國立交通大學

科技法律研究所

碩士論文

網路中立管制：
差別待遇之經濟效應及其合理性認定

A Study on Network Neutrality: Reconsidering Economic Impact as a Factor in Determining Reasonable Discrimination

研究生：劉昊恩
指導教授：王立達 博士
　　　　　李　淳 博士

中華民國一百零一年七月

網路中立管制：

差別待遇之經濟效應及其合理性認定

A Study on Network Neutrality: Reconsidering Economic Impact as a Factor in Determining Reasonable Discrimination

研究生： 劉昊恩　　Student: Hao-En Liu
指導教授： 王立達　　Advisor: Li-Dar Wang
李 淳　　Chun Lee

國立交通大學

科技法律研究所

碩士論文

A Thesis
Submitted to Institute of Technology Law
College of Management
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Master
in
Laws

July 2012

Hsinchu, Taiwan, Republic of China

中華民國一百零一年七月

網路中立管制：

差別待遇之經濟效應及其合理性認定

研究生：劉昊恩　　　　　　　　　　　　　　指導教授：王立達 博士

李淳 博士

國立交通大學科技法律研究所碩士班

摘　　要

歷年來網路端對端的架構，導致對於經由網路設備傳輸的封包之個別屬性及內容一無所知。此類型開放性的架構促使了網路前所未見的創新及競爭，其擔任的角色廣被推崇。然而，日新月異的電腦運算已構成威脅並足以改變現狀，它使網路業者不僅能辨識封包的屬性及內容，也能決定如何處理其內容。

為因應這項新威脅，呼籲透過立法管制業者行為規範的聲浪四起，期能藉以防止業者利用他們新具有的技術能力成為網際網路的看守者。近年來，美國聯邦通訊委員會也已初步研擬了一套規範, 希望能夠化解這些疑慮，但實施這些規範可能面臨一些挑戰。本文將著重於其中一項可能面臨的挑戰－ 第五修正案的徵收條款，其討論重點為如沒有足夠證據證明業者的行為將強力影響網路創新的潛力時，限制業者特定的成本控制或利潤開發的行為，可能會構成第五修正案所為無合理賠償之徵收行為。

關鍵字： 網路中立、網路差別待遇、端對端、開放網際網路、美國國家通訊委員會、徵收條款、頻寬費用

A Study on Network Neutrality: Reconsidering Economic Impact as a Factor in Determining Reasonable Discrimination

Student: Hao-En Liu

Advisor: Dr. Li-Dar Wang
Dr. Chun Lee

Institute of Technology Law

National Chao Tung University

Abstract

The end-to-end architecture of the Internet has historically resulted in an open network lacking in significant awareness of the nature of the content passing through it. This open architecture has been widely praised for its key role in fostering the unprecedented level of innovation and competition that is evident on the Internet today. However, technological advances in computing have threatened to alter the status quo, giving network providers not only the ability to identify the nature of all content passing through their wires, but also the ability to decide how to treat that content.

In response to this new threat, there have been calls to regulate the conduct of network providers, seeking to prevent them from using their newfound capabilities to act as gatekeepers of the Internet. In recent years, the Federal Communications Commission has tentatively introduced a new set of rules, hoping to allay these concerns. However, enforcement of these rules is likely to meet numerous challenges. This paper seeks to highlight one of those challenges in particular – the Takings Clause of the Fifth Amendment. It argues that absent significant proof that a network provider's conduct will significantly affect the innovative potential of the Internet, restrictions on certain cost-controlling or profit-seeking conduct may constitute a taking of property without just compensation.

Key words: Net Neutrality, Network Discrimination, End-to-End, Open Internet, Federal Communications Commission, Takings Clause, Bandwidth Costs

Acknowledgements

The timely completion of this dissertation is nothing short of a miracle, and would never have been possible in absence of the ridiculous amount of help and support I received from *so many* selflessly kind-hearted individuals, some of whom leaned over backwards to assist me, despite barely knowing me at all. To everyone who extended a hand to me throughout the way, I would like to take this opportunity to convey my sincere gratitude.

First and foremost, I would like to express my most heartfelt appreciation to my advisor Prof. Li-Dar Wang for his valuable guidance, support, and motivation, but most of all for his extreme patience and understanding. I would also like to thank my advisor Prof. Chun Lee, who could always take time out of his ultra busy schedule to provide me with valuable suggestions and point me in the right direction. Besides my advisors, I would like to thank all the entire ITL office staff, especially Andrea, Peiyu, Cindy, and Judy, who have constantly encouraged me and proactively looked out for me every step of the way, making sure I never miss a deadline, and taking care of all the nasty procedural stuff that without them I would never have been able to keep track of. Knowing that they had my back made all the difference, and I cannot be more grateful for their help.

I must also express my deepest gratitude to my classmate Alice Shiu, who went out of her way to assist me in my time of need, showing me the way when I was lost. I am most grateful for her introduction to Tsai-Wen Yang, her friend and classmate, who provided me with invaluable research materials that went on to become the core references for my thesis. Make no mistake that this thesis would not have been possible without the materials that Tsai-Wen *lavishly* prepared for me. “Lavish” is no exaggeration. I am greatly indebted to her.

I’d also like to thank Serena Wang for her timely assistance on my thesis formatting. Without her template and formatting tips, the process would never have been as smooth and painless as it was. I would like to thank my peers for encouraging me, believing in me, and giving me much needed suggestions. I would further like to thank my company PIXNET for allowing me to take on this endeavor while under employment. I’d like to thank my bosses Jenny and JR for their understanding, and all my colleagues for their support, especially in the last few weeks of the ordeal when I barely had time to work at all. I want to thank JR in particular for generously leveraging his personal connections to help me gain access to respected scholars in the field, like Prof. Ching-Yi Liu of NTU, whose valuable insight and comments gave me a boost of confidence at precisely the moment I needed it the most.

Last but not least, I would like to thank my parents and my sister for giving me the strength to see this through. I love you all, forever and always.

Table of Contents

List of Tables

List of Figures

I. Introduction

1.1 Introduction

For many younger generations, it is hard to imagine life before the Internet. This global system of interconnected computer networks – linking together millions of private, public, academic, business, and government networks through a broad array of electronic, wireless and optical networking technologies – has had a profound impact on our economies, societies, and our way of life, fundamentally changing the way people communicate with each other, the way we consume, share and interact with information, and basically every other facet of our daily lives.

The Internet was first conceptualized in writing by J.C.R. Licklider of MIT, who envisioned a "galactic" network of interconnected computers through which users could seamlessly collaborate and share information and network resources[1]. This vision eventually took shape in the form of the ARPANET, an experimental network commissioned by DARPA[2], a research branch of the United States Department of Defense. Initially, the ARPANET connected the mainframes of four universities on the west coast of the United States, allowing researchers to share the mainframes at any of the networked institutions – a practice known as time-sharing.[3] The ARPANET grew rapidly, adding new nodes to its network at a steady pace. At the same time, other networks began to emerge, each with their own protocols and network infrastructures. It was only a matter of time before the demand emerged for these networks to interconnect with each other.

[1] *See* J.C.R. Licklider and Welden E. Clark, *On-Line Man Computer Communication* (Aug. 1962), *available at* http://www.computer.org/csdl/proceedings/afips/1962/5060/00/50600113-abs.html.

[2] Then known as the Advanced Research Projects Agency (ARPA), later renamed Defense Advanced Research Projects Agency (DARPA).

[3] *See* Barry M. Leiner et. al., *A Brief History of the Internet* (Dec. 2003), *available at* http://www.internetsociety.org/internet/internet-51/history-internet/brief-history-internet.

In order to unify these networks, a new protocol was needed that would be able to support the wildly different network architectures in use. Since each of these networks was independently run, the goal was to create a protocol that would allow each stand-alone network to communicate with each other over its existing infrastructure, without needing internal changes, and with no centralized management necessary[4]. What emerged would eventually be known as the TCP/IP protocol, and the merged networks would later grow into the Internet we know today.

The open and decentralized architecture of this new "inter-network" based on TCP/IP meant that no one entity was in control of the Internet. Individual networks were free to introduce whatever uses they wished to the "inter-network," and they would remain interoperable with all other networks as long as they conformed to the protocol's standards. This spawned the broad deployment and adoption of a wide variety of uses for the network, including electronic mail, file transfer, and probably most importantly, the World Wide Web.

The Internet has grown immensely since its humble beginnings. In the last decade alone, the number of estimated Internet users has grown five-fold, from around 360 million to nearly 2.3 billion[5], and this number continues to rise at a steady rate. The meteoric growth of the Internet has fueled – and been fueled by – the extraordinary explosion in innovation that has brought about all kinds of ingenious uses and applications unimaginable at the time the Internet was conceived. Voice and video conferencing has enabled people to communicate over great distances, and work or collaborate remotely in real-time. Blogs and social networks allow individuals to make their voices heard around the world, creating "a forum for a true diversity of political discourse [and] unique opportunities for cultural development."[6]

[4] *See id.*
[5] *See World Internet Users and Population Stats, Internet Usage Statistics*, Internet World Stats, http://www.internetworldstats.com/stats.htm (last visited Jun. 3, 2012).
[6] 47 USC § 230(a)

E-commerce has transformed our economy, allowing businesses and individuals alike to buy and sell goods and services online. Audio and video streaming have revolutionized the entertainment industry, offering consumers a wealth of content at the click of a button. All these innovations were made possible by the open architecture of the Internet. Professor Lessig, a leading thinker on Internet policy, describes the Internet as one of the most important "innovation commons" the world has ever seen, and remarks that it forms this commons not just through norms, but also through its specific technical architecture[7].

Recent advances in computing, however, have led many to fear that the underlying architecture of the Internet may be on the brink of change. These changes threaten to affect the "neutrality" of the Internet, in a way that could significantly affect its innovative potential. In this paper, we address the nature of these changes, the concerns they raise, and the regulatory regime that has emerged in response.

1.2 Framework

This paper is divided into three sections. Part one is an introduction of the architectural principles that have underpinned the Internet. We take an in depth look at the design principles that helped shape the so-called "neutral" Internet, and how these principles have evolved over the years. We discuss the implications of these changes in technical terms, and how these changes threaten to alter the status quo. Part two focuses on the regulatory framework that has developed in response to these changes. We track the development of regulatory proposals over the years, and conduct an analysis of the rules that have developed. Finally, in part three we explore possible challenges to the rules, and highlight factors that the regulatory agency must take into account to ensure its enforcement of these new rules achieves its goals without running into constitutional concerns.

[7] *See* LAWRENCE LESSIG, THE FUTURE OF IDEAS 23 (2002), *available at* http://www.the-future-of-ideas.com.

II. Network Neutrality and the Evolving Internet

2.1 Framing Network Neutrality

The network neutrality debate is actually a conflation of several issues. There are generally three ways to frame the debate. The first approach is one that attempts to define neutrality. In other words, this approach focuses on what it means for a network to be "neutral", and what a "neutral" network entails for the different actors in the system. Much of the literature on the topic of "network neutrality" understandably starts off this way, either with a definition of the term, or an attempt to define it. This may seem like a reasonable thing to do, but there are several key shortcomings to such an approach of analyzing this particular issue. For starters, the term "network neutrality" can mean a whole host of different things to a whole lot of people, in a whole range of different contexts[8]. Commentators have recognized since the start that "neutrality" was an imprecise term. In fact, in one of his earliest papers on net neutrality, Tim Wu commented that "neutrality, as a concept, is finicky, and depends entirely on what set of subjects you choose to be neutral among."[9] Simply by choosing the range of subjects in defining the term, one inevitably risks betraying a bias that ends up framing the debate in a certain way.[10] Secondly, by starting out with a definition of the principle, one risks unintentionally setting the stage for a pro and against debate with

[8] *Compare* G. Knieps & P. Zenhäusern, *The Fallacies of Network Neutrality Regulation*, 9 COMPETITION AND REG. IN NETWORK INDUS. 119 (2008) ("network neutrality is basically a debate on how best to finance the construction and maintenance of a broadband network") *with* Robert W. Hahn and Scott Wallsten, *The Economics of Net Neutrality*, 3 THE ECONOMISTS' VOICE 8 (Apr. 2006) ("net neutrality is actually a friendly-sounding name for price regulation") *and* Nicholas Economides and Joacim Tåg, *Network Neutrality on the Internet: A Two-Sided Market Analysis*, 24 INFO. ECON. & POLICY 91 (Dec. 12, 2001) ("net neutrality is defined as a restriction that Internet Service providers cannot directly charge content providers for access to consumers"); *See also* Rachelle B. Chong, *The 31 Flavors of Net Neutrality: A Policymaker's View*, 12 INTELL. PROP. L. BULL. 147 (2008) ("Net Neutrality is like the Baskin-Robbins ice cream store. There are several flavors that appeal to various tastes. Whatever you want, we can serve it up in a Net Neutrality cone.").

[9] *See* Tim Wu, *Network Neutrality, Broadband Discrimination*, 2 J. OF TELECOMM. & HIGH TECH. L. 141 (2003).

[10] *But see* Daniel S. Isenberg, *Framing Network Neutrality Right*, DAVID S. ISENBERG'S MUSINGS ABOUT LOCI OF INTELLIGENCE AND STUPIDITY (Dec. 6, 2006). (opining that it is clear what net neutrality means, and that by not recognizing that fact, we are falling into the trap of telecommunications and cable companies, who want "to keep Network Neutrality amorphous and undefinable [so that] we can't pass a law against it").

strawmen arguments that ultimately end up distorting the debate[11].

A second approach focuses on the regulatory means to address net neutrality: how we should regulate, and the desirability of regulation[12]. This approach mostly focuses the question of how net neutrality laws should best be enforced – for example, in the form of ex-ante regulations or ex-post remedies – or under what regime or authority the rules should be enforced – e.g. under the FCC's ancillary jurisdiction, a reclassification of broadband as a Title II telecommunications service, or with new explicit mandates from Congress.

The third and last approach is one that addresses specific goals of net neutrality. This focuses on what we wish to achieve, and the values we wish to protect, beyond neutrality of the network per se. Instead of defining neutrality, this approach focuses on identifying specific problems, and finding an answer for those problems.

This paper adopts the latter approach. Hence, for the moment, we will refrain from assigning the term any definition at all, short of its literal reading. This paper also does not seek to answer the question of how network neutrality rules should best be enforced, or by whom. Instead, we look towards the technical and architectural origins of the Internet to help shape our view of what "network neutrality" actually is – and how well the current policies reflect that – rather than what activists and stakeholders want it to mean, or think it means.

2.2 The Underlying Architecture of the Internet

There are several features of the Internet that have led to it being hailed as a "neutral" network. Before we can truly understand the "neutral" nature of the original Internet, we must

[11] For example, in reality very few commentators disagree with the advantages of an open Internet, but rather, most "opponents" of net neutrality actually oppose the need for regulatory oversight.

[12] *See, e.g.*, Jonathan E. Nuechterlein, *Antitrust Oversight of an Antitrust Dispute: An Institutional Perspective on the Net Neutrality Debate*, 7 J. on Telecomm. & High Tech. L. 19 (2009) ("This paper focuses instead on the comparatively neglected institutional dimension of the debate: an inquiry into which federal agencies are best positioned to resolve net neutrality disputes when they arise.").

have a basic understanding of how the Internet works.

