BASIC DETAILS

Consultation title: Traffic Management and 'net neutrality

To (Ofcom contact): Stephanie Peat

Name of respondent: Rick Wadsworth, Director Investor and Government Relations

Representing (self or organisation/s): Sandvine Incorporated

Address (if not received by email):

CONFIDENTIALITY

Please tick below what part of your response you consider is confidential, giving your reasons why

Nothing	X	Name/contact details/job title	
Whole response		Organisation	
Part of the response		If there is no separate annex, which parts?	

If you want part of your response, your name or your organisation not to be published, can Ofcom still publish a reference to the contents of your response (including, for any confidential parts, a general summary that does not disclose the specific information or enable you to be identified)?

DECLARATION

I confirm that the correspondence supplied with this cover sheet is a formal consultation response that Ofcom can publish. However, in supplying this response, I understand that Ofcom may need to publish all responses, including those which are marked as confidential, in order to meet legal obligations. If I have sent my response by email, Ofcom can disregard any standard e-mail text about not disclosing email contents and attachments.

Ofcom seeks to publish responses on receipt. If your response is non-confidential (in whole or in part), and you would prefer us to publish your response only once the consultation has ended, please tick here. ☐

Name Rick Wadsworth Signed (if hard copy)

Comments Re:

Ofcom's Discussion Document on Traffic Management and 'Net Neutrality

September 9, 2010

Introduction to Sandvine

1. Sandvine appreciates the opportunity to provide comments in connection with Ofcom's *Discussion Document on Traffic Management and 'Net Neutrality* (the Discussion Document). Headquartered in Waterloo, Ontario, Canada, Sandvine was established in 2001 and employs over 400 people globally. Sandvine's solutions are used by more than 190 Internet service provider customers in over 80 countries, including over 40 in Europe and eight in the United Kingdom alone. Together, Sandvine's customers serve over 90 million fixed line broadband subscribers and more than 200 million mobile subscribers.

2. Sandvine is the global leader in network policy control solutions, which make the Internet better by protecting and improving the Internet experience for subscribers. The solutions comprise network equipment and software that help DSL, FTTx, cable, fixed wireless and mobile operators better understand network traffic, manage network congestion, create new services, mitigate traffic that is malicious or undesirable to subscribers, deliver QoS-prioritized multimedia services and increase subscriber satisfaction. In January 2010, Infonetics Research named Sandvine as the market share leader in the "Standalone DPI Market." Deep packet inspection, or DPI, is one of the enabling technologies of the Internet.

3. Sandvine is very familiar with the Network Neutrality debate. One of Sandvine's major customers, Comcast Corporation, used a Sandvine solution to enable the traffic management technique that was at the centre of the Network Neutrality debate in the United States. In late 2008, Comcast switched to another Sandvine solution, Fairshare Traffic Management, and Comcast still uses that solution to manage traffic today. In 2009, Sandvine made submissions to the United States' Federal Communication Commission's (FCC) Notice of Proposed Rule Making on the Open Internet[1] and the FCC's Public Notice on broadband measurement and consumer transparency in fixed line networks[2] and a similar Public Notice for mobile networks[3]. In Canada, Sandvine made submissions to the Canadian Radio-television and Telecommunications Commission's (CRTC) Review of Internet Traffic Management Practices[4].

4. In this document, Sandvine provides comments to the questions in Ofcom's Discussion Document and attempts to clarify some misconceptions about DPI technology, a core component of Sandvine's network policy control solutions.

i) How enduring do you think congestion problems are likely to be on different networks and for different players?

Network congestion happens

5. Network congestion occurs when more people or traffic use an Internet path than there is capacity to support that use (e.g. demand exceeds supply). Put another way, when incremental demands for network resources yield diminishing or no incremental 'throughput', congestion is present. All access networks share the property that they connect large numbers of points to an unknown set of

[1] Sandvine Incorporated. See http://fjallfoss.fcc.gov/ecfs/document/view?id=7020370020
[2] Sandvine Incorporated. See http://fjallfoss.fcc.gov/ecfs/document/view?id=7020352787
[3] Sandvine Incorporated. See http://fjallfoss.fcc.gov/ecfs/document/view?id=7020514594
[4] Sandvine Incorporated. See http://www.crtc.gc.ca/public/partvii/2008/8646/c12_200815400/1029527.pdf

destinations. Since network operators have no control over the destinations, new popular content or applications can create high congestion in otherwise uncongested networks. This problem is most pronounced in next-generation mobile data networks as the sector size shrinks, and as the number of users present in a sector is an unknown and variable (for example, if an otherwise empty field hosts a concert the next day).

6. Congestion is a product of peak usage, not average usage. At peak usage, congestion occurs at a small number of points in a network, a small number of times, and moves around largely out of the control of the network access provider. In all network technologies, past, present, future, this property persists. In other words, network congestion happens.

7. The Internet's broadband access networks are oversubscribed, such that the network demands of all subscribers could not be simultaneously supported, at a given moment in time. In this respect, broadband access networks are no different than other critical networks we rely on daily, such as phone, electricity, roads and transit networks. An oversubscription model allows for a critical service to be delivered to all at a reasonable price. In the case of broadband, the model gives subscribers the bandwidth they need when they need it, without paying the high cost of a dedicated connection for such bandwidth.

Congestion is not always predictable

8. While peak overall network usage is often reasonably predictable (typically occurring between 7:00 pm to 10:00 pm[5] on a consumer network) and can be planned for, service providers cannot adequately provision their networks for event or location driven surges in traffic. For example, there were widespread reports of mobile network issues during Barack Obama's inauguration as subscribers, called, texted and viewed video over their mobile devices [6]. Also, sudden changes in network demographics or in applications (e.g. YouTube switching to High Definition video), or sudden losses of capacity (such as from underwater network cable breaks) can overwhelm the network to immediately and unavoidably affect congestion.

9. The Internet is heterogeneous. The Internet's constituent access networks, such as cable, DSL, fibre, wireless, and satellite, have different characteristics that create susceptibilities to congestion and performance issues at different network locations at different levels of usage. For example, in cable networks, upstream capacity has traditionally been very limited, in DSL networks certain downstream links can be more subject to congestion and in mobile networks upstream and downstream bandwidth is at a premium as it is fixed by the physical properties of the underlying radio spectrum. DSL is oversubscribed by the number of users per access router typically. A DSLAM, the router which handles the termination of residential DSL lines, may support 1000 users, each with an offered 'up to 24Mbps' rate, for a total of 2.4Gbps of potential bandwidth, and have a 1Gbps overall link capacity, making it 2.4:1 oversubscribed.

10. The Internet is dynamic. Networks are deploying new technologies and expanding in capacity all the time. New applications and new protocols are being delivered over the Internet all the time – each

[5] Sandvine Incorporated. *2009 Global Broadband Phenomena*. See http://www.sandvine.com/downloads/documents/2009%20Global%20Broadband%20Phenomena%20-%20Full%20Report.pdf

[6] Macword.com. *Inauguration taxes mobile phone networks*. See http://www.macworld.com/article/138331/2009/01/cellphone_inauguration.html

with unique and unpredictable demands on network resources. Subscribers are always looking for new ways to use the Internet and are constantly demanding increased quality and dependability from applications. Sandvine's own studies demonstrate that from 2008 – 2009 subscribers' use of real-time entertainment, including bandwidth-intensive video applications, surged by almost 120% at peak network hours[7] – that's just in one year. Such rapid change is difficult to anticipate and will no doubt result in congested networks as operators adapt.

Access network operators cannot control congestion alone

11. The Internet is a commons. While the focus is most often on network access providers, all stakeholders – network operators, application and content providers and subscribers – contribute to network congestion.

12. Content providers enrich the Internet experience, but certain content providers (which may also be application providers) can also be disproportionate users of the Internet commons. For example, according to Sandvine's study a single website – YouTube – represents roughly 5% of all Internet traffic. Facebook, iTunes, Xbox Live and Xbox Live Marketplace all represent over 1% of network traffic[8]. These content providers want high network performance for the subscribers who access their services. They are not necessarily concerned about how their utilization of network resources affects other stakeholders, including other content and application providers, contending for the same shared resources. In fact, they optimise their utilization of the network to optimize the value they extract from it, with no concern for how that affects others.

13. Broadband applications have been developed for an ever-increasing set of uses, with great variation in the demands that they place on network resources. Massively adopted personal communication tools like email and instant messaging provide a high degree of value, yet they put a very light load on the network. Sandvine's study showed that residential email traffic consumes less than 1% of total bytes, globally. In other words, all the emails sent over residential Internet access networks in the world consume fewer bytes than any one of YouTube, Facebook, iTunes, or Xbox Live. By contrast bulk file-transfer or file-sharing applications (such as peer-to-peer (P2P) file-sharing, storage and back-up services and news groups) have typically been used by a much smaller portion of subscribers but represent a much higher portion of network traffic. According to Sandvine's research, P2P file-sharing and storage and back-up services are both top-five applications in terms of the amount of network bytes consumed: combined, they represent over 30% of network traffic[9]. These applications take advantage of weaknesses in Transmission Control Protocol (TCP), one of the core Internet protocols, to use 100% of available network capacity when available, without regard to fairness among users or applications. TCP provides what is called flow-fairness. In general, flow-by-flow, the network enforces some simple fairness. As a workaround, applications like P2P AND HTTP can open multiple flows simultaneously, thereby increasing their share of bandwidth consumption. While bandwidth is important, as will be discussed later, applications also differ greatly in the amount of latency, jitter and packet drops they can endure while still delivering a reasonable quality of experience. These other factors are often more important in determining the users' quality of experience for a given application.

---

[7] Sandvine Incorporated. *2009 Global Broadband Phenomena*. See http://www.sandvine.com/downloads/documents/2009%20Global%20Broadband%20Phenomena%20-%20Full%20Report.pdf

[8] Ibid.

