A Proposal for Network Neutrality

Tim Wu, Associate Professor of Law
University of Virginia Law School,
580 Massie Rd., Charlottesville, VA, 22903
timwu@virginia.edu / (202) 421-5445
June 2002

Summary

The growth of home broadband has created a new and difficult regulatory problem. Few dispute that broadband operators need the freedom to manage their networks to ensure maximum efficiency. Yet there is growing evidence that carriers can restrict the use of their broadband networks in ways that distort the market for internet applications, home networking equipment and other markets of public value.

The regulator is faced with a challenge: What principle can balance the legitimate interests of broadband carriers in administering their networks with the danger of harm to new application markets? And how can such a principle be translated into both clear legal guidelines and the practice of network design?

This proposal introduces the principle of *network neutrality* or *non-discrimination* as a tentative answer to these questions. As a general description, the proposal would strike a balance: it would forbid broadband operators, absent a showing of harm, from restricting what users do with their internet connection, while giving the operator general freedom to manage bandwidth consumption and other matters of local concern. The principle achieves this by developing "forbidden" and "permissible" grounds for discriminating among packets on its network. Generally speaking, the forbidden grounds are inter-network indicia, such as IP address or application type. Conversely, the allowable grounds for restriction are local indicia—particularly, bandwidth.

As the D.C. Circuit once said of the Bell system, the consumer has a right "reasonably to use his [connection] in ways which are privately beneficial without being publicly detrimental."[1] This tentative proposal represents an effort to find that balance.

[1] Hush-A-Phone v. United States, 238 F.2d 266, 268 (D.C. Cir. 1956).

I. The Problem Described

Broadband providers have some natural incentives to deliver unrestricted internet access. The unrestricted product is attractive to end users, and restrictions would serve to make a competing services (DSL or cable, respectively) more attractive. But recent experience has shown that broadband providers will impose, in addition to reasonable restrictions on end users, some restrictions that are troubling. Documented examples of troubling restrictions include limitations on the use of Virtual Private Networks (VPNs), limits on the types of equipment subscribers can attach to the network, and additional charges for certain forms of applications (as opposed to bandwidth).

A loss of consumer welfare through distortion of markets for new applications forms the primary cause for concern. In economic terms the problem can be expressed as an externality problem. It may narrowly make sense for a cable company to prohibit the use of a given application, either to keep its own costs low, or to protect its own, competing product. But the broadband operator will not be taking into account the externalized costs of such action—distortion of the market for the application.

For example, broadband operators may decide to prevent their connections from being used for Internet VPN services, because the additional revenue they derive from selling VPNs exceeds the money they lose from customers who refuse to subscribe to broadband for this reason. But the costs of prohibiting VPNs are felt elsewhere – in the retardation of the market for VPN services, and in the consequent loss of employee productivity nationwide.[2]

Usage restrictions can also be described, more informally, as an impediment to competition in application and home networking markets. At their worst, broadband usage restrictions can resemble the "foreign attachment" tariffs that lasted from 1913 into the 1970s, forbidding customers from attaching equipment not manufactured by Bell to the telephone network. There is little disagreement that such restrictions retarded the CPE market and disserved consumer welfare.[3]

All this makes the case against usage restrictions. But what makes the problem difficult is that that are also usage restrictions that are undeniably

---

[2] Some might argue that competition between broadband carriers would automatically solve this problem. But if the externality exists equally for all broadband carriers, first principles suggest that competition may not solve the problem. If, for example, in a given region, if both cable and DSL providers can make serve their own interests by banning VPN services, then they may not compete to allow such services. In economic terms, this resembles an oligopoly pricing problem.

[3] See Huber et al., Federal Telecommunications Law §8.4.1.1 (2d ed. 1999).

reasonable. Operators must also have the freedom to manage bandwidth, and prohibit uses of the network that damage the integrity of the network or seriously impinge the rights of other users. Such restrictions are necessary if broadband carriage is to be a viable business.

