Marsden, Christopher T. "Quality of Service: A Policy Primer." *Net Neutrality: Towards a Co-regulatory Solution*. London: Bloomsbury Academic, 2010. 56–82. *Bloomsbury Collections*. Web. 26 Apr. 2023. <<http://dx.doi.org/10.5040/9781849662192.ch-002>>.

Downloaded from Bloomsbury Collections, www.bloomsburycollections.com, 26 April 2023, 11:58 UTC.

Copyright © Christopher T. Marsden 2010. You may share this work for non-commercial purposes only, provided you give attribution to the copyright holder and the publisher, and provide a link to the Creative Commons licence.

CHAPTER TWO

Quality of Service: A Policy Primer

NGNs can be defined as networks with a packet-based architecture, facilitating the provision of existing and new/emerging services through a loosely coupled, open and converged communications infrastructure.

The advent of NGNs is bringing forward a series of innovative opportunities but also a greater array of challenges, touching upon competition, interconnection agreements and new business models.¹

OECD

This is the most technical of the chapters, though it is not particularly, even for policymakers with no technical background. I explain what NGNs are, how they will be deployed, the types of broadband available and the business cases for QoS. I then examine whether governments are creating a business case for DPI via intelligence requirements placed on ISPs, together with the laws on interception, data retention and behavioural advertising trials. I conclude by examining the recent breaches of European data privacy law in the United Kingdom, and the likelihood that the equipment for intrusive DPI is bought and compensated for by government, and therefore is likely to be employed by ISPs. In that sense, ISPs are ‘pushing against an open door’ in deploying DPI for QoS.

Given the move towards all-TCP/IP networks, described as NGNs, let us begin this chapter by considering broadband policy from a technical viewpoint.² Yoo summarizes the ‘original’ position of Internet routing:

Transmission Control Protocol/Internet Protocol routes packets anonymously on a ‘first come, first served’ and ‘best efforts’ basis. Thus, it is poorly suited to applications that are less tolerant of variations in throughput rates, such as streaming media and VoIP, and is biased against network-based security features that protect e-commerce and ward off viruses and spam.³

NGNs present enhanced opportunities for content, consumers and ISPs. The OECD stated above that NGNs allow for greater traffic prioritization and

control. Engineers appear unable to agree on whether QoS will be introduced in NGNs, as QoS is a very long-standing issue that as I write has never been implemented with commercial success on the public Internet (it is used for traffic management but not for advanced higher-priority commercial services). Content charging will rely on the type of QoS for the Internet, enabling network providers to discriminate between packets, and potentially even enabling content providers to offer better than best effort quality. There is an extreme net neutrality argument that the Internet should not develop QoS and therefore no filtering of packets or preferential increase in QoS should be allowed.⁴ As Internet engineering has, for many years, pursued the goal of increased reliability, speed and higher bandwidth, outlawing of QoS is vehemently opposed by Clark and Blumenthal⁵ and Crowcroft.⁶ Clark, as an author of the original paper describing the end-to-end design, has particular importance, and I take as given their argument that QoS should be permitted.⁷ The question is, therefore, under what circumstances such fast lanes should be introduced?

3GPP (3rd Generation Partnership Project), the standards body for 3G mobile telephony, has been working since 2000 on a set of standards called IMS.⁸ This is an operator-friendly environment intended to generate new revenue via DPI. In 2005, fixed-line carriers and equipment vendors created IPSphere, a new set of standards for network intercession in IP application flows.⁹ Both sets of standards support the ability to filter and censor by file type and content provider on the Internet. In an extreme case, one could degrade all content that is not tagged as paying a premium carriage fee. This enables the carrier to discriminate and decide which content to delay and which to permit to travel at normal speeds to the end-user. In wired networks, NGNs are where IMS and MPLS (Multiprotocol Label Switching) can create a 'single' traffic stream.¹⁰ John Waclawsky, then a Cisco Systems standards expert, explained that: 'This is the emerging, consensus view: That IMS will let broadband industry vendors and operators put a control layer and a cash register over the Internet and creatively charge for it.'¹¹

This can also lead to a type of arms race because P2P networks encrypt all traffic to prevent inspection in the same way that firewalls on Intranets were evaded using Port80 and other techniques.¹² Odlyzko and Levinson refute many of the arguments for fine-scaled charging that underlie the architecture of IMS and QoS. They note that:¹³

Technology appears to be making fine-scale charging (as in tolls on roads that depend on time of day or even on current and anticipated levels of congestion) increasingly feasible. Standard economic theory supports such measures, and technology is being developed and deployed to implement them. But their spread is not very rapid, and prospects for the future are uncertain ... the case for fine-scale charging is not unambiguous, and in many cases may be inappropriate.

I see no obligation to take any firm position on the issue. What is important in this discussion is the extent of such potential discrimination, and its justification.

So what will happen? Foresight exercises in broadband policy have been carried out in various territories, for instance, the Netherlands¹⁴ and European Commission¹⁵ exercises of 2002–3, and for the OECD.¹⁶ Broadband policy is under continual review as part of the i2010 strategy of the European Council.¹⁷ The need for a fresh approach considering economic and social issues as part of technology deployment has been emphasized by the National Science Foundation in the United States,¹⁸ with emphasis on QoS and security, an approach explained below.¹⁹ Broadband continues to require definition, as a term of art rather than science. It is defined by reference to its inverse: narrowband.²⁰ The Dutch government stated:

‘broadband’ is defined in terms of its functionality, not in terms of capacity or technology. Broadband is a continuously available connection suitable for good quality audio-visual applications and the exchange of large data files. With narrowband, one user at a time can use one service only. Super-broadband allows multiple users to use different services at the same time, via different platforms.²¹

It may be that in time ‘super broadband’ will be redefined as normal broadband. In the Netherlands, current broadband speeds are in the 2–20 Mbps range. In Japan, broadband is offered as standard at 100 Mbps.²² There is already a very wide range of broadband on offer, but it shares the characteristic of being ‘always-on’ and of sufficient speed to permit high-quality files (such as video) to be transferred. The interoperability and sector/service convergence of the broadband world should drive a convergence of

speeds. This is an important issue: it is yet unclear whether the world will separate into different-speed 'lanes' with crash barriers defined by different applications and ways of accessing data, or will converge on a common infrastructure/service offer, with application differences driven by, and responsive to, demand changes.

