TPRC, The 35th Research Conference on Communication, Information and Internet Policy, Sept. 2007

Net Neutrality through Open Interfaces[1]

SCOTT JORDAN
University of California, Irvine
Dept. of Computer Science
sjordan@uci.edu

Abstract

Net neutrality was the most contentious communications policy issue considered by Congress during 2005-2006. The issue is the result of a fragmented communications policy unable to deal with technology convergence. In this paper, we develop a net neutrality policy based on the layered structure of the Internet that gracefully accommodates convergence. Our framework distinguishes between discrimination in high barrier-to-entry network infrastructure and in low barrier-to-entry applications. The policy prohibits use of Internet infrastructure to produce an uneven playing field in Internet applications. In this manner, the policy restricts an Internet service provider's ability to discriminate in a manner that extracts oligopoly rents, while simultaneously ensuring that ISPs can use desirable forms of network management. We illustrate how this net neutrality policy can draw upon current communications law through draft statute language. We believe this approach is well grounded in both technology and policy, and that it illustrates a middle ground that may even be somewhat agreeable to the opposing forces on this issue.

1. Introduction

During 2005 and 2006, Congress worked on the most substantial rewrite of the nation's communications law in a decade. The issue initially motivating this task in both houses of Congress was video franchising, which attempted to streamline the laws regulating carriers offering broadcast video services, with the goal of enabling faster deployment of video products by telephone companies. The House bill included not only video franchising, but also a weak version of net neutrality, voice over IP (VoIP) interconnection, e911, and municipal broadband. The Senate bill included an even wider range of issues, including all of the issues in the House bill, plus universal service, broadcast flag, white spaces, interoperable emergency communications, DTV, wireless preemption, tax moratoriums, and many other issues. Among all of these communication issues, net neutrality was easily the most contentious. Lack of resolution of net neutrality was generally given credit for derailing the communications bill.

Net neutrality represents the idea that Internet users are entitled to service that does not discriminate on the basis of source, destination, or ownership of Internet traffic. The idea is rooted in the manner in which the Internet has historically operated, in which all traffic is forwarded as quickly as possible, with limited differentiation

[1] These ideas do not necessarily represent the views of the University of California, IEEE-USA, or Senator Ben Nelson.

based on the application and without any performance guarantees. There is great disagreement, however, about the future implications of this relatively simple idea as the Internet progresses and as the economic communications landscape changes.

Proponents of net neutrality (generally, application providers and consumer groups) argue that without a prohibition on discrimination, Internet Service Providers[2] (ISPs) may charge application providers discriminatory prices for access to dedicated bandwidth or for quality of service (QoS), or may outright block access to certain applications or websites, and that such activity will inhibit development of new Internet applications. Most proponents believe that ISPs should not be allowed to charge for priority access to the Internet portion of their service offerings.

Opponents of net neutrality (generally, ISPs) argue that there is no current problem, that competition is sufficient to ensure that commercially negotiated arrangements for bandwidth or QoS will not negatively impact consumers, and that any regulation will discourage investment in network infrastructure.

The topic is timely for a combination of technological and policy reasons. On the technology side, there is a steady convergence between the networks traditionally used to offer voice, video, and data services. Most networking technologists expect that the technical differences between telephone networks, video networks, cellular networks, and the Internet to diminish in future years, with all of these networks becoming capable of efficiently supporting a combination of voice, video, and data services. Competition between the carriers supporting the infrastructure of these various networks and application providers offering a wide range of services will thus intensify.

On the policy side, U.S. federal communications policy was separately developed for telephone networks, cable video networks, and cellular networks. In the absence of any explicit statutes regarding Internet access or services, the Federal Communications Commission (FCC) recently declared that Internet access is not subject to the same common carrier regulation that addresses telephone networks, which had the effect of removing from Internet access several prohibitions on discrimination that were included in common carrier regulation. This lifting of discrimination constraints triggered the push for net neutrality.

Net neutrality was a heavily lobbied issue in 2006. Estimates of the amount spent on net neutrality advertising and lobbying range up to $100 million[3]. The issue also attracted wide attention on the Internet and in the media. The term *net neutrality* appears on more than two million web pages[4]. Net neutrality was the topic of editorials, commentaries, and news articles in dozens of national and local newspapers and magazines, and garnered some attention on television, ranging from *Moyers On America* to *The Daily Show*.

The academic literature on net neutrality has been split on the issue. Openists (Bar et al., 2000; Lemley & Lessig 2001) believe that the Internet is best served by maintaining a *dumb network* that does not differentiate among different types of traffic. They support a policy based on *open access*, in which Internet infrastructure and applications can not be bundled using either technical or business mechanisms. They argue that vertical integration

[2] In this paper, we use the term ISP to represent Internet access providers, namely carriers who provide last mile access to residential and business customers. ISPs can provide two types of service - access to Internet infrastructure, e.g., DSL, and application services, e.g., e-mail. This distinction is made later in the paper when we introduce definitions to be used in statute language.

[3] See e.g., Bloomberg (2006).

[4] Google search on 6/14/2007.

harms consumers, that most innovation comes from application providers, and that open access will maximize social welfare. Openists support a strong version of network neutrality.

Deregulationists[5] (Owen & Rosston 2003; Yoo 2005) believe that ISPs are in the best position to determine the most beneficial evolution of the Internet. They expect that the Internet will become a *smart network* that uses traffic discrimination in order to accomplish product differentiation. They support a policy that deregulates the Internet in a manner that allows ISPs to vertically integrate, bundle services, and use traffic discrimination as they see fit. They argue regulation will hinder investment by ISPs and that in the absence of regulation ISPs will only vertically integrate in ways that maximize consumer welfare. Deregulationists oppose network neutrality.

There is a limited amount of academic literature that attempts to strike middle ground. Nondiscriminationists (Wu 2004; Peha 2006) believe that there are good and bad uses of traffic discrimination. They support a policy that allows vertical integration and traffic differentiation, but restricts their use to ensure that ISPs do not discriminate in a manner that extracts oligopoly rents. They argue that such a balanced approach will allow development of a smart network in a manner that does not restrict development of applications. They would support a limited version of net neutrality that falls short of open access, but have not yet fully formulated such a policy.

Finally, there is an academic literature that addresses net neutrality and related topics in the context of Internet architecture (Clark, Wroclawski, Sollins, & Braden 2005), broader revisions of communications policy (Bar & Sandvig 2000; Werbach 2002; Solum & Chung 2003), economic analysis (Speta 2000; Farrell & Weiser 2003; van Schewick 2007), and market analysis (Lehr, Gillett, Sirbu, & Peha 2006).

In this paper, we argue that net neutrality is, at its core, an attempt to address problems posed by a fragmented communications policy unable to deal with technology convergence. We adopt an approach jointly grounded in Internet technology and communications policy. We argue that the evolving layered Internet architecture supports the model of a smart Internet that allows only certain types of discrimination. We accept the premise that vertical integration between infrastructure and applications poses potential threats to a level playing field. We suggest that an important tool in solving such problems is a proper delineation of Internet infrastructure and Internet applications. We illustrate how such a delineation can be used to restrict an ISPs ability to extract oligopoly rents through discrimination, while simultaneously ensuring that ISPs can use desirable forms of network management. We further illustrate how this use of layering can appropriately limit the scope of regulation. Finally, we suggest that net neutrality can be addressed in a manner consistent with current Federal communications law, and we propose draft statute language on this basis.

Our analysis is heavily based on communications technology and historical communications policy. Our approach belongs in the Nondiscriminationist camp. Our solution essentially mandates the use of open interfaces, but does not mandate full open access.[6] Although we believe that there are many other interesting ways to address net neutrality, we hope that our convergence-inspired layered approach will illustrate a solution that is well grounded in both technology and policy, and that may even be somewhat agreeable to many of the opposing forces on this issue.

---

[5] The terms *openist* and *deregulationist* were introduced in Wu (2004).

[6] We do not take a position here on open access, as this is outside the scope of the paper.

The paper proceeds as follows. In section 2, we briefly discuss technology convergence, explain the layered Internet architecture model, and analyze evolving technological developments. In section 3, we review the relevant academic literature, present pro and anti net neutrality lobby arguments, and characterize the various congressional approaches to net neutrality. In section 4, we illustrate how a proper delineation of Internet infrastructure and Internet applications can serve as a foundation for a new net neutrality policy that restricts an ISP's ability to discriminate while allowing reasonable network management. In section 5, we delineate acceptable and unacceptable uses of network management. Finally, in section 6, we look to current Federal communications law for guidance, and translate our policy into draft statute language.

2. Communications Technology

2.1 Convergence

In operation since the 1890s, telephone networks have traditionally offered voice service. In the past 25 years, most telephone companies have deployed fiber in portions of their networks, and transitioned from analog to digital transmission in the network backbone. With the extra capacity and capabilities this provides, most telephone companies now offer voice service, video conferencing, and Internet access. Most of the Internet backbone runs over telephone networks. In addition, many telephone companies have recently announced plans to offer video service.

Cable video networks were the next significant entry into the field, and were designed to offer broadcast video service. Since their introduction in the 1970s, subscribership to cable and DBS networks has increased to about 86% of U.S. households. Earlier networks were analog and broadcast-only systems. Recently, many cable systems have migrated to digital transmission, which allow for a vastly increased number of channels, Internet access, and telephone service.