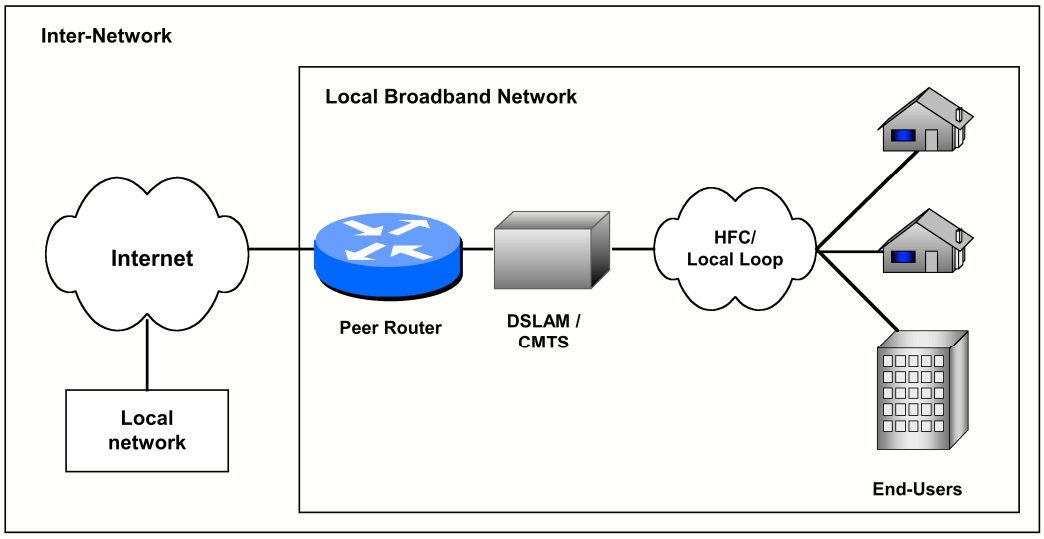
This proposal develops the principle of *non-discrimination* or *network neutrality to* balance the legitimate interests of broadband carriers in administering their networks with the danger of harm to new application markets.

The non-discrimination principle works by recognizing a distinction between *local network* restrictions, generally allowable, and i*nter-network* restrictions, viewed as suspect. The principle represents is ultimately an effort to develop *forbidden* and *permissible* grounds for discrimination in broadband usage restrictions.

II. Let Operators Police What They Own

Broadband carriers are members of two networks. They are members of a local network, which they own and manage individually. They are also members of the inter-network, which they collectively manage with other service providers.

Figure 1: **Broadband Carriers, Members of Two Networks**



Once we recognize that carriers are engaged in a collective management scheme, the origin of the externalized cost problem described above becomes clear. The effects of local network restrictions will, in general, affect only the network run by a single service provider. Such restrictions moreover, are

necessary for good network management. In contrast, by definition, restrictions at the internetwork layer or above will always affect the entire network, and are can create externality problems.

II. The Neutrality Principle & Its Legal Implementation

We can now state the inter-network neutrality principle, operationally, as a non-discrimination rule. First, we must clarify what is meant by "discrimination" in this context. It means to treat network traffic differently on the basis of certain characteristics; either with adhesive contracts that forbid users from receiving such traffic, or programming network equipment to block it. For example, an operator might "discriminate" against traffic from the game site "www.gamespy.com" if it blocked traffic from 127.12.23.1. It might also do the same by enforcing subscription agreements that bar access to game sites.

A total ban on network discrimination, of course, would be counterproductive. Rather, we need distinguish between forbidden grounds of discrimination—those that distort secondary markets, and permissible grounds—those necessary to network administration and harm to the network.

Reflecting the dual-network membership just described, generally it will be internetwork criteria of discrimination that cause concern. In technical terms, this means discrimination based on IP addresses, domain name, cookie information, TCP port, and others as we will describe in greater detail below. Hence, the general principle can be stated as follows: absent evidence of harm to the local network or the interests of other users, broadband carriers should not be allowed to discriminate in how they treat traffic on their broadband network on the basis of internetwork criteria.

The negative inference is that operators generally *may* discriminate in their treatment of traffic on the basis of *local* network criteria. In technical terms, this means imposing restrictions on the basis of what network engineers call "link" or "layer 2" information, like bandwidth, jitter, or other local Quality of Service indicia. In the next section, we can see how this distinction would work.

III. In Practice: The Example of Online Gaming

Popular online gaming applications[4] like *Everquest*, *Asheron's Call*, or *Online Quake* tend to be bandwidth intensive, particularly compared with episodic applications like email. Concerned broadband carriers have therefore been inclined to restrict the usage of such applications. However, with the neutrality principle in mind, we can distinguish between a "right" and a "wrong" way for this to happen.

[4] Also commonly referred to as "Massively Multiple Online Games," or MMOGs.

First, in today's environment, a broadband carrier could block traffic from gaming sites. It could do either by enforcing a contractual provision in a usage agreement, or more plausibly, use its control of the local network to block traffic from gaming sites based on either "layer 7" application information, or the IP address of the application provider.[5] Some carriers might elect, for a given supplemental fee, to remove the filter for specified users.