Domestic broadband connections are configured for downloading rather than uploading. The asymmetry of such connections is often a 10:1 ratio or higher (maximum download speed is 1–2 Mbps but upload speed is 100–200 kbps). This provides a quality-based 'entry barrier' – which is one reason why it persists, and why the price differentials depart from cost-justified tariffs, even accounting for congestion. Contention²³ and 'traffic shaping' affect quality. These ratios reflect the perception that home users are generally passive viewers of content developed by professionals. P2P networks and other user-generated content and application developments are eroding this distinction, but it remains a constant in asymmetrical consumer broadband.²⁴ For example, hosting of an e-commerce website or online marketing activities may lead to the need to upgrade to more symmetrical service. Price differences between symmetrical enterprise and consumer broadband products are very high.²⁵

In the future, where products will be offered at 100 Mbps and more, asymmetrical connections will continue to be an important issue. Technologies and charging models for traffic shaping and capacity management will also develop. Expect significant disparity in rates and capacity available within the developed and developing world, and between rural and urban locations. High-bandwidth services such as enterprise-class high-definition video could be provided via IP video; their efficiency will depend on the number of channels. How this efficiency relates to broadband bandwidth and provisioning architecture will drive investment and research and development decisions in fibre optics and backbone networks above 1 Gbps (1,000 Mbps).

The greater the network capacity, the lesser should be the incentive to shape the IP traffic so that asymmetrical connections are offered (just as a twelve-lane roadway can offer six lanes each way, whereas a three-lane roadway must choose which traffic takes precedence with two or even three lanes). However, building such roads is expensive and depends on a network to provide the traffic at the end points, especially in mobile. Bohlin et al. state: 'Fixed line ... will act as a complementary backhaul long distance and feeder access network.'²⁶ All networks, fixed and mobile, rely, to some extent, on incumbent cable and telecom backhaul services.

Wired Broadband

I explain the key components in the development of copper (DSL), fibre (FTTx) and coaxial (DOCSIS 3.0) fixed infrastructure. I do not consider satellite and powerline communications.²⁷ The focus is on the market development path for high-speed access in the 'local loop', from the local telephone exchange to the enterprise and home. The reason for this focus is that the costs of broadband in the NGN 'core', the network, are already extremely low and, in laboratories, optical fibre produces even more extreme broadband speeds, while last-mile access networks present a bottleneck, in terms of both cost recovery and available bandwidth.²⁸ The installed legacy networks often run at 655 Mbps, with fibre optics and Dense Wavelength Division Multiplexing (DWDM), which has been widely deployed since 2000, whereas the core NGN can run at 100–400 Gbps. Fibre to the local exchange, which is widely but not universally deployed, may not become universal. The area in which regulatory policy problems will remain in pure speed terms is therefore the local loop to the consumer, but pricing of backhaul will, in the short run, be more critical.

The ability to provide a broadband connection is a relatively simple technological challenge and may prove to be a logical marketing exercise.²⁹ The cost of the broadband connection is dependent on the range of support services that are mandated by law (as well as the demand-side priming by content and applications) and the cost of deployment: the roadworks and byway costs of installing new plant, and the types of technologies that are emerging from laboratory testing and are being experimentally deployed. Wireless Local Loop (WLL) for FTTx is a promising hybrid of wired and wireless technologies for less-developed geographies where fibre will not reach the premises, but currently fast connections rely on the expensive business of digging up the roads to install new plant.

Copper Networks

This network typically consists of twisted-pair copper in the local loop. This can be upgraded using a DSLAM (Digital Subscriber Line Access Multiplexer) in the local telephone exchange to offer ADSL (Asymmetric Digital Subscriber Line). This technology can offer up to 20 Mbps downloading speed or 2–4 Mbps symmetrical speed. It can be further upgraded successively to later generations of technology in terms of speed: ADSL2+, VDSL and VDSL2 with variations. This involves both more powerful DSLAMs and the placement

of backhaul fibre optical cable and the DSLAM itself closer to the customer premises (ideally in the street cabinet or even in the basement of the residential/enterprise's premises). At full capacity over short range, VDSL2 can provide 70–80 Mbps throughput.³⁰ This permits multiple high-definition video streams. DSL can have speeds perhaps doubled by Dynamic Spectrum Management (DSM), a technology which achieves 'noise' reduction on lines. Fourteen million lines have been installed in the United Kingdom.³¹ So a lot more speed can be squeezed out of DSL before investing in fibre for Next Generation Access (NGA).

It is often overlooked that the maximum speed of a broadband connection is by no means the only variable in effective file transfer. First, the 'burstiness' (i.e. alternation) of the connection can depend on the network load, but is less important in downloading for delayed usage (e.g. downloading to a Personal Video Recorder [PVR]). Second, networks for distributing files are increasingly efficient: P2P networks take advantage of distributed file sharing and thus prevent overloading at one point in the network. Third, the compression of files, particularly video files, is improving rapidly, with latest generation codecs (compression–decompression) several times more powerful than previous generations. Fourth, the degree of Digital Rights Management (DRM) on the file may delay its route through the network if rules are set up to intercept content that is not 'authorized' to travel from one end-user to another. Fifth, the power of processors both in the network and at the end devices has increased so that disassembly and reassembly of digital packets is far more efficient than with previous generations of technology. For all these reasons, the simple 'peak' speed of a broadband connection is not in itself an indicator of the efficiency of file transfer. The increasing cost-efficiency of file distribution is predicted to continue, but the increased storage cost-efficiency of multimedia-ready PCs and PVRs is likely to be the biggest change in consumers' homes in 2011.³² Given the early stage of consumer adoption of so many technologies, I am cautious about the prospects for rapid demand-led upgrade beyond the copper loop, although I offer snapshots of interesting potential upgrade paths in the sections that follow.