[9] Ibid.

14. Some individual subscribers are concerned about maximizing their own personal utility of the broadband service. There is no incentive for the subscriber to moderate their use of the network without some form of feedback via the service plan definition, cost, structure and enforcement. Sandvine's research [10] demonstrated that the top 1% of users is responsible for approximately 25% of total residential Internet traffic (measured in bytes) and 20% of users are responsible for 80% of that traffic. While peak usage, not total or average usage, is responsible for congestion, it is clear from these statistics that individual users consume the network very differently and are potentially capable of having very different impacts on network congestion.

15. The Internet Engineering Task Force (IETF: the open standards organization that works to develop and promote Internet standards, in particular those related to TCP/IP and the Internet protocol suite) has recognized the shared obligations with respect to congestion in their recent work on Congestion Exposure (ConEx). ConEx begins to investigate technical means to make all stakeholders accountable for their impacts on the Internet commons. ConEx was discussed at a meeting of the IETF, on November 10, 2009 in Hiroshima, Japan[11] and is described this way:

    > "Congestion Exposure (ConEx) is a proposed new IETF activity to enable congestion to be exposed along the forwarding path of the Internet. By revealing expected congestion in the IP header of every packet, congestion exposure provides a generic network capability which allows greater freedom over how capacity is shared. Such information could be used for many purposes, including congestion policing, accountability and inter-domain SLAs. It may also open new approaches to QoS and traffic engineering."
    >
    > "The Internet is, in essence, about pooling resources. The ability to share capacity has been paramount to its success and has traditionally been managed through the voluntary use of TCP congestion control. However, TCP alone is unable to prevent bandwidth intensive applications, such as peer-to-peer or streaming video, from causing enough congestion to severely limit the user-experience of many other end-hosts."
    >
    > "We believe these problems stem from the lack of a network-layer system for accountability -- among all parties -- for sending traffic which causes congestion. We propose a metric where IP packets carry information about the expected rest-of-path congestion, so that any network node may estimate how much congestion it is likely to cause by forwarding traffic. A network operator can then count the volume of congestion about to be caused by an aggregate of traffic as easily as it can count the volume of bytes entering its network today. Once ISPs can see rest-of-path congestion, they can actively discourage users from causing large volumes of congestion, discourage other networks from allowing their users to cause congestion, and more meaningfully differentiate between the qualities of services offered from potential connectivity partners. Meanwhile end-hosts may be freed from rate restrictions where their traffic causes little congestion."

16. So, the Internet "industry" itself recognizes the persistence of network congestion and the commons nature of the Internet that is characterised by the shared obligations of its stakeholders.

---

[10] Ibid.

[11] IETF. See http://www.ietf.org/mail-archive/web/int-area/current/msg02041.html

ii) What do you think are possible incentives for potentially unfair discrimination?

17. In paragraph 1.9 of the Discussion Document Ofcom aptly addressed this question by stating "At the heart of the traffic management and net neutrality debate is a concern that traffic management could be used as a form of anti-competitive discrimination." Certainly traffic management technology could be used for anti-competitive effect, but overwhelmingly, it has not been used for this purpose. Instead, traffic management has been used to efficiently maximize the quality of most network users' experience most of the time.

18. That said, Sandvine has observed that certain areas of the world have laws that most people in the United Kingdom, the EU, and North America would likely consider to have an anti-competitive effect. For example, traditionally in certain parts of the Middle East, network operators have been *required* by regulators to block VoIP traffic[12] . While the effect of the blocking would be seen by most as anti-competitive, it is a problem with the telecom regulations, not the traffic management practice. The traffic management practice is required to comply with the regulations.

19. Ofcom states in paragraph 1.9 that "To date Ofcom has received no formal complaints from industry that require investigation," which mirrors industry experience globally. Quite simply, despite the fact that traffic management policies have been actively deployed in networks for decades, there is a paucity of evidence that network operators have deployed anti-competitive policies. Meanwhile, there is an abundance of evidence that traffic management has been used for positive effect to reduce network congestion, improve the subscribers' quality of experience for time-sensitive traffic, mitigate malicious traffic, and introduce new service tiers. Ofcom acknowledges these benefits in paragraph 2.8 of the Discussion Document, and Sandvine has attempted to provide some examples of those benefits in its answers to question iii) below.

iii) Can you provide any evidence of economic and or consumer value generated by traffic management?

20. Sandvine has experience deploying traffic management solutions to almost 200 network operators globally. These deployments add economic and consumer value in a variety of ways, including:
    a) Improving the Quality of Experience of latency-sensitive applications, such as VoIP or online video gaming so that consumers receive the experience they expect for these highly demanded applications.
    b) Enabling the delivery of new services to consumers, which gives subscribers more choices and creates new bases for competition amongst network operators.
    c) Fairly distributing available network resources among users so that a small group of heavy users do not seriously harm the quality of experience of average or light users.
    d) Optimizing the use of existing network infrastructure to improve the economic return on network operators' existing and future infrastructure investments, which in turn can mitigate the need for subscription price increases.

[12] The Next Web, Middle East. *UAE VoIP still a mess*, March 16, 2010. See http://thenextweb.com/me/2010/03/16/uae-voip-mess/

e) Reducing the amount of malicious traffic on the network, benefitting both consumers and network operators by liberating network resources and improving the performance of valued applications and content.

21. Sandvine can provide some direct examples of these benefits.

Enabling new services

22. For competitive reasons Sandvine is rarely able to disclose the identity of its customers, though there are exceptions. One such exception is Cricket Communications in the United States, who are offering an innovative 3G prepaid mobile broadband data plan.

23. Prepaid mobile data is a very rare offering in North America, so the Cricket offering represents a major innovation for the market. Cricket offers its prepaid service package through Walmart and Best Buy stores in the United States. Subscribers simply purchase the package in store, go home, plug the included modem into their laptop, register and start enjoying high speed Internet.

24. Sandvine's solutions can detect network conditions that trigger policies within the network to help service providers deliver these new services. For Cricket, Sandvine is identifying customers whose service terms are nearing expiry then redirects them to a customer service site where subscribers can conveniently extend their service term.

25. With Cricket's solution, supported by Sandvine's traffic management, consumers can enjoy entirely new service choices which could potentially develop into a significant new market niche in the United States.

26. There are growing examples of such innovative solutions worldwide, such as Rogers Wireless' Social View service for "unlimited access to your favourite Social Networking sites." Demon in the UK just launched a new online gaming service tier, while Sandvine customer Starhub in Singapore had already done so years ago.

Ensuring fair use of network resources

27. In 2008, the FCC made a highly controversial decision to require Comcast Corporation, a Sandvine customer, to change its traffic management policies. While Comcast appealed the decision (and won in 2010), Comcast committed to change their policy by the end of 2008. Their new solution, Fairshare Traffic Management, was also supplied by Sandvine. This solution is described in the FCC filings, and also in an IETF draft[13].

28. According to Comcast's disclosures to the FCC in connection with its Fairshare deployment and the results from it, "the goal of Comcast's new congestion management practices will be to enable all users of our network resources to access a "fair share" of that bandwidth…"

29. Comcast described the solution as follows:

---

13 IETF. *Comcast's Protocol-Agnostic Congestion Management System.*
See http://tools.ietf.org/html/draft-livingood-woundy-congestion-mgmt-09

i. "Software installed in the Comcast network continuously examines aggregate traffic usage data for individual segments of Comcast's HSI network. If overall upstream or downstream usage on a particular segment of Comcast' HSI network reaches a predetermined level, the software moves on to step two.

ii. At step two, the software examines bandwidth usage data for subscribers in the affected network segment to determine which subscribers are using a disproportionate share of the bandwidth. If the software determines that a particular subscriber or subscribers have been the source of high volumes of network traffic during a recent period of minutes, traffic originating from that subscriber or those subscribers temporarily will be assigned a lower priority status.

iii. During the time that a subscriber's traffic is assigned the lower priority status, such traffic will not be delayed so long as the network segment is not actually congested. If, however, the network segment becomes congested, such traffic could be delayed.

iv. The subscriber's traffic returns to normal priority status once his or her bandwidth usage drops below a set threshold over a particular time interval."

30. After extensive testing, Comcast determined the appropriate thresholds for the Fairshare algorithm in its network. Comcast decided that, for step one, upstream port utilization would have to reach 70 percent for it to be identified as being in a "near-congestion" state.

31. For step two, when a subscriber uses an average of 70 percent or more of his or her provisioned upstream or downstream bandwidth over a particular 15-minute period, that user will be identified as a user that could disproportionately contribute to bandwidth and flagged for lower priority should the network location actually become congested.

32. For step four, it was decided that a user's traffic would be released from the lower priority state when the user's bandwidth consumption drops below 50 percent of his or her provisioned upstream or downstream bandwidth for a period of approximately 15 minutes.

33. Based on data collected by Comcast from the trial markets where the new management practices were tested, on average less than one-third of one percent of subscribers had their traffic priority status changed to the lower priority state on any given day. For example, in Colorado Springs, CO, the largest test market, on any given day in August 2008, an average of 22 users out of 6,016 total subscribers in the trial had their traffic priority status lowered at some point during the day.

34. At the date of filing their disclosures with the FCC, Comcast had not received a single customer complaint in any of the trial markets that could be traced to the new congestion management practices, despite having broadly publicized its trials.

35. On September 21, 2009, in a speech to the Brookings Institute that launched the FCC's Notice of Proposed Rule Making on the Open Internet, FCC Chairman Genachowski acknowledged the value of a Fairshare-like solution when he stated, "During periods of network congestion, for example, it may be appropriate for providers to ensure that very heavy users do not crowd out everyone else."

iv) Conversely, do you think that unconstrained traffic management has the potential for (or is already causing) consumer/citizen harm? Please include any relevant evidence.

