Notre Dame Law Review

Volume 79 | Issue 3 Article 1

4-1-2004

The Layers Principle: Internet Achitecture and the Law

Lawrence B. Solum

Minn Chung

ARTICLES

THE LAYERS PRINCIPLE: INTERNET ARCHITECTURE AND THE LAW

*Lawrence B. Solum**
Minn Chung†

INTRODUCTION

In this Article, we address the fundamental questions of Internet governance: whether and how the architecture of the Internet should affect the shape and content of legal regulation of the global network of networks. Our answer to these questions is based on the concept of *layers*, the fundamental architectural feature of the Internet. Our thesis is that legal regulation of the Internet should be governed by the *layers principle*: the law should respect the integrity of layered Internet architecture. In this introductory section, we provide a rough and ready introduction to the layers. We then preview the case for our thesis and situate it in relation to prior work on the question of the relationship between architecture and regulation. We conclude our introduction with a brief reader's guide to our Article. The ideas that we preview in the Introduction will be developed in much greater depth in the main body of the Article. We front-load the basic ideas in order to introduce a working conceptual vocabulary and to provide the reader with a roadmap to the architectonic of our argument.

* Professor of Law, University of San Diego School of Law.

† Minn Chung Consulting. We owe thanks to Randy Barnett, Vinton Cerf, Edward Felten, Michael Geist, Paul Geller, Greg Hamer, and Karl Manheim for their comments on an earlier draft of this Article. We also owe thanks to an anonymous reviewer and to the participants at faculty workshops at Boston University on October 2, 2003, and at the University of San Diego on November 20, 2002.

A. *The Layers of the Internet*

The Internet is a global network of networks that has been the platform for revolutionary innovation.[1] The role of the Internet in enabling innovation is not accidental; rather it flows from the Internet's architecture. The key innovation-enabling feature of Internet architecture is comprised of layers, narrowly understood as defined by code or broadly understood as functional components of a communications system.

What are the layers of the Internet? Viewed as a system of communication between users, the six layers that constitute the Internet are:

The Content Layer: The symbols and images that are communicated;

The Application Layer: The programs that use the Internet, e.g., the Web;

The Transport Layer: TCP, which breaks the data into packets;

The Internet Protocol Layer: IP, which handles the flow of data over the network;

The Link Layer: The interface between users' computers and the physical layer; and

The Physical Layer: The copper wire, optical cable, satellite links, etc. We flesh out this skeletal description in greater detail below.[2]

The layers are organized in a vertical hierarchy. When information is communicated via the Internet, the information flows down from the content layer (the "highest" level) through the application, transport, IP, and link layers to the physical layer (the "lowest" level); across the physical layer in packets; and then flows back up through the same layers in reverse order. Communication on the Internet requires that content be digitalized by an application, and that the digital information be broken into packets by the transport layer and addressed by the Internet protocol layer so that it can be passed on by the link layer to the physical layer. Having reached the bottom layer, information then moves horizontally. The physical layer transmits the individual data packets by copper, fiber, and/or radio by various way-

1 *See* MANUEL CASTELLS, THE INTERNET GALAXY: REFLECTIONS ON THE INTERNET, BUSINESS, AND SOCIETY 2–5 (2001); MANUEL CASTELLS, THE RISE OF THE NETWORK SOCIETY 2 (2d ed. 2000); MICHAEL L. DERTOUZOS, WHAT WILL BE: HOW THE NEW WORLD OF INFORMATION WILL CHANGE OUR LIVES 17, 25–54 (1998); PHILIP EVANS & THOMAS S. WURSTER, BLOWN TO BITS: HOW THE NEW ECONOMICS OF INFORMATION TRANSFORMS STRATEGY *passim* (2000); MICHAEL HAUBEN & RONDA HAUBEN, NETIZENS: ON THE HISTORY AND IMPACT OF USENET AND THE INTERNET *passim* (1997); LAWRENCE LESSIG, THE FUTURE OF IDEAS: THE FATE OF THE COMMONS IN A CONNECTED WORLD 5–23 (2001).

2 *See infra* Introduction B, C.

points to an endpoint or destination on the network. Once at its destination, the information then ascends vertically through the layers to be interpreted by an application as content.[3] In a nutshell, the fundamental architecture of the Internet is layered.

B. The Layers Principle

Once the layered nature of Internet architecture is understood, normative implications follow. Our thesis is that the design of legal rules should respect a fundamental principle of Internet architecture, which we shall call the *layers principle*.[4] At this stage, we can roughly formulate the layers principle as a rule of thumb for Internet regulators: respect the integrity of the layers.[5] This fundamental principle has two corollaries. The first corollary is the *principle of layer separation*: Internet regulation should not violate or compromise the separation between layers designed into the basic architecture of the Internet.[6] The second corollary is the *principle of minimizing layer crossing*: minimize the distance between the layer at which the law aims to produce an effect and the layer directly affected by legal regulation.[7] The second corollary has obvious implications for the evaluation of Internet

3 The brief sketch elides some important distinctions between these layers. All of the layers except the physical layer are code or software. Two of the layers, the transport layer and Internet protocol layer, are part of TCP/IP, the communications protocol that is central to the Internet. One of those, the Internet protocol layer, is sometimes said to be the "network" layer and, in a narrow sense, might be called the layer that defines the Internet. We discuss these issues more fully below. *See infra* Part I.

4 The layers principle is related to the end-to-end principle articulated by Larry Lessig and others. We discuss end-to-end *infra* Introduction D; Part I.A. An approach to layers that shares our emphasis on their fundamental importance but focuses on a different set of issues is found in Kevin Werbach, *A Layered Model for Internet Policy*, 1 J. TELECOMM. & HIGH TECH. L. 37 (2002). A sense of Werbach's overall argument is provided by the following passage from his conclusion:

> The layered model addresses . . . the shortcomings of the current structure in the age of the Internet. Focusing on vertical layers removes the assumption that service boundaries are clear, and are tied to physical network boundaries. It implies a more granular analysis within each layer, moving from overarching policy goals to specific cases rather than applying categories that bring with them laundry lists of requirements. It brings the issues of interconnection between networks, and between functional layers within those networks, to the forefront. And it recognizes the significance of network architecture as a determining factor in shaping business dynamics.

Id. at 67.

5 *See infra* Part II.A.1, 5.

6 *See infra* Part II.A.2.a.

7 *See infra* Part II.A.2.b.

regulations. The best regulations attack a problem at a given layer with a regulation at that layer. The worst regulations attack a problem at the content layer by imposing a regulation at the physical layer—or vice versa.

C. *The Case for the Layers Principle*

The aim of this Article is to state the case for the layers principle and its corollaries. The arguments that comprise our case will proceed in stages that include both a technical analysis of the architecture of the Internet and a policy analysis of a variety of particular problems of Internet regulation. Thus, our argumentative strategy has two dimensions. The first dimension might be described as *top-down.* We start with facts about the Internet and construct a relatively abstract and general argument: given certain widely accepted normative premises, the nature of the Internet requires the layers principle and its corollaries that follow. The second dimension of the argument might be described as *bottom-up.* We consider a variety of particular problems in Internet regulation. We show that the layers principle handles these problems in a way that coheres with reasonable judgments about the appropriate way to handle the particular problems. The bottom-up argument runs from specific examples to the same general policy recommendations generated by the top-down arguments. The convergence of the two lines of argument provides strong evidence that the layers principle provides a sound basis for Internet policy.

Although the full argument for the layers principle defies easy summary, two themes run throughout the argument. The first theme is based on the familiar idea of fit between the ends of regulation and the means employed to achieve those ends. The second theme is based on the idea that the transparency of the Internet is an essential prerequisite for low cost innovation. We will investigate each of these themes in some detail below. In these introductory remarks, we offer a brief preview of each of the two themes.

The first theme is grounded in a fact that arises from the layered nature of the Internet: regulations that violate layering (or cross layers) inherently interfere with substantial innocent uses of the Internet. The *fit thesis* is the claim that layer-crossing regulations inherently produce problems of fit between regulatory ends and regulatory means.[8] That is, if a regulation attacks a problem at one layer with a regulation that directly impacts a different layer, the nature of the Internet guarantees that the regulation will suffer from problems

8 *See infra* Part II.D.1.

of overbreadth (impacting substantial innocent uses) and underinclusion (failing to eliminate the targeted harm). This thesis is most intuitively and accessibly illustrated by regulations that aim to achieve a result at the content layer by regulating at the physical layer. For example, a nation-state might attempt to address the problem of pornography on the Internet by severing the physical link between that nation's Internet backbone and the rest of the global network of networks.[9] By severing the physical link, the flow of pornographic content might be reduced,[10] but the consequence is that a wide variety of innocent content, ranging from baseball scores to scholarly papers, is interdicted as well. We demonstrate that the stunning overbreadth of layer-crossing regulation illustrated by this example flows inevitably from the design of the Internet. Despite its overbreadth, this regulation will also be underinclusive: even if a nation's links to the Internet backbone are severed, the nature of the Internet makes it highly likely that other routes (using the international telephone system to gain dial-up access, satellite links, etc.) will allow users to reach the forbidden content.