Fibre Networks

The bandwidth advantages of photonics using fibre-optic cable have been clear since the 1980s. The commercial deployment of such technologies outside research establishments and corporate Ethernet networks has been much

slower, a 'textbook' example of the slow diffusion of broadband technologies. Demand studies over time have become more cautious and realistic about the prospects of widespread deployment of fibre,³³ with increased realization that the cost of the fibre itself and the switching gear is trivial compared to other services such as roadworks for installation, billing, network resilience, security, regulation, content and applications and so on.³⁴ By mid-2006, there were 6.3 million fibre subscribers in Japan, with typically 100 Mbps download/10 Mbps upload capacity.³⁵ In Sweden there were 500,000 subscribers.³⁶ Italy's FastWeb was also a leader in Europe.³⁷ Nevertheless, the low penetration thus far of fibre in Europe should make one cautious at predictions that in a decade there will be 'fibre everywhere'.³⁸ McKeown illustrates the enormous potential for photonics to enhance capacity in networks, suggesting that integrated photonics was at the same stage in 2005 as integrated microprocessors were in 1965.³⁹ The potential exists to revolutionize the backhaul and router network, as well as the well-known effects of DWDM in creating much greater efficiency in the network backbone.

Cable Networks

European cable operators have in general been slow in upgrading to broadband.⁴⁰ The attempts in the 1980s to introduce an all-fibre network were abandoned on cost grounds in favour of hybrid fibre coaxial, which is the current standard. It can carry broadband at up to 25 Mbps.⁴¹ The US DOCSIS 3.0 (Data Over Cable Service Interface Specification) standard has the potential to offer roughly fibre-matching speeds, as explained by Burstein in relation to US Passive Optical Network (PON) standards Broadband PON (BPON) and Gigabit PON (GPON) (speeds in megabit per second):

GPON can go 250/125⁴² ... BPON can raise speeds to 100/30 using similar techniques bandwidth sharing techniques ... BPON is 622/155 split up to 32 ways. That's considerably better than the low end DOCSIS 3.0 (160/120), and similar to the high end DOCSIS 3.0 (1000/100), shared to probably hundreds of homes.⁴³

In Europe, where cable companies are generally loss-making and owned by US investors, the standards set in the United States are likely to predominate.

NGNs have been deployed across Europe since 2007, with completion of migrations to all-IP networks scheduled from 2009. The change to NGNs

creates new opportunities for de- and co-regulation as well as new potential for incumbents to create bottlenecks in access.⁴⁴ For broadband technology deployment, the business case for both incumbents and competitor networks depends on their assessment of the various factors at play.⁴⁵ An OECD workshop focussed on these issues⁴⁶ and found that only Japan, for instance, had completed the negotiation of the new Reference Interconnect Offer for NGNs by end-2006.⁴⁷ Wright has illustrated the extent of the interoperability and standards challenges in NGN by reference to the 'iceberg' (an analogy he credits to Cable & Wireless) intended to illustrate how much interconnection depends on cooperation between industry parties.⁴⁸ He has also illustrated the challenges to the existing UK interconnection regime: with 110 logical interconnect points, and 5,500 local exchanges as points of handover, there is a substantially different network architecture for competitors – who access parts of BT's network – to adjust their plans. The environment within which competition is possible is rapidly changing.⁴⁹

Wireless Local Loop from FTTx

In countries and geographies without the medium-term prospect of fibre in the local loop, an alternative which is made commercially available is WLL offered at the point of fibre connection, for instance, the local exchange. This offers the prospect of much higher bitrates than WiMAX because the WLL node also offers backhaul to the NGN network. Because the telecom only has to supply fibre as far as the local exchange, it provides a much more cost-effective way of offering near-FTTx speeds without needing to invest in the 'last-mile'–wired local loop.⁵⁰

Business Cases for Technology to Aid Traffic Discrimination

Fixed wireless offers upgrade paths from current broadband to much greater speeds, which are already available in the core network but depend on upgrading the network to NGNs. In the local loop bottleneck, the options for wired deployment are expensive and depend on anticipated returns on investment, which includes factors such as regulation, security and the ability to discriminate. It is important to consider the technological implications of the various competitive and regulatory options. Return on investment is the key to network operators' decisions to invest in NGNs, and a critical part of that decision is the consideration of whether the network is offered

as a non-discriminating wholesale or retail network, or a ‘walled garden’. These decisions, whether taken as a competitive strategy by the operator or mandated by the regulator, are critical to network architecture.

In brief, networks can build for pure speed or safety/convenience/privacy, and there is a critical cost trade-off between these two poles. Consider the mobile telephony network: it offers almost complete mobility (seamless except on urban public transport), filters out spam and viruses and provides vertically integrated content, applications and services that are robustly matched to the platform. It also currently operates at sub-1 Mbps speeds where users share a base station (mobile networks are discussed in Chapter 7). By contrast, many wired operators offer Ethernet connections to enterprises in their premises. This may appear a crude comparison, but it illustrates the point that strategy determines technology, and not vice versa.

There are incentives for network providers to police the traffic by type, if not by content. It enables the network providers, many of whom also operate their own proprietary applications, to charge a different price to non-affiliated content owners than affiliated owners. This differential pricing could make the profitable operation of non-affiliated providers more difficult. On that basis, the ‘walled garden’ might become the more successful business model. This model not only makes bottleneck control much easier, but also prevents some of the interoperability and open access for users that has led to much Web 2.0 innovation for businesses. Note that (with relative technical ease) users can encrypt their video and P2P content, and thus evade the more basic types of policing carried out by the network provider.⁵¹

Raising the issue of business cases for broadband filtering technologies illustrates the main conclusion from the study of broadband technologies: it is not the speed of broadband which is a challenge, but the cost and quality of the applications and content provided over that broadband network. This is the type of cost–benefit decision that is likely to influence the deployment of next generation broadband, in particular, the incentives and disincentives in security and QoS in fixed-line networks. I note that technology developments will depend on modelling the incentive structure. Challenges and opportunities ahead include truly addressing incentives in designs.⁵² Market evolution is dynamic and complex: the availability and design of a suitable regulatory response must reflect this dynamism, and also the responsiveness of regulators and market players to each other. Therefore, legislation to entrench a particular regime for net neutrality appears

a premature response to the emerging environment. Instead, I will propose in Chapters 6 to 8 that regulators equip themselves with the skills and evidence base to rapidly investigate potential problems of unjustified discrimination.