36. As Ofcom pointed out in the Discussion Document, there have only been two cases where it was concluded that a traffic management practice resulted in consumer or citizen harm. In the Madison River case, a court decided that a local US telecoms operator denied access to VoIP services. In the Comcast case, the FCC decided that Comcast was blocking certain P2P traffic. For their part, Comcast pointed out that even for the most heavily used P2P protocols, more than 90 percent of the flows were unaffected by the congestion management and that most of their customers using P2P protocols to upload on any given day never experienced any delay at all. The regulator's decision was subsequently overturned by the court. Other cases, such as the CRTC proceedings on Bell Canada's traffic management practices concluded that there was no consumer harm.

37. While the existence of even a single case demonstrates the *potential* for harm, the paucity of cases demonstrates that harm is not actually being caused in any meaningful way. Sandvine alone has over 190 network operator customers and only one of these customers has ever been sanctioned for "unreasonable" traffic management practices, and that decision was highly controversial and ultimately overturned.

An Unmanaged Network is not a Neutral Network

38. So while little harm has ever come as a result of traffic management, it is clear that significant harm to consumers can result from an unmanaged network.

39. Sandvine issued its 2009 Global Broadband Phenomena study[14] in October 2009, based on network data gathered during September 2009. The study consisted of analyzing data from more than 20 cable and DSL service providers' networks totaling 24 million subscribers. The networks were distributed across five regions: North America, Europe, Caribbean and Latin America, Asia-Pacific and Africa. To Sandvine's knowledge, it is the most comprehensive and diverse study of its kind ever prepared. Sandvine intends to release an updated report in 2010.

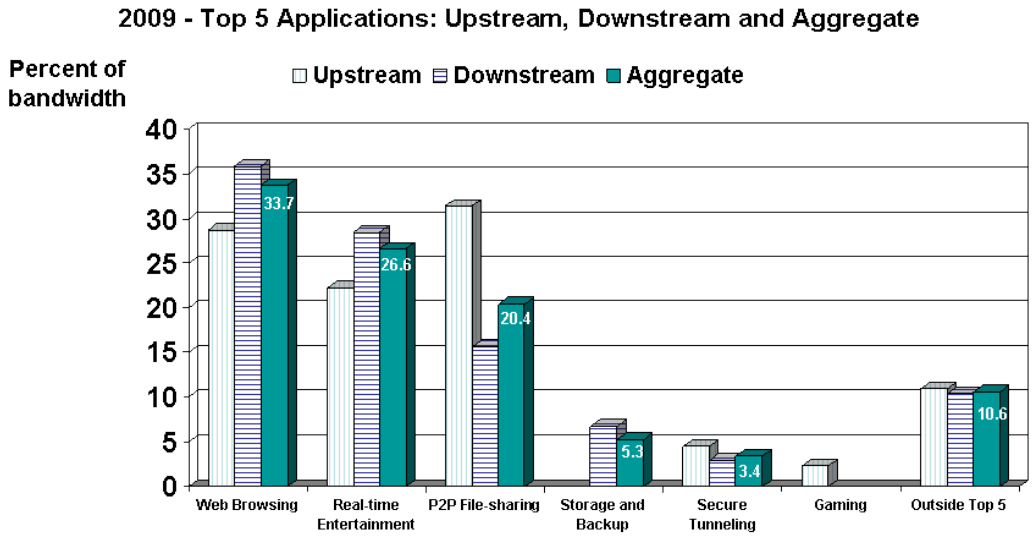
40. This report (and similar reports published by Sandvine in previous years) arrives at one inescapable conclusion: *an unmanaged network is not a neutral network*. There is tremendous differentiation in:
    - The demands placed on the network by different Internet applications;
    - Users' performance expectations of different Internet applications; and
    - The demands placed on the network by different users.

Application Demands are Differentiated

41. Sandvine's study showed that the top five categories of applications, by share of aggregate (upstream and downstream) bandwidth, throughout the day are:
    - Web browsing
    - Real-time entertainment (comprised of streaming audio and video, peercasting, place-shifting, Flash video)

[14] Sandvine Incorporated. See http://www.sandvine.com/downloads/documents/2009%20Global%20Broadband%20Phenomena%20-%20Full%20Report.pdf

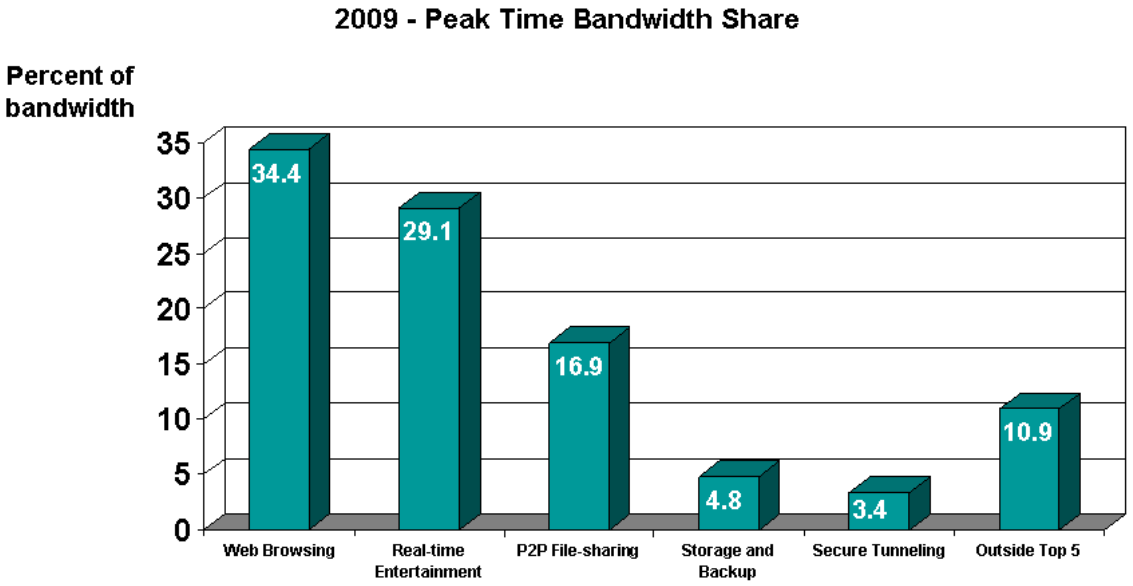
- P2P file-sharing
- Storage and back-up services
- Secure tunneling (e.g., virtual private network traffic)



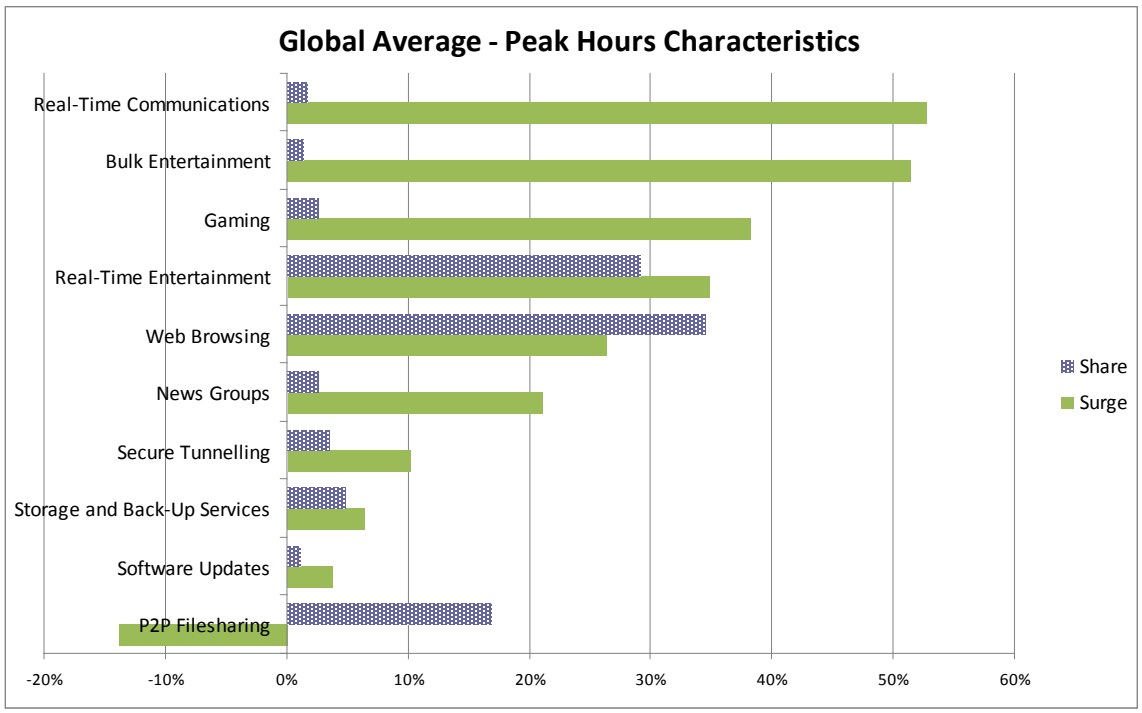
42. A couple of observations can quickly be made. First, users consume these applications in different ways and the applications themselves place different demands on the network. Real-time entertainment is consumed as it is received. For example, streaming video and audio is watched and listened to the instant it arrives at a user's computer. These applications are bandwidth-intensive but also time-sensitive. Delays in their delivery would be noticed (and not welcomed) by users. P2P file-sharing and storage and back-up services are also very bandwidth-intensive but are consumed very differently. Users initiate the process to download or upload files then can walk away from their computers, often overnight. Delays in delivery of the data would not be of particular concern to (or even recognized by) the user.