Cellular networks were introduced in the 1980s, and were initially designed to offer voice service. The first generation of cellular phones were built on an analog platform, but the second and third generations are digital. In the U.S., there is now an average of about 1.7 cell phones per household, and the number of cellular phones has recently eclipsed the number of wired phones. Future wireless networks are expected to transition to an architecture that more closely resembles the Internet, and wireless data services are expected to grow quickly.

The Internet, created in 1969 by the Department of Defense Advanced Research Projects Agency, was designed to offer file transfer service. Taking advantage of openly published rules of operation and freely distributed software, many research and educational institutions attached their computers to the Internet during the 1970s. The network has largely blossomed, however, due to the introduction of personal computers during the 1980s and the development of the World Wide Web in the 1990s. In the U.S., approximately 58% of households have Internet access[7]. The Internet is now in the process of transitioning toward an architecture that can more efficiently support real-time applications such as voice and video.

These four technologies (telephone networks, cable video networks, cellular networks, and the Internet) began to merge in the 1980s, with the introduction of optical fiber and advanced computer capabilities. Economic and regulatory forces have reinforced the technical trends. Recent trends are based on an increased desire for multimedia involving voice, data, and video. A flurry of mergers between communications companies has resulted, bringing together content providers, broadcasters, cable TV networks, local and long distance telephone networks, wireless networks, and Internet service providers.

[7] Approximately 73% of households that have Internet access use broadband.

Most networking technologists expect that the technical differences between telephone networks, video networks, cellular networks, and the Internet to diminish in future years, with all of these networks becoming capable of efficiently supporting a combination of voice, video, and data services.

Convergence is a principal cause for net neutrality's timeliness, since the motivation for differentiation between different Internet traffic lies in the carriage of voice and video traffic. We turn in section 2.3 to these specific developments. First, however, we must discuss layered Internet architecture.

2.2 Layered Internet architecture

Telephone networks, cable video networks, cellular networks, and the Internet are all based on the concept of a layered architecture. Each network device, and the network as a whole, is abstractly modeled as being composed of a number of vertical layers. Each layer provides certain functionalities.

Layering is a form of modularity. In modular design, a designer of a module need only understand the functionality and the *interface*, not the detailed operation, of other interoperating modules. Although designing a component in a modular fashion restricts the design space, the benefits usually outweigh the costs.

In a layered architecture, it is useful to think of two types of interoperation. First, within a network device, modules are only allowed to interoperate with other modules that are at most one layer above or below; typically, the higher layer requests the lower layer to accomplish some task, and the lower layer eventually responds with a message indicating success or failure. Second, modules that are part of two separate network devices are only allowed to interoperate with each other if they are members of the same layer. The benefit of these rules is that a very large number of network devices using a large number of network protocols can interoperate with each other. For instance, you and I can send each other e-mail even though we may use different e-mail programs.

The reference model for layered architectures is the OSI model, developed by the International Standards Organization. The OSI model is composed of 7 layers, as pictured in figure 1. It is useful to think of the physical connection, e.g., wire, as being located below the bottom-most layer (layer 1) and the user, e.g., you, as being located above the top-most layer (layer 7).

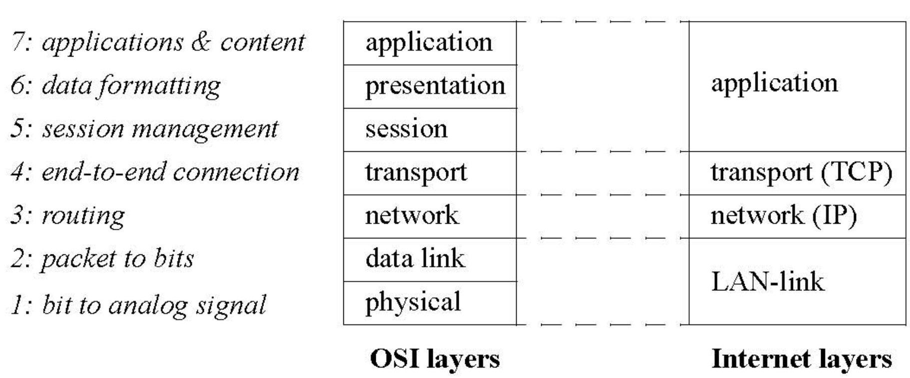

Figure 1: OSI and Internet layered models

OSI layer 1, called the *physical layer*, implements encodes a bit into a physical signal and vice versa. OSI layer 2, the *data link layer*, translates a packet into a set of bits and vice versa, and implements a set of rules (called a *protocol*) about which device can transmit when. OSI layer 3, the *network layer*, is concerned with routing a packet from one network device to the next. OSI layer 4, the *transport layer*, is concerned with functionality required to form a complete connection between a source and destination, including dealing with lost packets and responding to congestion. OSI layer 5, the *session layer*, manages an entire communication session, e.g., logging onto a service. OSI layers 6, the *presentation layer*, concerns data presentation, e.g., file or video compression. Finally, OSI layer 7, the *application layer*, deals with user applications and other high-level functionality, e.g., web browsing, e-mail, file transfer, file sharing, instant messaging, gaming, etc.

Not every network device contains all 7 layers. Personal computers do contain all 7 layers – the network interface card (e.g., Ethernet card) implements OSI layers 1 and 2, the operating system (e.g., Windows) implements OSI layer 3 and part of layers 4 through 7, and user-installed software implements the remainder of layers 4 through 7. A network router, however, often contains only layers 1 through 3. As a result, communication from source to destination follows a complicated path as pictured in figure 2.

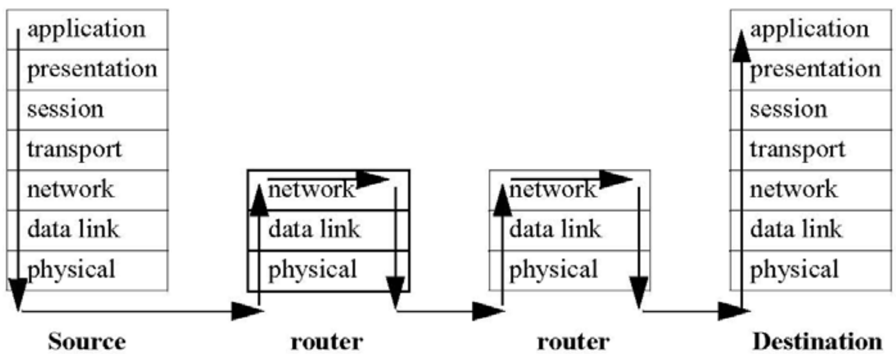

Figure 2: Routing

As an example, consider an e-mail sent from me to you. The e-mail is generated through my e-mail program on my computer (the source) at layer 7 (top left). Layer 7 hands the e-mail off to layer 6, which deals with encoding of the e-mail, e.g., html. Layer 6 then hands the e-mail off to layer 5, which logs onto my mailserver[8]. Layer 4 breaks the e-mail into a sequence of packets, and hands them one by one to layer 3. Layer 3 determines the first router on the path from me to you. Layer 2 decides when each packet can be transmitted, and layer 1 transmits each packet bit by bit. When these bits are received by the first router, layer 1 on that router translates the received signal into bits and layer 2 puts the bits back together into a packet, allowing layer 3 to look at the packet and determine the next router along the path. The second router acts similarly to the first router. Finally, layers 1 through 4 on your computer (the destination) assemble the received signal back together into the e-mail, layer 5 retrieves the e-mail from your mailserver, and layers 6 and 7 display the e-mail in your e-mail program.

Layers 1 through 3 can be thought of as network infrastructure, while layers 4 through 7 can be thought of as network applications. This distinction will be discussed in much more detail below as we consider the impact of layered architectures upon communications policy.

Although the OSI model serves as a reference for all network architectures, different networks have modified the model for their own use. As an example, the Internet uses a model with a reduced number of layers,

[8] For simplicity, we assume here that the mailservers reside on the source and destination computers.

as pictured in figure 1. OSI layers 1 and 2 are combined into a single Internet *LAN-link layer*. OSI layer 3 is also called the Internet *network layer*; it includes the *Internet Protocol* (IP). OSI layer 4 is also called the Internet *transport layer*; it includes the *Transmission Control Protocol* (TCP). OSI layers 5 through 7 are merged into a single Internet *applications layer*.

2.3 Technological developments

We turn now to a brief discussion of technological developments, due to convergence, that underlie the net neutrality issue. We start by considering the requirements of different types of applications. Applications differ in terms of the time scale on which users interact with the application. In real-time applications such as telephone calls or video conferencing, the information sent through the network must be received within a few tenths of a second after it is sent, or the perceived performance will be poor. In non-interactive applications such as e-mail, a delay of tens of seconds is usually still considered good performance. There are also applications with intermediate levels of interactivity, such as web browsing, in which a delay of a few seconds is usually acceptable. Applications also differ in terms of how they define good performance. Voice and video applications tend to base performance on *loss* (percent of packets never received) and *delay* (seconds from source to destination), while less interactive applications tend to base performance on *throughput* (bits per second received).

Traditional Internet transport uses *best-effort service*, in which the network attempts to transmit each packet from source to destination as quickly as possible. All users and all applications sharing a network link share the capacity of that link on a first-come first-served basis in the Internet network layer, without regard to the source or destination of the traffic and without regard to the type of application the packet supports[9]. The result is that congestion on a link degrades the performance of all users and all applications that pass through that link.