The second theme is based on the idea that the transparency of the Internet enables low cost innovation. We shall demonstrate that the transparency of the Internet is a product of layer separation. Substantial or systematic violations of the layers principle compromise the transparency of the Internet and increase the cost of innovation. The *transparency thesis* is the claim that layer-violating regulations inherently damage the transparency of the Internet.[11] We illustrate this thesis with the example of Tim Berners-Lee's development of the World Wide Web.[12] Without transparency, development of the World Wide Web would have been substantially more expensive. At worst, neither the Web nor a functional equivalent would have been developed; at best, the development of the Web would have been substantially delayed.

We show that the fit thesis and the transparency thesis combine to provide a compelling justification for the layers principle and its corollaries. Regulations that fail to respect the integrity of the layers preclude innocent uses of the Internet, cannot achieve their regula-

9 This abstract hypothetical has actual counterparts. *See infra* Part III.

10 For purposes of this Article, we do not discuss the question whether severing backbone links would effectively cut Internet users in a particular geographic region from the global Internet. Wire line telephony might, for example, be used to create narrowband pipelines; many such pipelines might be capable of substantial throughput. In this Article, we simply bracket these issues.

11 *See infra* Part I.B. and Part II.D.1.

12 *See infra* note 75 and accompanying text.

tory goals, and threaten the transparency of the Internet and consequently its ability to serve as the platform for innovation. We then show that these abstract arguments work in the context of specific examples of layer-violating Internet regulations.

D. Situating Layers Analysis

The layers principle and its corollaries synthesize two related notions. The first notion is drawn from software engineering and design. The Internet's architecture has been engineered through the use of layers—paradigmatically, the layers of Transfer Control Protocol/Internet Protocol (TCP/IP). Respecting the integrity of the layers is a fundamental principle of Internet design. The second notion is drawn from the work of Yochai Benkler on the regulation of communications; Benkler's analysis extends, generalizes, and abstracts the notion of layers—enabling the conceptualization of *content* as a distinct layer of the Internet viewed as a communications system.[13]

In addition, *layers analysis* builds upon and extends two fundamental insights that have been forcefully and eloquently presented in the work of Lawrence Lessig.[14] The first insight can be called the *code thesis*: the notion that the architecture of the Internet has profound implications for its legal regulation.[15] The second insight can be called the *end-to-end principle*, an idea from network engineering that Lessig applies to Internet regulatory policy.[16] Lessig argues that the end-to-end principle captures the key feature of the Internet architecture that enables the Internet to become an engine of innovation.[17] In this Article, we shall argue, *pace* Lessig, that the end-to-end principle does not fully and accurately capture the fundamental relationship between Internet architecture and sound or optimal regulation of the Internet. Layers analysis reconceptualizes the end-to-end principle, yielding a richer and more accurate model of the fundamental architecture of the Internet. The layers principle restates the normative implications of that analysis as a set of principles suitable for use by Internet policymakers. We argue that the layers principle captures

13 Yochai Benkler, *From Consumers to Users: Shifting the Deeper Structures of Regulation Toward Sustainable Common and User Access*, 52 FED. COMM. L.J. 561, 562–63 (2000). Benkler identifies three fundamental layers, which he calls "the physical infrastructure, logical infrastructure, and content layers." *Id.* at 568.

14 In addition to the sources cited below, *see supra* notes 15–17, see Lawrence Lessig, *The Architecture of Innovation*, 51 DUKE L.J. 1783 (2002); Lawrence Lessig, *The Law of the Horse: What Cyberlaw Might Teach*, 113 HARV. L. REV. 501 (1999).

15 *See* LAWRENCE LESSIG, CODE AND OTHER LAWS OF CYBERSPACE (1999).

16 LESSIG, *supra* note 1, at 34.

17 *See id.*

all of the content of the end-to-end principle, but that the layers principle does more, providing guidance for regulators where the end-to-end principle is silent or indeterminate. That is, the normative content of the layers principle is a superset of the normative content of the end-to-end principle.

The Internet is a global network of interconnected computer networks, and TCP/IP is the network communication protocol that enables the Internet to function as a network of networks.[18] Expressing the point more forcefully, without TCP/IP (or a functional substitute), there would not be an Internet.[19] Despite the critical and essential role that TCP/IP plays in the design and the functioning of the Internet, however, analysis of the architecture of the TCP/IP protocol and its implications is conspicuously missing in nearly all previous discussions of the interplay of Internet architecture and legal regulation of the Internet.[20] This Article ameliorates that deficiency. Our analy-

---

18 1 W. RICHARD STEVENS, TCP/IP ILLUSTRATED: THE PROTOCOLS 1 (1994).

19 Strictly speaking, we should say, "If there were no TCP/IP or functional counterpart, there would be no Internet or functionally equivalent communications system." The particular communications protocol TCP/IP could be replaced by a functional equivalent. Thus, TCP/IP is not, strictly speaking, necessary to the Internet; thus, the claim in text is simplified for ease of communication. The germ of truth in this textual claim is illustrated by the State of California's definition of the Internet, which makes direct reference to TCP/IP:

> "Internet" means the global information system that is logically linked together by a globally unique address space based on the Internet Protocol (IP), or its subsequent extensions, and that is able to support communications using the Transmission Control Protocol/Internet Protocol (TCP/IP) suite, or its subsequent extensions, or other IP-compatible protocols, and that provides, uses, or makes accessible, either publicly or privately, high level services layered on the communications and related infrastructure described in this paragraph.

CAL. BUS. & PROF. CODE § 17538(e)(6) (West Supp. 2002). For further discussion, see *infra* text accompanying notes 87–88.

20 *See, e.g.*, Johannes M. Bauer & Steven S. Wildman, *Rethinking Access: Introduction to the Symposium Theme and Framework*, 2002 L. REV. MICH. ST. U. DET. C.L. 605, 609 (2002) (discussing TCP as a feature of Internet architecture in the context of the digitalization of communications infrastructure); Viktor Mayer-Schönberger, *Impeach the Internet!*, 46 LOY. L. REV. 569, 580–81 (2000) (discussing TCP/IP as a feature of Internet architecture and the idea that code regulates the Internet); David McGowan, *Legal Implications of Open-Source Software*, 2001 U. ILL. L. REV. 241, 283 (discussing TCP/IP as a feature of Internet architecture in the context of open-source software); Dawn C. Nunziato, *Exit, Voice, and Values on the Net*, 15 BERKELEY TECH. L.J. 753, 755–56 (2000) (discussing TCP/IP as feature of Internet architecture and the idea that code is the most effective regulator of the Internet); Philip R. Principe, *Secret Codes, Military Hospitals, and the Law of Armed Conflict: Could Military Medical Facilities' Use of Encrypted Communications Subject Them to Attack Under International Law?*, 24 U.

sis will show that layers are the key architectural feature of TCP/IP and, hence, of the Internet.[21] And we show that this fact about Internet architecture yields powerful normative conclusions about Internet regulation.[22]

In the context of legal doctrine, layers analysis contributes vital information to familiar modes of legal analysis. First, in constitutional law, the idea of fit plays a key role in equal protection and free speech doctrine. The normative end of fit is facilitated by the familiar doctrinal concepts of underinclusivenss and overbreadth. Layers analysis provides the analytic tools needed to identify and explain overbreadth and underbreadth in the context of internet regulation. Second, layers analysis can identify a crucial policy consideration that traditional overbreadth analysis would miss: the preservation of transparency that is necessary for internet innovation. Layers analysis explains why certain regulations incur specific costs and benefits—which would otherwise be opaque to policymakers. Finally, by helping to identify the likely consequences of proposed Internet regulations, layers analysis can also be used to identify their affects on individual liberties, a consideration relevant to some rights-based legal theories. In sum, without the information provided by an analysis of layers, much that is of importance to the constitutionality, prudence, and justice of proposed Internet regulations may either be neglected or missed altogether.

E. Structure of the Article

We return, in Part I, to the layers model of Internet architecture, examining in detail the layers of the Internet, both in the narrow context of TCP/IP and in the broader context of the whole communications system that constitutes the Internet. Part II then explicates the layers principle and situates it in a broader jurisprudential context. Part III applies the layers principle to a variety of particular problems in Internet regulation. In this Part, we provide a detailed discussion of several real or hypothetical layer-violating or layer-crossing regula-

---

ARK. LITTLE ROCK L. REV. 727, 731 (2002) (discussing TCP/IP as a feature of Internet architecture in the context of the history of the Internet); Philip J. Weiser, *The Internet, Innovation, and Intellectual Property Policy*, 103 COLUM. L. REV. 534, 541–42 (2003) (discussing TCP/IP as a feature of Internet architecture in the context of explaining how the Internet works); Dina I. Oddis, Note, *Combating Child Pornography on the Internet: The Council of Europe's Convention on Cybercrime*, 16 TEMP. INT'L & COMP. L.J. 477, 483 (2002) (discussing TCP/IP as a feature of Internet architecture in the context of the history of the Internet).

21 *See infra* Part I.