Government Policy and Deep Packet Inspection

It is important to realize that traffic management takes place on all networks. Routers check packets as they enter the network, and most consumer ISPs try to identify spam email and filter that out of the traffic. Most websites rely on basic traffic measurement information provided by third parties, many of which carry out packet inspection to complete that activity. Trivial and essential though these activities may be, they are inspection – and by a strict measure of the law – interception. That's spying, and government is in the complex position of both making it a criminal offence and requiring it in certain circumstances.⁵³

Government legislation requires that ISPs perform various functions for network quality purposes, and to assist in crime fighting. Governments mandate many QoS standards for operators, and many more are contractually required within service level agreements between operators, and from operator to end-customer. Service level agreements typically offer a market-based equivalent to government requirements for network resiliency and emergency management: the need to keep the service running at capacity and guaranteeing to repair faults in good time. However, beyond these basic requirements, governments also set targets for universal service, number portability between carriers, emergency telephone number availability, power redundancy to maintain the system in case of a general electricity grid fault, billing system accuracy, customer service response and independent arbitration of disputes. At a local level, cable franchises are often required to offer a percentage of their channels for designated programming, and all local roadworks are subject to various requirements of repair, compensation and byway legislation. These are far greater costs than the comparatively trivial costs of running a broadband network at several gigabit per second.

To identify content that must be analysed by an ISP requires a type of control to identify the content as it enters the network: a type of very low 'walled garden'.⁵⁴ Network providers might be expected to argue that such a new control strategy would require extensive and expensive upgrades to their systems. While this is partially true, there are other reasons why providers may have this capability already:

1. DPI is required for government law enforcement and security purposes.⁵⁵
2. Network providers already provide filters against the more obvious types of ‘spam’ – unsolicited commercial communications.
3. Network providers cooperate with national security agencies in tracing potential terrorist activities via their file transfers on the Internet.
4. Network providers can trace non-encrypted VoIP communications and block these packets.
5. Network providers are increasingly adopting QoS architectures for their networks in order to prevent users from over-straining the network at times of peak usage and charge content owners for value-added high-volume services such as video files.

These new developments allow network providers to block file transfers, or more appropriately to charge the users a carriage fee for sending large files. This is the solution adopted by mobile operators and some network providers, and is generally termed as a ‘walled garden’ to denote the isolation of content on the network from other content on the wider Internet. I consider the cost and technology required for this policy below. I am not going to explain technical details of what has changed, except in the broadest terms, because computer scientists – especially Clark, Crowcroft, Clayton and Brown – do that so much better than any lawyer, especially this one. But one central issue stands out from the sometimes confusing technical discussion from a non-specialist point of view: ISPs are using ‘black boxes’ in their networks to look inside the packets that carry our communications, and to examine their content. This change from shallow to deep packet inspection is what lies at the heart of the change in technology and it carries very serious regulatory implications.⁵⁶

The range of network and information security requirements at European level, which must then be implemented as national law in the European countries, imposes far from trivial costs on the network. They are in addition to existing costs for spam filtering, protection against distributed denial of service (DDoS) attacks, phishing and other ‘malware’ that ISPs typically invest in to protect their subscribers from the worst excesses of IP traffic. Security is a growing problem as dependence on broadband (as a key element of the critical information infrastructure) grows and as the Internet moves towards pervasive computing, and the ‘Network of Things’.⁵⁷ There is an escalating arms race as criminal behaviour become more sophisticated.⁵⁸

The objectives and requirements are also changing on both sides: on the attacking side, the evolution from unauthorized access to data corruption, exposure or access denial; on the defending side, the change in data collection, storage, processing locations (centralized or not), data exchange and transfer of liability among buyers, sellers and ISPs. These are all part of the evolutionary play being performed in the market and regulatory 'theatre'. The responses are a mix of technological and 'soft' strategies, developed and deployed at both individual enterprise and wider levels. These include protected 'walled garden' environments and virtual private networks, both protecting consumers behind firewalls that block malware. Loss of Internet openness and end-to-end connectivity is one potential casualty of security concerns. Another is privacy, which in some ways is the mirror of security. An example of this trade-off is the Data Retention Directive 2006 (EC/2006/24).

The Data Retention Directive was approved by ministers in Brussels on 21 February 2006, concluding a lengthy debate inside and outside EU institutions.⁵⁹ It requires intermediaries in public communications networks to retain data on telephony, Internet, email and instant message communications for set time periods, imposing significant data storage costs. The implementation of this Directive in national law was due in August 2007, and has been delayed in several countries by the usual constitutional and parliamentary procedural lapses associated with transposition of European law into national law. Therefore this Directive began to impose costs on corporate actors in 2008. The Data Retention Directive was tabled after the Madrid bombings in March 2004 and then fast-tracked under the British EU Presidency after the London Underground attacks of 7 July 2005. Britain, France and Sweden stressed the need to retain data in order to trace terrorists using modern technology. According to the Directive, Member States oblige communication providers to store citizens' phone call data for six to twenty-four months, but the Directive does not stipulate a maximum time period: some Member States want longer storage periods. The data would only detail the caller's and receiver's numbers, not the actual conversations themselves, while so-called failed calls – calls that do not get through – will not be covered.

New legislation may only be passed into law following a mandatory impact assessment since 2003, as part of the better regulation agenda of the European institutions. Therefore, laws imposing costs on European businesses to provide rights for government or citizens to access data for security or data protection purposes must be costed and justified. The new rules for implementing

directives mean that a publicly available impact assessment was issued for the Data Retention Directive. It is an exemplar of the new approach to European legislation that will be adopted in future for legislative proposals. ETNO (the European Telecommunications Network Operators) indicated that for a one-year retention period, costs would be above €150 million for a large network and service provider (more than €100 million for retention of traffic data for software, server and security, and at least €50 million for annual overhead expense).⁶⁰ The European Commission indicated that:

Indirect benefits could thus include a possible reduction in cybercrime, and in any case a situation which is not worse than the current one for law enforcement efficiency. It will therefore be difficult to assess indirect costs before some years (sic) of implementation of the Directive, in 2010–12.