Best-effort service is often fine for applications whose perceived performance degrades slowly with increased congestion, e.g., e-mail, file transfer, and web browsing. In contrast, real-time applications such as telephone calls and video conferencing require a particular performance level (or better) to be usable; for these applications, best-effort service is often not sufficient to maintain acceptable performance.

In response, there have been recent technological developments in the Internet which provide differentiated service to different Internet traffic. One type of such differentiated service is to place a limit on the bandwidth used by certain Internet traffic. This approach, often called *traffic shaping*, is appropriate when there is a class of Internet traffic that tends to consume relatively large amounts of capacity, is viewed as low priority to the organization, and whose performance degrades smoothly with congestion.

Another type of such differentiated service is to label Internet traffic with distinct *priority* levels. At each router, in the Internet network layer, traffic with higher priority levels experiences better performance than traffic with lower priority levels. From the user's view, this priority approach is roughly similar to USPS PriorityMail service in that the packages go through the same system as First Class mail but experience lower loss and delay.

Finally, a third type of such differentiated service is to *reserve* network capacity for certain traffic classes and to limit the traffic in these classes. At each router, in the Internet network layer, such selected traffic shares a reserved portion of the network capacity; since this capacity is actively managed, the traffic experiences

[9] Technically, some applications are subject to congestion control measures in the Internet transport layer. However, these measures do not discriminate based on source, destination or application type.

guaranteed acceptable performance. From the user's perspective, this reservation approach is roughly similar to a toll or carpool lane[10].

The priority and reservation approaches are generically known as *Quality of Service* (QoS) mechanisms. Both QoS approaches give the selected traffic enhanced performance. In contrast, traffic shaping can be viewed as intentional *degradation* of the selected traffic class. It is worth noting that these QoS mechanisms are implemented at the Internet network layer; the consequence of this approach is that QoS must be provided at *every router* along the path from source to destination[11] in order to be effective.

In both traffic shaping and QoS, the decision of which traffic to enhance or degrade can be based on (1) the type of application, (2) the source, (3) the destination, (4) consumer payment, or (5) application provider payment. For example, if priority service is given to all voice and video traffic, or if traffic shaping is applied to file sharing traffic, then application type is the basis for the decision. If capacity is reserved for all voice and video traffic to/from consumers for a fee, then consumer payment is the basis. If capacity is reserved for all traffic to/from application providers for a fee, then application provider payment is the basis.

As traffic shaping and QoS mechanisms have become available in network equipment, many ISPs have adopted their use for various purposes. Some ISPs currently use QoS to guarantee acceptable quality for their own VoIP service[12]. Many universities use traffic shaping to limit file-sharing traffic to a small proportion of the organization's network capacity in order to protect the performance of other traffic viewed as more important to the organization's mission.

Use of these mechanisms is expected to grow with technology convergence. Most of the large carriers have announced plans to deploy QoS mechanisms, often in coordination with deployment of fiber and video service. In the near term, priorities or reservations will be used for selected traffic that both originates and terminates within the carriers network, or that transits onto the public switched telephone network. In this case, a carrier's own network management can provide acceptable performance to limited traffic classes using QoS techniques. In the long term, we expect that QoS will also be applied to selected traffic that originates or terminates (but not both) within the ISP's network and transits onto another carrier's portion of the Internet. In this case, acceptable performance may only be provided through cooperation with other carriers offering QoS.

We expect the initial use of QoS will be to support a carrier's own VoIP and video services. It is unclear whether carriers will offer QoS to competitors' applications, on the basis of application type, consumer payment, and/or application provider payment.

We believe these developments are at the core of the issue of net neutrality. Specifically, we are concerned about the basis on which differentiated service may be used to support real-time applications and applications that require relatively large amounts of bandwidth. In contrast, we are *not* concerned about applications that are less interactive and require little bandwidth, such as most web-browsing; we do not foresee best-effort Internet transport degrading to the point where such applications do not receive acceptable performance.

[10] Technically, it is similar to a toll lane in which the traffic is managed so that congestion never occurs in the toll lane.

[11] or at least at every congested router.

[12] See e.g., Cox (2004).

In addition, as networks converge, it is becoming less clear what portion of the integrated network is considered to be the public Internet and what portion is considered to be a private network. Specifically, in the case of traffic that does not transit onto another carriers portion of the Internet, the distinction between Internet traffic, VoIP traffic, and video traffic may be more a matter of user perception than of technical distinction. This causes a fundamental problem with separate regulation Internet traffic, telephone traffic, and video traffic.

3. Arguments For and Against Net Neutrality

3.1 Academic Literature

The academic literature on net neutrality is rooted in an earlier literature on open access, which in turn partially relies on an older literature on Internet architecture.

As discussed in section 2.2, the Internet is based on a layered architecture. As pictured in figure 2, the full 7 OSI layers only exist at the source and the destination (the endpoints), not at the intermediate routers. As a result, functions in OSI layers 1 through 3[13], such as routing, must be implemented in each router, as well as the source and destination computer. However, functions in OSI layers 4 through 7[14], such as e-mail upload and download, are only implemented in the source and destination computers, not in each router.

A closely related design paradigm is called the *end-to-end principle* (Saltzer, Reed, & Clark, 1984). The principle suggests that network functionality should be implemented in OSI layers 1 through 3, and hence in each router, only if it cannot be implemented effectively in higher layers. This principle has been followed in much (but not all) of Internet design to date. As a consequence, although there are a large number of protocols at the Internet LAN-link layer to accommodate transmission over various media (e.g., Ethernet, WiFi, DSL, cable modem, CDMA), and there are a very large number of protocols at the Internet application layer to support various applications (e.g., pop, smtp, http, p2p, IM), there are relatively few protocols at the intervening Internet network and transport layers (principally IP, TCP, and UDP).

The Openists believe that the end-to-end principle is responsible for the tremendous amount of innovation at the application layer. They argue that communications policy should mandate open access to the network infrastructure for application providers in order to assure continued innovation. The debate over open Internet access was precipitated by FCC consideration of whether to allow ISPs[15] to bundle Internet access with applications such as e-mail and web hosting[16]. Bar et al. (2000) suggest that policy for Internet access should be developed on the model of open access mandated for the telephone network in the *Telecommunications Act of 1996*, by mandating availability on cost-effective terms of key network components to service providers who do not own Internet access infrastructure. Lemley and Lessig (2001) support a policy in which Internet infrastructure and applications can not be bundled using either technical or business mechanisms. They argue that an open network encourages greater innovation than a closed network and that vertical integration harms consumers. They also demonstrate how open access leads to net neutrality, by using the end-to-end principle as an argument for

[13] or equivalently in the Internet LAN-link or network layers

[14] or equivalently in the Internet transport or application layers

[15] There is an unfortunate inconsistency in the use of the term *ISP*. In this paper, we use the term ISP to mean carriers who provide last mile access to residential and business customers. However, in the open Internet access debate, the term ISP was often used to mean providers of services such as e-mail and web-hosting who do not own Internet access infrastructure.

[16] The FCC later allowed such bundling. We discuss this in more detail below.

maintaining a dumb network that does not differentiate among different types of Internet traffic. Herman (2007) adds first amendment arguments to support net neutrality.

The Deregulationists believe that there is a trend away from the end-to-end principle. They believe that ISPs are in the best position to determine the most beneficial evolution of the Internet, and that this will likely entail a smart network that uses traffic discrimination in order to accomplish product differentiation. Speta (2000) argues against open Internet access. He postulates that vertical integration will not threaten competition among content providers, on the basis of an economic theory of network externalities that states that monopolists will only vertically integrate in ways that maximize consumer welfare. He argues that vertical integration may be necessary to justify investment in network infrastructure, and that ISPs will be motivated to allow competition to their own applications since this will add value to the broadband connection they sell. Owen and Rosston (2003) view net neutrality as a property rights issue. They claim that it is unknown whether welfare is maximized by assignment of property rights to ISPs or to application providers, and therefore that net neutrality regulation should not be imposed. Yoo (2005) extends the anti open Internet access argument put forward by Speta to net neutrality. He believes that the theory of network externalities undermines the need for mandated net neutrality, and furthermore that the lack of concentration of broadband access when measured at a national level predicts that ISPs will not have market power in the application space. Yoo also suggests that competition can be more effectively fostered between different physical networks that are matched to different types of applications. Yoo (2006) furthermore suggests that exclusive arrangements between ISPs and application providers can be justified on the basis of the efficiency of traffic shaping as a proxy for usage-based pricing. Finally, McTaggert (2006) argues that the Internet is not neutral, using examples of Internet portals, caching, routing, and QoS.

Some economists counter that application of the theory of network externalities to open access or net neutrality is not so straightforward. Farrell and Weiser (2003) discuss potential problems with application of the theory to open access, and suggest that uncertainty dictates further study of the effect of various types of Internet regulation in other countries. Van Schewick (2007) argues that a correct application of the theory demonstrates that ISPs do have an incentive to discriminate against competing application providers, and that the economic benefits of net neutrality outweigh the costs.