22 *See infra* Part II.

tions, including: (1) the Serbian Internet Interdiction Myth,[23] (2) Myanmar's "cut the wire" policy,[24] (3) China's Great Firewall,[25] (4) the French *Yahoo!* case,[26] (5) cyberterrorism,[27] (6) Pennsylvania's IP address blocking child pornography statute,[28] (7) port blocking and peer-to-peer file sharing,[29] and (8) the regulation of streaming video at the IP layer.[30] The final Part presents our conclusions.[31]

I. THE LAYERS MODEL OF INTERNET ARCHITECTURE

A. *The Code Thesis and the End-to-End Principle*

What is the relationship between layers analysis and the growing body of scholarship that argues for (or at least assumes) the proposition that an understanding of Internet architecture is essential to sound Internet regulation? The relationship between Internet architecture and the law has been discussed in the context of Internet regulation generally,[32] and with respect to Internet governance issues,[33]

---

23 *See infra* Part III.A.2.a.

24 *See infra* Part III.A.2.b.

25 *See infra* Part III.A.3.b.

26 *See infra* Part III.A.3.c.

27 *See infra* Part III.A.3.d.

28 *See infra* Part III.A.3.e.

29 *See infra* Part III.A.4.a.

30 *See infra* Part III.A.5.

31 *See infra* Conclusion.

32 *See, e.g.*, Julian Epstein, *A Lite Touch on Broadband: Achieving the Optimal Regulatory Efficiency in the Internet Broadband Market*, 38 HARV. J. ON LEGIS. 37, 42–43 (2001) (discussing Internet architecture in the context of broadband regulation); Barbara Esbin, *Internet over Cable: Defining the Future in Terms of the Past*, 7 COMMLAW CONSPECTUS 37, 49–57 (1999) (discussing Internet architecture in the context of regulation of cable access to the Internet); Harold Feld, *Whose Line Is It Anyway? The First Amendment and Cable Open Access*, 8 COMMLAW CONSPECTUS 23, 39–40 (2000) (discussing Internet architecture in the context of cable and broadband access regulation); Llewellyn Joseph Gibbons, *No Regulation, Government Regulation, or Self-Regulation: Social Enforcement or Social Contracting for Governance in Cyberspace*, 6 CORNELL J.L. & PUB. POL'Y 475, 487–89 (1997) (discussing Internet architecture in the context of Internet governance); Jack L. Goldsmith, *The Internet and the Abiding Significance of Territorial Sovereignty*, 5 IND. J. GLOBAL LEGAL STUD. 475, 477–90 (1998) (discussing Internet architecture in an analysis of the relevance of territorial sovereignty to regulation of the Internet); Jay P. Kesan & Rajiv C. Shah, *Fool Us Once Shame on You—Fool Us Twice Shame on Us: What We Can Learn from the Privatizations of the Internet Backbone Network and the Domain Name System*, 79 WASH. U. L.Q. 89, 130–41 (2001) (discussing Internet architecture in the context of privatization of Internet functions); Edward Lee, *Rules and Standards for Cyberspace*, 77 NOTRE DAME L. REV. 1275, 1322–28 (2002) (discussing regulation of Internet architecture); Mark A. Lemley & Lawrence Lessig, *Open Access to Cable Modems*, 22 WHITTIER L. REV. 3, 4–34 (2000) (discussing Internet architecture

including the relationship between Internet governance and democratic control of policy[34] and content.[35] In addition, the role of Internet architecture has been discussed in connection with a variety of more particular issues and fields, including antitrust law,[36] civil procedure (e.g., notice[37] and jurisdiction[38]), commercial law,[39] constitu-

---

in the context of broadband and cable regulation); Lewis E. Schnurr, *Media and Telecommunications Regulation and the Internet: Regulate or Strangulate?*, 8 SPG MEDIA L. & POL'Y 11, 12–14 (2000) (discussing Internet architecture in the context of regulatory policy); James B. Speta, *A Common Carrier Approach to "Internet Interconnection,"* 54 FED. COMM. L.J. 225, 243–47 (2002) (discussing Internet architecture in the context of regulation of the Internet backbone); Jonathan Weinberg, *The Internet and Telecommunications Services, Universal Service Mechanisms, Access Charges, and Other Flotsam of the Regulatory System*, 16 YALE J. ON REG. 211, 215–17 (1999) (discussing Internet architecture in the context of universal service issues in telecommunications regulation); Timothy Wu, *Application-Centered Internet Analysis*, 85 VA. L. REV. 1163, 1164 (1999) (proposing an analysis of Internet regulation focusing on the application layer of the Internet's architecture). *But see* Glen O. Robinson, *On Refusing to Deal with Rivals*, 87 CORNELL L. REV. 1177, 1227 n.212 (2002) ("[T]he analysis of the access issue gains no purchase from abstract theorizing about the appropriate architecture of Internet technology. Framing the question as one of Internet network architecture only diverts attention from the issues of economic policy by implying that some unique technological principle is driving the analysis.").

33 *See, e.g.*, Lyombe Eko, *Many Spiders, One Worldwide Web: Towards a Typology of Internet Regulation*, 6 COMM. L. & POL'Y 445, 453 (2001) (discussing Internet architecture in the context of institutions regulating the Internet); Philip J. Weiser, *Internet Governance, Standard Setting, and Self-Regulation*, 28 N. KY. L. REV. 822, 829–35 (2001) (discussing Internet architecture in the context of Internet governance).

34 Joel R. Reidenberg, *Yahoo and Democracy on the Internet*, 42 JURIMETRICS J. 261, 271–75 (2002) (discussing democratic control and Internet architecture).

35 *See, e.g.*, Julien Mailland, Note, *Freedom of Speech, the Internet, and the Costs of Control: The French Example*, 33 N.Y.U. J. INT'L L. & POL. 1179, 1198 (2001) (discussing Internet architecture's effect on French regulation of content).

36 *See, e.g.*, Christopher S. Yoo, *Vertical Integration and Media Regulation in the New Economy*, 19 YALE J. ON REG. 171, 269–71 (2002) (discussing Internet architecture in the context of antitrust regulation).

37 *See, e.g.*, Shaun B. Spencer, *Cyberslapp Suits and John Doe Subpoenas: Balancing Anonymity and Accountability in Cyberspace*, 19 J. MARSHALL J. COMPUTER & INFO. L. 493, 519 (2001) (discussing Internet architecture in the context of identifiability of litigants).

38 *See, e.g.*, Sanjay S. Mody, Note, *National Cyberspace Regulation: Unbundling the Concept of Jurisdiction*, 37 STAN. J. INT'L L. 365, 368–69 (2001) (discussing Internet architecture in the context of jurisdiction).

39 *See, e.g.*, Michael Lee et al., *Electronic Commerce, Hackers, and the Search for Legitimacy: A Regulatory Proposal*, 14 BERKELEY TECH. L.J. 839, 879–86 (1999) (discussing Internet architecture in the context of commercial transaction security).

tional law (e.g., the dormant Commerce Clause[40] and freedom of speech[41]), copyright law,[42] disability rights,[43] domain name and trademark issues,[44] gambling regulation,[45] privacy,[46] race and law,[47] securi-

40 *See, e.g.*, Jack L. Goldsmith & Alan O. Sykes, *The Internet and the Dormant Commerce Clause*, 110 YALE L.J. 785, 816, 826 (2001) (discussing Internet architecture in the context of the dormant Commerce Clause).

41 *See, e.g.*, Raymond Shih Ray Ku, *Open Internet Access and Freedom of Speech: A First Amendment Catch-22*, 75 TUL. L. REV. 87, 93–101 (2000) (discussing Internet architecture in the context of freedom of speech); Lawrence Lessig, *What Things Regulate Speech: CDA 2.0 vs. Filtering*, 38 JURIMETRICS J. 629, 670 (1998) (discussing Internet architecture in the context of freedom of speech); Lawrence Lessig & Paul Resnick, *Zoning Speech on the Internet: A Legal and Technical Model*, 98 MICH. L. REV. 395, 404–31 (1999) (discussing Internet architecture and freedom of speech); Alex C. McDonald, *Dissemination of Harmful Matter to Minors over the Internet*, 12 SETON HALL CONST. L.J. 163, 165–69 (2001) (discussing Internet architecture in the context of protection of children from harmful content); Karl J. Sandstrom & Janis Crum, *Is the Internet a Safe Haven for Corporate Political Speech?* Austin v. Michigan Chamber of Commerce *in the Shadow of* Reno v. ACLU, 8 COMMLAW CONSPECTUS 193, 194–97 (2000) (discussing Internet architecture in the context of regulation of corporate speech); Alexander Tsesis, *Hate in Cyberspace: Regulating Hate Speech on the Internet*, 38 SAN DIEGO L. REV. 817, 826–32 (2001) (discussing Internet architecture in the context of regulation of hate speech); Timothy Zick, *Congress, the Internet, and the Intractable Pornography Problem: The Child Online Protection Act of 1998*, 32 CREIGHTON L. REV. 1147, 1153, 1201 (1999) (discussing Internet architecture in the context of freedom of speech and regulation of child pornography).