Under the neutrality principle here proposed, this approach would be impermissible. Instead, a carrier concerned about bandwidth consumption would need to invest in policing bandwidth usage, not blocking individual applications. Users interested in a better gaming experience would then need to buy more bandwidth – not permission to use a given application.

The neutrality of such control would prevent a distortion in the market for Internet applications. If carriers choose to block online games in particular, this gives a market advantage to competing application that have not been blocked. But if broadband carriers only police bandwidth, the result is an even-playing field. It may be the expense of more bandwidth lead people to choose different ways to spend their money. But if so, that represents a market choice, not a choice dictated by the filtering policy of the broadband carrier.

III. Borrowing from Well-Established Categories

One advantage of the proposal is that it relies on well-established legal and technological criteria to achieve its consumer-welfare goals. Respectively, it borrows from principles of harm requirements and non-discrimination familiar to lawyers, along with a local / inter-network distinction that is fundamental to datacom networks.

The Harm Requirement

In the telephony context, the "foreign attachment" problem discussed above was addressed by a "harm" rule; that is, a rule barring the Bells preventing attachment of equipment unless harm to the network could be shown. Its origins are found in the *Hush-a-Phone* case, where the FCC ordered Bell to allow telephone customers to attach devices that "do[] not injure … the public in its use of [Bell's] services, or impair the operation of the telephone

[5] For an explanation of how a broadband carrier would do so, see, e.g., *The Cisco Content Delivery Network Solution for the Enterprise*, Cisco White Paper (April 2002) available at http://www.cisco.com; Cosine, Inc., Digital Subscriber Lines and Managed Network-based Services: A Perfect—and Profitable—Marriage, Cosine White Paper, available at http://www.cosine.com.

system."[6] As the D.C. Circuit stated in that litigation, the consumer has a right "reasonably to use his telephone in ways which are privately beneficial without being publicly detrimental."[7]

In the broadband context, it is discrimination against certain content and applications is the major problem. But the justification of requiring public harm to justify restrictions can be usefully employed.

Discrimination

The principle of permissible and non-permissible bases of discrimination is a familiar legal tool. In employment law, for example, employers may generally fire or refuse to hire individuals for a range of reasons, such as education-level, intelligence, and demeanor. The law recognizes that it is essential that the employer retain the freedom to fire incompetents and hire only those with necessary skills. On the other hand, criteria such as race, sex, or national origin are forbidden criteria of discrimination.[8]

While discrimination among Internet packets is a different context, the principle is the same. In the employment law context, as in the broadband context, it may often be the case that discrimination would actually serve the narrow self-interest of the employer in question. The reason for the ban, however, is public; the ban on discrimination serves broader economic and social interests. The general need to strike a balance between legitimate private and public interests in discrimination are shared in the broadband and employment context.

Local / Inter- Networking

Finally, on the technological side, the distinction between inter-networking and local networking is very well established in the datacom industry. While the distinction is best reflected and usually discussed in the context of the OSI network reference model (as the difference between layer 2 and layer 3 networks),[9] it is in fact independent of OSI. As a practical matter, different physical equipment and different protocols run the different networks. In a given network, "switches" run local network, while "routers" collectively manage the layer 3 network. Services can be offered at both levels -- for

---

[6] Hush-A-Phone Corp. v. AT&T, 22 FCC 112, 114 (1957). This led in turn to the broader *Carterphone* decision, 13 F.C.C.2d 420 (1968), and finally Part 68, which adopted a protective circuitry approach to protecting the telephone network, see 47 CFR §68 *et seq.*

[7] Hush-A-Phone v. United States, 238 F.2d 266, 268 (D.C. Cir. 1956).

[8] See generally 42 U.S.C. § 2000e *et seq.* (codification of Title VII of the Civil Rights Act of 1964).

[9] Cf. Andrew Tanenbaum, Computer Networks 10-18 (4th ed. 2002).

example, VPNs and telephony can be offered either as a layer 2 service or as a layer 3 service.

In addition, other schema used to describe network layers embody the same, fundamental, local / internetwork distinction. For example, the TCP/IP network model maintains a distinction between the "link" layer and the "network" layer. This is exactly the same distinction as the layer 2 / layer 3 distinction in the OSI model, and the local / internetwork distinction more generally. Again, this is no surprise, because virtual description simply reflects the physical network design. The existence and pervasiveness of the local / internetwork distinction makes it a natural dividing line for reasonable restrictions on use.

Of course, greater detail can be achieved by specifying forbidden criteria of discrimination. But the basic local / inter network distinction provides an existing line to draw upon.