Nondiscriminationists see more complexity in the net neutrality issue than the Openists or Deregulationists. Wu (2003) supports net neutrality but opposes open access. He is not opposed to vertical integration, since he believes that QoS may require it. He argues that open access is inferior to direct rules on discrimination that ensure that harmful types of discrimination are not used by ISPs. He suggests that such a policy might be constructed by allowing discrimination at the Internet LAN-link layer but not at the Internet network layer. As a result, he would allow bandwidth tiers and some forms of QoS, but discrimination on the basis of the type of application or on the source or destination would be prohibited. Peha (2006) surveys various technical methods of discrimination, and argues that some uses are beneficial and some are not. He discusses dangers from harmful uses of discrimination, focusing on oligopoly rents in upstream markets. He suggests that good policy might be based on lists of allowed and banned discrimination.

3.2 Lobbyist Arguments

The pro net neutrality lobby falls roughly into two categories - application providers (led by Google, Amazon, Yahoo!, eBay, and Microsoft) and consumer groups (led by moveon.org, Free Press, Consumers Union, and Common Cause).

The application providers are roughly in the Nondiscriminationist camp. They are primarily concerned that ISPs will charge them discriminatory prices for access to reserved bandwidth. When they examine carrier plans for deployment of fiber, they worry that carriers will sell reserved bandwidth to selected application providers rather than adding capacity to their Internet access offerings for shared use by all Internet traffic. The application providers are large business customers of ISPs, and they consider it acceptable for an ISP to charge their *own customers* for access to the Internet based on bandwidth. However, the application providers do not want to be forced to make contractual arrangements with the ISPs of each of the *application provider's customers*. They argue that ISPs have always charged only their *own customers*, not customers of *peering ISPs*. They are ok with ISPs charging their own customers different prices for different levels of bandwidth. Application providers are also concerned that ISPs will choose to strike exclusive deals for access to reserved bandwidth, tilting the playing field among application providers.

Application providers remain unconcerned that ISPs might refuse to offer access to reserved bandwidth to application providers who compete with the ISPs own service offerings. Such behavior is likely for some ISPs who deploy their own video service, particularly if the available bandwidth is not high enough to support a large number of video streams. Application providers, however, typically believe that net neutrality need only apply to Internet access, and do not attempt to apply to it what they view as private network services[17].

Some application providers are concerned that ISPs will not offer access to QoS mechanisms, or at least not without charging application providers. They worry that ISPs will use QoS to support their own real-time applications but refuse access to QoS to competing application providers' traffic. Alternately, they worry that ISPs will attempt to charge application providers, rather than the ISP's own customers, for QoS.

Application providers are concerned that lack of access to reserved bandwidth or QoS will hinder the development of new applications. They argue that small application providers will not be able to compete effectively.

The application providers want a form of net neutrality that prohibits discrimination on the basis of source, destination, or ownership of Internet traffic. Discrimination on the basis of application type or consumer payment, e.g., enhanced performance for all voice and video traffic, would be allowed. This prohibition would only apply to what is thought of as Internet traffic, as opposed to private network services. Application providers generally believe that such a prohibition would not allow ISPs to charge application providers for reserved bandwidth or QoS.

The consumer groups in support of net neutrality are generally Openists. They share many of the same concerns as the application providers. In addition, they are concerned that ISPs may block access to certain applications or websites. Some groups are worried that an ISP may block access to websites based on the content of the website, e.g., political beliefs, and interpret net neutrality as an issue of free speech. They are also worried about new charges that may increase the cost of Internet access.

The anti net neutrality lobby consists mostly of ISPs (led by AT&T, Verizon, and Comcast, but including many other wireline and wireless ISPs). They fall squarely into the Deregulationist camp. They argue that there is no current substantial problem, and hence no need for regulation. Furthermore, they argue that if substantial problems arise, then the FCC has the expertise and authority to effectively take action. The ISPs believe that net neutrality would reduce ISPs' incentives to invest in their network infrastructures by reducing the return they may earn on these investments.

---

[17] However, as mentioned above, we believe it will be increasingly difficult to determine which services are Internet services and which are private services.

The ISPs further believe that they should be allowed to choose whether or not to make available reserved bandwidth or QoS, to charge consumers for these services, and to make individual contractual arrangements with application providers. They argue that refusal to provide access to reserved bandwidth or QoS to application providers would not hinder the application provider's ability to deploy new services. Most ISPs expect to offer access to reserved bandwidth or QoS to application providers for a fee, but want the flexibility to do this through individual contractual arrangements. Such ISPs are typically focused in the short term on arrangements in which the application provider becomes an ISP business customer, rather than providing access through peering ISPs. Some ISPs also claim that net neutrality impinges on the ISP's ability to perform network management tasks, including traffic shaping for p2p traffic.

Finally, many ISPs fear that net neutrality would reestablish many other provisions of common carrier regulation. In particular, they worry that net neutrality's occasional use of the term *nondiscriminatory* will lead to imposition of other common carrier rules such as unbundling and open access.

3.3 Bills

In 2005-2006, Congress considered several approaches to net neutrality. The House bill, the *Communications Opportunity, Promotion, and Enhancement Act of 2006*, and the Senate Commerce Committee bill, the *Advanced Telecommunications and Opportunities Reform Act*, take similar approaches to net neutrality and essentially present the anti net neutrality position. They give the FCC authority to enforce the FCC's four principles (discussed below), but do not allow the FCC to issue additional rulemaking to further define or extend these principles. Blocking of web pages or applications is prohibited, but no further restrictions are placed on ISPs. ISPs remain free to offer or deny access to QoS mechanisms. Either degraded service or enhanced service can be based on application type, source, destination, consumer payment, and/or provider payment. Payments can be on the basis of commercial arrangements, meaning that there is no requirement that they be consistent or equitable between different application providers.

The *Internet Non-Discrimination Act of 2006* (sponsored by Sen. Wyden) essentially presents the most extreme pro net neutrality position among the bills. It prohibits blocking of web pages and applications. The bill includes a broad prohibition on discriminatory behavior; our interpretation is that any use of QoS is prohibited, since the bill prohibits discrimination in data transport, which presumably even includes differentiation based on application type. The bill explicitly prohibits charging application providers that are not customers of the ISP, although this seems redundant given the prohibition on discrimination. The bill prohibits degradation of traffic; our interpretation is that traffic shaping would fall into this prohibition. (These provisions are summarized in table 1.)

	House & Senate bills	Wyden bill	Markey & Sensenbrenner & Snowe-Dorgan bills	Snowe-Dorgan amendment
Prohibit blocking	Y	Y	Y	Y
Degradation:				
allow traffic shaping	Y	N	Y	Y
prohibit if not based on type	N	Y	Y	Y
QoS:				
allow based on type	Y	N	Y	Y
prohibit exclusive use	N	n/a	Y	Y
QoS fees:				

allow consumer payment	Y	n/a	N	Y
allow provider payment	Y	n/a	N	?
prohibit discriminatory fees	N	n/a	n/a	?
prohibit unreasonable fees	N	n/a	n/a	?

Table 1. Comparison of the net neutrality provisions in various Congressional bills

Three other bills – the *Net neutrality Act of 2006* (sponsored by Rep. Markey, defeated as an amendment to the House bill), the *Internet Freedom and Nondiscrimination Act of 2006* (sponsored by Rep. Sensenbrenner), and the *Internet Freedom Preservation Act* (sponsored by Sen. Snowe and Sen. Dorgan) - take similar approaches to each other and present refined pro net neutrality positions. They each prohibit blocking of web pages and applications, as do the other bills. They each also include a broad prohibition on discriminatory behavior; our interpretation is that use of QoS is allowed if based on the type of application, but prohibited if based on the source or destination of the traffic, e.g., based on the application provider. The bills explicitly prohibit charging for QoS, either in the form of customer or application provider payment. The bills also prohibit degradation of traffic on the basis of source or destination, but allow network management based on application type, e.g., traffic shaping.

Finally, Senators Snowe and Dorgan proposed a streamlined version of their bill as an amendment to the Senate bill; this amendment was defeated in the Committee markup on a tie vote. The amendment states a broad prohibition on behavior that discriminates on the basis of source, destination, or ownership of Internet traffic. This has the effect of prohibiting blocking of web pages and applications, allowing QoS or degradation on the basis of the type of application, and prohibiting QoS or degradation on the basis of the source or destination of the traffic. The amendment is silent on whether an ISP can charge for QoS. Our interpretation is that it allows consumer payment, but it is less clear whether it allows provider payment and if so whether such payment would be mandated to be reasonable and nondiscriminatory.

4. A Framework for a New Net Neutrality Policy

4.1 Challenges from Convergence

U.S. federal communications policy was separately developed for telephone networks, cable video networks, and cellular networks. As others have noted, convergence between voice, video, and data services is therefore a fundamental challenge to such *vertical regulation*. Common carrier regulation has focused on voice service, and does not effectively address video or data service. Similarly, cable services regulation has focused on video service, and does not effectively address voice or data service. Internet services are not effectively addressed anywhere in the Communications Act. Furthermore, when voice, video and data services are offered over a integrated converged network infrastructure, it is unclear how to apply conflicting regulations from different titles.

This vertical regulation approach started breaking down decades ago, as the telephone network progressed from a set of local monopolies to a somewhat competitive marketplace. Since the Telecommunications Act of 1996, services directly based on the infrastructure are classified as *telecommunication services*, and applications are classified as *information services*. This approach is an attempt at *horizontal regulation*.

The distinction between telecommunication services and information services has been an effective tool for communications policy. However, convergence has now progressed to the point where this distinction is no longer sufficient. Indeed, many of the topics in the 2006 communications bills stem from convergence, including video franchising, net neutrality, VoIP, municipal broadband, universal service, broadcast flag, and white spaces. Despite this trend, Congress has yet to consider creation of a communications policy that fundamentally recognizes convergence.

