International Journal of Communication 1 (2007), 567-579 1932-8036/20070567

net neutrality issue

Net Neutrality:
The Technical Side of the Debate ~ A White Paper

JON CROWCROFT
University of Cambridge

Network Neutrality is the subject of much current debate. In this white paper, I try to find the signal in the noise by taking a largely technical look at various definitions of network neutrality and the feasibility and complexity of implementing systems that support those ideas.

First off, there are a lot of emotional terms used in the context of the "net neutrality" debate. For example, "censorship" or "black-holing" rather than route filtering, fire-walling and port blocking; "free-riding" rather than overlay service provision to describe the business of making money on the Net; or "monopolization" instead of the natural inclination of an organization with a lot of investment trying to make revenue from it.

This paper[1] describes the basic realities of the net, which has never been a level playing field for many accidental and some deliberate reasons, and then looks at the future evolution of IP (and lower level) services, the evolution of overlay services, and the evolution of the structure of the ISP business space (access, core and other). Finally, I appeal to simple-minded economic and regulatory arguments to ask whether there is any case at all for claiming the Internet as a special case, different from other services, or utilities.

Jon Crowcroft: **jon.crowcroft@cl.cam.ac.uk**
Date submitted: 2006-06-21

[1] Much of what I have read on the subject of net neutrality by economists is technically naïve and simplistic. The best single reference I can give which references much other (good) work in the economics, technical and legal/regulatory side on net neutrality is the current entry on Wikipedia, which my colleague, Tim Griffin pointed me to: http://en.wikipedia.org/wiki/Net_neutrality. In preparing this paper, I would like to acknowledge the comments of Richard Mortier at Microsoft Research, Mark Handley, Ken Carlberg and Richard Gold of UCL, and from the Ph.D. students in SRG in Cambridge.

I. Introduction

Flan O'Brien, the celebrated Irish columnist wrote,

> "Recently I referred briefly to a new type of telephone patented by the Research Bureau. It is designed to meet an urgent social requirement. Nearly everybody likes to have a telephone in the house, not so much for its utility (which is very dubious), but for the social standing it implies. A telephone on display in your house means that you have at least some 'friend' or 'friends' that there is somebody in the world who thinks it worthwhile to communicate with you."

The lesson we might draw from this is that even *no network at all* is of some value.

In reality, networks are very complex systems with many stakeholders. The OSI model[2] captures some of the subtlety by describing the relationships, both horizontal, between *peer* layers (e.g., physical, link, network, transport, application, etc.), and vertical, between *provider* and *consumer* layers. There are at least four layers and two end points, so the simplest system one might consider -- with only two computers and one physical link interconnecting them -- still has ten places where tussels can potentially occur (three service API interfaces at each end, and four peer-to-peer relationships). The OSI model was intended to enable multiple vendors of the various components and is embedded deeply within the Internet Architecture, which further complicates matters by permitting all sorts of recursive application of the architectural ideas, such as tunnels, proxies and protocol translators. The key word in the OSI model is *Open*, which is intended to convey the idea that the architecture admits of a level playing field for vendors of the various components, whether as hardware, software or services. Reality has proved somewhat different.

II. Discussion

Let me try to illustrate the complexity and subtlety of the debate with a few, real stories from the last ten years of Internet Experience, each of which is chosen because it captures several facets of the problem space at once.

Priority Rights: Like many other people, I have 8Mbps Internet Access through an unbundled DSL broadband provider, which I share throughout my house using a $50 router to provide 10/100 Ethernet and Wireless access to a server and the family's laptops and media centers. I don't secure the net with WEP keys and access control, since I use secure end systems with host firewalls and virus checkers, although the router runs some useful filters to lower the background nonsense. When my phone line went down for three weeks earlier this year, my kids found three neighbors with open WiFi access to their broadband lines (luckily all still working -- indeed one cable, and two different DSL providers, one

[2] International Standards Organization (ISO) Open System Interconnection (OSI) model.

bundled, and one unbundled). We asked our neighbors if it was OK to use their net[3] and they said "sure," but tellingly also admitted that this was because our usage would have no impact on their usage since they all used routers which implemented priority queues (see, for example, http://openwrt.org/). While their nets were open, they had all independently discovered that it was possible to set higher forwarding priority for their own packets than everyone else, thus being socially friendly at the same time as not giving up any resource they paid for.