42 *See, e.g.*, I. Trotter Hardy, *Computer Ram "Copies": Hit or Myth? Historical Perspectives on Caching as a Microcosm of Current Copyright Concerns*, 22 U. DAYTON L. REV. 423, 430, 461 (1997) (discussing the relationship of Internet architecture to caching practices); Henry H. Perritt, Jr., *Property and Innovation in the Global Information Infrastructure*, 1996 U. CHI. LEGAL F. 261, 286–93 (discussing the impact of Internet architecture on the protection of intellectual property); Mathias Strasser, *Beyond Napster: How the Law Might Respond to a Changing Internet Architecture*, 28 N. KY. L. REV. 660 *passim* (2001) (discussing Internet architecture and copyright protection of digital content); Richard S. Vermut, *File Caching on the Internet: Technical Infringement or Safeguard for Efficient Network Operation?*, 4 J. INTELL. PROP. L. 273, 281–94 (1997) (discussing Internet architecture in the context of copyright analysis of file caching).

43 Patrick Maroney, *The Wrong Tool for the Right Job: Are Commercial Websites Places of Public Accomodation Under the Americans with Disabilities Act of 1990?*, 2 VAND. J. ENT. L. & PRAC. 191, 193–95 (2000) (discussing Internet architecture's impact on accessibility to the disabled).

44 G. Andrew Barger, *Cybermarks: A Proposed Hierarchical Modeling System of Registration and Internet Architecture for Domain Names*, 29 J. MARSHALL L. REV. 623, 647–49 (1996) (discussing the connection between Internet architecture and domain name policy); A. Michael Froomkin, *Form and Substance in Cyberspace*, 6 J. SMALL & EMERGING BUS. L. 93, 109–11 (2002) (discussing the role of Internet architecture in connection with domain name system control issues); Laurence R. Helfer & Graeme B. Dinwoodie, *Designing Non-National Systems: The Case of the Uniform Domain Name Dispute Resolution Policy*, 43 WM. & MARY L. REV. 141, 263 (2001) (discussing the connection between Internet architecture and domain name policy).

ties regulation,[48] tax policy,[49] telecommunications regulation,[50] and a variety of other topics.[51] Given that Internet architecture is already on the table for discussion, what value does layers analysis add? Our answer to this question begins with an explication of the code thesis.

---

45 Jenna F. Karadbil, Note, *Casinos of the Next Millennium: A Look into the Proposed Ban on Internet Gambling*, 17 ARIZ. J. INT'L & COMP. L. 413, 437–38 (2000) (discussing Internet architecture in the context of gambling regulation).

46 Carlos Perez-Albuerne & Lawrence Friedman, *Privacy Protection for Electronic Communications and the "Interception-Unauthorized Access" Dilemma*, 19 J. MARSHALL J. COMPUTER & INFO. L. 435, 445 (2001) (discussing Internet architecture in the context of privacy issues); Jonathan Weinberg, *Hardware-Based ID, Rights Management, and Trusted Systems*, 52 STAN. L. REV. 1251, 1259–63 (2000) (discussing Internet architecture in the context of privacy and security issues); Joshua S. Bauchner, Note and Comment, *State Sovereignty and the Globalizing Effects of the Internet: A Case Study of the Privacy Debate*, 26 BROOK. J. INT'L L. 689, 708–09 (2000) (discussing Internet architecture in the context of globalization and privacy).

47 Jerry Kang, *Cyber-Race*, 113 HARV. L. REV. 1130, 1160–61 (2000).

48 John G. Moon, *The Dangerous Territoriality of American Securities Law: A Proposal for an Integrated Global Securities Market*, 21 NW. J. INT'L L. & BUS. 131, 188–90 (2000) (discussing Internet architecture in the context of securities regulation).

49 Arthur J. Cockfield, *Transforming the Internet into a Taxable Forum: A Case Study in E-Commerce Taxation*, 85 MINN. L. REV. 1171, 1235–36 (2001) (discussing Internet architecture in the context of tax policy); Edward A. Morse, *State Taxation of Internet Commerce: Something New Under the Sun?*, 30 CREIGHTON L. REV. 1113, 1124–27 (1997) (discussing Internet architecture in the context of taxation of e-commerce transactions); W. Ray Williams, *The Role of Caesar in the Next Millennium? Taxation of E-Commerce: An Overview and Analysis*, 27 WM. MITCHELL L. REV. 1703, 1707 (2001) (noting that "[f]amiliarity with the Internet's architecture is critical to understanding the policy of taxing a transaction based on a nexus or a physical presence requirement"); Neal Harold Luna, Comment, *Implications of Singapore's Income and Consumption Tax Policies on International E-Commerce Transactions of Digitized Products*, 10 PAC. RIM L. & POL'Y J. 717, 723 (2001) (discussing Internet architecture in the context of Singapore's tax policy).

50 Henry E. Crawford, *Internet Calling: FCC Jurisdiction over Internet Telephony*, 5 COMMLAW CONSPECTUS 43, 43–44 (1997) (discussing Internet architecture in the context of the regulation of Internet telephony); Steve Kelley, *Liberating Our Digital Future: How the 1996 Telecommunications Act Definitions Are Hobbling Change*, 27 WM. MITCHELL L. REV. 2137, 2143–47 (2001) (discussing Internet architecture in the context of telecommunications policy); Christopher Libertelli, *Internet Telephony Architecture and Federal Access Charge Reform*, 2 B.U. J. SCI. & TECH. L. 13 *passim* (1996) (discussing Internet architecture in the context of regulation of Internet telephony); Tuan N. Samahon, Comment, *The First Amendment Case Against FCC IP Telephony Regulation*, 51 FED. COMM. L.J. 493, 501–02 (1999) (discussing Internet architecture in the context of regulation of Internet telephony).

51 Principe, *supra* note 20, at 727, 730–37 (discussing Internet architecture in the context of law of war).

1. The Code Thesis

The distinctive contribution of layers analysis is illuminated by comparison to two ideas advanced by Lawrence Lessig: the code thesis and the end-to-end principle. Before we begin our exposition, we wish to acknowledge our profound debt to Lessig's work. Although we disagree with Lessig in some important ways, our analysis operates within a framework that Lessig and others pioneered. We believe that the move to the layers analysis both extends and supports the fundamental normative thrust of Lessig's work.

Lessig has argued for a fundamental idea that provides a framework for analyzing Internet regulation—which we shall call the code thesis.[52] The code thesis is the claim that the nature of the Internet or cyberspace is determined by the code—the software and hardware that implements the Internet.[53] As a product of human endeavor, cyberspace lacks inherent natural properties that can be attributed to the various regions of physical space. This is a fundamental difference between the Internet and the vacuum of outer space or the liquid of the world's oceans. Although those environments can be modified by human activity, their fundamental nature is not created or shaped by engineering.

By contrast, *how* the Internet runs or cyberspace operates is completely dependent on the code that implements it. This point can be missed. One might argue that the Internet has an inherent nature, and that regulation of the Internet must respond to that nature.[54] So, for example, one might argue that the Internet cannot be regulated

52 *See generally* LESSIG, *supra* note 15.

53 *Id.* at 6.

54 This is not to say that the Internet has no essential characteristics. In this footnote, we digress briefly on the point. The term "Internet" is conventionally capitalized, indicating that it is a proper name (or in philosophical discourse, a "rigid designator")—the existing global network of networks unified by TCP/IP as well as the authoritative Internet Protocol (IP) address system and the authoritative Domain Name System (DNS), managed by the Internet Corporation for Assigned Names and Numbers (ICANN) and sometimes called the IANA functions. *See, e.g.*, INTERNET CORP. FOR ASSIGNED NAMES & NUMBERS, PROGRESS REPORT ON PERFORMANCE OF IANA FUNCTIONS (May–July 2000), *available at* http://www.iana.org/periodic-reports/progress-report-may-jul00.htm; Stuart Lynn, *What are the IANA Functions?*, *at* http://www.ripe.net/ripe/meetings/archive/ripe-39/presentations/icann/sld003.html (last visited Feb. 11, 2004) (displaying a slide show prepared by Stuart Lynn, then President of ICANN, and presented at the RIPE-NCC Meeting, Bologna, Italy, May 2001). We can contrast the actual Internet with "an internet," e.g. any possible system for creating a global network of networks. The Internet may evolve in various ways, but in order for *the Internet* to count as *an internet*, the Internet must remain a global network of networks and not something else.

by national governments, because, as a global network of networks, activity on the Internet can originate in any physical location, anywhere in the world.[55] This argument conveys a large measure of truth, but it misses an essential point—the Internet has this property because of the code or software that makes physical location irrelevant. That code could have been different or it could change in the future.

In this sense, code is the prime regulator in cyberspace—in Lessig's felicitous phrasing, "the Code is Law."[56] The point is *not* that that the software that enables the Internet is literally a law, in the sense that it would satisfy a theory of the nature of law.[57] Nor is it Lessig's point that the role of software in determining the properties of the Internet is the same as the role of the laws of nature in determining the properties of the Pacific Ocean.[58] Rather, Lessig's point is that software or code has regulative effects on human behavior. In this sense, Internet architecture is like the architecture of buildings and cities. Just as the architecture of a building enables and encourages humans to move and congregate in certain ways, so the architecture of the Internet enables some activities by users and regulators while discouraging others.