2.2.1 The Internet: a Network of Layers

In order to communicate over the Internet, a host computer relies on set of protocols known as the Internet protocol suite or TCP/IP, after its two main protocols, the Transmission Control Protocol (TCP) and the Internet Protocol (IP). The TCP/IP model enables seamless communication between different host devices running on a diverse array of hardware and software platforms, each connected to the network through a variety of types of physical mediums. In fact, TCP/IP was designed to be able to run over anything. It achieves this by employing a system of layers, whereby each layer implements a certain predefined set of functions via its own internal-layer actions, while relying on functions provided by the layer below. Listed from lowest to highest, the four layers of the TCP/IP protocol are:

A. The Link Layer

The link layer contains a wide variety of protocols responsible for physically transporting data packets across a point-to-point link. To transmit and receive data from the network, a host must implement a communication link to interface with the network. As previously mentioned, this link may be realized through a variety of types of physical media, including but not limited to coaxial cable, copper wire, fiber optics, or radio spectrum. Link layer protocols define the procedures for interfacing with network hardware and accessing the transmission medium, providing higher layers with a consistent and predictable data transport mechanism regardless of the underlying physical link.[13]

B. The Internet Layer

The purpose of the Internet layer is to select the best route through the network for data

[13] *See* BARBARA VAN SCHEWICK, INTERNET ARCHITECTURE AND INNOVATION 84 (2010).

packets to travel across the link layer to their destinations. It is the only layer of the Internet protocol suite with one single common protocol: the Internet Protocol (IP). All Internet transport protocols use the IP to carry data from one end host to another. To do so, the data must pass through a series of routers. It is the job of IP to manage the addressing and delivery of these raw data packets so they can travel successfully from router to router across the physical network. However, IP does not provide end-to-end delivery guarantees. It is a connectionless protocol, transmitting each and every data packet independently, based on each router's best guess as to where the packet should go next. Data packets may arrive damaged, out of order, or even be lost altogether. It is the job of higher layers to make sure data is correct and complete.[14]

C. The Transport Layer

The transport layer is responsible for the sending and receiving of data between different end hosts. On a sending host, it processes data from the application layer and determines where that data should be sent. On a receiving host, it receives incoming packets from the Internet layer, and determines which application it is destined for. It is also be responsible for making sure (if required by the application) that the data it sends or receives is complete and error-free when it arrives, and may implement mechanisms for retransmission of lost or damaged data. This extra layer shields applications from the unreliable nature of the IP protocol.**[15]**

D. The Application Layer

The application layer consists of the higher-level protocols applications use to communicate amongst each other. Well known application layer protocols include the

[14] *Id.* at 85.
[15] *Id.* at 86.

Hypertext Transfer Protocol (HTTP), the File Transfer Protocol (FTP), and the Simple Mail Transfer Protocol (SMTP). These protocols specify the rules and syntax according to which data should be formatted and transmitted, the request methods and responses, and the appropriate procedures and behaviors expected of the applications, so they can work with each other in a seamless manner.[16]

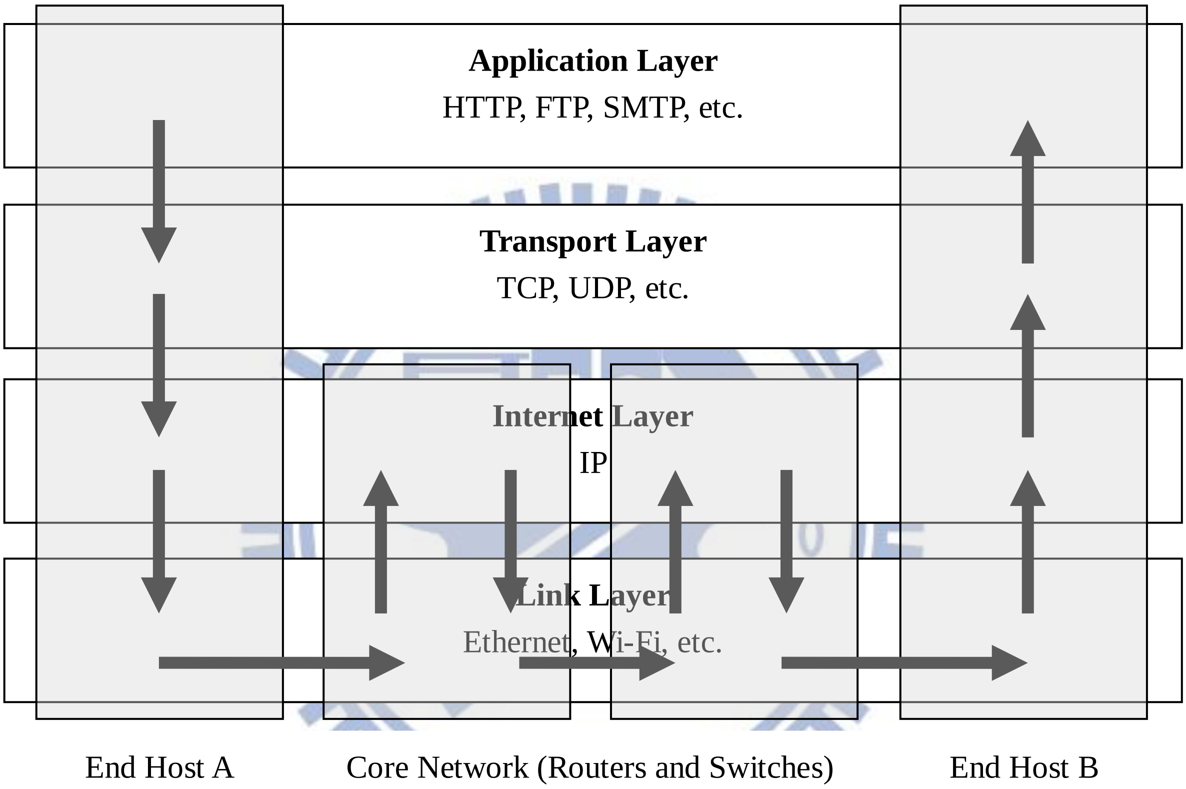


Figure 1: The Four Layers of the Internet Protocol Suite

Source: Adapted from Barbara van Schewick, Internet Architecture and Innovation

To put the four layers into a practical context, when an application on End Host A wants to communicate a piece of information to its counterpart on End Host B, it first formats the information according to the specifications set forth by the application's protocol of choice (e.g. HTTP) at the Application Layer, then submits a request to the Transport Layer to transmit that information to the designated destination on the network (End Host B). The

[16] *Id.* at 87.

Transport Layer will establish a connection with End Host B using TCP or some other Transport Layer protocol, and the data will be encapsulated into packets that will be routed using IP over the Internet layer, running over whatever physical media happens to be providing the Link Layer.

This layered approach realizes a separation of concerns among different components of the system, whereby each layer performs a clearly defined function, independent of the other layers. Each layer in turn has a set of "visible information," which allows it to interface with other layers in a predetermined and predictable manner. This modular design creates an architecture in which components can be designed and improved independently but still work together[17]. As long as the visible service interface provided by the lower layer does not change, the inner workings of the layer can be changed without breaking functionality at other layers in the system. This insulates applications on the end hosts from changes caused by the inevitable changes in the Internet's routing architecture. At the same time, since lower layers are not allowed to use the services of higher layers, the Internet layer and the link layer remain unaffected by innovations in the transport and application layers. Thus, the Internet Protocol can support an unlimited variety of application protocols at the higher layers, while taking advantage of all kinds of new physical network infrastructures and improved transmission and link technologies at the lower layers, without requiring corresponding changes at the higher layers.

Continuing with the previous illustration, this means the application does not need to be concerned with the inner workings of the lower layers, like how the data is being routed to the destination, or on what type of medium. As far as the application is concerned, everything below the transport layer is a black box. It communicates on an end-to-end basis directly across the Application Layer. The application on End Host A sends out a piece of data, and the

[17] For an overview of the concept of modularity in network architecture, *see id.* at 38.

application on End Host B receives that data. All it needs to know is that the lower layers will faithfully handle the data as it requests, so that it will be delivered to the destination. As for the lower layers, all they need to be concerned about is executing orders from the layer above them. A real-world analogy for this would be the postal service, where the mail senders do not need to know the inner workings of the post office, and the delivery guy does not need to know the content of people's packages.

On its face, the layering principle may seem pretty straight-forward, but it raises a critical question for network designers: At which layer should specific functionality be implemented on the network?

2.2.2 The End-to-End Argument

The "end-to-end" argument is a design principle that attempts to serve as a guide for resolving the issue of how to allocate functions among the layers. At its core, it argues for a framework for organizing the distribution of functionality within a network in a way that "intelligence" in the network be implemented at the "ends" of the network, where the higher layers of the network are. In the context of the Internet, this is the Application Layer at the end host. On the flip side, it calls for the lower layer communications protocols themselves to be as "simple and general" as possible, in order to maximize its utility for all applications.[18]

The end-to-end principle has implicitly guided the development of the Internet since its inception, but it was not explicitly recognized as a design principle until the early 1980s, in a paper entitled *End-to-end Arguments in System Design*[19], by Professors Jerome Saltzer, David Reed, and David Clark. In various subsequent papers, the same authors have, jointly and

[18] *See* Mark A. Lemley and Lawrence Lessig, *The End of End-to-End: Preserving the Architecture of the Internet in the Broadband Era*, 48 UCLA L. REV. 925 (Oct. 1, 2000).

[19] Jerome H. Saltzer, David. P. Reed, and David. D. Clark, *End-to-End Arguments in System Design*, 2 ACM TRANSACTIONS ON COMPUTER SYSTEMS 277-288 (Nov. 1984). An earlier version appeared in the Second International Conference on Distributed Computing Systems 509-512 (Apr. 1981).

independently, sought to refine and clarify the principle and what it means for the underlying architecture of the Internet. One general depiction of the principle is as follows:

> End to end arguments have … two complimentary goals: (1) Higher-level layers, more specific to an application, are free to (and thus expected to) organize lower level network resources to achieve application-specific design goals efficiently (application autonomy); (2) lower-level layers, which support many independent applications, should provide only resources of broad utility across applications, while providing to applications useable means for effective sharing of resources and resolution of resource conflicts (network transparency).[20]

The principle, however, is not without its own uncertainties. For one, it relies on the ability to distinguish clearly between application specific and non-application specific functions. At a more technical level, it is ambiguous as to how it should allocate potentially application specific functions that may be possible to "completely and correctly" implement at multiple layers. As such, it is not always clear what exactly the end-to-end principle entails in certain specific cases, resulting in the awkward situation in which both proponents and opponents of a technical implementation invoke the end-to-end-principle to back up their views.[21]

In her book, *Internet Architecture and Innovation*[22], Professor Barbara van Schewick is one of the first academics to it undertake a critical analysis of the inconsistencies in the different interpretations of the end-to-end principle. She finds that even in the writings of Saltzer, Reed, and Clark, there exist "two versions of the end-to-end arguments that represent different rules for architectural design."[23] The first version (the "narrow" version) states that "a function should only be implemented in a lower layer, if it can be completely and correctly implemented at that layer. Sometimes an incomplete implementation of the function at the

[20] Jerome H. Saltzer, David. P. Reed, and David. D. Clark, *Active Networking and End-To-End Arguments*, 12 IEEE NETWORK 66, 70 (May 1998).
[21] *See* VAN SCHEWICK, *supra* note 13, at 81.
[22] *Id.*
[23] *See id.* at 58.

lower layer may be useful as a performance enhancement."[24] The second version (the "broad" version) states that "a function or service should be carried out within a network layer only if it is needed by all clients of that layer, and it can be completely implemented in that layer."[25] Van Shewick notes that technical discussions tend to focus on the narrow version, whereas policy texts and descriptions of the Internet's architecture tend to focus on the broad version. Generally speaking, most of the literature that refers to the end-to-end arguments simply quotes one or the other.

The function…				
can be completely and correctly implemented at		is needed by all clients of the lower layer	Narrow Version	Broad Version
Higher layer	Lower layer		may be implemented at	
YES	YES	YES	Both layers	Both layers
YES	YES	NO	Both layers	Higher layer only
YES	NO	n/a	Higher layer, with additional implementations at the lower layer allowed for performance considerations	Higher layer only
NO	YES	n/a	Lower layer only	Lower layer only

Table 1: Differences between narrow and broad versions of the end-to-end argument

Source: Adapted from Barbara van Schewick, Internet Architecture and Innovation

From Table 1, we can see that in practice, the adoption of the narrow or broad version of

[24] *See* Saltzer, Reed, and Clark 1984, *supra* note 19, at 278.
[25] *See* Saltzer, Reed, and Clark 1998, *supra* note 20, at 69.

the principle can result in a different implementation in network architecture in two sets of circumstances:

(1) When a function can be implemented in both layers, but is not needed by all clients of the lower layer, the narrow version allows implementation at both layers, while the broad version allows implementation at the higher layer only.

(2) When a function can only be completely and correctly implemented at the higher layer, the narrow version allows for additional incomplete implementations at the lower layer for performance considerations.

The differences between the two versions may seem trivial, but the distinction is important for reasons that will later become apparent. The narrow version focuses on an end-to-end system where the sole emphasis is placed on "correctness"[26] — functions can be implemented at any layer where they may be "correctly and completely" implemented, whether at the ends, or at the core. The broad version, on the other hand, goes beyond the concept of correctness, and insists on implementations of any type of non-general, application specific functions being placed at the higher layers, which by extension means placing them at the end points, away from the core of the network.[27] Van Schewick lays out several key advantages to the latter approach, in terms of the network and the applications:

A. Network Evolvability

As a general purpose network, the Internet needs to have the flexibility to be able to support any kind of application. Since each type of application may have a different set of requirements, implementing functions at a lower layer to increase the performance of a certain

[26] *See* VAN SCHEWICK, *supra* note 13, at 79.
[27] *See id.* at 76.

type of application may increase the overhead for another, or even render the network unusable for some other type of application.

Van Schewick cites a classic example of network optimizations that ended up presenting unintended obstacles for subsequent application innovations: the use of load coils in traditional public switched telephone networks to boost the transmission of high frequency voice communications.[28] A side effect of the use of load coils was that frequencies above 3.4 kHz would be cut off. Since voice telephony did not use frequencies over 3.4 kHz at the time, network designers did not see this as a problem. However, this limitation later posed serious problems for the introduction of Digital Subscriber Line (DSL) services over the same lines, as DSL used higher 25 kHz frequencies that were effectively cut off by the load coils. The moral of the story is that optimizations that appear benign in the context of one type of application (voice telephony) could become catastrophic in the context of another application (DSL). By placing application-specific functionality in a higher-layer protocol at the end hosts, we avoid the possibility of lower layers becoming a bottleneck in future innovation. The network itself remains free evolve, as the lower layer continues to accommodate any kind of innovation that may come along.