We can unpack at least three lessons from this tale: firstly, it is literally child's play to build community wireless networks; secondly, it doesn't take technical experts to deploy priority services; thirdly, cooperation and selfishness are not necessarily orthogonal.

Content Re-Distribution: In the mid 1990s, the UK Academic Network provider, UKERNA, ran an experiment in usage charging for International Traffic. The goal was twofold: firstly, the charges might trickle down to real users and create a disincentive to misuse or similar carelessness in moving large amounts of data around unnecessarily; secondly, the goal was to raise more revenue to pay for upgrades to the international links. One important lesson from this experiment was that the second of these goals was far more successful than the first, however, another piece of the story is interesting.

A number of national research networks in Europe provided very large web proxy caches to create a positive alternative (lower latency and potentially higher throughput). At the same time, UKERNA allowed free access to web caches (very sensibly). However, other European countries (who had *not*} implemented the charging disincentive side of the story soon noticed that users in some UK universities were using their caches (especially in well provisioned areas with good UK connectivity such as Scandinavia). The Europeans rapidly introduced first priorities (better access for IP sources in their own Address Spaces (ASs) and eventually blocking of IP addresses outside their own networks.

Data and Digital Mobile Phones: I have a cell phone. It uses around 14kbps to carry voice, and provides a global service which is extremely pervasive and affordable. Indeed, there are more cell phones than Internet Hosts (2.5 billion active mobile phone numbers in the world at the time of writing). My cell phone provides data (GPRS, EDGE and 3G as it happens). The 3G service runs at around 384kbps in the UK, and seems to have pretty low latency -- I do not know the architecture of the backhaul network once the wireless segment of a route is terminated, but it seems to support pretty close to zero loss. I can run Skype or any vanilla VoIP system on this fairly easily. However, the volume and time tariff of the data service is set such that a normal pattern of voice calls made over it would cost more than the GSM service. This is fairly surreal (in fact, usually when I read my e-mail via my phone, I 'dial-up' over GSM as it is cheaper), but you can see that there are powerful reasons for the cellular network providers to stay in this regime for a while, or else have to explain a massive loss of revenue to their shareholders.

They key lesson here is that *legacy* service providers resist the pressure to become merely bit pipes.

---

[3] This is a UK legal requirement since recent precedents in *unauthorized* access to *open* WiFi nets being deemed an offence under the Misuse of Computers Act.

Digital TV and Fiber: Another UK example is that, as part of the agreement its privatization, BT was explicitly *not* allowed to carry broadcast TV. At one point fairly recently they offered to put fiber to every home and office in the UK if this rule was relaxed, but the government regulator rejected this offer. The argument was that it might create a monopoly of Internet TV, despite the fact that digital TV was until recently a near (non-IP) monopoly in the UK, without the benefit of fiber everywhere. Indeed, the UK's current near-monopoly commercial digital TV provider owns the box in your house, the channel and a large part of the content too, whereas an Internet alternative would create a *pair* of vertical bundles which would compete, but would also allow arbitrary bandwidth access to all the other Internet TV content in the world.

There are several lessons to tease out here: infrastructure and bundles are incommensurable; secondly, the timescales for regulation may often be wrong (both too short and too long), and need constant revision, possibly requiring smart regulated markets[4] rather than fixed franchises.

Preferential Treatment of Customers: The (possibly anecdotal) story of why Strowger invented the automatic telephone exchange is famous, but worth repeating in this context. Strowger ran a funeral business. He suspected that a rival business in town was getting more customers because the telephone operator was the sister of the owner of the rival business, and when asked to connect a bereaved client to a funeral service, would, of course, choose her brother's phone line rather than Strowger's. He automated the bias out of the system.

One lesson here is that a biased service may be entirely innocent at one level, but cause problems at another.

Summing up

Each of the preceding stories illustrates a different aspect of the neutrality argument, whether it is the basic IP service, an overlay service, a provider who owns a last mile and a distinct end-to-end legacy service like voice, or the bundling of content and service and network facilities. All of these arguments existed *before* the Internet existed, but the generality of the Internet allows for *convergence*, and the debate we are seeing is really just another result of the fallout of the final reality of convergence.[5]

In the rest of this paper I look at these aspects of the debate: the IP Service, including current and future evolution (section III); Access Networks (section IV); and Content and Bundling (section V). Finally, I discuss economic aspects briefly, and draw some conclusions (section VI).