In the absence of explicit statutes regarding Internet access or services, the FCC recently addressed the issue of whether Internet access is a telecommunications service or an information service. The decision by the FCC to classify Internet access (including DSL and cable modem service) solely as information services exempted these services from regulation under Title II of the Act, and left them only subject to the FCC's Title I ancillary authority. This decision means that Internet access is not subject to the requirement that charges and practices be just and reasonable (section 201), the prohibition on unjust and unreasonable discrimination (section 202), or the requirement that interconnection rates, terms, and conditions be just, reasonable, and nondiscriminatory (section 251). It is this decision that inspired the pro net neutrality coalition.

In 2005, the FCC issued a set of principles related to net neutrality. They express the sentiment that consumers should be entitled to connect devices and to access content and applications of their choice. The principles also state that consumers are entitled to competition among network and application providers. However, the principle regarding competition is too vague to be useful, and the principles are not binding.

4.2 Our Approach

We begin by strongly disagreeing with the FCC's decision to classify Internet access solely as an information service. Telecommunication services can be thought of as the lower layers of the network, and information services can be thought of as the upper layers. The immediate consequence of this approach, purely on a technical basis, is that the Internet clearly consists of both telecommunication services and information services.

We find it useful to separate the layers into two subsets. We consider OSI layers 1-3 (the Internet LAN-link and network layers) as network infrastructure and OSI layers 4-7 (the Internet transport and application layers) as network applications.

A number of other papers have suggested using layers as a tool to formulate communications policy. Lessig (2001) considers a model consisting of physical, logical, applications, and content layers. He argues that the physical and logical layers of the Internet have historically been neutral. He believes it is acceptable for the physical layer to be closed, but proposes that the logical layer should be open and act as a commons. Werbach (2002) uses a similar set of four layers: the physical layer corresponds to OSI layer 1, the logical layer to OSI layers 2-6, and the application and content layers share OSI layer 7. He argues that communications policy should be formulated around these layers, with open interfaces between them. Solum and Chung (2003) propose a six-layer model, and argue that communications policy should attempt to respect the integrity of layers and to place regulation at or near the layer where the problem occurs. Whitt (2004) suggests a four-layer model similar to Werbach's, and presents principles concerning how layers should inform policy formulation.

While all of these approaches support horizontal regulation, and many result in a wider applicability than considered here, most of these layered models combine OSI layers 3 and 4 (the Internet network and transport layers) into a single logical layer. We believe that these two OSI layers display quite different characteristics. The infrastructure layers (OSI layers 1-3) include functions that must be implemented *in every portion of the network* including the access network, e.g., routing, addressing, and QoS (if used). The application layers (OSI layers 4-7), in contrast, include functions that can be implemented *at any point within the network*, e.g., e-mail, web-hosting, and caching.

One of the strengths of this delineation is that the network infrastructure layers exhibit a *high barrier-to-entry*, while the application layers exhibit a *low barrier-to-entry*. The high barrier-to-entry of infrastructure has heretofore resulted in a small number of carriers offering service in any particular location. In contrast, the low barrier-to-entry of applications has resulted in a competitive market with a large number of application providers.

The distinction between network infrastructure and applications also helps with correct application of the end-to-end principle. Using the new terminology, we can view the principle as a suggestion that network functionality should be implemented in network infrastructure *only if* it cannot be implemented effectively in network applications. With the recognition that OSI layers 1-3 are implemented *in every hop* along the transmission path, whereas OSI layers 4-7 are only implemented *in the end devices*, this is equivalent to the suggestion that network functionality should be implemented in every hop only if it can not be implemented effectively solely at the endpoints.

With the Internet's layered architecture in mind, we now turn to the Communications Act for potential policy models. We then meld these models with our understanding of the architecture and functionality of the Internet to form broad policy goals.

The Communications Act broadly lays out three potential models to follow. First, if a communications market has sufficient competition, then minimal regulation is applied. This is captured in Title I of the Act, which applies forbearance from regulation unless required to ensure just and reasonable practices. The paradigm here is that a free market will regulate itself, and will promote investment in a fair manner that maximizes social welfare.

If sufficient competition is not present to ensure just and reasonable practices, Titles II and VI apply different models of regulation. Both models were initially formulated assuming a local monopoly over a portion of the corresponding network, but have since been applied to circumstances in which there is limited (but not sufficient) competition. Both models attempt to maximize social welfare by limiting monopoly behavior.

Title II (common carrier) uses the open access model. The network can be separated into two layers: infrastructure and applications. Local infrastructure, e.g., the local loop, is usually deemed to be lacking sufficient competition, due to the high cost of building this infrastructure. Applications, e.g., telephone service, in contrast are often offered by a large number of competitors. Title II essentially dictates open access to infrastructure, and requires that carriers make critical network elements available to application providers without unjust or unreasonable discrimination and without undue or unreasonable preference. Title II also mandates reasonable and nondiscriminatory rates, terms, and conditions for interconnection with other infrastructure providers. As a result, telephone carriers do not limit who a customer can call and do not provide differentiated service based on who a customer calls.

Title VI (cable services), in contrast, uses a paradigm of a closed network. In this model, the network cannot be naturally separated so that infrastructure and applications can be offered by separate entities. Vertically integrated carriers control both the infrastructure and the applications. Title VI recognizes this natural vertical integration, and applies limits to the cable carrier's behavior when it is also integrated with content providers. As a result, cable carriers do determine which content a customer can receive.

We believe that the Internet's architecture supports the need for some type of openness. However, we do not believe that net neutrality by itself requires open access. This intermediate approach places us loosely in the Nondiscriminationist camp. We suggest below that net neutrality can be effectively achieved by a properly regulated *open interface* from applications to network infrastructure. The requirement of an open interface captures the central tenet of a layered Internet architecture, and yet is less intrusive than the requirement of open access.

We also believe economic arguments support the need for some type of openness. We believe that carriers should offer an interface to network infrastructure to applications without unjust or unreasonable discrimination and without undue or unreasonable preference. We argue below that an open interface mandate can be used to limit monopoly or oligopoly behavior. We adopt the free market model, based on Title I forbearance, for the application portion of the Internet. We mandate an open interface, based on selected elements of Title II, for the local infrastructure portion of the Internet. We believe this combination is most likely to encourage continued development of the Internet in a manner that maximizes social welfare.

We believe that an open interface should attempt to provide a level playing field between application providers and ISPs who offer competing applications. Threats to a level playing field are potentially posed by vertical integration between Internet service providers and application providers. Certain *Internet applications* (e.g., video on demand, high quality VoIP, video conferencing) require specialized *Internet infrastructure services* (e.g., dedicated bandwidth or QoS). Our goals for net neutrality are:

Goal A) ISPs should be prohibited from refusing to provide enabling Internet infrastructure services to competing application providers in order to differentiate the ISP's own application offerings,

Goal B) ISPs should be prohibited from providing Internet infrastructure services to competing application providers at inflated prices in order to favor the ISP's own application offerings, and

Goal C) ISPs should be prohibited from making exclusive deals to provide enabling Internet infrastructure services to certain application providers.

Such threats to a level playing field are limited in scope, and therefore any remedy should be similarly limited in scope. First, vertical integration only affects applications that rely on specialized Internet infrastructure. A characteristic of such network services is that they must be provided directly in the *access network* (e.g., "the last mile"). Networks services that can be provided outside the access network (e.g., caching, search engines, web hosting) can be provided by competing ISPs or application providers, and therefore do not represent a threat to a level playing field. Second, when sufficient competition exists in the Internet access market, there is a lower danger that vertical integration would cause a non-level playing field.

Our approach, therefore, is to apply existing elements of the Communications Act that prohibit anti-competitive and unreasonably discriminatory behavior to ensure a level playing field. Our goal is to apply these elements, however, in a limited fashion:

Goal D) Limit the scope through new definitions that transcend the limitations of current telecommunication and information services definitions,

Goal E) assure forbearance where sufficient competition exists,

Goal F) ensure that ISPs have the right to apply network management mechanisms that do not threaten a level playing field, and

Goal G) ensure that regulation does not impede an ISP from making arrangements with consumers, application providers, and peering ISPs for Internet infrastructure services in a manner that does not conflict with the above goals.

4.2 Comparison to other approaches

Openists believe that the Internet is best served by maintaining a dumb network that does not differentiate among different types of traffic. They use the end-to-end principle to support this position. We disagree with both the reasoning and the conclusion. We believe that the end-to-end principle suggests implementing functionality in the application layers if feasible. However, as noted above, QoS mechanisms such as priority or reservations must be implemented at the Internet LAN-link and/or network layers in order to be effective. The end-to-end principle therefore does not ban their use, and therefore does not mandate a dumb network. Indeed, one of the original authors of the end-to-end principle was also one of the authors of an Internet protocol to implement QoS using reservations (Braden, Clark, and Shenker, 1994), and has since suggested that a new theory addressing network core functionality should be created to co-exist with the end-to-end principle (Blumenthal & Clark, 2001) and that the end-to-end principle should be generalized with a new form of modularity that can gracefully accommodate tussle along various competing players (Clark et. al., 2005). We believe that the introduction of QoS, and hence the creation of a smart network, will *increase* the range of applications that can be efficiently run on the Internet.