For our purposes, the architecture of the Internet is the regulating entity that allowed the explosion of innovation in cyberspace. The Internet is configured in a way that enables low cost innovation at the application layer. The architecture of the Internet is a function of the software (or code) and hardware that constitutes the Internet. Software and hardware are the bricks and mortar of the Internet.

---

55 *See* Justin Hughes, *The Internet and the Persistence of Law*, 44 B.C. L. REV. 359, 365–70 (2003).

56 *Id.* at 369.

57 Take, for example, H.L.A. Hart's theory that law is the union of primary and secondary rules as identified by a rule of recognition. For Hart, the sine qua non of law is a social rule. *See generally* H.L.A. HART, THE CONCEPT OF LAW (Joseph Raz & Penelope Bullock eds., 2d ed. 1997). Code is software, and although socially created, software is not a social rule. Likewise, code is not law in John Austin's sense. It is not the command of the sovereign backed by the threat of punishment. *See* JOHN AUSTIN, THE PROVINCE OF JURISPRUDENCE DETERMINED 18–37 (Wilfred E. Rumble ed., 1995).

58 Like the laws of nature and the ocean, the code that constitutes the Internet causally determines the properties of the Internet. Unlike the laws of nature, code is malleable. Humans cannot alter the specific gravity of H_2O, but they can rewrite TCP/IP.

2. The End-to-End Principle

What then is the architecture of the Internet? This is, of course, a very abstract and hence potentially ambiguous question. A narrower and less ambiguous question would be: Which feature or features of Internet architecture are responsible for the distinctive and valuable functionality that the Internet provides? Lessig suggests that the answer to these questions can be captured in large part by a single principle; he has argued that the primary characteristic of the Internet architecture that enables innovation is the end-to-end principle.[59] As Lessig explains, the end-to-end principle says to keep intelligence in a network at the ends or in the applications, leaving the network itself relatively simple. In short, the principle calls for a "stupid network" and "smart applications." The network simply forwards or routes the data packets and does not—and cannot by architecture—discriminate or differentiate traffic generated by different applications.

This point may require further explanation, even for readers who are very familiar with the operation of the Internet as end users. This is because most contemporary Internet users are only aware of one application, an application that is so ubiquitous that it can be confused with the Internet itself. That application is Hyper Text Transfer Protocol (HTTP), the software that enables the World Wide Web.[60] The Web is not the Internet. It is one application that communicates using the Internet. Despite the ubiquity of the Web, most Internet users have used two other applications, neither of which is the Web (although they are accessed through the Web in some cases). The first of these is File Transfer Protocol (FTP), the application that allows a file to be transferred from one server on the Internet to another.[61] You are likely to have used FTP if you have downloaded a file from a server on the Internet. Although FTP is available (with various

---

59 LESSIG, *supra* note 1, at 34. *But see* Christian Sandvig, *Shaping Infrastructure and Innovation on the Internet: The End-to-End Network that Isn't*, *in* SHAPING SCIENCE AND TECHNOLOGY POLICY: THE NEXT GENERATION OF RESEARCH (David H. Guston & Daniel Sarewitz eds., forthcoming 2004) (manuscript at 25), *at* http://www.spcomm.uiuc.edu/users/csandvig/research/Communication_Infrastructure_and_Innovation.pdf (last visited Feb. 24, 2004) (arguing that transparency, participation, and flexibility, and not end-to-end, should guide regulatory policy).

60 *See* Webopedia, *HTTP*, *at* http://www.webopedia.com/TERM/H/HTTP.html (last visited Feb. 13, 2004) (defining HTTP).

61 *See* Webopedia, *FTP*, *at* http://www.webopedia.com/TERM/F/FTP.html (last visited Feb. 11, 2004) (defining FTP). On the distinction between FTP and HTTP, see Webopedia, *The Difference Between FTP and HTTP*, *at* http://www.webopedia.com/DidYouKnow/Internet/2002/FTP_HTTP.asp (last visited Feb. 11, 2004).

interfaces) as a stand-alone program,[62] it is also integrated with web browsers, such as Internet Explorer, Opera, Mozilla, and Netscape Communicator, which are used to access the Web. The second application with which most users are familiar is Simple Mail Transfer Protocol (SMTP), the basis for e-mail.[63] Although the standard e-mail application can be accessed via the Web, most users access the application through a stand-alone program, such as Outlook or Eudora. Recently a new family of applications has become familiar to many users. Peer-to-peer (P2P) file sharing programs[64] such as the late lamented Napster, and its cousins KaZaA and Bearshare, are especially popular for the sharing of digital music (MP3) files. These programs are built around distinct applications.[65]

The Internet does not "know" whether a given packet of data is a web page (HTTP), an article downloaded from an online service[66] (FTP), an e-mail message (SMTP), or an MP3 file being shared by use of KaZaA.[67] This point can be illustrated by considering a hypothetical reader, whom we call Alice, who downloads this Article from the Social Science Research Network (SSRN). When we wrote this Article using Microsoft Word, the content (the semantic content of the Article expressed as syntactic content, e.g., letters, words, and sentences) was encoded as a digital file in a particular format (a DOC file). Adobe Acrobat then translated that file into a different format (a PDF file), preserving its semantic and syntactic properties. When that PDF file is loaded on a server by the Social Science Research Network (SSRN), a web browser can be used to locate the file and initiate a request to use the FTP application. The semantic and syntactic content of the Article has now been passed to the application layer. The next step is the crucial one. The digital file (the PDF) file is passed by the application (FTP) to the transport layer (the TCP part of TCP/IP). The transport layer breaks the PDF file into data packets. The

62 One popular implementation of FTP is CuteFTP. *See* Globalscape, *The World's Favorite FTP Client, CuteFTP!*, *at* http://www.globalscape.com/products/cuteftp/index.asp (last visited Feb. 11, 2004).

63 *See* Webopedia, *SMTP*, *at* http://www.webopedia.com/TERM/S/SMTP.html (last visited Feb. 13, 2004) (defining SMTP).

64 *See* Webopedia, *Peer-to-Peer Architecture*, *at* http://www.webopedia.com/TERM/P/peer_to_peer_architecture.html (last visited Feb. 13, 2004) (defining peer-to-peer architecture).

65 *See infra* Part III.A.4 (discussing various peer-to-peer applications).

66 Social Science Research Network, *Homepage*, *at* http://www.ssrn.com (last visited Feb. 13, 2004).

67 The discussion elides the role of port number assignments, which do allow TCP to discriminate between applications (but not content) on the basis of port numbers. For a discussion of port and the layers principle, see *infra* Part III.A.4.b.

transport layer cannot differentiate between PDF files and MP3 files. TCP does not distinguish the packets that carry a law review article by Cass Sunstein from the packets that, when reassembled, would be a popular song by Eminem. And, once the transport layer has broken the data into packets, the Internet protocol layer (the IP part of TCP/IP) cannot distinguish one packet from another on the basis of content. If Alice downloads our Article from SSRN, the software or code that makes up the Internet cannot distinguish the packets of data that comprise our Article from any other packets of data. The content only becomes accessible to the code after the transport layer on Alice's computer reassembles the packet into a PDF file. Once this has happened, the operating system of Alice's computer (Microsoft Windows) can recognize that the file is a PDF file associated with the Acrobat Reader program. That program in turn can read the digital file and display its syntactic content on Alice's monitor. Alice will then infer that the squiggly lines on her screen are letters forming words and sentences, and she will attempt to determine what we (the authors of this Article) meant.

This example illustrates what is sometimes called the "stupidity" of the Internet. The Internet doesn't know that FTP is trying to send a large file from one of our computers to Alice's computer. The Internet cannot coordinate the routing of packets so that they all arrive at Alice's computer at the same time. Ninety-five percent of the packets may arrive in the first minute, and the remaining five percent, traveling by a different route on the Internet, might arrive minutes later. The Internet cannot decide that delivery of PDF files is a high priority for an academic institution and delivery of MP3 files is a low priority. Nonetheless, the deaf, dumb, and blind Internet can sure transmit a mean PDF file. Compare the efficiency of searching and downloading works in progress from the Social Science Research Network with the old-fashioned system of striking up conversations at conferences, leading the conversation to what you are working on, triggering an invitation to send a draft, and then snail-mailing the draft in a bulky envelope to one reader at a time.

3. Transparency and the Economics of Innovation

The end-to-end principle is shorthand for this feature of the architecture of the Internet—the combination of a "stupid network" and "smart applications." The network does not discriminate between applications. The software at the transport and Internet protocol layers simply does not include code that would allow the Internet to asso-

ciate data packets with application file types.[68] We can call this characteristic of the Internet *transparency*. The Internet is transparent to applications. The transparent, nondiscriminatory nature of the Internet, Lessig has argued, is what has enabled the explosion of innovation and creativity on the Internet.[69] Because the network is nondiscriminatory or transparent to applications, and the intelligence is implemented at the ends by the applications, the innovation or creative activity is placed in the hands of the application creators. Thus, innovation is decentralized and the opportunity to devise new applications is available to millions of creative individuals with Internet access.