B. Application Autonomy

It is undisputable fact that applications will always know their own needs better than the network. Van Schewick notes that it is virtually impossible that lower-layer designers will be able to guess in advance all the features applications at the higher layers will potentially need, especially in the case of applications that have yet to materialize. However many features lower-layer designers attempt to cram into the network, applications will most likely end up having to implement application-specific services themselves anyway. Furthermore,

[28] *See id.* at 69.

additional features may not be suitable for all uses of the application, and may create extra overhead, ending up being more harmful than helpful in certain cases. Placing application specific functions at the higher layers ensures that applications have the freedom to determine their own actions, and most importantly, the consequences thereof.[29]

C. Reliability

By implementing functionality specific to certain applications at the lower layer, we may be introducing additional points of failure in the system for those applications that rely on these functions. Since these network functions are not under the direct control of the designer or user of the application, they have no means to correct those problems when they arise. By restricting the placement of application specific functions at the higher layers, we can ensure that all potential points of failure for an application can be addressed by the designer or user of the application, without needing intervention from the lower layers. At the same time, this approach reduces the complexity of the software that needs to be implemented on the hardware at the lower layers. This makes designing and maintaining the network easier, and less prone to malfunction. Together, this makes both the applications and the network more reliable.[30]

D. Lack of Application Awareness in the Core

This is not so much a "feature" of the broad version of the end-to-end principle, as it is a consequence of the architectural limitations set by the principle. Since all application specific functionality is systematically removed from the lower layers, this inevitably results in a network core that lacks any sort of application awareness. With all the "intelligence" placed at the ends, the network itself becomes a "stupid" network, responsible only for the transmission

[29] *See id.* at 71.
[30] *See id.* at 72.

of raw data packets, without regard as to the nature of the content in the packets. This effectively places control of how to use the network in the hands of the applications and the users of the end hosts.[31]

2.3 The Emergence of Application Awareness in the Core

The layering principle in conjunction with an adherence to a broad interpretation of the end-to-end principle necessarily results in an architecture where the core of the network is not able to distinguish between applications. However, for the most part of its history, the choice of whether to adhere to the narrow or broad version of the end-to-end principle did not make all that much of a difference. Computing resources were scarce, and any kind of application specific functionality would have added immense overhead to the routers in a manner that would have greatly impacted throughput. The performance trade-off meant that regardless of which version of the principle one chose to follow, it was generally more rational to leave such functionality at the higher layers on the end points, where computing resources where far more abundant, even if it would have been acceptable in principle to implement those functions at the lower layers.

As computing power continued to grow at an exponential rate, this began to change. At the turn of the century, technological capabilities had advanced enough that hardware equipment vendors began to introduce new network hardware with enhanced capabilities that could inspect data packets as they passed through the network. Later, new hardware would appear that not only allowed network providers to know exactly what was passing through their systems, but also gave them the capability to assign priorities to packets, and most importantly, change the way it handled them. This was a fundamental departure from the original Internet's lack of application awareness.

[31] *See id.*

2.4 Packet Inspection and the Growing Threat of Discrimination

Up until then, the only way network providers had been able to control the end user’s use of the network was through contractual usage restrictions and acceptable use policies. Now, through packet identifying technologies such as Deep Packet Inspection (DPI), they had the capability to directly control how users made use of the network.[32] Suddenly, the choice between adhering to a narrow or broad version of end-to-end made the world of a difference, and the prospect of network providers controlling the flow of information on their networks became so much more realistic. This development understandably had many stakeholders and policy makers worried. It signaled a return to the centralized architecture of the telephone system, where the network provider could act as a gatekeeper, and decide who would get what kind of treatment. In an article ominously titled *Deep Packet Inspection*, one commentator remarked on the implications of the technology:

> Operators can tag packets for fast-lane or slow-lane treatment – or block the packets altogether – based on what they contain or which application sent them…When a network provider chooses to install DPI equipment, that provider knowingly arms itself with the capacity to monitor and monetize the Internet in ways that threaten to destroy Net Neutrality and the essential open nature of the Internet.[33]

The ability to discriminate on the basis of content or application was definitely a legitimate concern. However, despite the long list of potentially unsavory uses, not all types of discrimination were inherently bad. In fact, there were equally as many ways discrimination could be used for the benefit of the user. In the following section, we take a look at the mechanics of network discrimination from a technical perspective, and how they may be used for good and problematic purposes.

[32] *See* Jon M. Peha, *The Benefits and Risks of Mandating Network Neutrality, and the Quest for a Balanced Policy*, 1 INT’L J. COMM. 644, 648-50 (2007).

[33] M. Chris Riley and Ben Scott, *Deep Packet Inspection* (Mar. 2009), *available at* http://www.wired.com/images_blogs/threatlevel/files/dpi.pdf.

2.4.1 The Mechanics of Network Discrimination

To understand how network discrimination works in practice, we must first understand how data traverses the network. Picture a scenario where end host A and B are respectively located on separate networks X and Y, which are connected through Z. When host A sends a packet to B, the data is transferred from network X, through Z, to network Y, via a series of routers and switches along the network. Whenever a router receives a packet, it must first determine which outgoing link to send it on. If the link is available, the packet is sent on its way. If the link is busy, the packet is queued in a buffer, and waits its turn to use the link. If the buffer is full, which happens when the network is overloaded, the packet may be dropped[34].

In the original application-agnostic Internet, all packets were transferred on a first come first serve basis. In an application aware network, the system has far more choices when it comes to deciding what to do with the packet. In the paper *Nuts and Bolts of Network Neutrality*[35], Edward Felten describes some of the different approaches network owners may take, which we adapt here.

A. Best Efforts or Absolute Non-Discrimination

Absolute non-discrimination is where the network does not discriminate at all between the single bits that pass through it. Every individual packet transmitted through the system is treated in exactly the same way, on a first-come-first-serve basis, regardless of its properties. This was referred to as a “best-efforts” service, whereby the network would attempt to deliver any packet based on its best guess and best effort as to how to get it to its destination. When a

[34] According to the TCP/IP protocol, a dropped packet signals to the sending end host that the link is congested, and a well behaved host will then back off and reduce the rate of transmission until the link returns to an uncongested state.

[35] Edward W. Felten, *Nuts and Bolts of Network Neutrality* (Aug. 2006), http://itpolicy. princeton. edu/pub/neutrality. pdf.

link buffer is full and a new packet comes in, the router has several choices: (1) it can drop the new incoming packet, or (2) it can allow it into the queue by dropping another packet in the queue, likely the oldest packet in the queue, if not some other packet at random. In such a scenario, any packet has an equal chance of being dropped.

B. Minimal Discrimination

There are, however, no rules requiring the router to drop packets in a certain way. In fact, a router can discard packets in any way it pleases. Minimal discrimination is a scenario whereby the network assigns priorities to packets in the queue. When necessary, rather than dropping packets at random, or based on their order of arrival, the router will drop packets with the lowest priority first. For example, whenever the buffer is full, the router may decide to drop P2P packets first. Felten calls this “minimal” discrimination[36], because it only discriminates against certain types of packets when the network is congested and therefore cannot serve all packets at once. Most of the time, when the network is not congested, there is no difference between treatment of higher and lower priority packets.

C. Non-Minimal Discrimination

There is another type of implementation, however, in which the routers may selectively discard low priority packets even if there is enough capacity on the network to deliver them. For example, the router may be set to reserve 50% of the network’s capacity for high priority packets. When the percentage of lower priority packets reaches the threshold, they may face being dropped, even if the remaining 50% stays idle. Felten calls this kind of discrimination “non-minimal,” because it artificially restricts certain packets to an arbitrary percentage of

[36] *See id.* at 2.

capacity. [37]

D. Delay Discrimination

Another type of discrimination possible is delay discrimination. This type of discrimination can happen in conjunction with minimal and non-minimal discrimination. Unlike the previous two types of discrimination, which are executed through the dropping of packets, this type of discrimination works through the reordering of packets. Just as the Internet Protocol does not specify what in what order packets should be dropped, it likewise does not specify the order in which they should be sent. While routers generally route packets on a first-come-first-serve basis, it is equally acceptable to send packets in a different order. For example, a router could allow high priority packets to always cut in front of the line, or advance through the queue at a faster pace. Low priority packets therefore experience an extra delay when passing through the router, much like humans do when people cut in line. This delay is known as “latency.” Another consequence of delay discrimination is that packets may be sent out of order, or experience different delays. This variation in delay is known as “jitter.” [38]

E. Absolute Discrimination

This is the most extreme type of discrimination, and in practice this is synonymous with blocking. What happens is that certain types of packets are categorically blocked when they pass through the router, regardless of whether or not there is a link available, or if there is a buffer queue. For example, a network provider with incentives to block Internet voice services could decide to drop all voice packets passing through their network, rendering the network unusable for VoIP.

[37] *See id.*

[38] *See id.* at 3.

2.4.2 Possible Purposes of Network Discrimination

So far we have discussed the means available to network providers for discrimination, but what would be their rationales for engaging in such practices? There are several purposes for which a network provider might choose to employ discriminatory practices on their networks. Some are mostly benign, others more problematic. We discuss some of the most common applications here:

A. Network Congestion Control

In the age of dial-up or “narrowband” Internet, congestion was not much of an issue for network providers. The voice networks that were provided over traditional copper wire could only support a maximum throughput of 56.6 kilobits per second for Internet access, setting a hard cap on the amount of bandwidth each user could use. Furthermore, most of these connections were rarely continuous for long periods of time. This was due to (1) the costs associated with dial-up, and (2) the cost of Internet access being billed in hours. This meant that the level of bandwidth in use at any one time was far lower than necessary to become a cause for concern at the backbone level, where bandwidth was comparably more plentiful.

Eventually, however, Internet access moved from voice to other more efficient modes of transport such as DSL and cable, and more currently fiber and wireless, bringing about highly increased access speeds and “always on” connections, often at flat rates. End users were no longer physically capped by the speeds of their access equipment, but instead by artificial limits set by the network providers. Since it was unlikely that all users would be using 100% of their bandwidth at the same time, ISPs generally adopted a practice known as “oversubscription” when setting such user bandwidth limits. Oversubscription is the practice of selling more bandwidth than you have capacity for, by planning for the typical demand rather than peak demand. This is a more efficient use of network resources, because it "supports the maximum amount of subscribers on the least amount of infrastructure."[39] It also allows users of the network to take advantage of a higher level of bandwidth at any given moment than would be possible in a non-oversold network.

However, usage trends gradually evolved, and users began to use the network in ways not previously envisioned by the network providers. As the Internet began to exhibit a shift in behavior towards more bandwidth-intensive, rich-media content, such as streaming video and peer-to-peer (P2P) file sharing, these bandwidth-intensive uses inevitably put an increasing strain on the backbone. This resulted in more frequent congestion as users began to use up more and more of their bandwidth allotments on a regular basis.

During periods of congestion, the network provider may have an incentive to employ network management policies to ensure fair and/or efficient allocation of network resources. Network management does not necessarily have to be employed in an application discriminating manner; however, there may be certain benefits to doing so. For example, instead of slowing down all the traffic from a user using a high amount of bandwidth, the network could have a congestion policy that slowed down a user's file downloads but not VoIP packets. This can be achieved through a combination of minimal discrimination and delay discrimination. By giving higher packet priority to VoIP packets, the network provider could ensure that user's connection could be used normally for voice communications even during periods of congestion. This may make the network more useful as a whole. On the other hand, application specific network management can also be used in more sinister ways, for example, by slowing down access to certain high bandwidth sites during periods of congestion. Determining what types of congestion management policies are acceptable is often a core issue of the net neutrality debate.

[39] *See* Tom Mitchell, *Avoiding the Pitfalls of Oversubscription in DSL Networks*, VISION2MOBILE, Apr. 2000, http://www.vision2mobile.com/articles/2000/04/avoiding-the-pitfalls-of-oversubscription-in-dsl.aspx.

B. Quality of Service Assurances

Another application of network discrimination is to provide Quality of Service ("QoS") assurances. In simple terms, QoS entails the prioritization of certain types of data over others based on their special requirements. We already discussed in the previous section how network providers could prioritize VoIP packets so voice communications could continue to function when the network was saturated. While this is especially important during periods of congestion, there are equally compelling reasons to employ such prioritization during periods of non-congestion.

Some applications are not sensitive to packet delay. For example, it does not matter as much if your file takes 10 more seconds to download, or if the packets for your website arrive in a slightly different order. However, the same may be detrimental for a real-time application such as VoIP or IPTV. Generally, there are two types of QoS guarantees a network may provide:

(1) *Bandwidth guarantees*

While most applications can adapt to available bandwidth by sacrificing either speed or quality, some network applications require a constant level of bandwidth. This is most often the case with streaming audio of video, which can sacrifice neither transfer speed nor quality. By reserving resources in the system and under-subscribing for those resources, the network can make sure certain applications will always receive a certain share of the link capacity, even during periods of congestion. In practice, this is a kind of non-minimal discrimination. [40]

(2) *Delay guarantees*

[40] *See Deploying Guaranteed-Bandwidth Services with MPLS* (2002), CISCO SYS., INC., http://www.cisco.com/warp/public/cc/pd/iosw/prodlit/gurtb_wp.pdf.

> Bandwidth guarantees alone may not ensure low latency or low jitter. For example, a longer route through the network may guarantee bandwidth, but result in a higher delay. It may not matter to you that your Internet video stream plays at a five second delay as long as it is continuous and clear, but that same five second delay may render applications such as voice communications or online-gaming unusable. This is even more crucial in the case of mission-critical applications, such as financial or medical applications, where the difference of a millisecond may be the difference between millions of dollars, or worse, life and death. By assigning a higher priority to latency sensitive data, the network can ensure the data arrives with the least amount of latency possible. This, however, has the consequence of imposing delay discrimination on other lower priority packets. [41]

There are clear benefits to providing QoS guarantees on the Internet, but the fact is that prioritizing one type of data necessarily means de-prioritizing some other type of data. How the network provider may decide which packets to prioritize (and by extension, what it may de-prioritize) is the subject of much controversy. QoS may be employed for public good, like in many of the examples above, or for self serving interests, like prioritizing services of favored partners only. Determining how the network provider may prioritize, and under what conditions it may prioritize, is another core issue of the net neutrality debate.

C. Blocking, Filtering and Censorship

Network discrimination may also be applied to block, filter and censor content or traffic. This may sound highly problematic in terms of net neutrality, but there are actually many non-sinister reasons a network provider might choose to do so. For one, the network might be required to filter or block content in accordance to the law. The one extreme example of such

[41] *See id.*

an application of network discrimination would be China. Of course, due to strong First Amendment protections built into the Constitution, there is currently no comparable situation in the United States, and in practice there is very little, if any, government mandated technical filtering[42]. However, there have been attempts to introduce proposals that could have brought precisely such requirements into law[43]. Regardless of whether such laws are desirable or not, it is generally not considered controversial for the network provider to comply with such legal obligations.

Another quite legitimate purpose for filtering is network security. There are many harmful and abusive uses of the network that can threaten the utility of the Internet for the public as a whole. For example, denial of service attacks can flood the network with bogus packets, clogging the network and rendering it unusable for other purposes. Viruses may attempt to replicate through the network, infecting other computers on the network through weaknesses in their systems. Malicious software or end users may attempt to exploit the network in a way that violates or interferes with standard protocols, monopolizing the resources or affecting other users' ability to use the network. Compromised machines may be remotely controlled by hackers to send spam and launch distributed denial of service attacks. An application aware network can easily filter out such communications at the lower level, employing absolute discrimination policies on harmful traffic, ensuring the network can continue to function properly, and providing end users with an extra level of protection. This is one of least controversial rationales for network discrimination.

Of course, there remains the very real possibility that network providers discriminate for self-serving purposes. For every legitimate filtering purpose, one can easily think of a hundred ways network providers could employ filtering for far more unfavorable purposes. For

[42] *But see* Children's Internet Protection Act of 1999, S. 97, 106th Cong. (1999).

[43] *See, e.g.*, Stop Online Piracy Act of 2011, H.R. 3261, 112th Cong. (2011); PROTECT IP Act of 2011, S. 968, 112th Cong. (2011).

example, network providers could block competing websites or applications, and censor content that it does not like, or demand a fee to deliver content to their users. These are all legitimate concerns that would likely raise opposition.

D. Intentional Degradation of Service

This is not so much a “purpose” of discriminatory network practices as it is a crude application of such practices absent of any of the legitimate purposes outlined above. It basically serves no goal at all other than the degrading of service per se. There are hardly any justifications for this type of discriminatory conduct, as it puts the discriminated entity at a great disadvantage, while benefiting no one, except maybe the network provider.