43. Second, where in the "Top 5" list of applications are email, instant messaging, VoIP and online gaming? Despite the enormous popularity of these applications among Internet users, they don't consume much bandwidth. Text-based data such as that in emails and instant messages are simply not bandwidth-hungry, neither is the voice data from a VoIP call nor the "move left", "move right" and "fire" commands in online video gaming. Yet the performance of these immensely popular applications can be impacted significantly by less popular applications (measured in number of users) that consume much more bandwidth, such as P2P file-sharing or storage and back-up services. Over 300ms of combined latency and jitter in a VoIP call and the callers are stepping on each others' words. 75ms of combined latency and jitter delay in a "fire" command could leave the shooter fired upon: game over.

44. Sandvine's study also demonstrated that during peak hours (shown by the study to be 7:00pm to 10:00 pm, normalized by time zones), when network congestion is most likely to cause performance issues, time-sensitive applications represent an even larger component of network traffic.



45. Further, as demonstrated in the following graph, at peak Internet hours the use of time-sensitive applications surge the most. The peak-time bump in traffic is almost completely attributable to the surging evening popularity of Real-Time Entertainment and Web Browsing – not only do both of these categories experience huge per-subscriber increases in bandwidth demand (rising by almost 35% and 26%, respectively), but these categories also make up a significant portion of the overall utilized bandwidth (29% and 34% of network traffic, respectively).

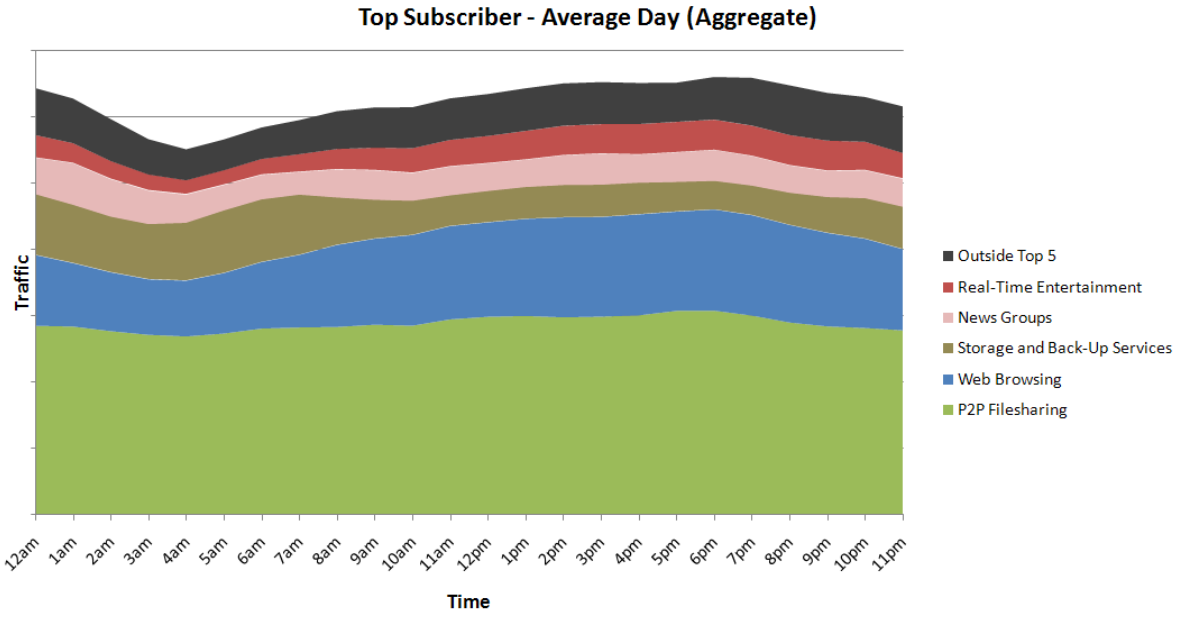


Subscriber Demands are Differentiated

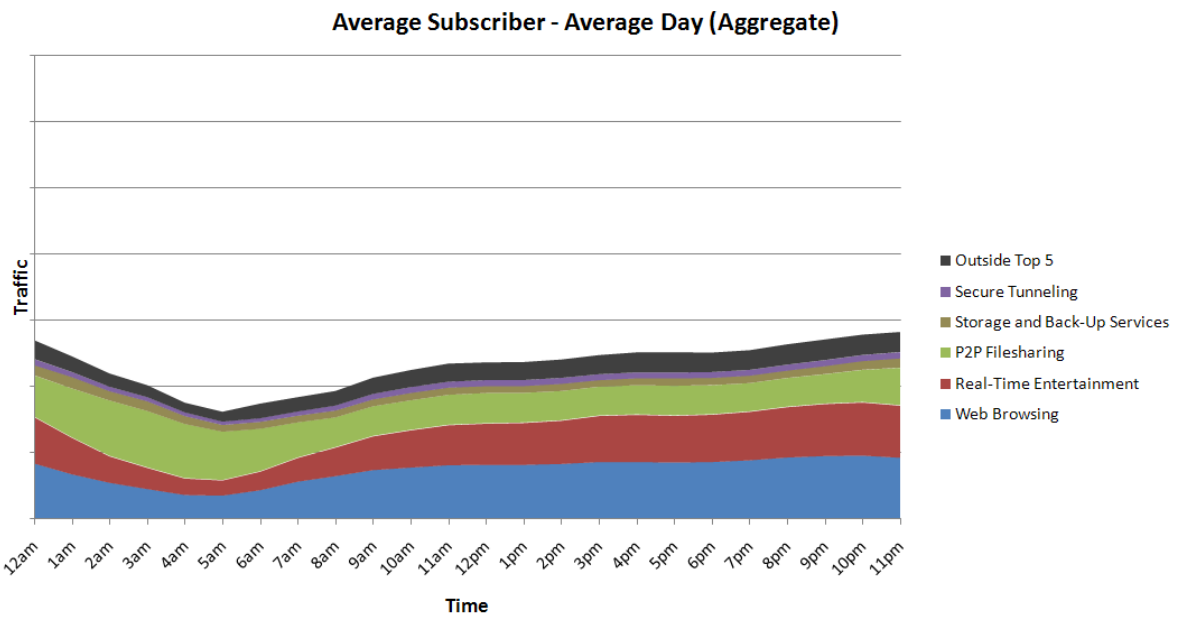
46. Just as applications make differentiated demands on the network, so too do users. Sandvine's research demonstrated that over a month the top 1% of users, by total consumption, account for 25% of total bytes on residential access networks and the top 20% of subscribers account for fully 80% of total bytes[15]. The study also demonstrates that the monthly data consumption of a heavy Internet user exceeds that of an average user by a factor of about 200.

47. Two key, related factors contribute to this enormous variation in individual network requirements between "top" and "average" subscribers:
    - The network's top users exhibit relatively little change in usage throughout the day. Conversely, average subscribers pop online and offline throughout the day and over the course of a month.
    - Top subscribers still rely heavily on bulk download applications like P2P file-sharing, storage and back-up services, and news groups – applications that are responsible for massive amounts of traffic volume with very little user involvement.

48. The graphs below show the top five categories for a top subscriber and an average subscriber over the course of the day. The two aggregate bandwidth graphs share a common y-axis scale, so visual comparisons between them are valid. Since the scales are consistent, we know that at any instant in time, a top subscriber is likely to be using more bandwidth for P2P file-sharing than an average subscriber uses in total for all categories.



[15] Ibid.



49. In contrast to a top subscriber (for whom P2P file-sharing accounts for almost half of all traffic), an average subscriber favours the on-demand nature of Web Browsing and Real-Time Entertainment. Also, while P2P file-sharing is still present in the average subscriber's profile, it accounts for less than a quarter of daily bytes.

Differentiated Treatment for Differentiated Demands

50. So, left unmanaged, certain applications and subscribers win the inherent competition for shared network resources. To deliver the unique quality of experience that subscribers expect from each application for the maximum possible number of subscribers for the maximum amount of time, the network needs to differentiate between the heterogeneous needs of individual applications and subscribers. *Again: an unmanaged network is not a neutral network.* Traffic management should be encouraged in order to improve the experience of Internet users.

51. Today, certain services and applications that consumers would value receiving over their Internet connection are currently not feasible absent traffic management that enables a minimum quality of experience. An example of such services would be telepresence, which is beyond the delivery capabilities of current networks but could be feasible with appropriate traffic differentiation.

52. Traffic management helps to protect subscribers' rights by fairly allocating scarce network resources in times of congestion so that (in the words of FCC Chairman Julius Genachowski) "very heavy users do not crowd out everyone else". By combining a subscriber-specific and application-specific approach a network provider could create a narrowly-targeted policy that affects:
    - *only* disproportionate users;
    - *only* application classes that contribute disproportionately to bandwidth consumption; and
    - *only* application classes that are not time-sensitive.

53. In the future Sandvine expects to be able to offer solutions that let the *subscriber* select which applications receive higher priority in the network in times of congestion. By definition, any network

management policy that is not only agreed to but *defined by* the Internet subscriber must be deemed reasonable from his or her viewpoint.

Sandvine's five Principles of Reasonable Traffic Management

54. The notion of "reasonable network management" has become a cornerstone concept in the developing Network Neutrality debate. Access to the Internet needs to be equitable – amongst subscribers, amongst application providers and amongst content providers – and management of the network is required to achieve that goal. In consultations with the FCC in the U.S., the CRTC in Canada, industry leaders such as the National Cable and Telecommunications Association, and network operators globally Sandvine has advocated the following criteria for "reasonable network management":

    1. Narrowly-tailored
       Traffic management is implemented only where congestion exists and when congestion is causing quality of experience issues for a large number of subscribers.

    2. Proportional and reasonable effect
       A traffic management policy has an effect on subscribers or applications that is proportional to the effect the user or application is having on the network. Policy applies the smallest reasonable intervention to alleviate congestion and improve quality of experience for the majority of subscribers.