The cost–benefit analysis of security legislation such as this must compare the long-term effect of legislation in countering crime versus increased costs for telecoms operators in implementing measures to comply with the law. In this case, the costs and benefits were not submitted to an independent evaluation for the impact assessment.⁶¹ It is obviously a case of having to compare apples and spanners, so different are the views of ISPs and security agencies.

These security requirements have increased hard and soft law pressure on ISPs to monitor traffic (together with pressure from copyright holders). Added to this is the greater co-regulation of Internet video in the Audiovisual Media Services (AVMS) Directive 2007.⁶² The broadcast regulators have relatively limited coordination with security agencies and telecoms policy divisions in governments, but all are pushing for greater regulation of their networks by ISPs, with ever-greater censorship in ‘public interests’ which each holds to override free speech and Internet freedoms. In the clichéd words of Washington deregulation lobbyists: ‘To a hammer, everything looks like a nail’ – to a regulator, control of the Internet would be much easier if it looked like telecoms or broadcasting. We return to this issue in Chapter 4, as increased knowledge may transform the liability of ISPs for content that travels over their networks.

Currently, neither is it a requirement for ISPs to notify customers when they block vital P2P-distributed applications, nor are the security reasons given within the remit of typical economic telecoms regulators. This governance

gap is partially overcome by the institutional arrangements in the European Commission.⁶³ Where the security reasons given by ISPs for blocking P2P traffic, which carries malware and other harmful content, is typically the concern of the Ministry of the Interior (in the United Kingdom, the Home Office) and occasionally the Ministry of Industry, the regulator defers to these senior agencies because it has little technically specific knowledge of data security.⁶⁴ In brief, it is the police and the spies who are making policy for network design.

Inside the Black Box: What and Why are They Inspecting?

There are security considerations claimed by ISPs in blocking certain types of content. Spam is routinely filtered by consumer ISPs; and certain types of unencrypted traffic are also allegedly throttled, though given the lack of official monitoring, proof of this remains circumstantial.⁶⁵ In fact, ISPs claim that P2P traffic contains a high proportion of malware, spam and spyware, and therefore it is filtered in the end-user's interest and in conformity with the terms of use for end-users.⁶⁶ Email spam and Web surfing are the vectors for malware, but the ISPs don't block such traffic. Doctorow explains succinctly that there is a standard Pareto distribution of Internet usage:⁶⁷

The reality is that network usage follows a standard statistical distribution, the 'Pareto Distribution', a power-law curve in which the most active users are exponentially more active than the next-most-active group, who are exponentially more active than the next group, and so on ... Think of it this way: there will always be a group of users in the 'top 2%' of bandwidth consumption ... But the real problem of per-usage billing is that no one – not even the most experienced Internet user – can determine in advance how much bandwidth they're about to consume before they consume it.

It is also a classic public goods problem: if no one rations your supply via price or other mechanisms, you will over-consume⁶⁸ (as we say in Liverpool: 'If it's free, I'll have two'). Users, with rare exceptions, do not know how much traffic they are using. This is the core of the problem – it is the P2P protocol that ISPs wish to police, not necessarily the users, but protocols do not wait for regulation. They are refined, encrypted and distributed; in short, they evade regulation. Therefore, ISPs can either throttle users by cutting off

their connections at peak times or once they have exceeded monthly quotas, or try looking inside the packets to see whether they are P2P or not. The latter becomes a very dangerous business to engage in because as we will see, governments are not only encouraging ISPs to look, they are actually subsidizing the DPI equipment to do so – and this sometimes in breach of both European and UK privacy and interception laws (the latter intended to prevent private spying, even if encouraged by government policy).

Phishers, identity thieves, spammers and child pornographers are in control of at least 100 million systems globally. This ‘botnet’ (hijacked and remotely controlled broadband-enabled robot computer) problem is getting worse.⁶⁹ Felten worries that regulators are used to standards bodies and classes of companies, when, for instance, BitTorrent is a protocol, not a company or a single standard.⁷⁰ Kulawiec states the common perspective amongst network architects:⁷¹

You can’t block by protocol, because those same protocols are used for lots of other things. (And even if you did, someone would just invent another protocol.) You can’t block by content, because no software method is even remotely close to reliable enough and all the ones involving humans are either biased, slow, or both. Besides, it’s easy enough to encrypt traffic ... The people who wish to profit by providing this material will have figured out very effective ways to bypass the filters.

Therefore, any attempt to block BitTorrent or P2P more widely will fail because the protocol will route around via encryption or other techniques. Regulators who think they can permit ISPs to ‘solve’ a congestion problem by slowing P2P traffic are therefore fooled into believing a solution can be found when none exists. The idea that the protocol will simply re-engineer itself and pop up elsewhere on the network with even more disruptive effect has been compared to the US concept of failing to control moles digging on your lawn. Unlike gardeners, ISPs have not yet been persuaded to adopt the permanent solution of gassing their moles.

Many assertions are made about the implications of certain types of traffic, but regulators have no basis for deciding if such assertions represent big or small problems. This is an example of how a baseline of traffic and usage would help the regulator understand the importance of claims made by stakeholders. Blocking and other forms of traffic shaping are controversial

because, under current network management tools, it is a blunt tool. For instance, all P2P traffic using a certain protocol may be blocked. Sieradski and Maxwell explain that:⁷²

The Comcast decision does not resolve the question of when DPI will be deemed reasonable or unreasonable. What justification is needed to show that DPI is used in a ‘narrowly tailored’ manner to serve a ‘critically important interest?’ Is it reasonable to invade customers’ privacy when screening for child pornography or pirated content, but not when screening for disfavored applications? Does it matter whether customers have given their consent to screening? This is closely related to questions in the US and in Europe regarding behavioral advertising and targeting, as well as potential collaboration between ISPs and content providers to screen and block file transfers that violate copyright laws.