2.5 Responding to the Threat of Discrimination

While the layering principle and the end-to-end argument have historically resulted in a network lacking application awareness at the core, technological advances have rendered that limitation irrelevant. The core network is now more intelligent than ever, whether we like it or not. We have presented several ways that intelligence can be used to discriminate against different types of traffic, and we have discussed how they can be employed to achieve several different types of objectives, some of which may be beneficial and others which may be harmful. The question, therefore, is whether network providers should be prevented from harnessing that intelligence.

The stakes are high in this debate. Some, like Lemley and Lessig, argue that the network providers should not be allowed to use that intelligence, because to do so would “compromise an important architectural principle that has governed the Internet since its inception …

Nothing less than the structure of the Internet itself is at stake in this debate."[44]

As we've seen, the broad version of the end-to-end principle calls for the network to be "dumb", and it presents several sound architectural justifications for doing so. But it is not simply the architectural justifications that matter the most for advocates of regulation. It is the "relationship between these architectural principles and the innovation of the Internet." Lessig writes:

> While the [end-to-end] design principle was first adopted for technical reasons, it has important social and competitive features as well. [end-to-end] expands the competitive horizon, by enabling a wider variety of applications to connect and use the network. It maximizes the number of entities that can compete for the use and applications of the network. As there is no single strategic actor who can tilt the competitive environment (the network) in favor of itself, or no hierarchical entity that can favor some applications over others, an [end-to-end] network creates a maximally competitive environment for innovation, which by design assures competitors that they will not confront strategic network behavior.[45]

The consequences of these architectural principles have indeed been profound, but they are still just that: architectural principles. In Internet Engineering Task Force (IETF) Request for Comments No. 1958, a "snapshot" of the then-current principles of Internet architecture, then-Chair of the IETF Brian Carpenter thusly wrote:

> In searching for Internet architectural principles, we must remember that technical change is continuous in the information technology industry. The Internet reflects this. Over the 25 years since the ARPANET started, various measures of the size of the Internet have increased by factors between 1000 (backbone speed) and 1000000 (number of hosts). In this environment, some architectural principles inevitably change. Principles that seemed inviolable a few years ago are deprecated today. Principles that seem sacred today will be deprecated tomorrow. The principle of constant change is perhaps the only

[44] *See* Lemley and Lessig, *supra* note 18, at 925.
[45] *See id.* at 931.

> principle of the Internet that should survive indefinitely.[46]

Few would dispute the very real benefits that the end-to-end architecture and the consequentially "neutral" Internet have created for society. However, the desirability of hard-coding those architectural principles into law is debatable. Some argue that while the threat is indeed possible, it is for the moment speculative at best. Others worry that hard-coding technological principles into law may have unintended consequences further down the road.

It is these differing responses to threat of network discrimination – rather than disagreements over definitions – that form the core of the net neutrality debate.

[46] *See* Brian Carpenter, *Architectural Principles of the Internet*, IETF RFC 1958 (Jun. 1996), http://www.ietf.org/rfc/rfcl958.txt.

III. The Current State of Net Neutrality Regulation

Over the years, regulators have taken into account the concerns and considerations of stakeholders, activists, and commentators from all sides of the spectrum, and have gradually attempted to formulate a framework for responding to the threat of network discrimination. On the legislative level, despite much political posturing and being prominently on the agenda for several years, not much has substantially changed. Numerous attempts have been made over the years to introduce net neutrality provisions into law, but every single one of those proposals has failed to pass. For now, the issue remains under the purview of the Federal Communications Commission (“FCC” or “the Commission”).

3.1 The FCC and the Telecommunications Act of 1996

The task of formulating national Internet policy has traditionally fallen to the FCC. Over the years, the FCC has successively adopted numerous positions and policy statements in support of net neutrality principles; however, its authority to regulate broadband has recently been thrown into question following the outcome of a D.C. Circuit Court decision. In this chapter we examine the source of the FCC’s statutory authority, and the developments that have shaped the course of net neutrality discourse and the current regulatory framework in the United States.

3.1.1 The FCC’s Statutory Authority for Internet Regulation

Established under the Communications Act of 1934 (“the Act” or “the Communications Act”)[47], the FCC was vested with broad powers to regulate interstate radio and wireline communications[48]. Obviously, the Internet did not yet exist at the time the FCC was created,

[47] Communications Act of 1934, 47 U.S.C. §§ 151-614 (2006).

[48] Section 1 of the Communications Act provides: “For the purpose of regulating interstate and foreign commerce in communication by wire and radio so as to make available, so far as possible, to all the people of the

but its nature as a form of wireline communications plausibly placed it under the jurisdiction of the Commission. The 1996 overhaul of the Act, formally known as the Telecommunications act of 1996 ("the 1996 Act")[49], brought with it the first official mentions of the Internet, specifically in Section 230(b) of the Act, which was amended to include the following provisions:

> (b) POLICY
> It is the policy of the United States—
>
> (1) to promote the continued development of the Internet and other interactive computer services and other interactive media;
>
> (2) to preserve the vibrant and competitive free market that presently exists for the Internet and other interactive computer services, unfettered by Federal or State regulation;
>
> (3) to encourage the development of technologies which maximize user control over what information is received by individuals, families, and schools who use the Internet and other interactive computer services;
>
> (4) to remove disincentives for the development and utilization of blocking and filtering technologies that empower parents to restrict their children's access to objectionable or inappropriate online material; and
>
> (5) to ensure vigorous enforcement of Federal criminal laws to deter and punish trafficking in obscenity, stalking, and harassment by means of computer.[50]

Additionally, the 1996 Act addressed the deployment of high speed broadband services

United States, without discrimination on the basis of race, color, religion, national origin, or sex, a rapid, efficient, Nation-wide, and world-wide wire and radio communication service with adequate facilities at reasonable charges, for the purpose of the national defense, for the purpose of promoting safety of life and property through the use of wire and radio communications, and for the purpose of securing a more effective execution of this policy by centralizing authority heretofore granted by law to several agencies and by granting additional authority with respect to interstate and foreign commerce in wire and radio communication, there is created a commission to be known as the 'Federal Communications Commission', which shall be constituted as hereinafter provided, and which shall execute and enforce the provisions of this chapter." *Id.* § 151.

[49] Telecommunications Act of 1996, 47 U.S.C. §§ 151, 216, 607-09 (2006).

[50] 47 USC § 230(b).

in Section 706, requiring the FCC to conduct annual reports on the state of advanced telecommunications services:

> SEC. 706. ADVANCED TELECOMMUNICATIONS INCENTIVES.
>
> (a) In General: The Commission and each State commission with regulatory jurisdiction over telecommunications services shall encourage the deployment on a reasonable and timely basis of advanced telecommunications capability to all Americans (including, in particular, elementary and secondary schools and classrooms) by utilizing, in a manner consistent with the public interest, convenience, and necessity, price cap regulation, regulatory forbearance, measures that promote competition in the local telecommunications market, or other regulating methods that remove barriers to infrastructure investment.
>
> (b) Inquiry: The Commission shall, within 30 months after the date of enactment of this Act, and regularly thereafter, initiate a notice of inquiry concerning the availability of advanced telecommunications capability to all Americans (including, in particular, elementary and secondary schools and classrooms) and shall complete the inquiry within 180 days after its initiation. In the inquiry, the Commission shall determine whether advanced telecommunications capability is being deployed to all Americans in a reasonable and timely fashion. If the Commission's determination is negative, it shall take immediate action to accelerate deployment of such capability by removing barriers to infrastructure investment and by promoting competition in the telecommunications market.[51]

As Congressional statements of policy rather than clear statutory mandates, none of these provisions did in fact constitute express statutory authority for the FCC to regulate the Internet. Instead, the Commission has historically relied on Section 4(i) of the Communications Act, which provides that “[t]he Commission may perform any and all acts, make such rules and regulations, and issue such orders, not inconsistent with this Act, as may be necessary in the execution of its functions.”[52] This so called “ancillary” jurisdiction has

[51] Telecommunications Act of 1996, Pub. L. No. 104-104, § 706(a), 110 Stat. 56, 153.
[52] 47 U.S.C. § 154(i).

historically given the FCC far-reaching powers to impose regulatory requirements on forms of communications far beyond the scope of services specifically regulated by the Act. For example, before the Communications Act was amended to cover cable services, the FCC was able to employ its Title I ancillary jurisdiction to enforce certain restrictions on cable services in order to ensure "the effective performance of the Commission's various responsibilities for the regulation of television broadcasting"[53]. It is upon this same ancillary authority that the FCC now bases its net neutrality policies.

3.1.2 The FCC's Evolving Approach to Net Neutrality

The concept of network neutrality has never been a new one for the FCC. For the longest periods of time, the principles of common carriage and non-discrimination had already been an integral part of communications policy. However, the Internet emerged during an era of aggressive deregulation, prompting the Commission to take a hands-off approach in its efforts to encourage deployment of broadband Internet[54]. In fact, not only did it choose not to impose non-discrimination principles on the nascent network, the early FCC actually went out of its way to ensure the broadband Internet access market would not be subjected to traditional common carrier restrictions[55].

This, however, did not mean the Commission did not find a neutral network an attractive proposition. In fact, maintaining the openness of the Internet and empowering consumers continued to remain a primary concern for the FCC[56]. Consistent with its deregulatory stance at that time, the Commission's initial foray into net neutrality was in the form of an

[53] *See* United States v. Sw. Cable Co., 392 U.S. 157 (1968); United States v. Midwest Video Corp., 406 U.S. 649 (1972).
[54] *See* Chong, *supra* note 8, at 149.
[55] *See* Appropriate Framework for Broadband Access to the Internet Over Wireline Facilities, 20 FCC Rcd. 14953 (2005) [hereinafter Wireline Broadband Order].
[56] *See* Michael K. Powell, Chairman, FCC, Remarks at the Silicon Flatirons Symposium on "The Digital Broadband Migration: Toward a Regulatory Regime for the Internet Age" (Feb. 8, 2004), *in Preserving Internet Freedom: Guiding Principles for the Industry*, *available at* http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-243556A1.pdf.

educational campaign rather than regulatory act. During a speech at the Silicon Flatirons Symposium in 2004, then-FCC Chairman Michael Powell introduced his vision of four "Internet Freedoms." These principles would later become the cornerstone of the Commission's Internet policy. The "four freedoms" were:

A. *Freedom to Access Content.*

First, consumers should have access to their choice of legal content.

B. *Freedom to Use Applications.*

Second, consumers should be able to run applications of their choice.

C. *Freedom to Attach Personal Devices.*

Third, consumers should be permitted to attach any devices they choose to the connection in their homes.

D. *Freedom to Obtain Service Plan Information.*

Fourth, consumers should receive meaningful information regarding their service plans.[57]

In August of 2005, as part of the broader Wireline Broadband Order[58] deregulating broadband Internet service, the newly appointed FCC Chairman Kevin Martin reaffirmed the principles laid out by his predecessor by incorporating them into the Commission's 2005 Internet Policy Statement ("the Policy Statement" or "the Policy")[59]. This was the Commission's first official position on net neutrality. The Policy Statement similarly outlined four principles, each preceded by the explicitly stated goal "[to] encourage broadband deployment and preserve and promote the open and interconnected nature of the public Internet":

(1) consumers are entitled to access the lawful Internet content of their choice.

[57] *See id.*

[58] *See* Wireline Broadband Order, *supra* note 55.

[59] *See* Appropriate Framework for Broadband Access to the Internet Over Wireline Facilities, 20 FCC Rcd. 14986 (2005) [hereinafter Internet Policy Statement].

(2) consumers are entitled to run applications and use services of their choice, subject to the needs of law enforcement.

(3) consumers are entitled to connect their choice of legal devices that do not harm the network.

(4) consumers are entitled to competition among network providers, application and service providers, and content providers[60].

Though fundamentally similar in nature, the new Policy Statement was significant in that it marked a first step by the FCC towards delineating the boundaries of Powell's four freedoms. Notably, in contrast to the original four freedoms, the new Policy Statement added an important exception to the rules, providing in a footnote that all the principles would be "subject to reasonable network management."

Although the Policy Statement itself did not amount to enforceable rules, it laid the groundwork for future rule-making proceedings that would implement the policy, and served as a guideline for the regulated entities regarding what would be expected of them by the Commission following the deregulation of broadband Internet access. The principles were subsequently adopted as de facto rules in the form of conditions for the approval of mergers between several telecommunications carriers, including the mergers of AT&T with SBC and MCI with Verizon[61].

3.1.3 The Madison River Case

In early 2005, the FCC launched an investigation into the practices of Madison River Communications LLC. The rural telephone carrier was accused of blocking VoIP services for

[60] *See id.*

[61] *SBC Commc'ns, Inc. and AT&T Corp. Applications for Approval of Transfer of Control*, Memorandum Opinion and Order, 20 FCC Rcd. 18290, 18392, ¶ 211 (2005); *Verizon Commc'ns Inc. and MCI, Inc. Applications for Approval of Transfer of Control*, Memorandum Opinion and Order, 20 FCC Rcd. 18433, 18537, ¶ 221 (2005); *AT&T Inc. and BellSouth Corp. Application for Transfer of Control*, Memorandum Opinion and Order, 22 FCC Rcd. 5662, 5663, ¶ 2 (2007).

users on its network. This was the first time that the FCC had ever investigated an alleged net neutrality violation. Although there were in fact no enforceable rules at the time, in an effort to avoid a lengthy inquiry and possible disciplinary action, Madison River entered into a Consent Decree with the FCC in which it agreed to stop discriminating against VoIP in exchange for the FCC dropping the inquiry. In addition, Madison River agreed to make a voluntary contribution to the Unites States Treasury in the amount of $15,000, as a form of pseudo fine.[62]

Since the investigation was terminated, the FCC did not formally charge Madison River from violating any laws or regulations. Also, as the case was resolved without adjudication, it did not set a precedent for subsequent cases. Nevertheless, the case for the first time signaled the Commission's willingness to use its power to enforce open Internet principles. In a press release, Chairman Powell commended the swift action of the enforcement bureau, and remarked that "[the] industry must adhere to certain consumer protection norms if the Internet is to remain an open platform for innovation."[63]

3.2 The Comcast BitTorrent Case

In 2008, cable broadband provider Comcast was accused of "selectively targeting" and "interfering" with the use of peer-to-peer (P2P) applications by its users, in contravention of the Internet Policy Statement's rule that "consumers are entitled to access the lawful Internet content of their choice… [and] to run applications and user services of their choice." In response to complaints, the FCC launched an inquiry into the alleged practices. Here we discuss the findings of that investigation, and the consequences of the FCC's subsequent

[62] Madison River Comms., LLC, 20 FCC Rcd. 4295, 4297 (2005).

[63] *See FCC Chairman Michael K. Powell Commends Swift Action to Protect Internet Voice Services* (Mar. 3, 2005), http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-257175A1.pdf

actions.[64]

3.2.1 FCC Findings and Ruling

Initially Comcast publicly denied the charges[65], declaring that they did not block or throttle any traffic. However, subsequent tests conducted by the media and advocacy groups concluded that the network provider was indeed “actively [interfering] with attempts by some of its high-speed Internet subscribers to share files online.”[66] The Associated Press also found that the method used by Comcast was “difficult to circumvent and [involved] the company falsifying network traffic.”[67] Specifically, Comcast routers were forging RST packets appearing to come from their users’ computers, signaling their peers to terminate the connection, effectively preventing the transfer of data. This is akin to a telephone operator joining a phone call between you and a friend, and informing you in your friend’s voice that she has to hang up.[68]

Following the tests, Comcast admitted that it did in fact target P2P traffic for management, but asserted that its conduct was necessary to ease congestion on the network and was only employed during periods of peak network congestion. However, further tests by users contradicted Comcast’s claim that the interference practices were limited to “times of congestion”, concluding instead that the discriminatory practices were in effect constantly “regardless of the time of day or night, regardless of the day of the week, and [despite] the presumable differences in network congestion during prime time and non-prime time hours of use.” Confronted with this evidence, Comcast changed its tune once again, finally conceding

[64] *See* Formal Complaint of Free Press and Public Knowledge Against Comcast, 23 FCC Rcd. 13028 (2008) [hereinafter Comcast Order].

