response to Zittrain's 2008 book "The Future of the Internet").

Thaler, D. (2008). Evolution of the IP model. http://kathrin.dagstuhl.de/files/Materials/08/08242/08242.ThalerDave.Slides.pdf. Presentation at Dagstuhl Perspectives Workshop: End-to-End Protocols for the Future Internet, June 2008, Dagstuhl, Germany.

van Schewick, B. (2007). Towards an economic framework for network neutrality regulation. *Journal on Telecommunications and High Technology Law 5*(2), 329–392. http://www.colorado.edu/law/jthtl/articles_0502/0502_web_vanschewick.pdf.

Windhausen, Jr., J. (2006). Good fences make bad broadband: Preserving an open internet through net neutrality. Public Knowledge white paper, Public Knowledge, Washington, DC. http://www.publicknowledge.org/pdf/pk-net-neutrality-whitep-20060206.pdf.

Zittrain, J. (2008). *The Future of the Internet—And How to Stop It*. Yale University Press. http://ssrn.com/abstract=1125949.