The economic significance of transparency is that it dramatically lowers the investment required to produce a given innovation. Given a transparent Internet, an innovator need only invest in innovation at the application layer.[70] If the Internet were opaque, then new applications would also require investment in changes at lower layers (the transport, IP, link, or physical layers).

Moreover, transparency reduces the cost of innovation to consumers. Given a transparent Internet, a consumer need only invest in an application itself in order to make use of it. For example, in order to use KaZaA, I need only download the KaZaA program; I do not need to ask my network administrator to reconfigure the network to permit KaZaA to run—with one caveat.[71] If the Internet were opaque,

---

68 A program called a packet sniffer can examine the content of packets on a network. *See* Webopedia, *Sniffer, at* http://www.webopedia.com/TERM/S/sniffer.html (last visited Feb. 11, 2004). An example is "Sniff'em," which is described by its manufacturer as "[a] Packet Sniffer . . . that captures, monitors and analyzes network traffic, detecting bottlenecks and other network related problems. Using this information, a network manager can keep traffic flowing efficiently. A packet sniffer can also be used legitimately or illegitimately to capture data being transmitted over a network." Sniff'em, *What is a Packet Sniffer or Network Sniffer?, at* http://www.sniffem.com/sniffem.shtml (last visited Feb. 13, 2004).

69 *See* LESSIG, *supra* note 1, at 120–34.

70 The cost of innovation at the application layer is affected by the constraints imposed by the transport and IP layers. Simplicity at those layers reduces the cost of innovation at the application layer. Complexity at those layers would increase the cost of innovation at the application layer. For this reason, layer separation by itself does not guarantee that the cost of innovation at the application layer will be low. Many other factors affect those costs, including complexity introduced at lower layers.

71 There is an exception to this rule. If a network administrator closes the port utilized by a particular program, that port must be opened to permit the application to run. *Cf.* Internet Assigned Numbers Auth., *Port Numbers, at* http://www.iana.org/assignments/port-numbers (last updated Mar. 31, 2003); searchNetworking.com, *Definitions: Port Number, at* http://searchnetworking.techtarget.com/sDefinition/0,,sid7_gci212811,00.html (last visited Feb. 13, 2004).

consumers would be required to invest in changes to their network infrastructure in order to utilize an innovative application. Because networks are complex and are used by multiple users, network reconfiguration may be relatively expensive. Because the Internet is even more complex and is used by hundreds of millions of servers, reconfiguration of the whole Internet to enable a new application would involve a very large investment.

This point about transparency and adoption costs leads to another important concept, *networking effects*.[72] The economic value of some innovations depends on networking effects. The value of an application like SMTP (e-mail) is a function, in part, of the number of adopters; the more users of e-mail, the more valuable it is.[73] For some applications, there may be a tipping point,[74] at which the number of adopters reaches critical mass resulting in a discontinuous and large increase in value from networking effects or from similar phenomena with respect to awareness of (or belief in) networking effects. Once this tipping point is reached, the application becomes sufficiently valuable to broaden its appeal from early adopters to ordinary users. Reducing the cost of adoption for consumers increases the likelihood that these networking effect gains will be realized. If adoption costs are too high for early adopters, the tipping point may never be reached. For this reason, transparency may be required to enable innovations that would be efficient even given the large costs associated with reconfiguring the Internet. Without transparency, the benefits of the innovation would never become apparent and hence the investments would never be made.

The economics of Internet innovation are illustrated by the development of the World Wide Web described by Tim Berners-Lee and

---

72 *See* Mark A. Lemley & David McGowan, *Legal Implications of Network Economic Effects*, 86 CAL. L. REV. 479 (1998). *See generally* OZ SHY, THE ECONOMICS OF NETWORK INDUSTRIES 1–6 (2001); Philip H. Dybvig & Chester S. Spatt, *Adoption Externalities as Public Goods*, 20 J. PUB. ECON. 231 (1983); Michael L. Katz & Carl Shapiro, *Network Externalities, Competition, and Compatibility*, 75 AM. ECON. REV. 424 (1985); S.J. Liebowitz & Stephen E. Margolis, *Network Externality: An Uncommon Tragedy*, 8 J. ECON. PERSP. 133 (1994).

73 For purposes of illustration, we set aside the possibility that there may be a point at which additional users of e-mail actually reduce its utility by creating sorting problems and encouraging "spam."

74 The idea of a tipping point is frequently used to describe norm cascades. *See* Laurence R. Helfer, *Overlegalizing Human Rights: International Relations Theory and the Commonwealth Caribbean Backlash Against Human Rights Regimes*, 102 COLUM. L. REV. 1832, 1846–47 (2002) (citing Martha Finnemor & Kathryn Sikkink, *International Norm Dynamics and Political Change*, 52 INT'L ORG. 887, 896, 901 (1998)).

others.[75] The core idea of the Web is the hypertext link: by clicking here, on this word, you go there, to that word. By clicking here, on this link, you go there, to that web page. Although Berners-Lee developed this idea independently, the same idea had been developed before, on several occasions—all of them before the rise of the Internet.[76] What the transparency of the Internet made possible was the implementation of Berners-Lee's HTTP application without any need for cooperation from or modification of the communications system on which it was implemented. That is, the transparent nature of the Internet dramatically lowered the investment that Berners-Lee and others needed to make in order to produce the Web.

The World Wide Web sits on top of the Internet. From the application user's point of view, the Internet is simply invisible—except, of course, when network congestion calls it to our attention. If the architecture of the Internet had been opaque, then the counterfactual equivalent of TCP/IP would have required modification for HTTP to run on the Internet. And, if that had been the case, there is good reason to believe that there would never have been a World Wide Web. The networking effects that transformed the World Wide Web from a merely great idea into an enormously useful communications tool would never have begun cascading if the platform on which the Web runs hadn't already been widely accessible. Berners-Lee faced real obstacles in getting HTTP accepted, even though he was giving it away and it could be run for free (at least from the point of view of short-run marginal costs).[77] Had Berners-Lee been required to convince managers of the Internet to expend resources to enable the Internet to run HTTP, we have every reason to believe that he would have failed. Indeed, he might have been unable to convince his own employers to allow him to make the necessary modifications in the network at CERN for the experimental work that developed HTTP.

The economics of Internet transparency and networking effects are very powerful. HTTP enabled what we now think of as "the Internet revolution." But the dot com boom and bust, eBay and Amazon.com, homepages and browsing—the cyberspace phenomena that have become the furniture of our day to day ordinary existence—all of these are the product of a single innovative application—HTTP. The peer-to-peer phenomenon (from Napster to KaZaA and beyond)

75 *See* TIM BERNERS-LEE, WEAVING THE WEB: THE ORIGINAL DESIGN AND ULTIMATE DESTINY OF THE WORLD WIDE WEB *passim* (2000).

76 See Dan Connolly, *A Little History of the World Wide Web from 1945 to 1995, at* http://www.w3.org/History.html (last visited Feb. 18, 2004) (presenting a timeline with precursors to the World Wide Web's hypertext features).

77 BERNERS-LEE, *supra* note 75, at 30–34.

is the product of a small number of innovative applications. We have no crystal ball, and Kenneth Arrow's work on valuing innovation[78] tells us that we cannot reliably estimate the discounted present value of the future Internet innovation. Nonetheless, if past is prologue, we have reason to believe that the potential gains are very large indeed. If you insist on a number of dollars, we submit that it could easily have fifteen or even sixteen digits.[79]

B. Transparency and the Layers

Lessig's analysis of Internet architecture is essentially correct. Lessig has properly focused our attention on the key innovation enabling feature—the transparency of the network to applications. Moreover, Lessig's attribution of this transparency to the end-to-end principle does capture an important, indeed crucial, part of the full story as to how the architecture of the Internet enables transparency. But, it is only part of the story. Transparency is not a direct result of the end-to-end principle, but rather is a built in characteristic of the layered architecture of the Internet. That is, *layers* are the key, central characteristic of the Internet architecture. The end-to-end principle emerged from the layers model as an articulation and abstraction of implicit ideas inherent in the layers model. In this section, we describe in some detail the end-to-end principle and the layered architecture of the Internet communication protocol, the TCP/IP, in order to clarify these concepts. Then we present an argument that the layers model is the key characteristic of Internet architecture. Finally, we discuss how layers and end-to-end relate to the innovation and regulation on the Internet.

1. Layers and End-to-End Principle

The end-to-end principle derives from pioneering work in the theory of network design by Jerome Saltzer, David Reed, and David Clark. As they put it, end-to-end is an argument (or a class of arguments) that says "[t]he function in question can completely and correctly be implemented only with the knowledge and help of the application standing at the endpoints of the communication system.

78 Kenneth Arrow, *Economic Welfare and the Allocation of Resources for Invention, in* THE RATE AND DIRECTION OF ECONOMIC ACTIVITY 609–26 (Nat'l Bureau Comm. for Econ. Res. eds., 1962).