[65] *See* Marguerite Reardon, *Comcast Denies Monkeying with BitTorrent Traffic*, CNET NEWS BLOG (Aug. 21, 2007), http://www.news.com/8301-10784_3-9763901-7.html (last visited July 2, 2012).

[66] Peter Svensson, *Comcast blocking some subscriber Web traffic, AP tests show*, ASSOCIATED PRESS (Oct. 19, 2007), *available at* http://seattletimes.nwsource.com/html/businesstechnology/2003962029_webcomcast19.html.

[67] *Id.*

[68] *See* Courtney Erin Smith, *Net Neutrality, Full Throttle: Regulation of Broadband Internet Service Following the Comcast/Bittorrent Dispute*, 50 SANTA CLARA L. REV. 569 (2010).

that its P2P management practices were triggered regardless of the level of network congestion.[69]

The FCC issued an Order censuring Comcast for its behavior, concluding that Comcast's "discriminatory and arbitrary [practices] unduly [squelched] the dynamic benefits of an open and accessible Internet and [did] not constitute reasonable network management." It also added that "Comcast's failure to disclose the company's practice to its customers [had] compounded the harm." The Commission ordered Comcast to disclose the details of its network management practices, and submit a compliance plan describing how it would cease its unreasonable management practices by the end of the year.

3.2.2 Comcast v. FCC[70]

Although Comcast complied with the order and subsequently ceased its discriminatory practices against P2P traffic on its network, it appealed the FCC ruling, arguing that the Commission had no statutory authority to oversee its network management practices. The Commission argued that it did in fact have such authority, relying on the Supreme Court's ruling in *Brand X*[71], which affirmed the Commission's authority under Section 4(i) "to impose special regulatory duties on facilities-based ISPs under its Title I ancillary jurisdiction."

The D.C. Circuit court applied a two-part test for ancillary authority, as laid out in *Am. Library Ass'n v. FCC*[72], maintaining that the FCC can only exercise its ancillary jurisdiction if it can show that "(1) the Commission's general jurisdictional granted under Title I [of the Communications Act] covers the regulated subject and (2) the regulations are reasonably ancillary to the Commission's effective performance of its statutorily mandated

[69] *See* Comcast Order, *supra* note 64, ¶ 9.
[70] Comcast v. Fed. Commc'ns Comm'n, 600 F.3d 642 (D.C. Cir. 2010) [hereinafter Comcast Case].
[71] National Cable & Telecomms. Ass'n v. Brand X Internet Services, 545 U.S. 967, 996 (2005).
[72] American Library Ass'n v. Fed. Commc'ns Comm'n, 406 F.3d 689, 692 (D.C. Cir. 2005).

responsibilities."[73]

As Internet services undoubtedly qualified as "interstate and foreign communication by wire," the first prong of the test was clearly satisfied. However, in a unanimous decision, the Court found that the Commission had failed to present any "statutorily mandated responsibilities" that its oversight of network management practices could have been deemed "reasonably ancillary to." The court had the following to say about each of the sources of authority advanced by the Commission on appeal:

A. *Section 230(b)*

   Policy statements cannot be considered statutorily mandated responsibilities, and thus cannot be sources for ancillary authority.[74]

B. *Section 706*

   A prior order by the Commission[75], which was still in force, explicitly concluded that it had no ancillary authority, as such the Commission was estopped from advancing this argument[76]

C. *Section 256*

   The language of the statute explicitly rejected the possibility for an expansion of authority, restricting its use for ancillary jurisdiction.[77]

D. *Section 257*

   FCC could have ancillary authority over certain issues under this statute, but not on the issue of network management.[78]

E. *Section 201*

[73] *Id.* at 691-92.
[74] *See* Comcast Case, at 654.
[75] *See* Deployment of Wireline Servs. Offering Advanced Telecomms. Capability et al., Memorandum Opinion and Order and Notice of Proposed Rulemaking, 13 FCC Rcd. 24012 (1998).
[76] *See* Comcast Case, at 658-59.
[77] *See id.* at 659.
[78] *See id.*

The court declined to examine this claim, since the FCC did not advance the same argument for it on appeal as it did in the original Order.[79]

F. *Title III*

There court found no source for this argument in the original Order and similarly declined to rule on it.[80]

G. *Section 623*

Ancillary authority for this statute is sharply limited, and the Commission did not base its authority on that narrow grant of power.[81]

Absent an explicit statutorily mandated responsibility to rely on, the Court declined to find a basis for the Commission's ancillary authority. It maintained that doing otherwise would "virtually free the Commission from its congressional tether," giving the FCC free rein to impose regulations on Internet service providers as it saw fit.[82]

3.2.3 Implications of the Comcast Decision

The Comcast decision had enormous implications for the Commission's policy objectives. From a narrow perspective, the FCC now had to be very clear on its source of power when it wanted to issue an Order. From a broader perspective, the FCC's ability to advance any of its Internet related policies (e.g. National Broadband Plan, consumer privacy protections, publication of cost information, etc.) would now be in great jeopardy, as was its general authority to regulate Internet services in general. Every Order issued by the Commission would now likely be subject to challenge in court.

However, the Comcast decision did not foreclose all options for the FCC to assert its

[79] *See id.* at 660.
[80] *See id.*
[81] *See id.*
[82] *See id.* at 655.

authority. The Court left several possible avenues open, such as through appropriate arguments based on the statutes the Court had not ruled on, or through an overruling or clarification of the FCC's prior interpretation on Section 706 authority, which is in fact what the FCC subsequently did[83]. If all failed, there was always the option of reclassifying broadband Internet services as "telecommunications services" under Title II, bringing ISPs directly under the authority of the Commission. Ironically for network providers, if the FCC chose to take the latter approach, disputing the Commission's ancillary authority could have had the opposite effect of bringing themselves under stricter regulation.

3.3 FCC Open Internet Order of 2010

The Comcast decision was a huge setback for the FCC, but it did not deter the Commission from continuing to advance its policy for net neutrality. In fact, the decision may have had the effect of accelerating the Commission's adoption of formal regulations. In December of the same year, the FCC formally introduced a set of new rules through the issuing of a new Open Internet Order of 2010[84]. The order outlined the first set of official net neutrality rules promulgated by the FCC, building upon many of the concerns and arguments the Commission had discussed in the Comcast Order.

3.3.1 Open Internet Rules

The new regulations specified three rules to be followed by network providers:

(1) Transparency

> A person engaged in the provision of broadband Internet access service shall publicly disclose accurate information regarding the network management practices,

[83] *See* Preserving the Open Internet Broadband Industry Practices, 25 FCC Rcd. 17905 ¶ 119 (2010) [hereinafter Open Internet Order].

[84] *Id.*

performance, and commercial terms of its broadband Internet access services sufficient for consumers to make informed choices regarding use of such services and for content, application, service, and device providers to develop, market, and maintain Internet offerings.[85]

(2) No Blocking

A person engaged in the provision of fixed broadband Internet access service, insofar as such person is so engaged, shall not block lawful content, applications, services, or non-harmful devices, subject to reasonable network management.

A person engaged in the provision of mobile broadband Internet access service, insofar as such person is so engaged, shall not block consumers from accessing lawful websites, subject to reasonable network management; nor shall such person block applications that compete with the provider's voice or video telephony services, subject to reasonable network management.[86]

(3) No Unreasonable Discrimination

A person engaged in the provision of fixed broadband Internet access service, insofar as such person is so engaged, shall not unreasonably discriminate in transmitting lawful network traffic over a consumer's broadband Internet access service. Reasonable network management shall not constitute unreasonable discrimination.[87]

The new FCC rules can be considered a consolidation and natural evolution of the original four freedoms and the open Internet principles advanced by the 2005 Internet Policy Statement. While the original principles outlined a general policy, the new regulations set the

[85] 47 C.F.R. § 8.3
[86] 47 C.F.R. § 8.5
[87] 47 C.F.R. § 8.7

stage for a credible enforcement regime. In this respect, the major development is the elevation of the reasonable network management exception from a tiny footnote of the principle to an integral part of the rule itself. Determining a rational standard for “reasonableness” will likely become one of the most significant tasks for this new policy.

Also significant is the way the rules create a distinction between fixed and mobile broadband. Under the new rules, absent of reasonable network management justifications, fixed broadband providers may not block any type of lawful traffic[88]. Mobile broadband providers, on the other hand, are only prohibited from blocking websites, and are free to block other applications and uses unless they compete with its own voice and video telephony services. This signals that the Commission recognizes there are differences between the two types of networks. Although this is an important aspect of the rule, for the purposes of this paper, we will not be delving much into its implications, but will rather be focusing on reasonable network management provisions and the standard for reasonableness.

3.3.2 Setting a Standard for Reasonableness

The FCC order employs the term “reasonableness” with relative impunity. The term and variations thereof appear no less than eleven times in the substantive rules alone. We first turn our focus to the order’s “no blocking” and “no unreasonable discrimination” rules, where seven of these occurrences are found. The first rule prohibits the blocking of lawful traffic, subject to reasonable network management. The second prohibits the unreasonable discrimination of lawful traffic, while making clear that reasonable network management would not constitute unreasonable discrimination. It is noteworthy that the Commission chose to adopt a distinct construct in the two rules for the same exception. Assuming the choice of

[88] The Open Internet Order clarifies that “[the] phrase ‘content, applications, services’ refers to all traffic transmitted to or from end users of a broadband Internet access service, including traffic that may not fit cleanly into any of these categories.” *Id.* ¶ 64.

wording is deliberate – and there is no reason to believe otherwise – the discrepancy seems to signal a difference in the way the exception will be applied.

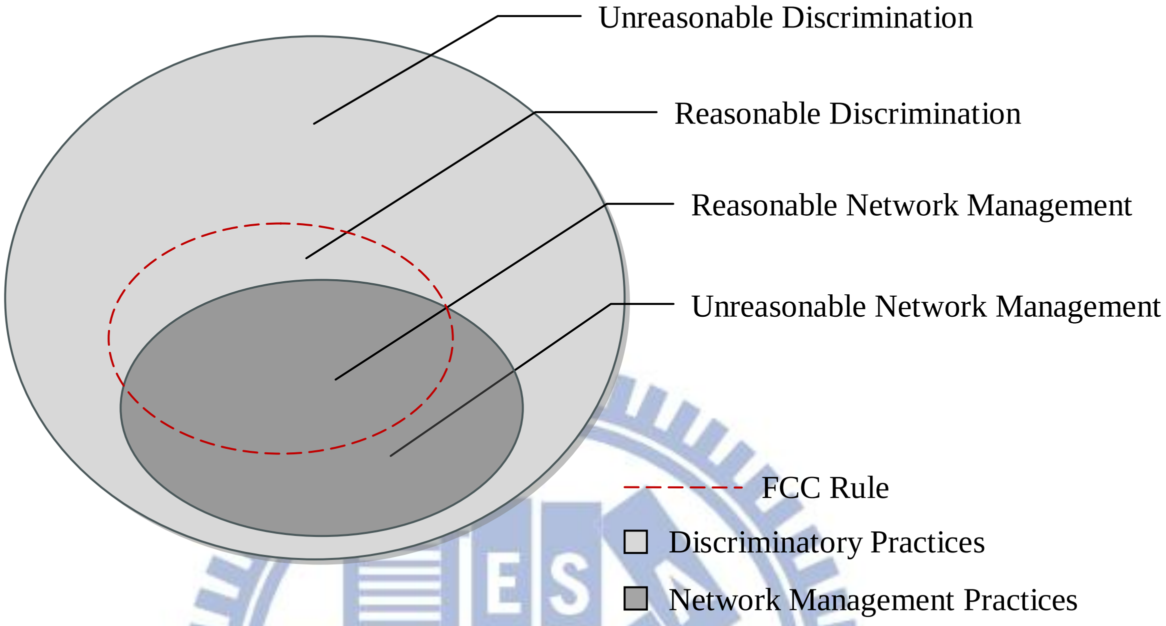

Unreasonable Discrimination
Reasonable Discrimination
Reasonable Network Management
Unreasonable Network Management
FCC Rule
□ Discriminatory Practices
▨ Network Management Practices

Figure 2: Topography of Discriminatory Network Practices According to FCC rules.

According to a strict interpretation of the rules, while reasonable network management is the only justification for the blocking of lawful content, it seems there may be more leeway for other types of discriminatory treatment. We can see in Figure 2 a deconstruction of the rule, which reveals a topography outlining two classes of discriminatory network practices that could be considered reasonable under the rules. While reasonable network management is always considered a reasonable discrimination, the figure shows that reasonable discrimination does not necessarily have to be a "reasonable network management practice". The order defines a "reasonable network management practice" as "[a] network management practice [that is] appropriate and tailored to achieving a legitimate network management purpose, taking into account the particular network architecture and technology of the broadband Internet access service."[89] This would seem to leave open the possibility for "reasonable" discriminatory practices that do not have legitimate network management purposes. Legitimate network management purposes explicitly outlined by the Commission include:

(1) Ensuring network security and integrity, including by addressing traffic that is harmful to the network;
(2) Addressing traffic that is unwanted by end users (including by premise operators), such as by providing services or capabilities consistent with an end user's choices regarding parental controls or security capabilities; and
(3) Reducing or mitigating the effects of congestion on the network.[90]

This presents two questions: (1) are there other legitimate purposes that the Commission does not identify, and (2) what discriminatory practices not for these purposes are allowed? Although the FCC order does not elaborate on what these "reasonable discriminatory practices", it provides certain clues. For example, it mentions pay-for-priority as a practice that is "unlikely" to survive scrutiny.[91] But does that leave the door open for some that might?

In the next chapter, we will explore the current limits of "reasonable network management practices" and "reasonable discrimination" according the open Internet rules, and additional factors the FCC needs to incorporate into its decision-making process.

[89] *Id.* ¶ 6.
[90] *Id.* ¶ 82.
[91] *Id.* ¶ 76.

IV. Possible Challenges to Open Internet Rules

5.1 Net Neutrality and the Fifth Amendment

Although the Commission dismisses constitutionality concerns in its order, it would nevertheless be prudent for it to tread a cautious path in developing and enforcing its rules. While the order mentions First Amendment and Fifth Amendment concerns, here I wish to focus my attention on the latter alone.

The Fifth Amendment, more famous for bestowing the right to remain silent, also contains a lesser known provision limiting the government's power to exercise "eminent domain" – the taking of private property for public use. This "Takings Clause" of the Fifth Amendment requires that the government provide "just compensation" for the taking of private property. Commentators have suggested that open Internet rules may constitute a "virtual taking" of private property, and may fall under the purview of this Constitutional right. Drawing a parallel to real property law, Daniel Lyons notes that "the rights granted to content and application providers are akin to a virtual easement to traverse broadband providers' networks."[92] At the basic level, this "virtual easement" seeks to restrict the network provider's "right to exclude," which the Court has remarked "has traditionally been considered one of the most treasured strands … of property rights."[93] However, the bigger concern is that the owner of the network could lose its ability to use its own network. Lyons writes:

> By surrendering permanent access to third parties, broadband providers also lose the ability to control the *use* of their networks. At a base level, a broadband provider physically cannot use for its own purposes bandwidth that has already been occupied by a third party. Nor may it sends its own signals through the

[92] Daniel Lyons, *Virtual Takings, The Coming Fifth Amendment Challenge to Net Neutrality Regulations*, 86 NOTRE DAME L. REV. 65 (2011).