    3. Legitimate and demonstrable technical need
       Congestion and/or quality of experience issues can be demonstrated to exist in the network and management's technical remedies are effective in achieving its targeted goals.

    4. Transparent disclosure
       Network operators need to disclose traffic management policies and changes thereto in a simple, useful and predictable manner.

    5. Auditable
       Network operators can demonstrate that the above requirements are met through the auditing and reporting capabilities of its traffic management solution.

v) Can you provide any evidence that allowing traffic management has a negative impact on innovation?

55. Sandvine is unaware of any evidence that traffic management has a negative impact on innovation. To the contrary, by helping each application get the resources it needs for the best possible user experience in a contended network, application-specific differentiation of network traffic helps encourage investment and innovation in applications and related content. Further, it is important to remember that innovation occurs at every level of the Internet value chain, including innovation by the network operators that invest in innovative traffic management solutions to improve network quality.

56. Traffic management has also demonstrated to some application developers that there is a need to optimise for network utility. An example would be BitTorrent's development of uTP[16] as a non-congestion-causing transport protocol.

Innovation and Investment Occurs at the Network Level

57. Thanks to past investments and innovation by network operators and equipment vendors, traffic management solutions have already made the Internet much more intelligent and capable of delivering a wider variety of services than at any previous time.

58. Network providers are just beginning to explore the use of traffic management practices to help them create service offerings that are more attractive to consumers in an increasingly competitive Internet access market. In the fixed line broadband market, high-speed Internet services are largely offered in the form of flat-rate, monthly plans. Consumers may be interested in other types of service plans that better reflect the unique ways that they use their Internet connections. Such plans would necessitate the ability to differentiate between the traffic of individual subscribers, and between applications.

59. For example, "light" Internet users may be interested in a service package that ties their fees to the bytes they consume on the network. But would these consumers want to pay for malicious traffic that affected their usage in a month, or visits to the service provider's web service portal to address service issues? A user- and application-specific traffic management policy would be required to manage the plan.

60. By contrast, disproportionately heavy users likely don't want to pay "by the byte", but they may be interested in a service plan that provided a financial incentive to shift their activity to non-peak network hours. Such a plan would help all users by freeing up more capacity at peak times, when network congestion and application degradation is most likely to occur.

61. Other consumers may value their Internet connection by the quality of experience they receive for their favourite applications, like latency-sensitive Internet video gaming or VoIP. Network providers could offer a Premium Video Gaming or Premium VoIP service plan that delivers exactly the type of Internet experience these consumers want. As mentioned previously, StarHub in Singapore and Demon in the United Kingdom have introduced such plans, which are supported by application-specific and user-specific traffic management policies. As discussed above, in the mobile broadband market, Canada's Rogers Communications' offers Social View for "unlimited access to your favourite Social Networking sites" and for "an integrated experience of your favourite social networks with embedded functions on your device."

62. Will these innovations be successful? Likely, some will and some won't. But all offer consumers new choices that are impossible without some form of traffic management.

[16] See http://en.wikipedia.org/wiki/Micro_Transport_Protocol)

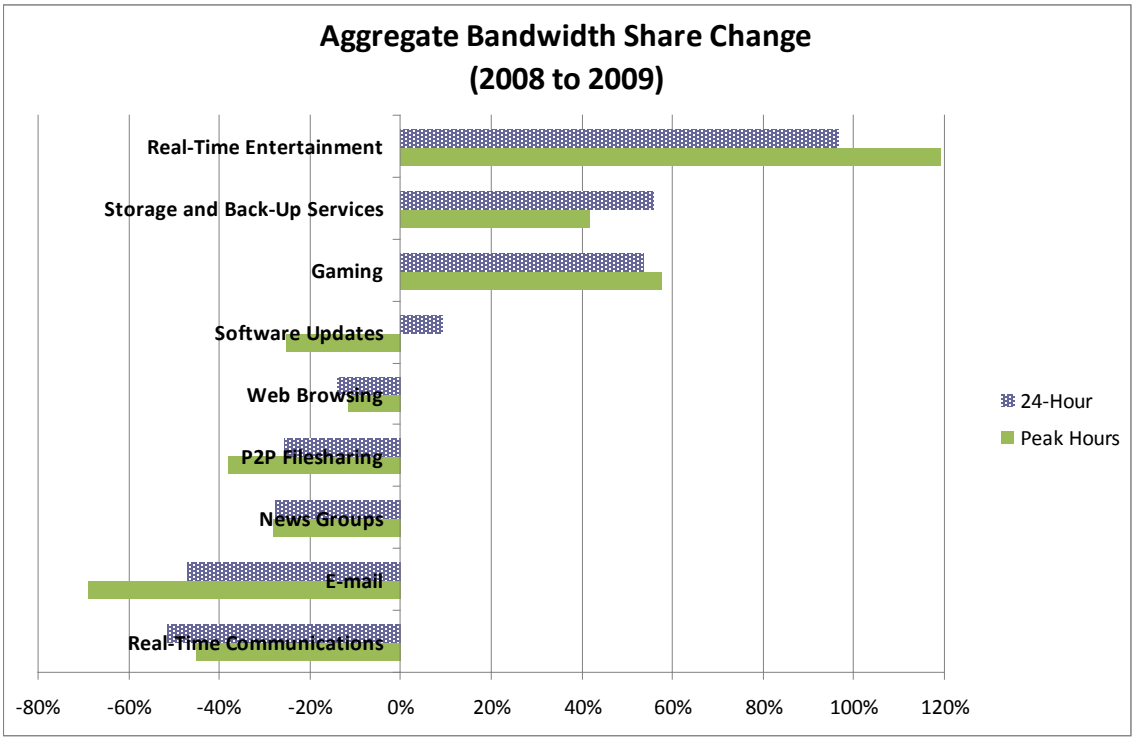
vi) Ofcom's preliminary view is that there is currently insufficient evidence to justify *ex ante* regulation to prohibit certain forms of traffic management. Are you aware of evidence that supports or contradicts this view?

63. Sandvine believes that it could be appropriate to have an ex-ante rule prohibiting any traffic management that supports an anti-competitive practice. Otherwise, ex ante rules are problematic given the dynamic nature of Internet traffic and the heterogeneous, dynamic nature of its constituent networks. An "unreasonable" traffic management practice today may be reasonable tomorrow (or vice versa) and a reasonable traffic management practice on a fixed line network today may be insufficient or unreasonable on a mobile network. "Reasonable" must be measured by the impact the traffic management practice has on the majority of subscribers and that impact can only be measured ex post, on a case-by-case basis, not in advance for all networks and situations.

Internet Traffic is Dynamic

64. In the North American Network Neutrality debate some have called for ex ante prohibition of application-specific traffic management, i.e., applying differentiated treatment to different application traffic. As Sandvine argued in its answer to question iv) above, such differentiation is necessary to make the network neutral and to provide the most subscribers with the best possible performance for as many possible applications for the maximum possible amount of time. Furthermore, the Internet is so dynamic that the popularity of applications is changing all the time. Traffic that did not require management yesterday may need to be managed today.

65. Internet traffic is dynamic – this is another inescapable conclusion from Sandvine's 2009 Global Broadband Phenomena study. In 2009 there was a massive shift in subscriber behavior from a reliance on "download now, use later" content acquisition to an on-demand mentality where bytes are consumed as they arrive. The graph below depicts the percentage change between 2008 and 2009 in bandwidth share of various applications over a typical day and at peak hours only.



66. The graph shows a significant shift towards real-time entertainment and online gaming applications in 2009 at the expense of traditional bulk data acquisition, most notably P2P file-sharing and news groups. During peak hours, when network congestion is greatest, the explosion in popularity of real-time entertainment and gaming applications is even greater than for the 24-hour average. Such shifts have significant implications for network development and management.

67. In such a dynamic environment, it is critical that network providers retain the flexibility to innovate with new traffic management practices that recognize the differentiated needs of applications – both those popular today and those as-yet-unknown applications with as-yet-unknown demands on network resources that will become popular tomorrow.

Access Networks are Heterogeneous and Dynamic

68. The Internet's constituent access networks, such as cable, DSL, fibre, wireless, and satellite, have different characteristics that create susceptibilities to congestion and performance issues at different network locations at different levels of usage. For example, in cable networks, upstream capacity has traditionally been very limited, in DSL networks certain downstream links can be more subject to congestion and in mobile networks upstream and downstream bandwidth is at a premium as it is fixed by the physical properties of the underlying radio spectrum. In bandwidth-constrained mobile networks, issues related to latency, jitter and packet loss also become exacerbated. So, mobile networks are particularly susceptible to congestion and quality-of-service issues, and such limitations are already being noticed by users of some of the world's largest mobile networks[17] despite still-modest data usage.

[17] ArsTechnica. *AT&T CTO downplays role of iPhone in network's issues*. See http://arstechnica.com/apple/news/2009/10/att-cto-downplays-role-of-iphone-in-networks-issues.ars?utm_source=rss&utm_medium=rss&utm_campaign=rss

69. A few bulk file-sharing or file-transfer sessions are unlikely to have a crippling effect on users' enjoyment of other applications in today's fixed line networks. The same may not be true for a given mobile network. A few similar sessions on a cell site could seriously impair the web surfing, voice call and gaming experience of all users' connected to that site, for example. To protect the user experience for these popular applications, it would be reasonable to create a policy that began to manage bulk application traffic at a lower threshold (and/or manage it in a different way) than for a fixed line network. In fact, these applications, and others like Slingbox (which "slings" bandwidth-intensive television signals to Internet devices, such as a Smartphone), may have such a detrimental effect on network performance for all applications that blocking them could be deemed a reasonable practice *in a given situation at a given time*. Similarly, if managing mobile data traffic on a subscriber-specific basis, it might be necessary to start managing "disproportionate users'" consumption at a threshold level that would be much lower (or managing it in a different way) than for fixed line networks. A case-by-case analysis would have to be performed to know.