The E-Commerce Directive (ECD), as we will see in Chapter 4, gives ISPs limited liability where they act as ‘mere conduits’ but not where they have constructive or actual knowledge of illegal content. Their traffic is thus something of a Pandora’s box – if they look inside, all liabilities flow to them, from child pornography to terrorism to copyright breaches to libel to privacy breaches. The Canadian Privacy Commissioner made a submission to the CRTC in the context of its proceedings on traffic management practices, expressing concern that:

DPI can look into the content of the message sent over the Internet. To use a real-world example, using DPI is akin to a third part opening an envelope sent by surface mail, and reading its contents before it reaches its intended destination.⁷³

The submission continues: ‘it is not clear that examination of content is necessary for network management and may constitute an unreasonable invasion of an individual’s privacy’.⁷⁴

Let us briefly review what ISPs already do to filter web traffic. The Internet was designed to allow the efficient transmission of information between networks around the world. Its basic functionality does not include censorship. However, as Brown explains: ‘In a piecemeal fashion, courts and governments in France, Germany,⁷⁵ Switzerland, Finland, the United

Kingdom and Italy have ordered ISPs to filter their users' access.⁷⁶ In a celebrated case, the Superior Court of Paris in 2000 ordered Yahoo Inc. to block the sale of Nazi items to French users. The court found that by providing access to French citizens, Yahoo US was subject to French law, that nothing in the First Amendment to the US Bill of Rights prevented Yahoo from being selective about its auctions and that blocking access to French citizens was technically difficult but not impossible.⁷⁷ Yahoo has attempted to comply, but also obtained a Californian court declaration that ruling had no effect on a US corporation. Court experts gave evidence that Yahoo could block French users from certain auctions with 70–80% efficacy.

ISPs required to block access to specific websites have relied on three crude mechanisms:⁷⁸ IP address filtering, Domain Name System (DNS) poisoning and keyword searching. A fourth possibility is to use a hybrid which combines two or more of these systems. The simplest filtering mechanism is for ISPs to block traffic to and from lists of websites specified by their IP address. Any packets of data with a destination or source address on this list will be dropped by the routers within ISP networks, especially those that exchange traffic with overseas networks. Skilled and determined users can evade such filters by accessing blocked sites using overseas Web proxies, intermediate machines that retrieve Web pages on behalf of users for a number of purposes such as increased efficiency and privacy protection.⁷⁹ Less advanced and motivated users find their access curtailed by IP address filtering. Because WWW servers typically host many (sometimes many thousands) of individual sites, a block on one of those sites will mean none of the other sites hosted on that server will be accessible. When in 2003 the Indian government ordered ISPs to block access to a specific Yahoo! Group, many simply blocked access to the entire domain, cutting access to around 12,000 groups. Pennsylvania's Internet filtering law was struck down in *CDT v. Pappert* (2004) partially because of such overblocking, with no evidence that the Act 'has reduced child exploitation or abuse'. The blocking of 400 sites prevented access to over 1.1 million other sites, whilst being easily circumvented, thus achieving the unintended consequence of censoring ordinary Internet users' experience yet failing to achieve its declared objectives.

DNS servers translate human-readable domain names such as *chrismarsden.blogspot.com* into the numerical IP address equivalent, allowing Web browser software to connect directly to Web servers. ISPs can remove or change the addresses for blocked sites on their own DNS

servers, and block user attempts to connect to DNS servers elsewhere on the Internet.⁸⁰ Like IP address filtering, DNS poisoning will lead to overblocking of sites that share domain names (such as Yahoo! Groups and YouTube videos). It can also block access to non-Web services on the targeted domain such as email and chat. It is trivially circumvented using proxies, and by users requesting Web pages using IP addresses. Both IP address filtering and DNS poisoning need government-compiled or backed lists of servers that should be blocked. Given the speed at which new content appears on the Internet, this is a time-consuming process.

Routers and government-run Web proxies can filter individual pages based on lists of forbidden keywords such as 'Moussavi' in the case of Iran.⁸¹ Search engines have also been pressured by China to filter search results that contain certain keywords such as 'free Tibet'.⁸² The quantity of pages blocked by keyword filters is unlikely to be acceptable outside totalitarian states. However, they can also be used to block access to specific Web pages, rather than entire websites, as with IP address filtering and DNS poisoning. This type of filtering is much more resource intensive than IP address filtering. Keyword filters in routers can be circumvented using proxies that encrypt data sent back to the requesting user, avoiding their detection. Clayton, Murdoch and Watson also found that the specific mechanism used in Chinese networks to block access to pages based on keywords could simply be ignored by Web browsers and servers.⁸³ Hybrid filtering systems have been developed that combine one or more of the filtering techniques described in the preceding text. BT's 'Cleanfeed' system redirects requests for Web pages on a list of specific servers to a keyword filter that blocks access to specific Web pages hosted on those servers. This combines the efficiency of IP address filtering with the precision of keyword filtering applied to specific pages. However, Clayton showed that the BT system could be reverse engineered to find child pornography sites contained on the secret filtering list.⁸⁴

A further problem with most filtering technologies is that they are targeted at content distribution systems, particularly the Web, which are becoming decreasingly popular in terms of total Internet traffic. P2P systems such as BitTorrent are estimated to carry up to 60% of data on large Internet networks. Many P2P systems are designed in response to music and movie industry attempts to block the sharing of copyrighted works and are hence much more robust in the face of censorship. All of these systems avoid reliance on single machines that can be shut down or blocked, and on DNS names

that can be poisoned, and are able to exchange encrypted files that cannot be keyword-filtered by intermediaries.⁸⁵

Interception and the Law

The continued attempts by ISPs to intercept communications on their own networks are by themselves legal under the law of interception. However, they may not allow others to intercept on their behalf or grant to others the right to intercept for their own purposes. The law appears clear on this point, in the United Kingdom at least, as follows. Interception of communication is subject to the Regulation of Investigatory Powers Act 2000 (RIPA) which states in Section 2(2):

For the purposes of this Act, but subject to the following provisions of this section, a person intercepts a communication in the course of its transmission by means of a telecommunication system if, and only if, he –

- (a) so modifies or interferes with the system, or its operation,
- (b) so monitors transmissions made by means of the system, or
- (c) so monitors transmissions made by wireless telegraphy to or from apparatus comprised in the system,

as to make some or all of the contents of the communication available, while being transmitted, to a person other than the sender or intended recipient of the communication.