[93] Loretto v. Teleprompter Manhattan CATV Corp., 458 U.S. 419, 435 (1982).

> network if doing so will disproportionately “degrade” third party content (for example, by adversely rerouting third-party data packets in a way that would cause delays or packet loss). Indeed, broadband providers even lose the ability to control how third parties use the network, insofar as the rules prohibit providers from prioritizing certain third-party packets for faster delivery.[94]

In response to these concerns, the FCC cites *Penn Central Trans. Co. v. City of New York*[95], in which the Supreme Court laid out a set of guidelines for evaluating whether a regulatory action by the government could constitute a “taking” of private property. The balancing test weighs the “economic impact of the regulation” and the degree of interference in the owner’s “investment-backed expectations” against “the character of the government action.”[96] The FCC argues that in light of “the history of broadband Internet access services” there is no reasonable basis for an investment-backed expectation “[relying] on a policy regime in which providers are free to conceal or discriminate without limit.” Furthermore, it argued that the rules adopted “should not impose substantial new costs on broadband providers,” negating the economic impact of the regulation.[97]

The Commission, however, may be too quick to dismiss these concerns. Unless it leaves ample wiggle room in its rules to allow for cost considerations, the assertion that the rules do not “impose substantial new costs” could fall flat on its face. Similarly, if the rules over-reach and end up restricting certain potentially profitable practices without a proper justification, the “lost profits” from those practices could most plausibly be considered a factor in the economic impact of the regulation. While this in no way implies that the open Internet rules themselves are unconstitutional, it does demonstrate that certain enforcement policies could risk failing Constitutional scrutiny.

In order to prevent its rules from running afoul of the Constitution, the Commission will

[94] *See* Lyons, *supra* note 92, at 95.
[95] Penn Central Trans. Co. v. City of New York, 438 U.S. 104 (1978) [hereinafter Penn Central].
[96] *Id.* at 124.
[97] *See* Open Internet Order ¶ 150.

need to be very careful in tailoring its rules to achieve its goals without significantly burdening other kinds of practices. In the following sections, we will discuss two situations where these concerns might arise. Before we can do that, we must first analyze the Commission's order, to determine the goals the regulations seek to achieve.

5.2 Identifying the Legitimate Interests for Enacting Open Internet Rules

In order for the Commission's rules to survive Constitutional scrutiny, the government must show that it has legitimate or compelling interests to achieve, and that the regulations are necessary and narrowly tailored to achieve those goals. As we discussed previously[98], the clamor for net neutrality regulations began as a response to the threat of discrimination on the Internet. We outlined the means of discrimination[99], and identified several primary purposes for which a network provider could choose to adopt those discriminatory practices, including (1) congestion management, (2) QoS assurance, (3) blocking, filtering and censorship, and (4) intentional degrading of service[100]. Almost all of the purposes – except arguably the last one – could likely be used for both beneficial and harmful purposes. What exactly, though, do these harmful purposes harm? What values of the non-discriminatory Internet do we seek to protect? Based on an analysis of the concerns elaborated in the FCC's order, we can break the Commission's issues down into three categories.

A. Freedom of Speech

The FCC mentions the importance of freedom of expression several times in its Open Internet Order. The overwhelming fear is that network providers could use their discriminatory capabilities to act as gatekeepers for content on the Internet, either blocking or filtering content it does not like, or slowing it down to make it more difficult to access. During

[98] *See* discussion *supra* Section 2.5.
[99] *See* discussion *supra* Section 2.4.1
[100] *See* discussion *supra* Section 2.4.2.

his campaign for presidency, then-Senator Barack Obama embodied the essence of freedom of speech concerns in a podcast on the issue of net neutrality, stating:

> [It is] because the Internet is a neutral platform that I can put on this podcast and transmit it over the internet without having to go through some corporate media middleman. I can say what I want without censorship.[101]

Advocates for this cause would likely be concerned with the blocking, filtering and censorship aspects of network discrimination, as well as intentional degradation of service. The open Internet rules address these concerns by prohibiting the blocking of lawful content. It also bars impairing or degrading particular types of traffic so as to render them effectively unusable, holding that “degrading traffic can have the same effects as outright blocking.”[102]

B. Innovation and Competition

Apart from First Amendment concerns, much of the contemporary discourse advocating codification of restrictions against network discrimination is based on the argument that it was the non-discriminatory nature of the Internet which brought about the unprecedented level of innovation and progress that pushed the Internet to become what it is today. Advocates of this view argue that the ability of the network to discriminate risks allowing the network provider to become a gatekeeper, deciding winners and losers in the market, and inhibiting innovation in the application market. The Commission emphasizes three key ways the open, non-discriminatory nature of the network has been essential for encouraging innovation – and by extension, competition – in the content and application market:

(1) Innovation without permission – Although the Commission does not elaborate much on this feature of the open Internet, it repeats that point consistently as a main

[101] See Barack Obama, *Network Neutrality – Podcast* (2006), *available at* http://obamaspeeches.com/076-Network-Neutrality-Obama-Podcast.htm.
[102] *See* Open Internet Order ¶ 66.

purpose and justification for its rules[103]. In literature, much emphasis has been placed on the fact that you do not need permission from anyone, even the network provider, to engage in innovation on the network. This is based on the premise that "on the Internet, network providers don't have to do anything to enable new applications to run."[104] If you have a bold and crazy idea, you do not need to persuade anyone other than yourself that the idea is worthwhile to pursue. Whatever the merit of your ideas, the Internet will connect you to a worldwide audience where, in the end, it is the consumers and end-users that will ultimately decide whether your idea was good enough or not. As with freedom of speech concerns, which is basically the same concept sans innovation, the open Internet rules effectively address these concerns by prohibiting the blocking and degrading of traffic, ensuring that anyone can reach their target audience through the internet.

(2) Low barrier of entry – In addition to the ability to innovate without permission, the FCC has also consistently emphasized the importance of the Internet's low barrier of entry on fostering innovation and competition. This is premised on the idea that anyone with a connection to the network can become an innovator at a very low cost. This low barrier of entry has resulted in a highly competitive market with a huge number of application and content providers. Advocates fear that the cost of innovation could be raised by discriminatory practices, discouraging potential innovators from entering the market. One of the concerns frequently raised is that network providers could demand a fee to deliver content to their users, and either refuse to deliver traffic from non-paying application or content providers (which

[103] *See id.* ¶¶ 3, 10, 13, 78.

[104] *See* Barbara van Schewick, *Opening Statement at the FCC Workshop on Investment, Innovation and the Open Internet* (Jan. 13, 2010), *available at* http://www.law.stanford.edu/display/images/dynamic/publications_pdf/van%20Schewick%20Opening%20Statement.pdf.

would also conflict with innovation without permission), or relegate them to a slow lane with degraded service.[105] This would increase the costs of innovation, by forcing innovators to pay up in order to have decent levels of service, reducing their profits, and their incentives to innovate. The order further points out that “many [of the] new entrants are new or small ‘garage entrepreneurs,’ not large and established firms, [and] are particularly sensitive to barriers to innovation and entry, and may have difficulty obtaining financing if their offerings are subject to being blocked or disadvantaged by one or more of the major broadband providers.”[106] In addressing these concerns, the FCC not only bars the blocking and degrading of traffic, it also explicitly disfavors pay-for-priority arrangements, mentioning that it would be unlikely such arrangements would be considered reasonable discrimination. This would seem to rule out paid QoS services. It notes that “pay-for-priority arrangements could raise barriers to entry on the Internet by requiring fees from edge providers, as well as transaction costs arising from the need to reach agreements [with] broadband providers.”[107]

(3) Level playing field – While keeping the cost of market entry low is important for attracting innovators, it is the ability to innovate on an equal footing that is essential for allowing the small innovators to compete. A non-discriminatory network treats you the same no matter who you are, whether a “large and established firm” or a “garage entrepreneur”. A new entrant to the market can always innovate with the confidence that the traffic for his or her application will be treated no differently from that sent by Google, or anyone else. Applications can then compete on their merits, rather than on their relationships (commercial or otherwise) with network providers.

[105] *See* Open Internet Order ¶ 26.
[106] *See id.*
[107] *See id.* ¶ 76.

> This reinforces the Commission’s stance on viewing pay-for-priority negatively. In addition to raising a barrier for entry, pay-for-priority skews the market by unbalancing the playing field, granting fast lanes to some companies but not others. Similarly, the Commission also has significant concerns about network providers prioritizing their own traffic, or that of affiliates. While such prioritization may not raise the cost of entry for other entrants, it may nevertheless tilt the playing field for other innovators.

C. Consumer Empowerment Concerns

Throughout the order, the Commission constantly emphasizes the importance of consumer choice and end-user control. The FCC states that consumers should be able to “make their own choices about what applications and services to use and are free to decide what content they want to access, create, or share with others.”[108] This is as much about free expression as it is about promoting competition and innovation. In a way, it is the flip-side of the level playing field, whereby only through giving the consumer complete control over how they use the Internet can they make the ultimate choices as to which products and services they want to use, and by extension, which innovations succeed or fail. The Commission quotes van Schewick in concluding that “letting users choose how they want to use the network enables them to use the Internet in a way that creates more value for them (and for society) than if network providers made this choice,” and “is an important part of the mechanism that produces innovation under uncertainty.”[109] The open Internet rules attempt to preserve this value through the implementation of the transparency requirement, requiring that network providers make clear to their users the terms and limitations on their services, giving the consumer the ability to make “informed choices regarding the purchase and use of

[108] *See id.* ¶ 3.
[109] *See id.* ¶ 71, quoting Barbara van Schewick.

broadband service"[110]. At the same time, the Commission explicitly outlines both "transparency" and "end-user control" as major factors in evaluating reasonableness.[111]

Sure enough, the rules seem substantially competent to achieve their goals. However, the question at hand now is no longer simply whether they achieve those goals, but whether in accomplishing them they might possibly be restricting other conduct that does not substantially harm those goals. More specifically, in the context of a Fifth Amendment challenge, the issue will be whether the Commission's rules will restrict conduct that could also result in substantially negative economic impacts on the network provider. In the upcoming section, we study two types of discriminatory practices that may just fall into such categories: P2P throttling, and pay-for-priority.

5.3 On the Economic Impact of Restrictions on P2P Throttling

P2P throttling has been at the heart of the net neutrality debate ever since the Comcast case. Not only is a P2P case one of the only two instances of the Commission attempting to enforce open Internet principles, but the Commission also cites in its order "the blocking, slowing and degrading of P2P traffic" as primary evidence of the actual existence of discriminatory practices by network providers to limit the openness of the Internet.[112] The FCC presents this as a rebuttal to the argument that the threats to openness are "speculative" or "theoretical." However, the order does not elaborate on what the reasonable management exception entails for the management of P2P traffic. Indeed, it explicitly declines to determine "whether any of these practices violated open Internet principles," only noting that "they have raised concerns among edge providers and end users, particularly regarding lack of transparency."[113] However, could the inherent cost-structure of P2P traffic make a case for it

[110] *See id.* ¶ 53.
[111] *See id.* ¶ 6.
[112] *See id.* ¶ 35.
[113] *See id.* ¶ 36.

being treated differently from other traffic?

5.3.1 Peering and Transit: An Analysis of Bandwidth Costs

The Internet is a network of networks. However, it is virtually impossible for a network to be directly connected to all other networks on the planet. In order to be able to reach the entire Internet, a network relies on “peering” and “transit” agreements with other networks.[114] These arrangements affect the cost of bandwidth in a variety of different ways, directly influencing the network provider’s operating costs, and by extension, profitability.

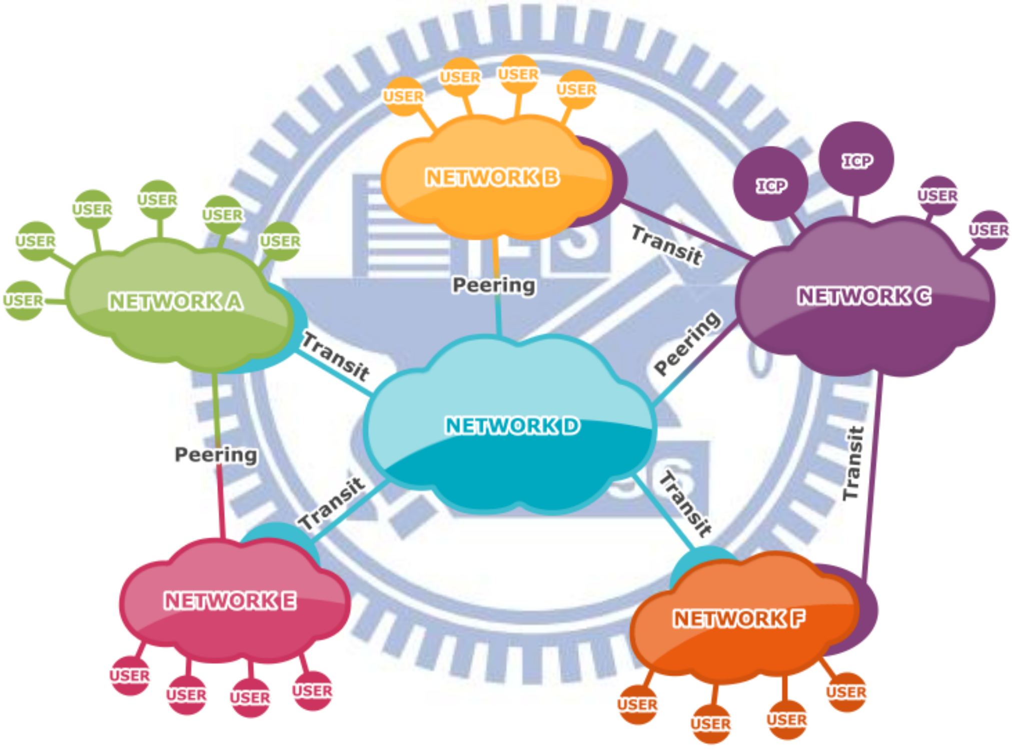


Figure 3: A Network of Networks Interconnected Through Peering and Transit.

“Transit” is when a network contracts another network to carry its traffic to and from other parts of the Internet with which it is not directly interconnected (peered). This way it does not have to be connected to every other network directly. The transit network will charge

[114] For a more detailed explanation of peering and transit, *see* Rudolf Van Der Berg, *How the 'Net Works: An Introduction to Peering and Transit*, *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1443245 (last visited Mar. 12, 2012).

a fee for this service, generally calculated based on the amount of traffic carried, or bandwidth used.[115] A network will usually have various transit providers, giving it multiple paths to access the entire Internet, likely at different costs and speeds, so it can choose a most cost effective path for every destination.[116]

"Peering," on the other hand, is when a network interconnects directly with another network to exchange traffic for mutual benefit, usually for free. Peering partners cannot usually transit through each others' networks, but traffic bound for the other network's customers can be sent to its destination without needing to pass through a third network, eliminating the need to pay transit costs.[117]

Unsurprisingly, transit comprises a good bulk of a network provider's operating costs. Networks will often attempt to reduce their transit bills by bypassing transit routes as much as possible. One way they can do this by finding suitable peering partners to interconnect with. When two networks both determine that the costs of establishing direct interconnections are lower than the costs of buying transit to each other, they'll have an economic incentive to peer. The cost savings of such an interconnection are dependent on the prices each of the networks pay for transit, so not all networks may have an incentive to peer[118].