70. Mobile networks are also the newest entrants in the market for broadband access so user behaviour is rapidly evolving. Consequently, the nature of data traffic traversing the mobile network is more dynamic than for any other access network class. More time is required to understand how users will consume the Internet over mobile devices and what network management policies may be appropriate.

71. Even within an access network category, no two service providers have networks that are identically architected. A particular technical approach that achieved reasonable results for users in one network may have a very different effect on users in another network – no bright lines can be drawn. Traffic management solutions are highly configurable. Both subscriber-specific and application-specific approaches can result in reasonable or unreasonable effects for users. The determination cannot be made ex-ante, only ex-post after seeing the actual effect on users.

vii) Ofcom's preliminary view is that more should be done to increase consumer transparency around traffic management. Do you think doing so would sufficiently address any potential concerns and why?

72. If a market is competitive (as it is in the United Kingdom) and if due to transparency the consumer fully understands the significant terms (including the impacts of significant traffic management policies) of the service to which they intend to subscribe, then by contracting for these services they have defined the traffic management practices as reasonable. Reasonability needs to be judged from the subscriber's perspective, and transparency is critical in making sure that perspective is well-informed.

73. However, to fully address concerns about traffic management, Sandvine believes that one additional characteristic is necessary: the results must be auditable. While transparent disclosure protects subscribers entering a contract, the requirement for an auditable traffic management trail protects subscribers who have already entered contracts from concerns that the transparently disclosed practices are being followed by the network operator.

viii) Are you aware of any evidence that sheds light on peoples' ability to understand and act upon information they are given regarding traffic management?

74. Sandvine has not done any research in this area, however there are ample cases where operators have implemented different peak/off-peak traffic management policies (including bandwidth caps) and consumers have adapted to them. One case would be Virgin Media. The traffic management policies have been disclosed in plain language[18] on Virgin Media's site and many users have reacted constructively in discussion forums, such as forum.utorrent.com, by suggesting ways to schedule their activities to work well with the traffic management practice.

ix) How can information on traffic management be presented so that it is accessible and meaningful to consumers, both in understanding any restrictions on their existing offering, and in choosing between rival offerings? Can you give examples of useful approaches to informing consumers about complex issues, including from other sectors?

Transparent Disclosure of Fixed Services Network Performance

75. For network performance disclosures to be meaningful, subscribers need to know whether a network can reliably deliver the expected quality of experience for their favourite applications – based on the service level they have contracted for with their network operator. Network performance disclosure should cover peak and off-peak times and times when traffic management is in effect. To achieve these goals, network operators need to report the capabilities of their network at the subscribers' location by application and compare that to the minimum performance requirements for each application class.

Application requirements

76. Applications differ with respect to the amount of bandwidth, latency, jitter, and packet loss that they require or can tolerate in order to be delivered at an expected quality of service level.
    - Bandwidth: traffic volume over time. It is usually measured over a short time, such as bits/second or megabits/second (Mbps), which is 1,000,000 bits/second.
    - Latency: the delay for a message to get from one communications end point to the other, e.g., the time it takes for a voice-over-IP (VoIP) data packet to leave the speaker's mouth and arrive at the listener's ear. It is typically measured in milliseconds.
    - Jitter: the variation in the latency of one message to another, typically measured in milliseconds (e.g. if the first message takes 1ms and the second message takes 10ms, then there is 9 ms of jitter).
    - Packet Loss: occurs when one or more packets of data traveling across a computer network fail to reach their destination. Packet loss is usually measured in percentage terms. It can be caused by a number of factors, including signal degradation over the network medium, oversaturated network links, corrupted packets rejected in-transit, faulty networking hardware, faulty network

[18] Virgin Media Traffic Management FAQs. See http://www.virgin.net/allyours/faqs/trafficManagementFAQ.html,

drivers or normal routing routines. Packet loss is expected in an IP network since it is a mechanism of Transmission Control Protocol (TCP) used for rate-control and congestion control.

77. While bandwidth gets most of the attention, adding bandwidth is not always the answer to improving the user's quality of experience for an application. The other factors can play a critical role. An application can be classified into one of three categories (bulk, interactive or paced/burst-paced), based on its requirements of the network across these four characteristics.

Application Categories and Requirements

78. Bulk. These applications include P2P filesharing (e.g., BitTorrent, FastTrack, etc), web surfing, usenet news (NNTP), and file transfers over FTP or HTTP, for example, and will go as fast as the network will permit, accelerating until packet loss occurs. TCP is designed to achieve the maximum communication rate possible. In practice bulk applications will go as fast as the thinnest part of the network between the client and server. In the case of the server collocated within the ISP network (e.g. a content-delivery network, a cache), this will be bound by the access equipment speed. In the case of a server which is located farther away, this may be bound by transit (connection to all worldwide public networks) or peering (connection to other nearby private networks) performance. Typically servers of bulk applications (e.g. Speedtest.net, Rapidshare.com, Megaupload.com) will saturate the download speed of the consumer's modem, as they typically download-only. In the case of P2P filesharing, it is bi-directional so it can also have the same affect in the upstream direction.

79. Most bulk applications can run unattended by the user. File transfers are initiated by the user, who may then walk away – often for hours or even overnight – while the process completes. The content is typically for offline consumption. Bandwidth is the primary determinant of transfer speed and performance will generally improve linearly with increases in bandwidth. As a result, latency and jitter matter much less – users likely would not even notice their effect. Packet loss is used by the network to control the maximum achieved speed.

80. Web surfing represents an exception in the Bulk category. "Web 2.0" sites have introduced interactive components to web surfing – typically the user interacts with the website and expects near-immediate response. Data is traveling bi-directionally as users have started to be content providers in their own right, by posting videos to YouTube, for example. Increases in bandwidth do not translate linearly to increased performance because it takes several "round trips" between a personal computer and the related web servers to load a website – typically at least four: the Domain Name Server (DNS) lookup two for the three-way handshake established by TCP and one to retrieve the content. Each of the four round trips is subject to the latency in the network, and when added together this delaying effect becomes the limiting factor in the transmission. Consequently, additional bandwidth does not dramatically improve loading times for a website.

81. Interactive. These applications are paced by the consumer. In the case of VoIP, bandwidth largely depends on silence suppression and the codec bandwidth chosen, but it is typically 8-30Kbps. The bandwidth requirements of interactive applications are often modest (though in the case of video conferencing the rates are significantly higher: 200-500Kbps is common), but they typically require very low latency, jitter and packet loss to achieve a satisfactory quality of experience. For example, a VoIP user can perceive latency of 150 milliseconds on a call, and delays greater than 300

milliseconds render the call unusable[19]. As with web surfing, adding bandwidth will not necessarily address quality of service issues. In general, because of the sensitivity of Interactive applications to latency, jitter and packet loss it is particularly important to protect the quality of service for these applications.

82. Paced/Burst-paced. Streaming applications such as YouTube and SHOUTcast fall into this category. The media involved has a natural bit rate based on the content's encoding, and the connection tries to achieve this rate on average over its lifetime. Though for short durations the media will 'burst' to provide buffering on the client to allow for packet loss on the network (YouTube, because it uses TCP, will attempt to transmit at line rate when possible to build the buffer then reduce to the natural rate). So, these applications can be modeled by the media they carry. For typical Internet streaming today, rates of approximately 300-400Kbps are common. Hulu, YouTube, and others are starting to shift to higher definition video, for which the rate can increase to 1-7Mbps of bandwidth.

83. With paced/burst-paced applications it is important that a network sustain the minimum bandwidth requirements, but because of the buffering involved additional bandwidth only marginally improves performance, by making the applications less sensitive to latency, jitter and loss in the network.

84. The following table provides some representative benchmarks to achieve a minimum quality of service for certain popular applications.

Application Category	**Application Class**	**Minimum Bandwidth**	**Maximum Latency**	**Maximum Jitter**	**Maximum Loss**
Bulk	P2P	19Kbps	n/a		
	Web surfing	1Mbps (Web 2.0)	166ms (latency + jitter)		n/a
	Email	60Kbps	n/a		
	Usenet news	195Kbps	n/a		
	FTP file transfers	195Kbps	n/a		
Interactive	VoIP	16Kbps	300ms (latency + jitter)		< 0.5%
	Video gaming	50Kbps	75ms (latency + jitter)		< 0.5%
	Video Conferencing	250Kbps	300ms (latency + jitter)		< 0.05%

[19] See http://voip.about.com/od/glossary/g/latency.htm, or T. Blajic, D. Nogulic, M, Druzijanic, *Latency Improvements in 3G Long Term Evolution*, p. 1-2, available at http://www.ericsson.com/hr/about/events/mipro_2007/mipro_1137.pdf, or http://www.telephonyworld.com/training/brooktrout/iptel_latency_wp.html.

Application Category	Application Class	Minimum Bandwidth	Maximum Latency	Maximum Jitter	Maximum Loss
Paced (and burst-paced)	Video streaming	300Kbps, to not have much of a wait time	< 1s for "channel change"	<50ms	<0.05%
	High def video	1-3Mbps depending on quality of HD.	< 1s for "channel change"	<50ms	<0.05%
	Audio streaming	Audio: 128Kbps for CD quality. 56Kbps for radio	< 1s for "channel change"	<50ms	<0.05%

85. Sandvine submits that each Internet application provider should report the minimum bandwidth and maximum latency, jitter and packet loss required for satisfactory delivery of their applications. This way, network operators could have access to averages for an application class for their own network reporting purposes (as described below) and consumers could make better decisions about how to use their own network connections.