Openists support a policy based on open access, including a broad ban on vertical integration. We agree with the need for an open interface between network infrastructure and applications, but do not believe that net neutrality by itself is threatened by vertical integration that crosses other layer boundaries.

Deregulationists present economic arguments that regulation will hinder investment by ISPs and that in the absence of regulation ISPs will vertically integrate only in ways that maximize consumer welfare. However, other economists argue that the same theory of network externalities indicates that ISPs do have an incentive to discriminate against competing application providers, and that the economic benefits of net neutrality outweigh the costs. We accept here this latter claim, and agree with the goal of providing a level playing field. Although we agree that development of applications drives network infrastructure investment, we do not believe that vertical integration violating an open interface between infrastructure and applications is necessary to foster new applications.

Some deregulationists have also suggested that competition can be more effectively fostered between different physical networks that are matched to different types of applications. We strongly disagree, and believe that 20 years of convergence provides strong evidence for the efficiency of integrated networks over the use of separate infrastructures.

Finally, many deregulationists equate QoS with other types of discriminatory network behavior (e.g., caching or priority listing of search results) and use this to argue that QoS should not be regulated. We strongly disagree. We are concerned with the use of prioritization when there is limited competition (e.g., QoS); we have faith that free market dynamics will adequately address the use of prioritization when there is sufficient competition (e.g., caching or priority listing of search results). Equating the two types of prioritization ignores the distinction between network infrastructure services (e.g., QoS) which must be implemented in the access network and network applications (e.g., caching) which can be implemented in other ISPs' networks.

We agree with the Nondiscriminationist camp's arguments that there are good and bad uses of traffic discrimination. Although we do not believe that QoS requires vertical integration, we support a policy that allows vertical integration and traffic differentiation, but restricts their use to ensure that ISPs do not discriminate in a manner that extracts oligopoly rents. However, whereas Wu (2003) would ban discrimination at the Internet network layer and hence ban QoS based on the type of application, we believe that an open interface is a cleaner and more effective solution. This difference will become clear as we analyze what types of ISP network management should be considered unacceptable.

5. Acceptable and Unacceptable ISP Network Management

In this section, we use our approach to delineate which types of network management should be considered acceptable and which unacceptable. We consider a wide range of network management actions related to net neutrality. For each, we decide whether the action should be allowed or prohibited by considering the goals presented in the last section as well as the general intent of the Communications Act as applied to other technologies. We focus on the case in which sufficient competition does not exist in the Internet access market; where sufficient competition exists we choose to apply forbearance. When doubt remains as to the desirable policy goal, we choose the action most likely to encourage a level playing field.

5.1 Blocking

First, consider *blocking* of either a web page or an application. Currently, blocking (when used) is usually done on the basis of the *type of application*. This type of blocking can be beneficial, e.g., firewalls that block security intrusions. However, it can also be anti-competitive, e.g., an ISP that blocks all VoIP traffic in order to limit competition to its own voice service. In some instances, blocking based on application type can be viewed as either beneficial or harmful depending on user preferences, e.g., a firewall that blocks all file-sharing traffic.

Is it acceptable for an ISP to block applications, on the basis of the application type? In the absence of sufficient competition in infrastructure, we believe such blocking is not acceptable because it may violate goal A stated above. While some ISPs block certain applications to limit bandwidth use, we believe that network management should be transparent and that tiering is a superior solution. Even if sufficient competition exists, then the layered architecture of the Internet strongly suggests that such blocking should be rare. Our answer is that blocking based on application type should be a decision made by the consumer, not the ISP.

It is less common to block web pages or applications based on the *source and/or destination*. However sometimes this type of blocking is used for security purposes. Examples include parental control software and firewalls that block all traffic from unknown addresses. Is it acceptable for an ISP to block applications, on the basis of the source and/or destination of the traffic? We believe this is fairly straightforward, and that blocking based on source and/or destination should also be a decision made by the consumer, not the ISP.

5.2 Degradation

Next, we consider *degradation* of Internet traffic. Currently, intentional degradation is common in a few scenarios. First, it is common to place a limit on the bandwidth of a broadband user's access link that is lower than the speed the link could accomplish. Such limits, referred to as *tiering*, are often used to differentiate service offerings by a selection of access rates, and to charge different prices for these different service offerings. Although tiering is a form of intentional degradation, neither the pro or anti net neutrality lobbies oppose its use. Second, it is increasingly common for an ISP to place bandwidth limits on certain classes of applications, e.g., traffic shaping. Currently, the most common such limits are on file sharing traffic, which otherwise may consume a large proportion of the available bandwidth and cause congestion for all other applications. The bandwidth limits degrade the performance of the applications to which they are applied, but do not block the applications entirely. Both tiering and traffic shaping are intentional degradation based on the type of application (and in the case of tiering on consumer payment). In contrast, degradation of traffic based on the source and/or destination is both technically difficult and uncommon[18].

Is it acceptable for an ISP to intentionally degrade traffic? We believe that degradation based on application type, e.g., tiering and traffic shaping, is acceptable since it does not bias the network toward certain providers. Indeed, we believe that such techniques can be crucial elements of network management, support goal F, and should not be discouraged. However, we believe that any degradation based on source or destination should also be a decision made by the consumer, not the ISP, since it would violate goal A in much as the same manner as blocking based on source or destination.

5.3 QoS

Finally, we consider QoS mechanisms such as reservations or priorities. QoS is often used by carriers to provide acceptable performance for their own VoIP or video conferencing traffic as part of business packages that offer combined voice, video, and data transport. Although use of QoS for residential service is currently less common, we expect its use to increase with the growth in residential VoIP, video conferencing, and video streaming.

A traffic stream receives enhanced performance through QoS if the stream requests it and the carrier grants the request. A carrier may grant QoS requests: (1) to only its *own applications* (e.g., only its own VoIP subscribers), (2) to *particular applications types* (e.g., VoIP packets from all providers), or (3) on the basis of *contractual arrangements* (e.g., consumer and/or application provider payment).

A central component of net neutrality policy is the decision of how to treat QoS mechanisms. Is it acceptable for an ISP to implement QoS in its network? Some net neutrality proponents argue that the Internet should only provide best-effort service so that all Internet streams are treated identically; this group opposes any use of QoS. Although some carriers may choose alternate methods to provide acceptable performance to voice and video applications, we believe that a broad ban on use of QoS is a Luddite instinct. Furthermore, we do not see any precedent in the Communications Act for a similar ban on a set of technologies. We therefore believe use of QoS can be acceptable. The real issue, we believe, is how QoS is used.

Is it acceptable for QoS to be based on application type? Currently, the Internet adopts different types of transport[19] for different types of applications. QoS would fall into the same category. We do not see a threat to goals A-C through this use of QoS, and our open interface approach would allow it. As an example, if a carrier chooses to deploy QoS mechanisms in their network and to make these mechanisms available to all VoIP traffic, regardless of the application provider supporting the service, we believe this approach should be allowed. The anti net neutrality lobby supports this position. The pro net neutrality lobby, excepting those who believe QoS should not be allowed in any manner, also supports this position.

Is it acceptable if a carrier deploys QoS in its network to support some of its own applications, but does not offer these mechanisms to competing application providers? The anti net neutrality lobby argues that refusal to provide access for QoS to application providers would not hinder the application provider's ability to deploy new

[18] We do not consider congestion control measures (e.g., TCP) which treat all users similarly as being intentional degradation based on source or destination.

[19] TCP versus UDP.

services. We disagree. QoS should be classified as infrastructure, not application, because it must be implemented *in every portion of the network* in which congestion occurs in order to be effective. As a result, there is no effective option available to application providers who wish to provide enhanced performance to voice and video applications; they must receive the same QoS as the carrier uses for its own similar applications in order to compete effectively. In support of goals A and C, we conclude that it is not acceptable for a carrier to deploy QoS in its network to support some of its own applications, but not to offer these mechanisms to competing application providers. Therefore, we believe that QoS on the basis of source or destination is not acceptable.

The remaining issue with use of QoS is whether it is acceptable for QoS to be based on consumer and/or application provider payment. The anti net neutrality lobby argues that this approach should be allowed without constraint. Such contractual arrangements based on consumer payment are likely to be combined with tiering, e.g., for a small additional fee all VoIP traffic to/from a particular residence will receive QoS. Alternately, contractual arrangements for QoS could be based on provider payment, e.g., all VoIP traffic to/from a particular provider will receive QoS. Both approaches are technically feasible. The fee for such service could be a flat monthly fee, a per-minute charge, or a per-byte charge. Any consumer fee would almost certainly be based on a publicly available fee schedule, and therefore be nondiscriminatory. We see no threat to goals A through C from QoS based on consumer payment, and believe allowing this approach supports goal G. We also note that most pro net neutrality groups would agree.

Application provider fees for QoS, however, could potentially be subject to *commercially negotiated* agreements, as are commonly used when carriers offer virtual private networks to business customers. Such agreements open the door to potential discrimination between various application providers. Under the policy developed above, we have already chosen to prohibit an ISP from offering QoS to only selected application providers, so the issue here is whether an ISP can charge different fees to different application providers. We believe the open interface approach suggests that any such discrimination must be reasonable, e.g., based on the differences in cost of providing QoS. We conclude that similar constraints should be placed on QoS based on provider payment.