---

[4] As is the case with pollution credits and 3G spectrum resale markets.

[5] "Convergence'' is a telecom term for the merging of telephony, television and data services onto a single infrastructure.

III. IP Service: History and Evolution

The Internet provides a Universal Service.[6] However, the service that IP provides is merely connectivity at the network layer, whereas the PSTN (in analog, digital and wireless forms) also specifies delay bounds, minimum capacity, and availability. Parts of the Internet are engineered to provide resources (link speeds) that exceed the needs of new applications, but it is not part of the service or protocol specification, nor is there a forum for agreeing what might be such a part of a service, since *services and protocols are dealt with by different communities*, unlike the combination offered by the ITU.

Recently, there have been a number of concerns about the future evolution of what I might term the Universal IP Service. It has become apparent that for a variety of reasons, the core *connectivity* service is not as Universal as once thought. This is usually to do with security concerns (e.g., about appropriate content or activities) where there are differences across organizations about what is appropriate. Sometimes the IP level connectivity is there, but higher level mechanisms prevent access to sites (e.g., firewalls blocking transport ports). Sometimes, for performance reasons, such filtering is more easily done simply by blocking IP addresses. Sites sourcing much spam or DDoS attacks[7], or other malicious traffic may be black-holed[8] by providers, despite the fact that the cause might involve exploits of an "innocent'' users' vulnerable machines. This leads to a great deal of work in ISPs' call centers, handling requests from these users to be "re-connected.'' In the PSTN, it is much harder for a provider to disconnect a user or exchange unilaterally, without due warning, due to the legal obligations on them to provide, at least, emergency phone call services.

A. IP Service: History

The Internet has been around for 30 years, and the core best effort IP service interconnects around 1 billion devices in the world today. There are a large number of corporate, private and government Intranets, as well as a significant number of interconnected commercial Internet Service Providers (ISPs), which inter-work at the lowest common denominator level, which is to say that connectionless Internet datagram packets can be routed from end system to end system through a collection of intra-domain and inter-domain routers (and associated firewalls, NATs and other devices).

From the very first, the optimizations in routers driven by the statelessness of the IP level has meant that it is hard to introduce enhancements to the core service model such as Quality of Service

---

[6] "Universal Service'' is an ITU term for the minimum set of functions that all public telephony providers must provide within and between their networks, just as the Public Switched Telephone Network (PSTN) provides a Universal Service.

[7] Distributed Denial of Service (DDoS) attacks occur when multiple coordinated calls flood the bandwidth or resources of a target a system (e.g., a website server).

[8] Traffic is simply dropped.

(QoS). The scale of the system is such that now that any alteration to the model must retain backwards compatibility for a significant period. Examples of enhancements such as multicast, Mobile IP and IPv6 succeed or fail on the ability of the new feature to work in parallel with the existing services.

The core service model supports a very simple definition of performance, which is to say that there is none. Instead, it is implicit in provisioning in each segment of the network (and varies with time, and with source/destination) what basic performance one might see. As the Internet evolved, efforts concentrated on maintaining core connectivity with a low cost method to allow applications to co-exist in a shared resource that was based on congestion avoidance and control by end systems. Over time, managing connectivity and congestion in the face of a globally growing net has been met with increasing complexity that generally shows up in two places: (1) Below IP: mapping packets onto various link technologies; and (2) Above IP: in transport (TCP, RTP/UDP) and application (HTTP, P2P[9]) protocols. In addition, a number of practical middle box services have appeared which intermediate network access. Historically, these go back to gateways between different protocol worlds. Now they are used to provide programmatic ways of controlling access between heterogeneous segments of the net for a variety of reasons.

Technology moves on, and as it does, it diffuses through the research, academic, and then commercial networks. This process of innovation is continuous, and has an impact on services.