Network performance measurements

86. Sandvine submits that network performance should be measured on a per-application class basis because satisfactory application performance is of central importance to the user's experience. For an individual user, the measurement of his network performance will be in part determined by the applications he uses. For example, if a subscriber only uses his Internet connection for online video gaming, which typically demands bandwidth of approximately 50Kbps, then his measured bandwidth performance over a given period will be 50Kbps, even though his service tier may promise and could in fact deliver much more. The shortfall would be as a result of the subscriber's preferred usage of the connection, not necessarily any limitation in the network connection itself to deliver speeds up to the promised throughput. The user's experience for gaming is, in fact, better defined by the latency, loss and jitter.

87. Sandvine submits that network providers should measure their network's performance for *each subscriber*, *by application class*, on the following metrics:
    - Average achieved peak bandwidth at peak hours and off-peak hours.
    - Average latency during peak and off-peak hours.
    - Average jitter during peak and off-peak hours.
    - Average loss during peak and off-peak hours.

88. The network measurements could be taken monthly over a one-minute interval in peak and off-peak times. Sandvine's own study has shown that peak hours are from approximately 7:00 pm to

10:00 pm globally[20]. These hours could be reliably used to define "peak" and "off-peak" for the purpose of these measurements.

Reporting

89. The network measurements should be reported to consumers and Ofcom as averages for a service area, by access-type, and service plan for both peak and off-peak hours. Here is one potential presentation of the data:

Service Area X
DSL Platinum Service
Average Network Performance for January 2010
Peak Hours

Application Class	Minimum Bandwidth	Maximum Latency	Maximum Jitter	Maximum Loss
P2P				
Required	195Kbps	n/a	n/a	n/a
Delivered	250Kbps	n/a	n/a	n/a
Web surfing				
Required	1Mbps (Web 2.0)	166ms (latency + jitter)	n/a	
Delivered	2Mbps	160ms (latency + jitter)		
Email				
Required	60Kbps	n/a	n/a	n/a
Delivered	120Kbps	n/a	n/a	n/a
Usenet news				
Required	195Kbps	n/a	n/a	n/a
Delivered	250Kbps	n/a	n/a	n/a
FTP file transfers				
Required	195Kbps	n/a	n/a	n/a
Delivered	250Kbps	n/a	n/a	n/a
VoIP				
Required	16Kbps	300ms (latency + jitter)	< 0.5%	
Delivered	50Kbps	185ms	0.1%	
Video gaming				
Required	50Kbps	75ms (latency + jitter)	< 0.5%	
Delivered	75Kbps	45ms	0.1%	
Video Conferencing				
Required				
Delivered	250Kbps	300ms (latency + jitter)	< 0.05%	
	265Kbps	185ms	0.3%	
Video streaming				

[20] 2009 Global Broadband Phenomena Study, Sandvine Incorporated, http://www.sandvine.com/downloads/documents/2009%20Global%20Broadband%20Phenomena%20-%20Executive%20Summary.pdf

Application Class	Minimum Bandwidth	Maximum Latency	Maximum Jitter	Maximum Loss
Required	300Kbps	< 1s	<50ms	<0.05%
Delivered	425Kbps	0.5s	22ms	0.25
High def video streaming				
Required	2Mbps	< 1s	<50ms	<0.05%
Delivered	1.8Mbps	0.5s	22ms	0.25
Audio streaming				
Required	128Kbps	< 1s	<50ms	<0.05%
Delivered	200Kbps	0.5s	22ms	0.25

Green: Network **can** reliably deliver required application performance

Orange: Network **may** reliably deliver required application performance (average performance better than required performance by 10% or less)

Red: Network **cannot** reliably deliver required application performance (average performance worse than required performance)

90. The reported averages can be based on sampling of the full set of subscriber data measurements as long as an appropriate level of statistical reliability is achieved. In order to provide some idea of network trends, the data could also be presented in a time series (e.g., for the last six months) to graph changes in application performance on the network over time.

91. Consumer reporting of network performance measurements should be easy to find, understand and compare between network providers. Most users don't understand the network demands (bandwidth, latency, jitter, loss) of their favourite applications. So, any disclosure by the network provider about its network's ability to deliver an application has to be made in simple terms, i.e., "yes, we can deliver the application reliably", "no we can't deliver the application reliably", or "maybe can deliver the application reliably". Additionally, simple definitions of each performance characteristic could be provided for interested users.

92. The information should be made available on the network provider's public website and (to the extent it is practical) identified on the website in a standard way across network providers (e.g., "Network performance, by application") and to the extent feasible in a standard location (e.g., as a link from the operator's Terms of Service). Such standardization would assist consumers' ready access to the data.

93. With respect to the effect of traffic management practices on network performance, a similar table to that illustrated above could be presented for the network when traffic management policies are in effect.

Mobile Broadband Services Measurement and Disclosure

94. Conceptually the same principles that apply to measurement and reporting of application requirements versus network performance apply to mobile networks as well. However there are some significant technical differences between fixed and mobile networks that require special consideration. Additional or different measurements may need to be taken and reported on. For Sandvine's full discussion on this topic, please refer to http://fjallfoss.fcc.gov/ecfs/document/view?id=7020514597

x) How can compliance with transparency obligations best be verified?

95. Sandvine has not done any research in this area, but believes that Ofcom's initial work with SamKnows offers a good starting point and has helped to push the transparency issue to the fore globally. However, Sandvine believes that the operators should supply the data to Ofcom for more universal network coverage.

xi) Under what circumstances do you think the imposition of a minimum quality of service would be appropriate and why?

96. Sandvine believes that a minimum quality of service should be an optional part of a contract, with an appropriate tariff. Otherwise, a minimum QoS absent specific contract or consumer demands should only be sufficient to prevent disenfranchisement, such as citizens' rights of voting etc, as more and more public information and services are shifted online. Any minimum QoS requirement should exempt service providers from events that are out of their control such as unforeseeable network outages or changes in applications.

Deep Packet Inspection – correcting the misconceptions

97. A core part of many of Sandvine's solutions is technology known as "deep packet inspection," or "DPI." DPI is, from a network engineering and architectural perspective, the act of any network equipment which is not an endpoint of a communication using any field other than the layer 3 destination IP address for any purpose. DPI itself is a passive inspection technology, nothing more or less.

98. DPI inspects packets to identify the underlying applications, which information can be used for a number of different uses and by a number of different network applications. DPI has been used for decades in providing differentiated treatment of network traffic, enhancing network security, and providing data on network usage.

99. Frequently, participants in the Network Neutrality debate have created misleading or outright erroneous impressions about DPI technology. Some have called for an outright ban on the technology. With these comments, Sandvine intends to correct some of those errors and to encourage Ofcom to be technology-neutral when considering the issues related to traffic management and Network Neutrality.

100. Specifically, Sandvine hopes to communicate that:
- DPI is a necessary, time-tested, ubiquitous technology that fuels competition and innovation;
- DPI-supported traffic management solutions don't inspect content;
- Applications of technologies – not the technologies themselves – may raise privacy concerns;
- DPI-supported network management solutions offer consumer more choices and create more competition;
- DPI-supported network management solutions use IETF-approved techniques.

DPI is a Necessary, Time-tested, Ubiquitous Technology

101. DPI is necessary for the identification of traffic today because the honour-based port system of application classification previously used for this purpose no longer works. Under a port-based system, applications identify themselves by the TCP/UDP port they travel on. However, authors of these applications quickly started masking the activity of their applications under different port numbers rendering the system obsolete. As an illustration, in its filing under the Canadian Radio-television Telecommunications Commission's Review of Internet Traffic Management Practices, Telus reported that P2P file-sharing traffic represented as little as 3% of network traffic[21]. Telus was relying on port-based data alone. Sandvine's own studies at the time and the aggregated results of Canadian service providers filed under the CRTC review, suggested that total P2P file-sharing traffic levels (upstream and downstream traffic combined) were at the very least ten times higher. Today, DPI technology represents the only effective way to accurately identify different types of applications. The identification of applications has been and continues to be a critical component to the efficient functioning of the Internet.

[21] Telus Communications Company, "Response to Interrogatory PN 2008-19," see response to question 1 b) at http://www.crtc.gc.ca/public/partvii/2008/8646/c12_200815400/1005778.zip

102. Banning or limiting the use of DPI, as some participants in the global Network Neutrality debate have suggested, would have far-reaching and damaging consequences across the Internet, where the technology is used extensively. The router in your home uses DPI to act as a firewall and network address translation device for protocols such as VoIP. Firewalls, some built right into popular PC operating systems, use DPI to analyze packets for malicious intent like viruses, trojans, and Spam. Libraries, schools and government institutions rely on their firewalls to protect themselves and their users from attacks. Those firewalls use DPI technology. Load balancers and routers, indispensable hardware that distribute traffic on the Internet and private networks, use DPI to identify where a given packet or URL should be routed and what priority it should be given.

103. DPI is also a key part of the innovation in allowing a migration from IPv4 to IPv6, allowing a network operator to convert from one to the other using a carrier-grade network-address-translation and keeping protocols such as VoIP operational.

104. For several years, service providers have used DPI as part of the toolset to gain insight into trends in network traffic for better network planning, manage congestion in their networks and mitigate malicious traffic. Today, many are just investigating ways to use the technology to create new service plans that are better tailored to subscriber preferences. Such new plans would create more choices for subscribers and a new basis upon which service providers can compete.

105. So, DPI has been beneficially deployed in broadband networks for many years within many network devices. Yet certain participants in the global Network Neutrality debate have counter-intuitively concluded that the ubiquity of DPI somehow represents a threat rather than a benefit. As evidence, they point to *potentially* harmful applications that could incorporate DPI technology, while simultaneously ignoring the concrete track record of the technology's success to date. Overwhelmingly, DPI has been positive for networks and their subscribers.