5.4 Summary

The decisions made in above with regard to what types of network management should be allowed or prohibited result in an answer of yes for every row in table 2. All of the bills in Congress prohibit blocking, as would we. With regard to degradation, our decision to allow traffic-shaping puts us at odds with the Wyden bill, whereas our decision to prohibit degradation on the basis on source or destination (unless elected by the consumer) puts us at odds with the House and Senate bills. On use of QoS, our decision to allow QoS based on application type puts us at odds with the Wyden bill, whereas our decision to prohibit exclusive use of QoS puts us at odds with the House and Senate bills. Finally, our decision to allow either consumer or provider payment for QoS puts us at odds with the Markey, Sensenbrenner, and Snowe-Dorgan bills, whereas our mandate that such fees be reasonable and nondiscriminatory puts us at odds with the House and Senate bills.

To illustrate these policy decisions, consider their effect upon an ISP who offers both broadband Internet service and a separate VoIP service. The effects of our policy goals are shown in table 2.

Is it acceptable if …	Answer
… a carrier *blocks* a competitor's VoIP traffic?	no
… a carrier doesn't block a competitor's VoIP traffic, but doesn't offer QoS to *competitor's* VoIP subscribers while using QoS for its *own* VoIP subscribers?	no

… a carrier gives a *broadband subscriber* who uses a competitor's VoIP service the choice of (a) best-effort transport of their VoIP traffic as part of the basic broadband package, (b) enhanced performance for their VoIP traffic for an additional 1¢/min paid by the subscriber, or (c) enhanced performance for up to 500 minutes of their VoIP traffic for an additional $5/month paid by the subscriber?	yes
… a carrier gives a *VoIP competitor* the choice of (a) best-effort transport of their VoIP traffic as part of the subscriber's basic broadband package or (b) enhanced performance for their VoIP traffic for an additional 1¢/min paid by the VoIP provider?	yes
… a carrier charges different VoIP competitors *different prices* for QoS?	no
… a carrier charges VoIP competitors a uniform price for QoS, but a different price than charged to *its own affiliates*?	no

Table 2. Effect of policy goals upon an ISP offering VoIP.

6. Statute Language

Goals A-G, as presented above, are implemented in draft statute language in the following subsections. The new definitions of Internet infrastructure services, Internet application services, and access network are presented in section 6.1, which helps implement goal D. Sections 6.2-6.4 guarantee consumers the right to an open interface between infrastructure and applications, implementing goals A-C. Section 6.5 guarantees ISPs the right to use network management techniques in a pro-competitive manner, implementing goals E-G.

In each section, we alternate between presenting draft statute language (indented and written in *italics*) and discussion of the statute language.

6.1 Definitions

As discussed above, we strongly disagree with the FCC decision to classify Internet access solely as an information service. We present new definitions of Internet infrastructure and applications to rectify this, and to introduce terms that are meaningful in both telephone networks and the Internet. These terms will be used as a foundation for the language in the following sections:

SECTION 1. DEFINITIONS.

In this title:

(1) *INTERNET INFRASTRUCTURE SERVICES. The term 'Internet infrastructure services' means all services- (A) over a network that uses a public right-of-way; and (B) that reside at or below the network layer or are required to manage the network.*

(2) *INTERNET APPLICATION SERVICES. The term 'Internet application services' means all services- (A) over a network that uses a public right-of-way; (B) that are not infrastructure services; and (C) that do not fall under Title VI of the Communications Act.*

(3) *NETWORK LAYER. The term 'network layer' means the third layer of the 7-layer Open Systems Interconnection Model, responsible for message addressing and for routing information within the* *network, including routing within the telephone network and including the Internet Protocol within the Internet.*

These terms agree with the intent of the older terms *telecommunication services* and *information services*. Internet infrastructure services do not change the content of information, similar to telecommunication services. In contrast, Internet application services create, store, or change the presentation of information, similar to *information services*.

Maintaining a distinction between the Internet infrastructure and applications is critical to formulation of good communications policy. Internet infrastructure services can only be provided by carriers, and must be provided by each carrier on their portion of the network. Such services include multi-user sharing of a wire or frequency (e.g., Ethernet or WiFi), routing (e.g., IP), and address assignment (e.g., DHCP). Internet infrastructure services require large investments into loops or wireless spectrum and switches or routers. Since Internet infrastructure services often require substantial hardware and must be implemented in each portion of the Internet, they reflect a high barrier-to-entry. Internet application services can be provided by carriers or by many other application providers on the Internet. Such services include e-mail, web-hosting, caching, voicemail, and the portions of VoIP and IPtv that can be offered by independent application providers. Since Internet application services are typically software-based and can be placed at many locations within the Internet, they usually reflect a low barrier-to-entry.

Among Internet infrastructure services, a further geographic distinction is worthwhile. *Access networks*, commonly called *the last mile*, are the portions of the Internet that must be transversed to communicate with other Internet locations. Access networks currently often have limited competition. The portion of the Internet that resides outside of access networks is commonly called the *Internet backbone*. The backbone currently exhibits substantial competition. Section 1 presents a definition of access networks:

(4) *ACCESS NETWORK -The term 'access network' means the portions of the Internet service provider's network which must be transversed to form routes from the Internet to its subscribers.*

These definitions effectively limit the scope of regulation presented in other sections, and implement goal D.

6.2 Principles

Many of the bills addressing net neutrality incorporate some of the FCC principles. We start by implementing the first three FCC principles, with language based roughly on the Senate Commerce bill. These principles guarantee a subscriber's right to use and offer content and applications, and guarantee a subscriber's right to connect devices.

SECTION 2. PRINCIPLES

Except as otherwise provided in this title, with respect to Internet services, each Internet service provider shall allow each subscriber to-

(1) *access, use, post, receive, or offer any lawful content, application, or service of that subscriber's choosing;*

> ***(2)*** *connect any legal device of that subscriber's choosing to the Internet access equipment of that subscriber, if such device does not harm the network of the Internet service provider;*

We next implement part of the fourth FCC principle, with language taken from the Snowe-Dorgan amendment to the Senate Commerce bill. Unlike the Snowe-Dorgan amendment, however, this provision is subject to allowed network management techniques as specified in section 5. In conjunction with section 5, this guarantees ISPs the right to implement bandwidth allocation and QoS and to sell such service, and it guarantees subscribers the right to purchase such service that discriminates only on application type, broadly supporting goals A-C.

> ***(3)*** *receive service that does not discriminate in the carriage and treatment of Internet traffic based on the source, destination, or ownership of such traffic; and*

Finally, we implement a often proposed 5th principle, the consumer's right to information, using language taken from the Senate Commerce bill:

> ***(4)*** *receive clear and conspicuous information, in plain language, about the estimated speeds, capabilities, limitations, and pricing of any Internet service offered to the public.*

6.3 Discrimination

Although principle *(3)* above is an elegant statement of the pro net neutrality lobby goals, we do not believe it is sufficient to fully implement goals A-C. In particular, we do not believe that it alone prohibits unjust or unreasonable discrimination in charges. To reinforce these goals, section 3 of the draft statute language applies to ISPs, in a restricted manner, the portion of the Title II of the Communications Act that places limits on use of discrimination and preferences:

> *SECTION 3. DISCRIMINATION AND PREFERENCES*
>
> *For purposes of sections 202 and 206 through 209 of the Act, an Internet service provider shall be treated as a common carrier, and Internet infrastructure service shall be treated as a communications service.*

We apply the limits specified in section 3 only to services that can be provided only by the ISP, using the new term *Internet infrastructure service*. We also apply these limits only when a competitive market is absent, through coordinating provisions in section 4 below. Section 3 does not apply other requirements on common carriers, e.g., tariffs, interconnection, or unbundling, to ISPs.

6.4 Competition

The Internet has developed in a manner in which ISPs often also serve as application providers. To directly address such vertical integration, we draw upon other provisions from Titles II and VI of the Communications Act that address vertical integration in telephone and cable networks.

First, however, we implement goal E by limiting application of net neutrality to situations in which sufficient competition does not exist. This can be accomplished by ensuring that the forbearance provisions currently in Title I of the Communications Act apply to ISPs offering Internet infrastructure service:

SECTION 4. COMPETITION

(a) *For purposes of section 10 of the Act, an Internet service provider shall be treated as a telecommunications carrier, and Internet infrastructure service shall be treated as a telecommunications service.*

This provision therefore allows forbearance from the limits placed through section 3 when there is a competitive market.

We turn next to the responsibilities of vertically integrated ISPs to *their own subscribers and peers*. When an ISP offers applications that rely on Internet infrastructure services, we require the ISP to make available to competitors the same Internet infrastructure services at the same prices:

(b) *An Internet service provider shall make available to subscribers and other Internet service providers - on the same prices, terms, conditions of sale, and delivery - any Internet infrastructure services provided on its access networks as the Internet service provider offers to Internet application services provided by itself or its affiliates.*

(c) *An Internet service provider shall provide Internet infrastructure service to subscribers and other Internet service providers, that is at least equal in quality to that provided by the Internet service provider to itself or its affiliates.*

Provision (b) is based on section 628(c)(2) of the Communications Act, which governs program access in vertically integrated video programming and delivery. Provision (c) is based on section 251(c)(2)(c), which addresses quality of service responsibilities for interconnecting carriers.