One element of intercepting is that making available some or all of the contents of the communication to a person other than the sender or intended recipient is not permitted. It is enough that some portion of the contents of the communication is so made. Whether some of these contents (via the channels) are made available to anyone other than the ISP or a third party, they are *available* to someone other than the sender/recipient. The UK test is strict and requires both parties (sender and receiver) to consent.

This bar on interception has not stopped the British intelligence services from continuing activities designed to monitor Internet traffic. Packet-sniffing schemes such as Carnivore, a system implemented by the Federal Bureau of Investigation that was designed to monitor email and electronic communications, have been active since at least 1997. It used a customizable packet sniffer that can monitor all of a target user's Internet

traffic.⁸⁶ A larger-scale operation was built by various Western governments, called Echelon, which was investigated by the European Parliament in a report released on 5 September 2001.⁸⁷ Since the 11 September attacks on New York and Washington DC that year, intelligence agencies' surveillance has vastly increased. To give an example of the scale of the intelligence monitoring of networks, consider the latest UK government programme, explained by a government minister:

the objective of the Intercept Modernization Programme (IMP) is to maintain the UK's lawful intercept and communications data capabilities in the changing communications environment. It is a cross-government programme, led by the Home Office, to ensure that our capability to lawfully intercept and exploit data when fighting crime and terrorism is not lost.⁸⁸

The IMP programme was originally intended to comprise a central database of all UK citizens' private communications, at vast expense estimated at £12 billion in 2008 values, to be authorized by primary legislation in a Communications Data Bill proposed for the 2008–9 Parliament. Following a re-examination by the then-Home Secretary Jackie Smith, the central database was abandoned in favour of data retention by ISPs, which could be accessed by the police under a RIPA order. The ostensible reason for abandoning £12 billion UK database by government was user privacy: 'Most of the proposed 10-year budget for the system would be spent on deep packet inspection equipment that would allow ISPs to tap into third party communications data carried by their networks.'⁸⁹ Jackie Smith stated that providers will be refunded the cost of collecting and processing the data by the government: 'I agree that what we're asking the industry to do is something that will put a burden on them.' The government plans to spend £2 billion for ISPs to intercept details of their customers' emails, VoIP calls, instant messaging and social networking. Under the proposals, mobile and fixed-line operators will be required to process and link the data together to build complete profiles of every UK internet user's online activity. Police and the intelligence services would then access the profiles, which will be stored for twelve months, on a case-by-case basis.

Having explored the Data Retention Directive, RIPA and IMP, let us look at a more commercial use of DPI: behavioural advertising.

Behavioural Advertising: Phorm Trials

The most controversial of all recent attempts by UK network owners to secure revenue from users has been the experiments conducted by the behavioural advertising company Phorm with the UK's largest ISP, BT (and discussions with the next two largest, TalkTalk and Virgin Media⁹⁰), and Korea's largest ISP, Korea Telecom.⁹¹ Phorm employs a user-tracking system by which BT and other ISPs intend to target users more effectively than Google. A variant of this technology was first deployed widely in the United States on wireless ISPs.⁹² Phorm operates a behavioural advertising system somewhat like that of NebuAd in the United States,⁹³ intending to offer its ISPs and website clients a more accurate tracking of customers' Internet use, in order to more closely target advertising and other marketing via that data.

Phorm uses DPI to take a copy of ISP subscribers' Web browsing, in order to insert targeted advertising, as Clayton explains: 'This enables their systems to inspect what requests were made to the website and to determine what content came back from that website. An understanding of the types of websites visited is used to target adverts at particular users.' What this does in advertising terms is provide much more highly targeted adverts based on user's particular browsing histories. This is a highly attractive proposition to advertisers and could make a great deal of money for Phorm and its partner ISPs, if users were notified and agreed in advance to such tracking of their browser histories – it is much more granular than Google's AdWords, for instance, and Google made \$21 billion in advertising revenues in 2008.⁹⁴

The Center for Democracy and Technology (CDT) released a US legal analysis on the legality of Phorm under the Electronic Communication Privacy Act, claiming that 'Based on what we know so far, this new advertising model appears to defy reasonable consumer expectations and may violate communications privacy laws.'⁹⁵ If that process were to be successfully implemented, ISPs in particular could offer websites a more precise method of advertising than Google and other search engines. It could help pay for the next generation of access technologies, in the view of then-UK Communications Minister Lord Carter and the regulator Ofcom (Office of Communications Regulation). Carter told Parliament on 11 March 2009 that Phorm was an interesting new business model and that both it and QoS could help develop faster broadband.⁹⁶

It is both a marketers' dream and a privacy advocate's nightmare because the details that would be accessed by ISPs and Phorm may be illegal even

with the subscribers' consent. Committees of both the US Congress and the UK Parliament are carrying out inquiries into behavioural advertising in 2009.⁹⁷ Article 15 ECD⁹⁸ has also required Member States not to impose undue restrictions on ISPs since 2002:

No general obligation to monitor

1. Member States shall not impose a general obligation on providers, when providing the services covered by Articles 12, 13 and 14, to monitor the information which they transmit or store, nor a general obligation actively to seek facts or circumstances, indicating illegal activity.
2. Member States may establish obligations for information society service providers promptly to inform the competent public authorities of alleged illegal activities undertaken or information provided by recipients of their service or obligations to communicate to the competent authorities, at their request, information enabling the identification of recipients of their service with whom they have storage agreements.

This Article 15 is continually causing Member States to either derogate from ECD in the interests of crime fighting and anti-terrorism law or simply ignore the provision altogether. To an academic, this may appear scandalous, but UK civil servants have privately explained that so many features of wire-tapping and anti-terrorism law have been passed or amended since 2001 that there would by now be several thousand derogations across the Member States, certainly more than the small unit in the European Commission could deal with.