The differences between peering and transit result in there being different costs associated with different kinds of traffic passing through the network. Traffic that stays on the network is the cheapest, since the costs of running the network are generally fixed. Once the infrastructure is in place, the marginal cost of sending an extra bit on the network is zero.[119] Meanwhile, traffic over peering connections costs slightly more, but only in the sense that

[115] *See id.* at 1.

[116] *See id.* at 3.

[117] *See id.*

[118] For an in depth look at the motivations, financial justifications and the process of peering, *see generally* William B. Norton, *A Business Case for ISP Peering*, Equinix White Paper (2002), *available at* http://arneill-py.sacramento.ca.us/ipv6mh/ABusinessCaseforISPPeering1.2.pdf

[119] *See* Van Der Berg, *supra* note 114, at 3.

additional lines and ports must be installed in order to peer. Once the peering connection has been established, however, the marginal cost of sending an extra bit is also zero.[120] Transit traffic, on the other hand, is comparatively expensive. Since transit providers usually charge based on the amount of bandwidth used or total traffic carried, every extra bit sent may represent an extra expenditure for the network. [121]

There are two important implications that stand out. The first is that bandwidth costs money. The second is that not all types of bandwidth cost the same. It is important to recognize that bandwidth is neither a free nor unlimited resource, and some forms of bandwidth usage impose very real and much higher costs on the provider.

5.3.2 The Legitimate Concerns of Bandwidth Costs

Network providers have been concerned with controlling bandwidth costs since long before the advent of net neutrality concerns. The shared nature of most consumer bandwidth – due to the practice of oversubscription[122] – means that network providers have to make accurate estimates regarding the likely usage of bandwidth on its network at any time, in order to calculate the bandwidth it needs to deploy and acquire, as well as the rates it needs to charge to be profitable[123]. When end users veer from anticipated uses of the system, they may render these estimates highly inaccurate. It is unsurprising, therefore, that network providers have historically maintained acceptable use policies restricting bandwidth overuse, or conduct such as reselling bandwidth.

[120] *See id.*

[121] *See id.*

[122] *See* Mitchell, *supra* note 39; Fred Goldstein, *The Dismal Reality of Internet Management*, TECH JOBS, Mar. 5, 2008, http://jobs.tmcnet.com/topics/broadband-comm/articles/22237-dismal-reality-internet-management.htm.

[123] The higher the over-subscription ratio, the lower the price network providers need to charge each individual consumer. This is also why commercial broadband is much more expensive than consumer broadband – they are priced at a much lower over-subscription ratio in anticipation of a higher level of regular usage. *See* Goldstein, *supra* note 122; *see also* Richard N. Clarke, *Cost of Neutral/Unmanaged IP Networks*, 8 REV. NETWORK ECON. 61 (2009) highlighting the prices a consumer would have to pay for a dedicated amount of bandwidth with zero oversubscription.

Since consumer bandwidth speeds have generally been asymmetrical, offering much downstream bandwidth than upstream bandwidth, network providers have historically been particularly sensitive about upstream usage. For example, most early contractual restrictions included prohibitions on running servers on consumer broadband Internet[124], since servers used substantially more upstream than normal. But network providers have also equally concerned with inordinate amounts of downstream usage. This manifested itself in the form of restrictions on home networking, VPNs, and even Wi-Fi[125]. While inconvenient to end-users, such restrictions were not totally illegitimate. Since early homes rarely had more than one computer, network providers had estimated bandwidth usage based on the normal usage of a single machine. As the cost of computers came down and consumers started connecting multiple machines to their networks, this wreaked havoc with the network providers' estimates.

Over the years, restrictions on home networking, VPNs, and Wi-Fi have generally been relaxed, even in the absence of regulatory intervention. The main force behind these changes have been the fact that continuing restrictions of such mainstream usage needs would not be in the best interests of the network provider. Instead, network providers adapted by changing their estimates accordingly.

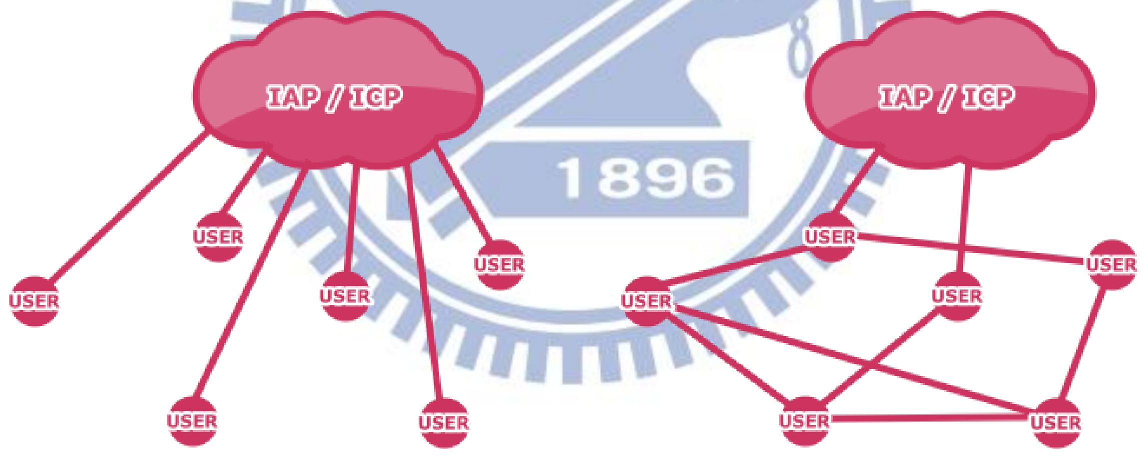
As use of P2P applications began to grow, the corresponding high-bandwidth usage attributed to the technology made it an obvious target for restricting. P2P technology was particularly problematic due to its nature of fully using both downstream *and* upstream bandwidth in a continuous fashion. Since the bandwidth was not allocated in a way that could allow for users to use fully and continuously, this development created major problems for network providers. However, one may ask how P2P is different from any other type of high

[124] *See* Wu, *supra* note 9, at 158.
[125] *See id*.

bandwidth usage of the network, and why is deserves to be singled out for discussion. To answer this, we first need to understand how P2P file sharing works.

5.3.3 The Cost-Shifting Nature of P2P Protocols

In a traditional client-to-server architecture, every end user (client) wishing to download a file from a content provider would retrieve it directly from the content provider’s source (server). With peer-to-peer technology, instead of acting as an exclusive source, the content provider may instead act as an initial “seed” for the content. Once clients around the world begin retrieving that content, those clients themselves become sources for the data as well. In other words, every client is also a server, simultaneously downloading from some peers, while uploading to others.



Client-to Server Architecture

Peer-to-Peer Architecture

Figure 4: Comparison of Traditional Client-to-Server v. Peer-to-Peer Structures.

Suddenly, instead of being constrained by the bandwidth and performance of a single server, every client has multiple sources to download from. The P2P application determines which of these available peer sources to download from by identifying the closest and fastest to transmit from. By virtue of this design, P2P applications will generally use as much of the bandwidth available to it as possible. By choosing the most efficient peers to source its data from, a client can maximize the utility of its available bandwidth. This may be greatly beneficial for increasing the efficiency of downloads, but it does so at the cost of occupying significant upstream bandwidth usage. This upstream usage is generally of neutral value to consumers, since they don't directly benefit from its use, nor suffer a significant cost from it. Indeed, since users generally pay flat rates for consumer broadband, this increased upstream bandwidth use imposes no marginal costs, giving them little reason to be concerned about it. In fact, many less-advanced P2P users do not even know they are actively uploading.[126]

This is, of course, great for content providers, who can benefit considerably by offloading their bandwidth needs onto consumers. In fact, once a content provider has seeded an entire file so every part of it is on the network, it can even hypothetically stop seeding entirely, relying on the network peers to continue act as sources of the data. Vuze, Inc., a company that runs an entertainment platform based on P2P protocols, describes the technology in these terms:

> Torrent technologies make use of resources – bandwidth, storage, and processing power – on a decentralized basis, allowing large data transfers to be made more efficiently and cost-effectively than ever before. Torrent technologies leverage the power of many individual computers by enabling each computer interested in a piece of content to obtain small pieces of it from multiple other computers, and simultaneously play the same role to others who seek the same content in the future.
>
> Accordingly, a distributor of content need not have many large central servers to store and send a file each time an Internet user is interested in a particular piece of content; instead, the content distributor need only have a handful of servers that operate as initial "seed servers" for the content, and can

[126] Many institutions have instituted programs alerting students detected to be using P2P software that their usage of such applications could lead to copyright infringement due to uploading activities. *See, e.g.*, *Iowa State University BAYU: Be Aware You're Uploading*, IOWA STATE UNIVERSITY, http://bayu.its.iastate.edu (last visited 20 Jun. 2012); *Northwestern University BAYU: Be Aware You're Uploading*, NORTHWESTERN UNIVERSITY, http://www.it.northwestern.edu/security/nubayu (last visited 20 Jun. 2012).

> then rely on the distributed computing capacity of all of the individual user computers (the “swarm”) that have that have agreed to be used as a “seed” for others.[127]

This can greatly lower the cost of distributing large files for Vuze. However, this “cost-saving” form of data distribution is not so much “cost-saving” as it is “cost-shifting.” The costs “saved” by the content provider are instead “shifted” onto consumers. Since consumers pay a flat-rate that does not reflect that cost, the cost is instead borne by the ISP. Brett Glass, founder of LARIAT, a rural non-profit telecommunications cooperative, describes this as “what would happen if a third party were to encourage customers to smuggle food out of an ‘all you can eat’ buffet.”[128] Suddenly, Ed Whitacre’s little tirade about upstarts wanting to “use [his] pipes for free” doesn’t seem so far-fetched after all.[129]

5.3.4 Cost Management as a Legitimate Network Management Purpose

In light of these features of P2P protocols, a regime that does not allow for the reduction of P2P bandwidth usage based on cost considerations would likely risk tilting the *Penn Central* scales in favor of the property owner. At the very least, the rules should probably allow for the reasonable throttling of inter-network P2P traffic to be considered a reasonable network management practice, even in the absence of congestion. On the other hand, since intra-network traffic has far lower costs, it would be conceivable that a practice of on-network P2P throttling could attract a higher level of scrutiny without Fifth Amendment concerns. Still,

[127] *See* Vuze, Inc. Petition To Establish Rules Governing Network Management Practices by Broadband Network Operators, WC Docket No. 07-52 (Nov. 14, 2007).

[128] *See* Comments of Laurence Brett Glass, D/B/A LARIAT, ET Docket No. 04-35, WC Docket No. 05-271, GN Docket Nos. 09-47, 09-51, and 09-137, filed Aug. 2, 2010, at 2.

[129] In 2005, then-SBC (now AT&T) CEO Ed Whitacre famously proclaimed: “How do you think they're going to get to customers? Through a broadband pipe. Cable companies have them. We have them. Now what they would like to do is use my pipes free, but I ain't going to let them do that because we have spent this capital and we have to have a return on it. So there's going to have to be some mechanism for these people who use these pipes to pay for the portion they're using. Why should they be allowed to use my pipes? The Internet can't be free in that sense, because we and the cable companies have made an investment and for a Google or Yahoo! or Vonage or anybody to expect to use these pipes [for] free is nuts!” Interview by Roger O. Crockett of Edward Whitacre, CEO, SBC Commc'ns, in Chicago, Il. (Nov. 6, 2005).

the issue of congestion remains a legitimate concern even for intra-network usage. For evaluating that kind of network management, one would have to return to determining the Commission's standard for "reasonableness."

In the Comcast case, there were many factors that the FCC considered in deciding to censure the company for unreasonable discrimination. The major reason was that the company failed to disclose its practices, and in fact attempted to conceal and deny it. Furthermore, Comcast's invasive management practice was a form of throttling that actively "interfered" with the protocol, rather than just passively slowing it down. There was also the issue that Comcast's throttling was continuous, without regard to the congestion level of the network.

Since the FCC declined to articulate what sort of P2P management practices would be considered acceptable, it is rather possible that it aims to leave the door open for other less intrusive forms of bandwidth management. It is also likely that a full disclosure could go a long way to justifying even slightly more brazen practices, provided the network provider provided a reason for it. For now, without any precedents to follow, network providers will undoubtedly adopt a more conservative approach.

While this may seem desirable in theory, there are certain implications that may spell bad news for other high-bandwidth applications. If network providers choose not to throttle bandwidth in order to avoid disciplinary action, their alternate solution in the long term may be to move towards usage based pricing, so that the cost of every bit is accounted for. In fact, many providers are already moving towards such a model. This trend is most evident in wireless networks, where three out of four of the nation's largest wireless broadband carriers have already moved towards a capped-bandwidth model with overage fees[130]. Fixed networks

[130] *See AT&T Data Plan - 4G LTE Data Plans with Wi-Fi & Tethering from AT&T*, AT&T, http://www.att.com/shop/wireless/plans/dataplans.html (last visited Jun 17, 2012); *Mobile Broadband Data*

have also started adopting data caps, though at comparatively higher levels.

Metered usage or bandwidth caps may in fact turn out to be the worst of all worlds for high-bandwidth applications. With the rise of Internet video services like YouTube, Hulu and Netflix, as well as cloud computing and storage solutions such as Dropbox, BitCasa, iTunes, and iCloud, consumer bandwidth usage is set to go through the roof. A capped connection could mean a user would have to be very selective as to the use of services.

One advertisement by a nation-wide carrier for its unlimited wireless data plan depicts the passive-aggressive rant of a young school-girl, whose father apparently decided to download an app rather than a video of his daughter's acting debut. "I know," the little girl gripes, "You're close to your data limit and had to choose. So my play lost out to a micro-strategy app. I don't even know what that is. Well whatever it is I hope you like it."[131] While tongue firmly in the cheek, the ad is likely an accurate depiction of what is to come if the trend towards capping continues. If that scenario plays out, the rules created to promote consumer access to applications could turn out to have done quite the opposite instead. This brings us to our next topic, which is precisely the issue of high-bandwidth applications.

5.4 On the Economic Impact of Restrictions on Pay-for-Priority

A second issue for special consideration in light of possible Fifth Amendment concerns is the Commission's seemingly blanket disapproval of all pay-for-priority arrangements. This section presents an argument against a blanket ban, focusing especially in the context of high-bandwidth applications.

Plans, T-Mobile, http://mobile-broadband.t-mobile.com/plans (last visited Jun 17, 2012); *Plans - Verizon Wireless*, Verizon Wireless, http://www.verizonwireless.com/b2c/plan-information/?page=mobileBroadband (last visited Jun 17, 2012).

[131] *Little Girl - Unlimited from Sprint*, YouTube, http://www.youtube.com/watch?v=qFp0P_Wuftc (last visited Jul. 1, 2012).

5.4.1 Innovation in High-Bandwidth Applications and Services

The primary justification for the Commission's de facto ban on pay-for-priority arrangements is that they represent a major hindrance to innovation and competition. The Commission highlights several qualities of the Internet that have been essential for innovation, which we have distilled down to three major concerns: (1) low barrier of entry, (2) level playing field, and (3) the ability to innovate without permission. The FCC believes pay-for priority threatens the low barrier of entry, by possibly requiring "fees from edge providers"[132] in order to maintain an acceptable level of service. It also threatens to unbalance the level playing field, by increasing the performance of certain edge providers but not others.

The open Internet rules, consistent with its approach of promoting end-user control, allows for the network provider to charge its customers for tiered-services, but not edge providers. However, since cash flow on the Internet generally tends to flow towards edge providers, rather than towards consumers, edge providers are much more predisposed to acquire services for money. Prohibiting network providers from extracting profits from edge providers may mean cutting off a considerably large chunk of potential revenue. This may be fine if the restriction is appropriately tailored to achieve the FCC's goals – e.g. promoting innovation and competition – but may fail to survive scrutiny in cases where the goals are not actually substantially furthered by the restrictions. It is argued here that not all types of innovations are equally sensitive to or affected by pay-for priority, and that for high-bandwidth applications in particular, the ban may not serve to substantially further innovation and competition concerns. The problem with a blanket ban on pay-for-priority is that it makes two incorrect assumptions with regards to high-bandwidth applications: (1) that the barrier of entry is low to begin with, and (2) that the playing field was fair for all innovators to begin with.