Finally, we turn to the responsibilities of vertically integrated ISPs to *subscribers of other ISPs*. Without regulating peering agreements, we wish to prohibit a vertically integrated ISP from using service level agreements with its peers to favor its own applications. We use language based on section 628(b) of the Communications Act, which places similar restrictions on vertically integrated video programming and delivery:

(d) *It is unlawful for an Internet service provider to engage in unfair methods of competition, unreasonably discriminatory conduct, or unfair or deceptive acts or practices, the purpose or effect of which is to hinder significantly or to prevent any Internet application provider from providing content, applications, or services to consumers.*

Provisions (b), (c), and (d) reinforce goals A and B.

6.5 Network management

Goal F requires a delineation of allowed network management practices. The Senate Commerce bill included most of the required provisions:

SEC. 5. NETWORK MANAGEMENT.

An Internet service provider may-

(1) *protect the security, privacy, or integrity of the network or facilities of such provider, the computer of any subscriber, or any service, including by- (A) blocking worms or viruses; or (B) preventing denial of service attacks;*

(2) *facilitate diagnostics, technical support, maintenance, network management, or repair of the network or service of such provider;*

(3) *prevent or detect unauthorized, fraudulent, or otherwise unlawful uses of the network or service of such provider;*

(4) *block access to content, applications, or services that Federal or State law expressly authorizes to be blocked, including child pornography;*

(5) *provide consumers Parental Control applications, devices, or services, including- (A) blocking access to websites with obscene or adult content; (B) blocking display of video content based on a common rating; or (C) offering a family friendly tier of service;*

(6) *allow a subscriber to elect to have content, applications, or services blocked at the request of such subscriber;*

We add a 7th provision to explicitly guarantee ISPs the right to alleviate congestion by treating all traffic similarly, or by treating all applications of the same type similarly:

(7) *alleviate congestion in a manner that does not distinguish based on the source or ownership of content, application, or service;*

Finally, we suggest two provisions to support goal G. Both work in conjunction with provision *(3)* in section 2. Provision *(8)* guarantees ISPs the right to sell reserved bandwidth and QoS to both their residential and business subscribers. Provision *(9)* guarantees ISPs the right to discriminate in the carriage of Internet traffic based on peering arrangements with other ISPs. Such peering arrangements would allow for reserved bandwidth and QoS to be provided cooperatively by multiple ISPs.

(8) *offer directly to a subscriber Internet service at different prices based on defined levels of bandwidth, quality of service, or the actual quantity of data flow over a user's connection, and discriminate in the carriage and treatment of Internet traffic based on such contract with that subscriber; and*

(9) *enter into contracts with other Internet service providers, and discriminate in the carriage and treatment of Internet traffic based on such contract with that Internet service provider.*

6.6 Further Study

Both the House and Senate bills required the FCC to conduct further review related to net neutrality. We suggest here a study more tightly focused on potential anti-competitive problems, using the new definition of Internet infrastructure services:

SECTION 6. FCC REVIEW

(a) *IN GENERAL. Beginning 1 year after the date of enactment of this Act, the Federal communications commission shall report annually to the Committee on Commerce, Science, and Transportation of the Senate and the Committee on Energy and Commerce of the House of Representatives regarding-*

(1) *the developments in Internet infrastructure services, including Internet architecture, peering and interconnection, routing, and quality of service;*

(2) *contracts between Internet service providers and subscribers for Internet service, including levels of bandwidth and quality of service, and any discrimination in the carriage and treatment of Internet traffic based on such contracts;*

(3) *contracts between interconnecting Internet service providers, and any discrimination in the carriage and treatment of Internet traffic based on such contracts; and*

(4) *how such developments impact the development of Internet applications that rely on evolving Internet infrastructure services.*

(b) *DETERMINATIONS AND RECOMMENDATIONS. The Federal Communications Commission shall make such recommendations under subsection (a), as the Commission determines appropriate.*

7. Conclusion

We have developed a net neutrality policy based on a delineation of Internet infrastructure and Internet applications. We argued that the net neutrality issue was caused by technology convergence, and that any solution should mirror the architecture of the Internet. We believe that the central problem is the ability of an ISP with market power to use Internet infrastructure to favor its own applications or otherwise discriminate in a manner that extracts oligopoly rents. We illustrated how communications policy can prohibit such anti-competitive behavior without restricting desirable forms of network management. This approach differs with many of the pro net neutrality bills in that it allows nondiscriminatory network management techniques and QoS and it allows consumer or provider payment for QoS; it differs with many of the anti net neutrality bills in that it prohibits discriminatory use of Internet infrastructure and discriminatory degradation and mandates reasonable and nondiscriminatory fees for QoS.

Our proposed statute language is not complete. In particular, enforcement must be addressed. Furthermore, net neutrality is not the only issue resulting from a fragmented communications policy; a broader consideration of the effect of technology convergence upon communications policy is warranted.

References

Bar, F., Cohen, S. S., Cowhey, P., DeLong, B. J., Kleeman, M., & Zysman, J. (2000). Access and innovation policy for the third-generation Internet. *Telecommunications Policy*, *24*, 489-518. (Available at http://repositories.cdlib.org/brie/BRIEWP137)

Bar, F., & Sandvig, C. (2000). Rules from truth: Communication policy after convergence. In *Proceedings of the 28th Telecommunications Policy Research Conference.*

Bloomberg News (2006). AT&T, Comcast Rout Google in `Net Neutrality' Battle. *Bloomberg News, July 20, 2006.*

Blumenthal, M. S., & Clark, D. D. (2001). Rethinking the design of the internet: the end-to-end arguments vs. the brave new world. *ACM Transactions on Internet Technology*, *1* (1), 70-109.

Braden, R., Clark, D., & Shenker, S. (1994). RFC 1633 - Integrated Services in the Internet architecture: An overview. *IETF Network Working Group*.

Clark, D., Wroclawski, J., Sollins, K., & Braden, R. (2005). Tussle in cyberspace: Defining tomorrow's Internet. *IEEE/ACM Transactions on Networking*, *13*, 462-475.

Cox Communications (2004). Cox Communications VoIP Whitepaper. May 2004. (Available at http://www.fcc.gov/oet/tac/7.28.04_TAC_Cox_VoIP_whitepaper.pdf)

Farrell, J., & Weiser, P. (2003). Modularity, vertical integration, and open access policies: Towards a convergence of antitrust and regulation in the Internet age. *Harvard Journal of Law and Technology*, *17*, 85-134. (Available at http://ssrn.com/abstract=452220)

Herman, B. D. (2007). Opening bottlenecks: On behalf of mandated network neutrality. *Federal Communications Law Journal*, *59*. (Available at http://ssrn.com/abstract=902071)

Jordan, S. (2007). Implications of Internet architecture upon net neutrality*. ACM Transactions on Internet Technology*, *forthcoming*.

Lehr, W. H., Gillett, S. E., Sirbu, M. A., & Peha, J. M.(2006). Scenarios for the network neutrality arms race. In *Proceedings of the 34th Telecommunications Policy Research Conference.*

Lemley, M. A., & Lessig, L. (2001). The end of end-to-end: Preserving the architecture of the Internet in the broadband era. *UCLA Law Review*, *48*, 925-972. (Available at http://ssrn.com/abstract=259491)

Lessig, L.(2001). The Internet under siege. *Foreign Policy*, *127*, 56-65.

McTaggart, C.(2006). Was the Internet ever neutral? In *Proceedings of the 34th Telecommunications Policy Research Conference.*

Owen, B. M., & Rosston, G. L. (2003). Local broadband access: Primum non nocere or primum processi?: A property rights approach. *Stanford Law and Economics Olin Working Paper No. 263*. (Available at http://ssrn.com/abstract=431620)

Peha, J. M. (2006). The benefits and risks of mandating network neutrality, and the quest for a balanced policy. In *Proceedings of the 34th Telecommunications Policy Research Conference.*

Saltzer, J. H., Reed, D. P., & Clark, D. D. (1984). End-to-end arguments in system design. *ACM Transactions on Computer Systems*, *2* (4), 277-288.

Solum, L. B., & Chung, M. (2003). The layers principle: Internet architecture and the law. *U San Diego Public Law Research Paper No. 55*. (Available at http://ssrn.com/abstract=416263)

Speta, J. B. (2000). Handicapping the race for the last mile?: A critique of open access rules for broadband platforms. *Yale Journal on Regulation*, *17*, 39-91.

van Schewick, B. (2007). Towards an economic framework for network neutrality regulation. *Journal on Telecommunications and High Technology Law*, *5*. (Available at http://ssrn.com/abstract=812991)

Werbach, K. D. (2002). A layered model for Internet policy. *Journal on Telecommunications and High-Tech Law*, *1*. (Available at http://ssrn.com/abstract=648581)

Whitt, R. S. (2004). A horizontal leap forward: Formulating a new communications public policy framework based on the network layers model. *Federal Communication Law Journal*, *56*, 587-672.

Wu, T. (2003). Network neutrality, broadband discrimination. *Journal of Telecommunications and High Technology Law*, *2*, 141-178. (Available at http://ssrn.com/abstract=388863)

Wu, T. (2004). The broadband debate: A user's guide. *Journal of Telecommunications and High Technology Law*, *3*, 69-95. (Available at http://ssrn.com/abstract=557330)

Yoo, C. S. (2005). Beyond network neutrality. *Harvard Journal of Law and Technology*, *19*, 1-77. (Available at http://ssrn.com/abstract=742404)

Yoo, C. S. (2006). Network neutrality and the economics of congestion. *Georgetown Law Journal*, *94*. (Available at http://ssrn.com/abstract=825669)