In contrast to this avalanche of control measures, there are European laws which are meant to protect citizens' privacy and liberty. Directive 95/46/EC⁹⁹ is the main law giving Member States responsibilities and citizens data protection rights against corporate actors. This European law sets a high standard for data protection, arguably higher than that in the United States. National data protection agencies have a permanent joint working group (the Article 29 Working Group) and are required to implement the Directive as uniformly as possible. The European institutions are also required by law to consider the Opinions issued on prospective legislation by the European Data Protection Supervisor, established in 2002. Directive 2002/21/EC lays down the tasks of telecoms NRAs, which include cooperating with each other

and the European Commission in a transparent manner to ensure the development of consistent regulatory practice, contributing to a high level of protection of personal data and privacy and ensuring that the integrity and security of public communications networks are maintained.¹⁰⁰ Directive 2002/58/EC (the ‘Electronic Privacy Directive’)¹⁰¹ includes measures intended to prevent spam, supplemented by a 2004 Communication¹⁰² on spam.¹⁰³ The critical test in both 2002/58/EC and 1995/46/EC is that subscribers have to opt for arrangements that may otherwise infringe their personal privacy, and that sensitive data must not be passed to third parties unless authorized and anonymized.

Unfortunately the original Phorm system trials by BT in 2006 and 2007 did not inform users or ask for their permission.¹⁰⁴ It is becoming increasingly clear that the government department responsible for interception of electronic communications was aware of, and tried to provide helpful regulatory guidance on, the trials and the behavioural advertising system. It emerged in April 2009 that the department, when contacted by Phorm in August 2007, had responded by asking ‘If we agree this, and this becomes our position do you think your clients and their prospective partners will be comforted?’¹⁰⁵ It appears that the consultations between the department and Phorm were extensive and amounted to forming a collaborative view of the law, with comments such as ‘My personal view accords with yours, that even if it is “interception”, which I am doubtful of, it is lawfully authorized under section 3 by virtue of the user’s consent obtained in signing up to the ISPs terms and conditions.’ In an email dated 22 January 2008, a Home Office official wrote again to Phorm and said: ‘I should be grateful if you would review the attached document, and let me know what you think.’ The publication of this history of emails resulted in a debate in the House of Lords in 2009. Baroness Miller stated that:¹⁰⁶

The fact the Home Office asks the very company they are worried is actually falling outside the laws whether the draft interpretation of the law is correct is completely bizarre.

As a result of the legal controversy that followed when the trials were made public in early 2008, the ISPs and Phorm itself agreed to insert both notification and consent into any future trial or deployment of the technology, and BT did so for its third trial in December 2008.

In legal terms, the system is not just contrary to permissions required in European privacy law under the 1995 and 2002 Directives, but also unlawful interception under the exclusively UK RIPA. In March 2008, the Foundation for Information Policy Research (FIPR) wrote to the Information Commissioner arguing that Phorm's system involved illegal interception contrary to RIPA.¹⁰⁷ Clayton, FIPR's treasurer (a security expert at Cambridge University), presented a report on the system, which Phorm responded to ensure technical accuracy.¹⁰⁸ Clayton stated: 'Examining the detail makes it crystal clear that our [FIPR] earlier letter came to the right conclusion. Website data is being intercepted. The law of the land forbids this.' Bohm, General Counsel for FIPR, stated: 'the illegality stems not from breaching the Data Protection Act directly, but arises from the fact that the system intercepts Internet traffic.' According to FIPR's legal analysis, BT appear to ignore the fact that they can only legalize their activity by getting express permission not just from their customers, but also from the Web hosts whose pages they intercept, and from the third parties who communicate with their customers through Web-based email, forums or social-networking sites.¹⁰⁹ Bohm's further analysis suggests that Phorm also infringed the database right for some website owners and almost all website owners' copyright, and none of the statutory exceptions in the Copyright Designs and Patents Act 1988 would be applicable.¹¹⁰

The European Commission is tasked with monitoring Member States' implementation of European law, in this case Directive 2002/58/EC, the Electronic Privacy Directive. In response to UK citizens' complaints that ICO was failing to prosecute Phorm and BT for breaching the Directive in not asking consent for the original trial, the European Commission formally asked the UK government to explain why action had not been taken. When the response received was unsatisfactory, it repeated its request for information in stronger terms. When that second response was unsatisfactory, the Commission responded in January 2009 by threatening the United Kingdom with legal action.¹¹¹ In April 2009, the European Commission initiated formal legal action against the UK government for ineffective application of the Directives (note this does not apply to interception laws).¹¹² Commissioner Reding stated:

We have been following the Phorm case for some time and have concluded that there are problems in the way the UK has implemented parts of EU rules on the confidentiality of communications. I call on

the UK authorities to change their national laws ... This should allow the UK to respond more vigorously to new challenges to ePrivacy and personal data protection such as those that have arisen in the Phorm case.

The British government responded to this infringement action within the two-month time limit on 14 June 2009, though as I write the response has not been made public.¹¹³

Conclusion: DPI and Club Goods

It is not only business decisions on QoS that decide the outcomes and deployments of DPI and other tracking or storage equipment in ISP networks. I will argue one simple conclusion holds: there is a ‘perfect storm’ of confluence of these other regulatory instruments driving ISPs to install DPI and other ‘packet-sniffing’ equipment. More than that, there are commercial and regulatory finance incentives for ISPs to install that equipment. Once you have the equipment in place, abandoning net neutrality becomes a no-brainer. The motive is simple: do it because you can. Governments have introduced at the bequest of their security services various measures, including anti-terrorist, anti-paedophile and anti-pirate legislation, increased controls on ISPs and creeping erosion of ‘mere conduit’ status – including examples such as co-regulatory requests for filtering on all ISP consumer users. We will explore this further in Chapters 4 and 6.

Equipment manufacturers – notably those of ultra-fast blade servers and DPI equipment – are using a joint-purpose technology that can be used not only for legitimate traffic management and law enforcement, but also for blocking of content as requested by government. It may only be marginally stretching the truth to describe this as equipment ‘made for China, imported to Europe’. If Buchanan’s insight into the excessive use of unmetered club goods¹¹⁴ holds true for P2P downloaders, perhaps it also holds true for ISPs. If government is willing to pay you for your DPI equipment, dual-purpose technology that can be used for much more than law enforcement purposes, do you have an incentive not to use that equipment to its maximum effectiveness? Government has given ISPs the tools to eliminate net neutrality. How they use these tools therefore becomes a matter of pressing and legitimate public policy.

This page intentionally left blank