[132] *See* Open Internet Order ¶ 76.

5.4.2 The Myth of the Low Barrier of Entry

As we've previously discussed in more detail, advocates of net neutrality frequently claim that the non-discriminatory nature of the Internet resulted in an abundance of innovation by the little guys. One of the factors they claim allowed innovation to flourish was the zero-cost innovation made possible by an open Internet. It is commonly pointed out that the many of the major players in the market today were built out of dorms and garages, with minimal outside investment. Examples cited include Google, Yahoo, Facebook, Amazon, and eBay. These and many other innovative startups that have sprung up on the Internet have undoubtedly qualified as low cost innovations. However, that is only because these types of innovation used a negligible amount of bandwidth.

Indeed, the zero-cost scenario fails to take into account that bandwidth is often one of the most significant costs of a startup. Some innovations require minimal bandwidth – sometimes just a spare computer and a connection to the Internet. Others, like video, require massive amounts of bandwidth. To put things into perspective, a consumer broadband connection could easily have supported several thousand users for sites like the original Yahoo and Facebook. The same pipe would easily have maxed out for even just a couple of users watching Netflix.

High-bandwidth applications and services were never low cost innovations to begin with. YouTube, for example, was still making a loss when it sold to Google in 2006. As of 2009, YouTube was still projected by analysts to lose as much as US$470 million[133]. That's hardly

[133] *See* Spencer Wang & Kenneth Sena, *Deep Dive Into YouTube; 1Q09 Preview*, CREDIT SUISSE EQUITY RESEARCH, Apr. 3, 2009*; See also* Eric Krangel, *Analyst: YouTube Will Take Half A Billion Off Google's Bottom Line This Year*, BUSINESS INSIDER, Apr. 3, 2009, http://www.businessinsider.com/analyst-youtube-will-take-half-a-billion-off-googles-bottom-line-this-year-2009-4; Farhad Manjoo, *Do You Think Bandwidth Grows on Trees?*, SLATE MAGAZINE, Apr. 14, 2009, http://www.slate.com/articles/technology/technology/2009/04/do_you_think_bandwidth_grows_on_trees.html.

zero-cost! The fact of the matter is, some innovations were always more costly than others.[134] It is the view of this paper, that with regards to these high cost innovations, the goal of regulation should not be to keep costs low, but to make sure that pay-for-priority does not result in degraded performance for non-paying edge innovators, forcing them to pay when not previously required to. This can be achieved, for example, by mandating the reservation of a certain level of bandwidth for non-prioritized traffic, and the disclosure of peak and average bandwidth speeds and over-subscription ratios. Meanwhile, innovators that are willing to pay a premium should not be prevented from acquiring better qualities of service, for example, in the form of QoS bandwidth guarantees. As long as innovation previously possible is still possible without permission, there is no legitimate reason to prevent pay-for-priority for high-bandwidth applications.

5.4.3 The Myth of the Level Playing Field

The second incorrect assumption that a blanket ban relies on is the idea that the playing field was level for all applications. This cannot be farther from the truth. Even on the original Internet, traffic that would otherwise be the same could travel across the network at different costs, speeds, and levels of reliability. The original Internet may have been neutral from the technical perspective, but it was never neutral in the sense that all innovators could receive the same speed and quality of service. In fact, at the most basic level, purchasing different levels of bandwidth *is* in itself paying for different levels of service. An innovator who pays for a dedicated 100Mbps service clearly has a massive advantage over an innovator who only has 10Mbps on a shared bandwidth line.

[134] *But see* Alex Veytsel, *YouTube Google's Phantom Loss Leader*, RAMPRATE, Jun. 17, 2009, http://www.ramprate.com/2009/06/youtube-googles-phantom-loss-leader/ (commenting that analyst estimates are probably wrong, and that Google is only likely losing a fraction of that amount, because it has private peering for 73% of its traffic, and uses unprecedented bulk purchasing power to secure favorable wholesale bandwidth rates. While this may or may not be true, it does not change the fact that independent innovators can only dream of having such arrangements.)

Furthermore, in addition to bandwidth speeds, there have always been a variety of other ways for rich application and service providers to enhance their network performance, for example, server farms, content distribution networks, edge caches, etc.[135] As long as one has the financial means, there have always been plenty of ways to enhance performance. Paying to gain an advantage has never been a new thing.

The goal of regulation, therefore, should not be to force all innovators to stand on an equal footing – which they arguably had never been on anyway – but to make sure that anyone who wishes to pay for that priority may do so if they so choose to. In other words, the focus of maintaining the level playing field should be on the “ability to acquire” same levels of service, rather than on keeping everyone’s level of service the same.

5.4.4 Bandwidth Guarantees as a Form of Reasonable Discrimination

Not all forms of QoS affect non-prioritized traffic in the same way. Bandwidth guarantees, for example, act as a form of non-minimal discrimination. The effects of this type of discrimination on other non-guaranteed traffic are only felt when the network is at capacity. If a network has a capacity of 100Mbps, and increases capacity to 120Mbps, while reserving 20Mbps for guaranteed traffic, non-guaranteed traffic would not observe any difference with regards to bandwidth on the network. In comparison, QoS assurance in the form of a maximum delay guarantee would work by allowing certain packets to cut to the front of a queue. This means that every other packet in the queue will feel the discrimination in the form of delay, regardless of network capacity.[136]

For this reason, it is this paper’s view that open Internet rules should recognize a difference between allowing pay-for-priority in the context of bandwidth QoS and delay QoS.

[135] *See* Christopher S. Yoo, *Network Neutrality or Internet Innovation*, 33 REGULATION 22, 25 (2010).

[136] *See* discussion *supra* Section 2.4.1 Part D, comparing the effects of delay QoS on non-prioritized packets to humans waiting longer in a queue when others are allowed to cut in front of them.

While paid delay QoS should not be allowed, bandwidth Qos should not be prohibited, as long as there is an acceptable standard of bandwidth available for non-prioritized applications and content. The critical issue is to ensure a network's implementation of pay-for-priority bandwidth guarantees is not achieved by encroaching on bandwidth originally available for non-prioritized traffic, so non-minimal bandwidth discrimination would not impose a substantial disadvantage to innovators. This should be enforced by requiring full disclosure on bandwidth levels, as well as requiring the network provider to publish average and peak throughput rates for its consumers, as well as over-subscription ratios.

Finally, the network provider should offer the same pay-for priority arrangements in a non-discriminatory fashion, similar to the fashion it currently offers different speed tiers. While pay-for-priority bandwidth QoS is not a hindrance to innovation and competition on its own, offering the priority service to some, but not to others, *would* be considered a distortion of competition, and an un-leveling of the playing field. Discrimination of this sort would most likely tilt the *Penn Central* scales towards allowing regulatory intervention, in light of its harm to public interests.

There are several advantages to allowing the network provider to fund capacity expansion through pay-for-priority arrangements. While the network provider certainly could implement non-minimal reservation of bandwidth, it could also allow that bandwidth to "bleed" back into the total available bandwidth, meaning that when there is extra reserve, all applications benefit. Another advantage is that users may enjoy previously unavailable innovations that require, or work substantially better, with bandwidth guarantees.

5.5 Summary

We have discussed in this chapter two forms of practices with economic implications that the FCC should make special considerations for – one in the form of network management, and one in the form of reasonable discrimination. In determining the reasonableness of network management practices and other discriminatory practices, the Commission must make factor in the economic impact of restricting those practices, and weigh that impact against the public interest.

V. Conclusion

The original architecture of the Internet has resulted in an ecosystem that has allowed innovation and competition to flourish. Advances in technology have threatened to change the status quo, and triggered an impassioned response from stakeholders, calling on the protection of the values that made the Internet so successful. However, considering the myriad of goals and concerns surrounding the issue, and the divergent policy approaches preferred – often all legitimate in their own ways – formulating a policy on the issue is an unenviable job. The FCC's road in formulating net neutrality rules has been an arduous and thankless one, since it is virtually impossible that any FCC policy will be able to please everyone.

As soon as they were published, the FCC's open Internet rules were met with criticism by some for not doing enough to protect consumers and innovators. At the same time, it has been denounced by others for "interventionist overreach,"[137] and erecting barriers to infrastructure investment and innovation. It is my view that the FCC has actually done a relatively good job of allaying network discrimination concerns, while at the same time providing enough flexibility for network providers to effectively manage the networks. It provides a good platform to move forward upon, while keeping options open for future changes.

However, in light of the cloud of uncertainty over the actual legality of the rules, the FCC will have to be extra cautious in pushing its agenda. Not only is the enforceability of the substantive rules in question, the way they are applied may also be cause for Constitutional concern. As we've seen, there are certain situations where the Commission may find it not only sensible, but also necessary – for legal purposes – to grant greater flexibility in its rules.

[137] *See* Ryan Singel, *FCC Net Neutrality Rules Slammed From All Sides*, WIRED BUSINESS (Dec. 20, 2010), http://www.wired.com/business/2010/12/fcc-rule/

The FCC must ensure its enforcement regime does not overly burden network providers by restricting them from certain cost-controlling or profit-seeking practices when they do not substantially affect the openness of the Internet.

Furthermore, the Commission should be especially vigilant as to whether its limitations may push network providers towards business models that could ultimately be worse for innovation. It might even be wise for the Commission to encourage profit seeking practices, insofar as a non-explicit approval alone might not be enough to allay network provider concerns of being censured. The permitting of these cost-controlling and profit-seeking practices would still need to be premised on the principles of the open Internet – that is, practices that do substantially harm public interests could still be allowed – likely through the requirement of full disclosure, equal treatment, and minimum service standards.

References

1. Books

(1) Benjamin, Stuart Minor, et. al., *Telecommunications Law and Policy*, Carolina Academic Press, 2 ed. 2006.

(2) Lessig, Lawrence, *The Future of Ideas*, Random House, 2002.

(3) Marsden, Christopher T., *Net Neutrality: Towards a Co-Regulatory Solution*, Bloomsbury Academic, 2010.

(4) Van Schewick, Barbara, *Internet Architecture and Innovation*, MIT Press, 2010.

2. Journals and Essays

(1) Atkinson, Robert D. & Phil Weiser, "*A 'Third Way' on Network Neutrality*", 13 The New Atlantis 47 (2006)

(2) Bonner, Elizabeth Austin, "*Network Neutrality Disclosures: More and Less Information*", 8 I/S: J. L. & Pol'y for Info. Soc'y 179 (2012)

(3) Chong, Rachelle B., "*The 31 Flavors of Net Neutrality: A Policymaker's View*", 12 Intell. Prop. L. Bull. 147 (2008).

(4) Downes, Larry, "*Unscrambling the FCC's Net Neutrality Order: Preserving the Open Internet – But Which One?*", 20 CommLaw Conspectus 83 (2011).

(5) Felten, Edward W., "*Nuts and Bolts of Network Neutrality*" (Aug. 2006), available at http://itpolicy.princeton.edu/pub/neutrality.pdf.

(6) Lemley, Mark A. & Lawrence Lessig, "*The End of End-to-End: Preserving the Architecture of the Internet in the Broadband Era*", 48 UCLA L. Rev. 925 (Oct. 1, 2000)

(7) Lyons, Daniel, "*Virtual Takings, The Coming Fifth Amendment Challenge to Net Neutrality Regulations*", 86 Notre Dame L. Rev. 65 (2011).

(8) Peha, Jon M., "*The Benefits and Risks of Mandating Network Neutrality, and the Quest*

for a Balanced Policy", 1 Int'l J. Comm. 644 (2007).

(9) Renda, Andrea, *"I own the pipes, you call the tune. The net neutrality debate and its (ir) relevance for Europe"*, Center for European Policy Studies (2008), available at http://works.bepress.com/andrea_renda/1.

(10) Roberts, Christopher E., "*Can I Still Google My Yahoo - Reframing the Net Neutrality Debate - Why Legislation Actually Means Deregulation"*, 77 UMKC L. Rev. 765 (2008).

(11) Ruane, Kathleen Ann, *"The FCC's Authority to Regulate Net Neutrality After Comcast v. FCC"*, Congressional Research Service, October 27, 2011, available at http://www.fas.org/sgp/crs/misc/R40234.pdf

(12) Smith, Courtney Erin, "*Net Neutrality, Full Throttle: Regulation of Broadband Internet Service Following the Comcast/Bittorrent Dispute"*, 50 Santa Clara L. Rev. 569 (2010).

(13) Van Der Berg, Rudolf, "*How the 'Net Works: An Introduction to Peering and Transit"*, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1443245 (last visited Mar. 12, 2012)

(14) Van Schewick, Barbara, "*Towards an Economic Framework for Network Neutrality Regulation"*, 5 J. On Telecomm. & High Tech. L. 329 (2007)

(15) Van Schewick, Barbara, "*Network Neutrality: What a Non-Discrimination Rule Should Look Like"* (September 20, 2010). Stanford Public Law Working Paper No. 1684677; Stanford Law and Economics Olin Working Paper No. 402. Available at SSRN: http://ssrn.com/abstract=1684677

(16) Wu, Tim, "*Why Have a Telecommunications Law? Anti-Discrimination Norms in Communications"*, 5 J. On Telecomm. & High Tech. L. 15, 28-35 (2006)

(17) Wu, Tim, "*Network Neutrality, Broadband Discrimination"*, 2 J. On Telecomm. & High Tech. L. 141 (2003)

(18) Wu, Tim, "*The Broadband Debate: A User's Guide"*, 3 J. On Telecomm. & High Tech. L. 69 (2004).

(19) Wu, Tim & Christopher S. Yoo, "*Keeping the Internet Neutral?: Tim Wu and Christopher Yoo Debate*", 59 Fed. Comm. L. J. 575 (2007).

(20) Weiss, Philip F., "*Protecting a Right To Access Internet Content: The Feasibility of Judicial Enforcement in a Non-neutral Network*", 77 Brook. L. Rev. 383 (Fall 2011).

(21) Yoo, Christopher S., "*Network neutrality and the economics of congestion*", 94 Geo. L. J. 1847 (2006).

(22) Yoo, Christopher S., "*Would Mandating Broadband Network Neutrality Help or Hurt Competition? A Comment on the End-to-End Debate*", 3 J. on Telecomm. & High Tech. L. 23 (2004)

(23) Yoo, Christopher S., "*Network neutrality, consumers, and innovation*", 25 U. Chi. Legal F. 179 (2008).

(24) Yoo, Christopher S., "*Network neutrality after Comcast: Toward a case-by-case approach to reasonable network management*" (February 1, 2009), NEW DIRECTIONS IN COMMUNICATIONS POLICY, Randolph J. May, ed., pp. 55-84, Carolina Academic Press, 2009.

(25) Yoo, Christopher S., "*Beyond network neutrality*", 19 Harv. JL & Tech. 1 (2005).

(26) Yoo, Christopher S., "*Innovations in the Internet's architecture that challenge the status quo*", 8 J. on Telecomm. & High Tech. L. 79 (2005).

(27) Yoo, Christopher S., "*Network Neutrality or Internet Innovation*", 33 Regulation 22 (Spring 2010).

3. Websites

(1) Net Neutrality – Cybetelecom, Federal Internet Law & Policy: http://www.cybertelecom.org/ci/neutral.htm