One of the key areas of evolution in terms of differences between ISPs has been that of Service Level Agreements (SLAs): Many ISPs offer statistical guarantees of performance (above and beyond a simple bland statement of "Best Effort"). For example, zero packet loss is offered by some tier-1 ISPs, while 95th percentile delay guarantees are given by others. Few offer this to traffic transiting to other ISPs. Inevitably, there is a tendency, under competition, to "level up" to the better offerings, as tools for provisioning and traffic engineering become more widely available, and as capacity prices have continued to fall, making the feasibility of pure statistical multiplexing based guarantees easier to achieve (even for VoIP traffic). However, there is little evidence of anyone using the same techniques to "level down'' -- indeed, the sheer numbers of ISPs (e.g., 300+ in the UK alone peering at the LINX) means that any such effort is doomed to lose customers to competitors quickly.

This type of economic dynamic (introduction of new services piecemeal, followed by widespread adoption) seems to have been missed by many commentators on net neutrality.

Next, I describe some of the current realities of the net, which has never been a level playing field for many accidental and some deliberate reasons.

The Internet was never really a level playing field. Recently, many areas of the Internet have tilted so far as to stress the system a little, but the idea that the network is innately fair (for whatever definition of fairness you wish to choose, whether proportional, max/min, or other), is fairly bogus. Some examples of accidental favoritism, effectively wired into the Internet Protocol Suite, include:

---

[9] Peer-to-peer.

- *End-to-end service*: Most traditional Internet applications run on TCP. The throughput you get from TCP depends crucially on (at least) four constraints. First, your bottleneck capacity may be the link speed or system I/O capacity on either end of the connection. Second, the throughput is limited by any other user's TCP flows traversing a shared bottleneck. Third, your capacity is a function of advertised window size, maximum segment size (MSS), and so on. Finally, and most arbitrarily, your capacity is a function of the round trip time and packet loss probability on a link (the latter may simply be a function of the other users' load, but not always). The dependence on round trip time is inverse: the further you are from a sender, the less capacity you get.

- *Inter-domain Routing*: The Internet is rich in numbers of service providers. To reach a site on another service provider's net, your traffic must traverse at least 1 border router. This introduces additional delay, but also, if the path (as often is the case) traverses multiple ISPs, it may be that the return path is not the same. This has a different effect on your traffic than others (e.g., users in the far end's domain, or at different intermediate ISPs). This is not directly intentional -- it is a side effect of the business relationships of ISPs: they are not targeting *you* personally.

- *NATs* (Network Address Translation): We are all too well aware of the whole middle-box debate, so I will not rehearse it here. However, I would say that anyone behind a NAT is not providing a service, so they are not on a level playing field.

- *Firewalls*: I guess the division of the Internet into those places reachable by a first TCP SYN packet to port X, and those not, is another form of balkanization. Of course, the network can always route around damage, but the net-cost of having to implement the superset of damage-avoidance rules may make it infeasible for most mortal users.

- *Proxies*: Caches, as I explained above, are put there to distribute load, and improve users' experience in terms of download delay (in fact, caches are simply a precursor to p2p and torrent ideas). However, caches (and many replication systems) implement rules to control the performance seen by the overall set of users. Indeed, many popular news and software distribution websites now implement *admission control* algorithms to control the perceived performance. The net effect is that users during a "slashdot'' event, see messages that are analogous to call blocking in under-provisioned (or overloaded, e.g., during flash crowd) telephone networks.

What happens when favoritism, or differentiation, is made a network layer first class service? Let us look at that next.

B. IP Service: Evolution

The basic IP service has no real definition (well, there is a definition of Best Effort as part of the Per-Hop Behaviour (PHB) but this doesn't define an end-to-end service). However, many ISPs and some Internet Exchange Points define Service Level Agreements (SLAs) which derive from related thinking in telecom networks of yore. In circuit markets, you buy facilities to connect points with certain characteristics. For example:

- *Isolation*: My traffic is not impacted at all by yours.
- *Protection*: My circuit is backed up to the nth degree by failover paths.
- *Throughput*: I get the capacity I pay for, point-to-point (see later, end-to-end).
- *Delay*: Whatever pattern of packet timings I send with is preserved (i.e., jitter) at the far end, and I see non time-varying delay.

The generality of the Internet has led away from a purely TCP-based (and associated Best Effort tolerant) applications. Now we have a very significant and growing number of users of network applications such as VoIP, IPTV, videoconferencing and networked games. Note that each of these applications has user expectations associated both with performance, *and* with being charged. We are not averse to paying for phone calls, for watching some TV programs, for being charged a lot for (legacy ISDN) videoconferencing, or for paying to be in a game (or even for objects in the game). Internet users now expect to see some of the properties of circuits.