79 Ira Magaziner estimated that "almost two-thirds of the real growth of the U.S. economy [during the mid- and late-1990s came] from the Internet economy." Ira Magaziner, *At The Crossroads of Law and Technology: Keynote Address, October 23, 1999*, 33 LOY. L.A. L. REV. 1165, 1166 (2000). This translates into trillions of dollars.

Therefore, providing that questioned function as a feature of the communication system itself is not possible."[80]

Stated in a generalized form, the end-to-end argument counsels against low-level function implementation.[81] What does this mean? Avoiding low-level function implementation means that a functionality desired by an application should be implemented at the level of the application rather than at a lower level. The version of end-to-end quoted above is a statement of the end-to-end principle for the engineering of the Internet and the software that uses the Internet.

For example, the functionality of ensuring data integrity for a file transfer application is better implemented at the application level than at the network communication level. Even if the network delivered data without corruption, the reasoning goes, the file transfer application still needs to do the data checking because the data corruption could happen before the data reach the network—due to a bad disk, for example.

The simple example of error checking illustrates two general features of the end-to-end argument: *vertical hierarchy* and *information mismatch.* Each of these two features must be explicated and clarified in order to illuminate the meaning and significance of the end-to-end principle in relation to the layers model.

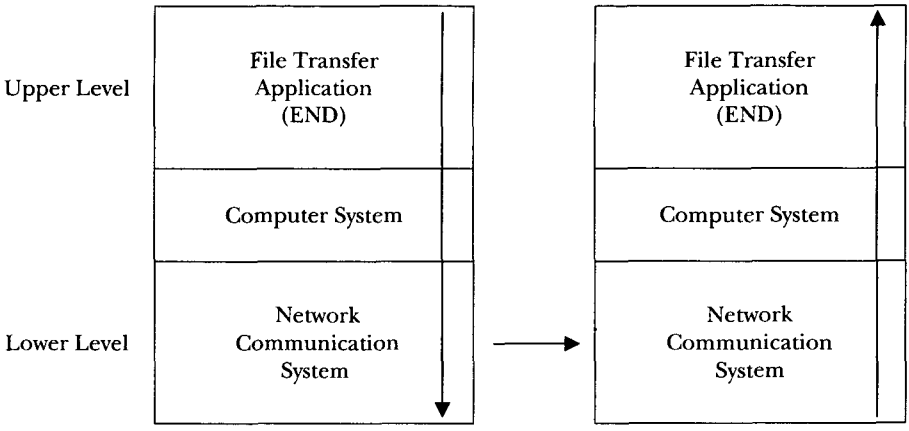
a. The Layers Concept is Implicit in the End-to-End Argument

The end-to-end argument presupposes some sort of vertical hierarchy—that is, it assumes lower and upper levels. In the above example, the file transfer application is at an upper level with respect to the network communication system, which, in turn, lies at a lower level with respect to the application. The vertical hierarchy in the above example is illustrated in Figure 1:

80 Jerome H. Saltzer et al., *End-To-End Arguments in System Design*, 2 ACM TRANSACTIONS ON COMPUTER SYS. 277, 278 (1984).

81 *Id.*

FIGURE 1. VERTICAL HIERARCHY



The fundamental insight of the end-to-end principle is that any given functionality should be implemented at a level where it is needed—the "end" point. Thus, an essential task of understanding and applying the end-to-end principle is identifying and understanding the relevant levels. Where is the "end"? "End" with respect to what? That is, identifying and understanding the relevant vertical hierarchies is essential to appreciate the end-to-end argument in specific circumstances. In the context of Internet architecture, the relevant vertical hierarchy implicit in the end-to-end argument are the layers of the TCP/IP protocol.

b. Information and Layer Mismatch

Given the relevant vertical hierarchy, the core of the end-to-end argument is that it is impractical, if not impossible, to provide a given functionality at a level lower than the level where the function is used—the "end" layer—because the lower layer may lack the information necessary to fully perform the function.

In the file transfer example above, the data integrity function at the network level is incomplete and redundant because data corruption—the key information in this functionality—can take place at a higher level and the network has no reliable information on this matter. Only the application knows or can know what the uncorrupted data look like, and it is impractical, if not impossible, to convey this information to the lower levels. The subtleties of the application requirement in many cases are not easily communicable to the lower levels. Furthermore, the requirements may not be static, but may change over time.

It should be noted that the general principles of the end-to-end argument are applicable in many areas outside the network system design. In fact, the creators of the end-to-end argument prefer to call it a class of arguments applicable to many areas, including, but not limited to, the design of encryption system, reliable storage system, and the CPU architecture.[82]

2. The Layers Model of the TCP/IP Protocol

a. The TCP/IP Protocol is the Code of the Internet

In general, a network protocol is a set of rules and conventions by which computers can communicate with each other.[83] Because a computer network is nothing but a set of computers connected with each other, the network architecture is determined by the architecture of the network protocol.[84] The TCP/IP suite is the network communication protocol for the Internet.[85] Thus, the architecture of the Internet as a network is determined by the architecture of the TCP/IP protocol.

The fundamental goal of the initial Internet architecture was to create a network of networks by interconnecting various computer network systems already in existence at the time.[86] In order to meet this goal, TCP/IP was designed to be a software-only protocol, independent of any particular computer and network hardware. That is, TCP/IP is "pure code." It is *the* code of the Internet that determines the architecture of the Internet.[87]

Many protocols other than TCP/IP are involved in Internet communications, such as Ethernet for Local Area Networks (LANs) and Frame Relay or Asynchronous Transer Mode (ATM) for Wide Area Networks (WANs). The key point is that, without a common set of protocols, these various networks will not be able to communicate with each other, and thus cannot form an interconnected network of networks. For the Internet, TCP/IP is the common protocol that ties together all these otherwise disparate networks into a functional network of networks. TCP/IP is *the* protocol that makes the Internet possible as a network of networks. Without TCP/IP there can be no

---

82 *Id.*

83 ANDREW S. TANENBAUM, COMPUTER NETWORKS 17 (1996).

84 *Id.* at 18.

85 1 STEVENS, *supra* note 18, at 1.

86 David D. Clark, *The Design Philosophy of the DARPA Internet Protocols*, 1988 PROC. SIGCOMM 88, *reprinted in* 18 ACM COMPUTER COMM. REV. 106 (1988).

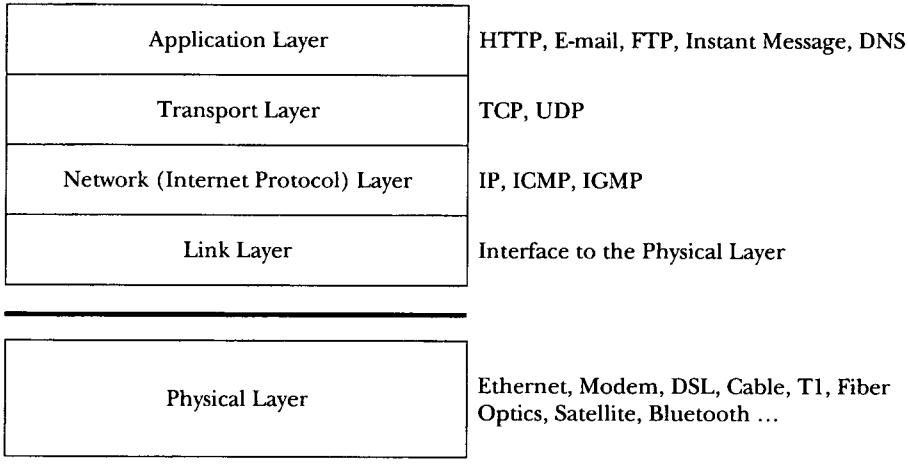
87 *See supra* note 19.

Internet. Thus, in this very important sense, TCP/IP *is* the Internet or, perhaps, is "the Internet as we know it."[88]

b. Layers Model of the TCP/IP Protocol Suite

TCP/IP is a layered network communication protocol that consists of four independent layers—the application, transport, network, and link layers—as illustrated in Figure 2:[89]

FIGURE 2. LAYERS MODEL OF TCP/IP



i. The Physical Layer

The physical layer is not a part of the TCP/IP protocol suite. It is the physical medium over which the actual transfer of bits takes place—e.g., Ethernet, modem, DSL, cable, T1, fiber optics, satellite link, etc. However, since no communication will take place without some physical medium, the physical layer must be considered as a part of the networking system to get a complete picture of the communication system. Importantly, the physical layer is rarely an architectural issue for TCP/IP, because of the protocol's independence from hardware.

ii. The Link Layer

The link layer handles all of the details of physically interfacing with the computer hardware and network hardware. As such, this layer is responsible for TCP/IP's freedom or independence over hard-

88 *See supra* note 19.

89 1 STEVENS, *supra* note 18, at 1–2.

ware. If new hardware—say a high-speed satellite link—is to be used for Internet communication, all that needs to be done is to implement the appropriate interface details at this layer.