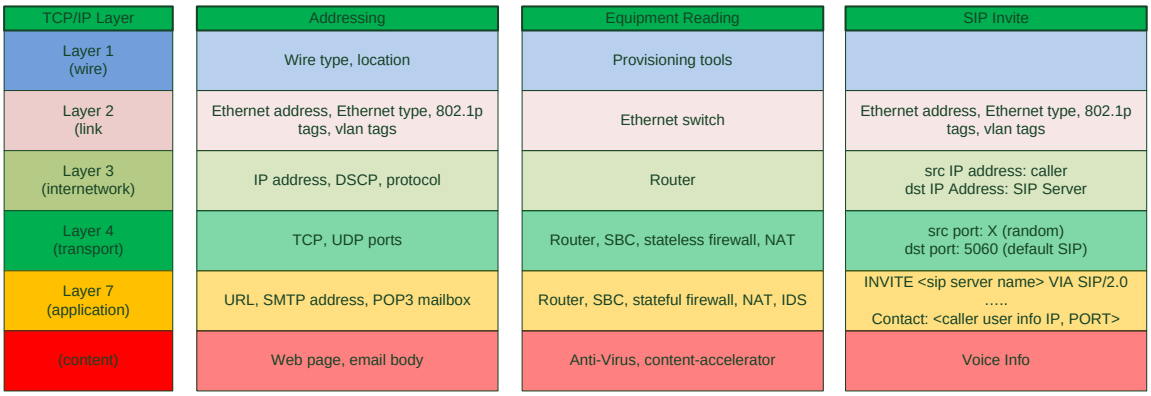
DPI-supported network management solutions don't inspect content

106. The ability of DPI to identify the type of a packet by looking at payload data has often been confused as inspection of "content" and therefore de facto invasion of privacy. In fact, each layer in the Open System Interconnection (OSI) Reference Model has a header and payload, all the way up through Layer 7 and beyond, and networking equipment has always read Layer 7 payload data. For example, mail servers route mail based on the e-mail address, which is located in the Layer 7 payload data. Session Initiation Protocol (SIP) is a signaling protocol widely used for setting up and tearing down multimedia communication sessions such as VoIP. SIP needs to look in the Layer 7 payload data to find both phone numbers involved in a VoIP conversation, then set up the data (voice) flow. Routers/firewalls look at the Layer 7 SIP exchange to extract this flow information to let the data through. If they don't, the voice component is blocked.

107. Not all "payload" is content. Sandvine submits that the true content of an Internet transmission is represented as the body of your e-mail message; the music or movie you are downloading; the video you are streaming; the words in your VoIP call, etc. Sandvine's traffic management solutions, including those that employ DPI, do not inspect content as the content is not relevant to traffic management solution. To be clear, they:

- Do not read your e-mail;
- Do not listen to your voice calls;
- Do not watch the video you are streaming, etc.

108. With DPI's signature-based inspection, a library of known application "signatures" is compared to a packet to identify matches. By way of example, the diagram below shows the breakdown of a SIP VoIP packet against the OSI Reference Model.



109. In most cases for SIP, the caller will contact the SIP server over port 5060. However, a SIP server can, and often is, configured to work on different ports. Thus, to accurately identify this traffic, a signature based on the IETF RFC standard for the SIP protocol is applied to the packet. In this example, and as shown in the fourth column of the diagram, at Layer 7 the solution looks for the presence of "INVITE" followed by some server address then "VIA SIP/2.0". By examining this protocol's header, the solution is able to determine whether the flow is SIP. The solution does not look at the flow's contents, i.e., the voice information, as it is not required to make the protocol identification.

110. If there is no match, then the solution immediately forgets the inspected data and compares the next signature definition in its library to the packet being inspected. The entire packet is not scanned, as if browsing through a magazine. Instead, only those locations that hold potentially identifying signature characteristics are inspected and only to the extent necessary to see if there is a match with the signature profile in the library.

111. In either case, the device never records any of the information past the life of the detection, other than the identity of the protocol, and it only uses this information as an input to decide whether it is relevant for the application of a network policy, such as enforcing minimum quality of service levels for a given application. The process is similar to a mail-sorting machine: the address is matched, the decision is made and the address is then forgotten.

112. Once identification has occurred further inspection not only stops, but the attributes examined in the process of arriving at that identification are discarded. For signature-based inspection, identification can typically happen in the first couple of data packets in a stream. More often than not, those first few data packets don't contain data that would typically be considered the content of a transmission, such as the text in an e-mail or the voice in a VoIP call, etc. For example, for a SIP-based VoIP call the first two data packets would be part of the "control flow", which is used to establish call permissions and locations, etc., to initiate the call. Data from the actual conversation would only appear in subsequent packets.

Applications of technologies – *not the technologies themselves* – may raise privacy concerns,

113. Despite the fact, that DPI does not inspect or retain subscriber "content" some suggest that the mere presence of DPI-based technology raises privacy issues, and have called for an outright ban on any such technology. Imagine if this approach were applied to other technologies. Single Lens Reflex (SLR) technology underlies cameras that can be used to take photos at family birthday parties or for surveillance of individuals and public spaces. One use of the technology raises privacy issues, the other does not. Nobody questions the value or legitimacy of SLR technology. So why question DPI technology? Privacy concerns properly attach to applications or uses of technologies, not to the technologies themselves.

114. Sandvine recognizes that certain solutions which are unrelated to traffic management (such as lawful intercept, copyright enforcement, and targeted advertising) may raise personal privacy considerations and are in high demand from consumers, governments and society. Such solutions are achieved through a variety of technologies, not just or even predominantly DPI.

115. Targeted advertising provides a good example. This type of solution can enhance the Internet experience for subscribers by presenting them with more relevant advertising information. Typically, targeted advertising solutions monitor users' Internet activities, such as detailed analysis of website visits, to create a more complete user profile for enhanced marketing. The collection and storage of that type of profile information has privacy implications for users.

116. DPI technology *can* comprise a component of targeted advertising solutions, but rarely has. Instead, other technologies have dominated. Google is one of the leaders in behavioral targeted advertising, but to Sandvine's knowledge Google's behavioral targeted advertising solutions do not use DPI. According to Google's own Advertising and Privacy notice in connection with its enormously popular Gmail e-mail application, Google reads your mail to make decisions on targeted advertising.

> "Google scans the text of Gmail messages in order to filter spam and detect viruses. The Gmail filtering system also scans for keywords in users' emails which are then used to match and serve ads.[22]"

117. Other Google applications have privacy implications. According to Google's Web History Privacy Notice, the Web History service:

> "Web History records information about the web pages you visit and your activity on Google, including your search queries, the results you click on, and the date and time of your searches in order to improve your search experience and display your web activity. Over time, the service may also use additional information about your activity on Google or other information you provide us in order to deliver a more personalized experience[23]."

[22] Google, "Google Privacy Center, Advertising and Privacy," see http://www.google.vg/intl/en/privacy_ads.html
[23] Google, "Web History, Web History Privacy Notice," see http://www.google.vg/searchhistory/privacy.html

118. Google's PageRank service operates by "sending Google the addresses and other information about sites at the time you visit them.[24]" According to Google's Privacy FAQ, Google stores certain information about your searches for as much as 18 months prior to anonymizing it[25]. Again, to Sandvine's knowledge, none of these solutions use DPI.

119. Lawful intercept provides another example of how privacy-sensitive solutions can be enabled by a wide variety of technologies. In the United States under the Communications Assistance for Law Enforcement Act (CALEA), service providers are required to identify and intercept criminal data traffic under a lawful warrant provided by law enforcement agencies. DPI technology could be used in a solution designed to support the collection of that data, but so too could a home computer "tapped" into the communications of the individual that is the subject of the warrant.

120. Again, applications of technologies may raise privacy concerns, not technologies themselves. Accordingly any privacy rules should attach to the application, not the underlying technology. Sandvine urges Ofcom to take a technology-neutral approach to its consideration of traffic management and Network Neutrality.

DPI-supported network management solutions offer consumers more choices and create more competition

121. As discussed in Sandvine's answer to question number three of the Discussion Document, network operators are just beginning to explore the use of traffic management practices to help them create service offerings that are more attractive to consumers in an increasingly competitive Internet access market. Many of the new service offerings, both those already implemented and those just being considered, require the use of DPI technology.

DPI-supported network management solutions use IETF-approved techniques

122. The Internet Engineering Task Force (IETF) is the open standards organization that works to develop and promote Internet standards, in particular those related to TCP/IP and the Internet protocol suite. DPI is an inspection technology, and while there are no IETF standards for inspection of Internet traffic many IETF standards implicitly require the use of DPI, such as RFC 3489, "Simple Traversal of User Datagram Protocol (UDP) Through Network Address Translators (NATs)"[26], and RFC 2766, "Network Address Translation - Protocol Translation (NAT-PT)"[27].

---

24 Google, "Install or uninstall: Toolbar Privacy Notice," see http://www.google.com/support/toolbar/bin/static.py?page=privacy.html&hl=&v=

25 Google, "Privacy FAQ," see http://www.google.com/privacy_faq.html

26 See http://www.faqs.org/rfcs/rfc3489.html

27 See http://www.faqs.org/rfcs/rfc2766.html

Also, IETF's RFC 4594[28], "Configuration Guidelines for DiffServ Service Classes," suggests that there should be different prioritization for different applications depending on their sensitivity to delay, loss and jitter. They suggest the following categories of services: telephony, telephony/video signalling, multimedia conferencing, real time interactive, broadcast video, low latency data, high throughput data, and low priority data. The RFC continues to discuss the sensitivity of each of the application categories to network conditions. Only through the use of DPI could these applications be reliably identified and therefore receive appropriate treatment in the network.

28 See http://www.ietf.org/rfc/rfc4594.txt