A number of technologies have emerged to support services that look a bit like circuits in the Internet, although most are only deployed within a single ISP, and often, mainly for corporate customers so far.

Differentiation: The IETF community has been struggling with a variety of concepts for introducing Quality of Service mechanisms to the Internet for 15 years or more. The IETF has been steadily tracked by the research community working on better signaling, admission control, and fair queuing algorithms, as well as simplifications of models that allow for ideas such as core-stateless fair queuing, and measurement- and probe-based admission control. Fairly recently, the IETF also was directly trying to address provisioning of priority services for emergency use of the Internet. Finally, we have a simple, but effective technique, which some ISPs have deployed, principally (as far as this writer is aware) to support the legacy services on IP such as VPNs and VoIP backbones for a national telephony service. However, these are good proofs of concept and there are plenty of customers interested in more dynamic service enhancements.

Provisioning: Any technology for QoS assurance of any kind is deployed coupled with a detailed knowledge of the topology of the network, the workload and traffic matrix, and its variation over the day, and a detailed model of all source behaviors. These are then fed into some provisioning model which also contains the traffic engineering mechanisms that the ISP is deploying. This could be based on a tool such as network algebra[10] or an emulation or simulation that is used to compute whether a new user or service can be admitted. The timescales of this are rather different than what was used in traditional admission control for telephone calls, but that is because we have more headroom in today's networks, and we have better tools to comprehend aggregate behaviors in the core.

Note that in the previous discussion I used the word *core*. Of course, the Internet as a whole has no core. It is built out of many Address Spaces (ASs) by many ISPs. Each may have a core network and may use intra-domain provisioning, but the case for inter-domain QoS has yet to be solved.

In the broader global scope, several proposals have been around for a while, including the old Internet 2 and Abilene idea of Brokers, extensions to BGP, and even the use of int-serv/RSVP to allocate inter-domain slices within which differentiation is done. All of this is subject of future work (or breakthroughs!). The inter-domain space is largely "valley free,'' which means that paths traverse up and down the ISPs in a hierarchy of tiers. So in some sense, one could imagine a "core'' at the AS level -- however, tools to reason about performance at this level are not yet available even in research.

Horizontal relationships: As has been observed by BGP experts, the inter-domain routing space has evolved to support a number of business models relating the ISPs either side of a border (and by implication, further afield). Usually, the dominant relations are termed: customer/provider and peering. There are other more complex ones, rarely published.

Vertical Relationships: Application Service Providers and Content Service Providers may have a wholesale relationship with ISPs. For example, a typical content acceleration service has to acquire rack space in data centers, typically co-located with higher tier ISPs' points of presence (POPs). The price for the rack space plus capacity (and other hidden benefits such as secure power supplies, reliable air conditioning, anti-DDoS systems, etc.) may be priced as a bundle. Indeed, buying redundant Internet access for reliability as well as performance (load balancing and lower latency access by having multiple sites around the world) may warrant a bulk discount. However, such agreements are rarely, if ever, published.

In summary, therefore, not all customers are equal.

Piecewise deployments can be seen as potentially applicable to other changes to the core IP service model, such as:

- *Security*: As hinted above, some ISPs provide firewall services in addition to Network Address Translation (NATs) to protect users from unwanted access. In some cases this

[10] See, for example, Le Boudec, J. and P. Thiran (2001) *Network Calculus: A Theory of Deterministic Queuing Systems for the Internet*, Springer LNCS, 2001; or, Cruz, R. (1991) "A Calculus for Network Delay," IEEE Transactions on Information Theory 37 (1): 114-141, January 1991.

may go further and include black-holing of sources of SPAM, and of DDoS attacks (sometimes only on request). The idea of providing more sophisticated security services (e.g., signed/authenticated and approved system distribution for sites) is already common place in private networks, and one can imagine ISPs requiring (and providing) approved systems and system patches to remove vulnerabilities (especially ones demonstrated to allow other sites to own and misuse a customer's machines).

- *Mobility*: The last few years has seen the emergence of Wireless ISPs (WISPs), offering pay-per-use wireless hotspots. Quite a few of these provide *roaming* arrangements, whereby credit on one service can be used on another.