IV. The Technical Meaning of the Local / Internetwork distinction

At this point, we can spell out what the anti-discrimination principle means if a basic local / internetwork distinction were implemented through regulation. As we can see from the following table, local network restrictions tend to be necessary to good network management, while internetwork distinctions have the potential to retard the market for consumer applications

Table 1: Local and Inter-Network Usage Restrictions

Local Network Restrictions
Limits on Bandwidth
Quality of Service within the service providers' network
Reaching local devices
Local broadcast, multicast
Policing of Ethernet/Frame Relay equipment

Inter-Network Restrictions
Blocking of specified Internet addresses
Limits on acting as a server
Bans on Internet VPN services
Bans on certain applications by TCP port number
Bans or limits on applications (based on cookie information).

V. Objections

Before concluding, it will be useful to consider some objections and challenges to the neutrality principle. We consider (1) whether it overly interferes with broadband carriers' ability to earn a return on their infrastructure investment; (2) whether local restrictions can be used to achieve the same problems as internetwork control, and (3) whether the principle interferes with administration of internet addressing.

The Return on Investment Question

First, does the neutrality principle restriction overly impinge the ability of broadband carriers to earn a return from their infrastructure investments? While a full analysis of broadband economics is beyond the scope of this proposal, we can nonetheless suggest that the neutrality principle is unlikely to interfere with the special advantages that a carrier gains from building their own infrastructure.

The simple answer is that investing in a local network infrastructure creates its own rewards, as it creates particular advantages in the offering of network services. We can see this clearly by considering the particular example of Virtual Private Networks under the neutrality principle. A broadband operator who owns the local infrastructure has a natural advantage in offering local VPN services. The advantage comes from the fact that they can offer service level guarantees that cannot be provided on a shared network. Proof that this is the case comes from the continuing growth of Frame Relay and ATM services, both of which are premised on ownership of a local network that supports VPN services.[10] Nothing in the neutrality principle would prevent a broadband operator from being in the unique position to sell such services.

But the principle would prevent operators from blocking use of internet VPNs – that is, VPNs that used the internet to reaches sites that no single local network can encompass. For example, a home user on the East Coast to connect to his business on the West Coast will almost certainly need to use an internet VPN. In offering this service, a broadband operator is in the exact position as any other internet VPN provider. Restricting use of internet VPNs should therefore not be allowed, to preserve undistorted competition for this application.

Can Local Control Disrupt Application Markets?

Some might observe that the local and internetwork are interdependent in certain ways. Won't broadband operators simply use their control over the local network to achieve the same distortion of application markets?

---

[10] See Yankee Group, Continued Growth in Frame Relay & ATM (March 2001).

No rule can perfectly stamp out all undesirable behavior. The point of the network neutrality principle is to make interference with the application markets much harder. Without the ability to discriminate on the basis of the origin of a packet or the application being used, the broadband carrier is left with the far blunter tools of local restrictions.

It might be argued that the address resolution protocol (ARP)[11] could be used to achieve the same goals as IP-address filtering, since the job of ARP on a typical network is to convert IP addresses into Ethernet MAC addresses. But in fact a broadband carrier manipulating ARP could only succeed in making his own users unreachable. The ARP-cache only holds the information to match up local physical addresses with local IP addresses. ARP has no idea how to stop a user from reaching a specific IP address, other than making that user unreachable. The example shows, in fact, the power of limiting a broadband carrier to local control.

The Need to Administer IP

Finally, some might point out that broadband carriers must have some control over the internet protocol side of their network. They must, for example, be able to allocate static and dynamic IP addresses, maintain routing tables, and so on. Does the network neutrality principle interfere with this?

The point of the neutrality principle is not to interfere with the administration of the internet protocol side of a broadband carrier's network. It is, rather, to prevent discrimination in that administration. Since it is phrased as a non-discrimination principle, a negative inference is that most aspects of IP administration can be conducted without concern. For example, the allocation and administration of IP addressing should not pose any discrimination problems, so long as the administration of such addresses in an even-handed manner.[12]

Conclusion

The neutrality principle here proposed would allow consumers to reach any internet application or operate any kind of home network while also preserving the ability of operators to police network abuse. Grounding the distinction in long-established aspects of legal doctrine and network design

[11] Described in IETF RFC 826.

[12] In today's environment, the scarcity of IPv4 addresses does appear to justify a form of discrimination: charging more for static addresses, than dynamic addresses. This forms a good example of "permissible" discrimination.

creates categories much more difficult to manipulate, and more likely to provide clear guidance.