- *Multicast*: IPTV is starting to take off with content problems being resolved, and net performance finally exceeding the threshold necessary to offer reasonable quality real-time TV. However, some live events may be of primary interest to large groups, and we may see pay-per-use IP multicast finally take off. On the other hand, P2P TV is also emerging as a model which doesn't stretch the ISP at all, but meets end-user requirements provided enough up-link capacity is available from participating customers. The ISP might in either case, broker the content and rights.

The key argument in the neutrality debate about differentiation lies in the question: does one level up or down? When offering a new service with higher performance, clearly any serious business will price and provision things so that the lower tariff offers lower performance. But what is the trend? Is the additional income used to provide more capacity so that the "poor'' do better, while the "rich'' do even better? Or is the capacity shared in a different way, so the rich win at the "expense'' of the poor? The jury is out, but you can bet your life it is a zero sum game at any instant.

IV. Access Networks

An entirely different version of the net neutrality debate concerns the access network. Here, there is some evidence that we are re-playing the arguments that led to the divestiture of AT&T all those years ago. It is worth remarking that the competition in local loop in different parts of the world varies enormously, and so one has to be very careful whether this is really a general debate, or one that reflects lack of competition in the local loop. As I hinted in the introduction, this sort of debate can also be held concerning wide area wireless (cellular) access, and has been noted in the previous section, it could also apply to WiFi pay-per-use hotspots.

Legacy services with vertical bundles (e.g.,, the PSTN, with phone line which happens also to be the last mile access for IP and cable TV) are crucial to many users of the Internet. The operators who own these local loops are quite heavily regulated in many parts of the world, in terms of telephony, and in terms of allowing competition access to the *exchange* end (or *head end* in the cable case) of the lines. Whether the line/access are bundled or unbundled is crucial.

The costs associated with maintaining 100s of millions of phone lines are quite high. The cost of deploying ever increasing speed DSL kit at the exchange ends is also high, and many incumbents would like to offset this by increasing charges. The cost of providing an alternative is also high, although fixed wireless broadband is a possibility looming on the horizon, as is the replacement of the entire access net with fiber in highly developed parts of the world such as Korea and Japan.

However, if the operator that owns the last mile also still owns significant long haul networks, and wishes to capitalize on both, there is a strong incentive to provide some modest level of *walled garden* by offering improved access link speed, provided some bundle of higher level service is subscribed to. This is entirely familiar to telephone users, digital TV users, and cellular telephone subscribers.[11]

The real question here is whether the last mile needs to be regulated, for example when there is a near monopoly *and* the provider behaves monopolistically. If that occurs, regulation can ensure performance and bundles are transparently measurable and priced, and alternatives (or potential alternatives) are evaluated on a level playing field by regulators and understood by consumers. This is one area where it seems to me the current regulatory frameworks (especially where this writer works, in the UK) have many of the right components, and there may not need to be any new definitions of neutrality. The Internet is just another service using these potential bottleneck facilities.

V. Content and Bundling - Overlay Services

One of the grand challenges to net neutrality noted by a number of companies in the U.S. net neutrality debates is the concern that some ISPs might seek to block or lower performance to certain applications *en masse*. The statements made by some ISPs implied that overlay services that are crucial to many users of services such as VoIP and Web Search engines (specific examples of course being Skype and Google) were *free riding*.

This emotive term was used almost certainly by marketing people, since it has connotations of illegal file sharing and piracy. However, most large scale overlay systems buy significant quantities of Internet access at very high speed, and (more importantly) buy it from many ISPs in data centers in POPs (as discussed in the previous section) so that they can offer a global application service. In other words, they are not riding for free at all. Nevertheless they make a lot of money, and ISPs that only offer IP packet transport are unsurprisingly jealous of that revenue.

Let us think about that for a bit because it is really quite amusing. An ISP is not forbidden from also being a content service provider (modulo certain special cases such as the BT TV example I mentioned earlier). An ISP that has data centers could build its own VoIP call-out service, and its own search engines. Indeed, it might be able to pinpoint "click-through'' far more accurately than a

---

[11] It is worth noting that the regulatory/industry situation is *very* different in Europe, the U.S. and Asia with regard to joint versus separate ownership of access and core networks, which also leads to confusion in this part of the debate.

search/lookup service at lower cost simply by monitoring network access patterns. However, what is the effect of "taxing'' the profit from overlay service providers? Well there are two possible outcomes: firstly, the service cost is passed on to the consumer (and the net profit decreases); or the service provider leaves the network (analogous to Google not indexing Belgian Newspapers as per a recent event). The effect is to *damage* the ISPs core business. The point is that there *is already a value chain* between clients, websites and search engines, and between broadband Internet clients and VoIP service providers and the ISP. The profit made by the overlays is *not independent* of the profit made by an ISP. Of course if the ISP is not making a profit, and the customers are, then the ISP should simply raise its prices transparently. Why would you want the market not to be free? The fact that they don't raise prices, and some ISPs don't make a profit speaks to some other problem.

A completely separate neutrality argument arises concerning the different kinds of content filtering, or *censorship* carried out at various levels (IP and above, e.g., by search engines) in different parts of the world. Technically, I do not feel competent to comment on this, but I would observe simply that the same rules are applied to postal service (e.g., for books, DVDs etc), and that the customer can work around those rules but takes the risk of breaking the law. Most cases I have read about in this area are merely reasonable observance of local variation in what is legal (e.g., pornography laws in the UK are more strict than most of the rest of Europe, holocaust denial is illegal in several, but not all countries in Europe, etc.).

VI. Economics and Neutrality

Many of the economists arguing about neutrality have observed that the Internet has been an engine for innovation unsurpassed by earlier playgrounds. The consumer has seen a remarkable improvement in wealth through expanded services, increased performance *and* cost reductions all at the same time.

The neutrality proponents argue that the role of the Internet as a platform for innovation is good and needs to be preserved, but the neutrality opponents argue that we are reaching the limits of this part of the Internet evolution. As with other industries (e.g., the car tire industry), after a period of quality improving evolution, one sees a shift to process engineering, where optimization focuses on the details of how a service is operated, rather than discovering whole new service offerings.

This matches our experience in the lower layers, where the core IP service saw a fair amount of evolution in the 1980s and very early 1990s, but then the action moved on up the stack to TCP and RTP/UDP evolution, and eventually HTTP and Web Service evolution, and now on to multi-party application evolution (e.g., P2P, games, etc).

Now that the Internet has grown up, determining how valuable preserving the ability to innovate further is difficult. What further innovations are of greatest value to consumers? To the economy? What is the replacement of digital TV by IPTV worth, or VoIP instead of GSM? Or P2P movie distribution instead of Netflix?

Part of the debate is about trying to define what the IP service is so that regulations or legislation can be properly focused. As I have outlined in the paper so far, the neutrality debate ranges up and down the Internet Architecture and this architecture is, itself, not a static thing. Any appropriate neutrality definition would need to address the different senses in which the Internet may (or may not) be neutral and how this is changing over time. As a strawman, I propose the following multi-part meta-definition for discussing the various components of the net neutrality debate:

- *Connectivity Neutrality* must be defined with respect to end-to-end service at each and every layer.

- *Performance Neutrality* must define rules for SLAs (existing ones and new ones with appropriate delay bounding services for IP TV) in a measurable, comprehensible and transparent fashion.

- *Service Neutrality* must define rules for availability of new net services such as multi-home, multicast, mobility, etc. in a way that allows cross-provider/cross-platform differences to exist until these services have sufficiently matured.

- *Cross Layer Neutrality* must define how combinations of services are built and how the consumer gets to choose between them.

However, having defined neutrality thus, I believe that these are Platonic ideals to which we might strive, but never attain. The system of innovation in the Internet community depends both technically and economically on differences, and the static models of neutrality fail to capture the essential living dynamics.

VII. Conclusions

The net neutrality argument is a debate between stakeholders with radically different goals. Libertarians and Liberals both argue in terms of welfare. The Internet is an evolving platform that must serve the dual masters of competitive market dynamics and ongoing innovation. The Internet's evolution has thrived on differences that through one lens may appear as non-neutral treatment while through another may appear as vigorous and healthy competition. Regulators who seek to impose a rigid or static definition of net neutrality would be well-advised to heed the lessons of Internet history and the examples cited herein. If a definition is appropriate, it must be one like the meta-definitions I propose above that articulates ideals but is robust to change.

In conclusion then: We never had network neutrality in the past, and I do not believe we should engineer for it in the future either.