Typically, the link layer is implemented as a device driver for the particular network hardware. Supporting new hardware is accomplished by swapping in a new device driver that provides the link between TCP/IP and the hardware. The upper layers do not need to be changed at all. In fact, the upper layers have no idea over which physical media the communication will take place. Thus, the link layer gives the hardware independence to the TCP/IP communication and consequently frees TCP/IP from the bondage of a variety of computer and network hardware.

iii. The Internet Protocol Layer

The Internet protocol, sometimes called the network layer, Internet layer or IP layer, handles the movement of data packets around the network. The encoding of IP addresses and the routing of the data packets for packet switching takes place here.

At this point, we need to dispel some unfortunate confusion about the naming of this layer. Because the IP layer was originally called the "network layer" or "Internet layer," many original design papers talk about "network design" or "Internet design" when what is in fact discussed are the design issues relevant only to this layer.[90] In order to avoid the confusion, it is a common practice among network software developers to call this layer the "Internet Protocol layer" or "IP layer." We shall adopt this naming convention for the remainder of this Article.

iv. The Transport Layer

The transport layer provides a flow of data between two hosts for the application layer above. This is where the data received from the application layer is broken up into data packets to be handed over to the network or IP layer, and the data packets received from the IP layer are assembled into a data stream to be delivered to the application layer.

---

90 *See, e.g.*, Memorandum from B. Carpenter, Network Working Group, to the Internet Community (Feb. 2000), *at* http://www.ietf.org/rfc/rfc2275.txt [hereinafter Internet Transparency]; Memorandum from B. Carpenter, Network Working Group, to the Internet Community (June 1996), *at* http://www.ietf.org/rfc/rfc1958.txt.

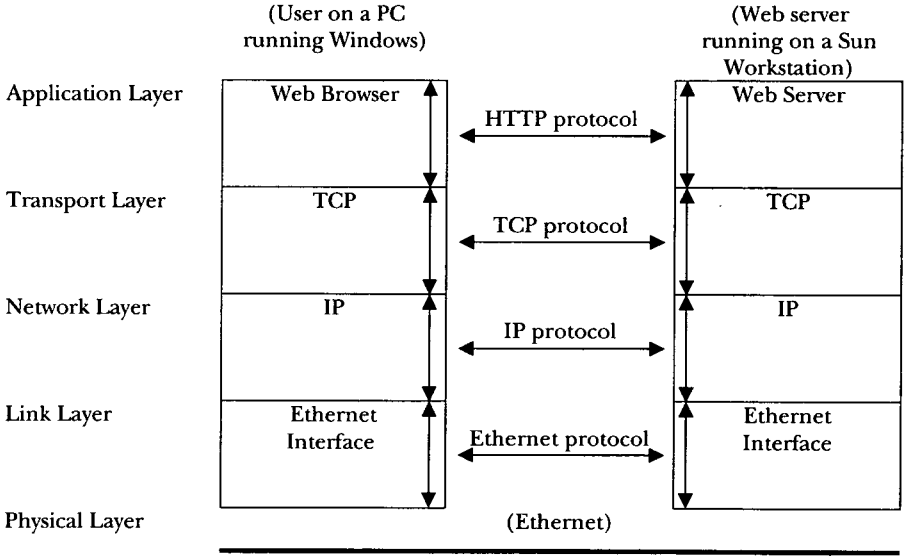
v. The Application Layer

The application layer handles the details of the particular application. Examples of the application layer protocols are—HTTP for web communication, SMTP for e-mail, FTP for file transfer, and the Domain Name System (DNS) for mapping IP address numbers to easily recognizable character strings.

c. An Example: Web Communication over Ethernet

To illustrate how these layers work together, let us consider a simple example of web communication over Ethernet, illustrated by Figure 3. The user launches a web browser to access a web server that is connected to the user's computer via an Ethernet link.

FIGURE 3. ETHERNET WEB COMMUNICATIONS



At the application layer, the web client (the browser) communicates with the web server using HTTP (the protocol of the World Wide Web) to exchange data. Although the horizontal communication between the web client and server takes place completely within the application layer, the communication mechanism is handing and receiving the data to and from the lower layer—the TCP layer. The same process is repeated down and up the layers.

d. The Horizontal Protocols

A key aspect of the TCP/IP protocol is complete and independent horizontal communications—independent of the lower layers as

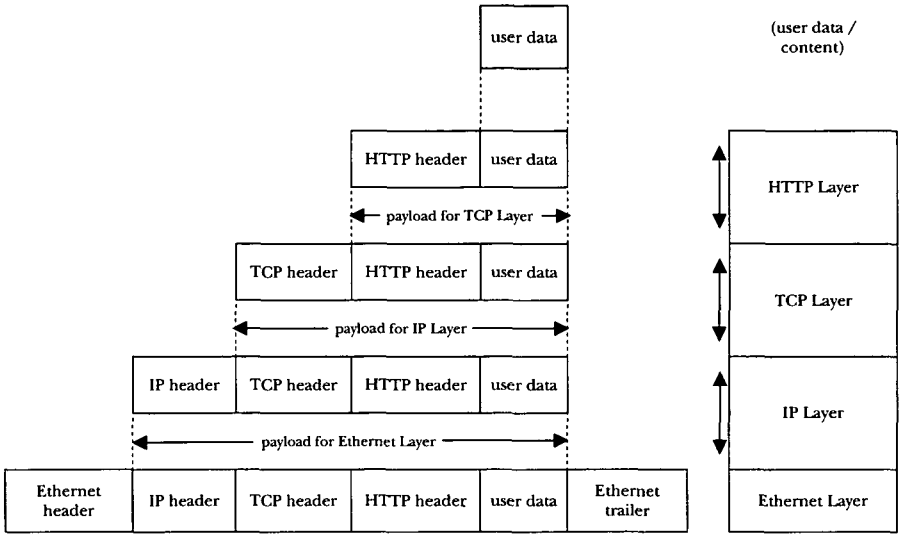
well as any computer hardware or operating system it runs on. As a communication protocol specification, HTTP does not know nor depend on which computer or network medium it runs on. The HTTP protocol specification is complete within the application layer—it is "just software." As long as the web server gets a protocol string that says something like "get me so-and-so document," the web server will send the file over to the requesting party without caring or ever knowing from what type of computer or over what kind of network medium the request came. Each of the other layers works in a similar fashion.

e. Vertical Protocols—Encapsulation of Data

Another key aspect of the TCP/IP protocol is how upper and lower layers communicate with each other. The vertical communication between the layers is achieved by a technique called encapsulation. A lower layer treats the data passed down from the upper layer as structureless "pure data" (or "payload" in network terminology), and puts header and/or trailer around the payload—i.e., "encapsulates" the payload data.

All of the information necessary for the horizontal communication at a specific layer is contained within the headers. When data is received from the lower layer, the upper layer looks for its layer header, and performs its layer function according to the information found in the headers. When the layer is satisfied that it has received a complete unit of data ready to be passed up, it then strips out the header and passes the payload up to the upper layer. The process is illustrated in Figure 4 for the web communication over Ethernet example above:

FIGURE 4. VERTICAL COMMUNICATION BETWEEN LAYERS



f. IPv6

Before proceeding further, we pause briefly to discuss IPv6 or version six of the Internet Protocol, the IP in TCP/IP. The Internet Engineering Task Force (IETF) designed IPv6 to update Internet Protocol, IP version 4 (IPv4), which is now nearly twenty years old. The most important feature of IPv6 is that it vastly expands the number of available IP Addresses—addressing the current shortage. The IETF expects IPv6 to gradually replace IPv4, with a period of several years during which the two protocols will coexist.[91] There are IPv6 task force organizations in North America, the European Union, India, and Taiwan.[92]

Most of the issues raised by IPv6 are outside the scope of this Article, but one point requires a very brief discussion. Because of the global shortage of IP addresses, many networks use Network Address Translation (NAT). "NAT allows two networks to be joined together, and is typically used to join a network of machines with non-routable IP addresses to the global internet."[93] NAT impinges on transparency for technical reasons.[94] Because IPv6 may eventually end the global

91 *See* IPv6, *Information Page*, *at* http://www.ipv6.org (last visited Feb. 11, 2004).

92 *See* IPv6 Task Force, *TF Around the World*, *at* http://www.ipv6tf.org (last visited Feb. 11, 2004).

93 Posting by Michael, michael@slashdot.org, to Slashdot, The Fight for End-to-End: Part One (Dec. 6, 2000), *at* http://slashdot.org/articles/00/12/06/1613233.shtml (last visited Feb. 11, 2004).

94 *See id.*

shortage of IP addresses, it should mitigate this particular threat to transparency.[95]

3. Layers and Transparency

From the above discussions, the key features of the TCP/IP protocol suite can be summarized as: (1) layered protocol, (2) horizontal protocols with complete and independent horizontal communication, and (3) vertical protocol with encapsulation. A key architectural principle implicit in these concepts is the separation of layers. The end-to-end principle and the transparency of the Internet to applications flow from the principle of separation of layers.

a. Separation of Layers and End-to-End Principle

Under TCP/IP design, various network functions are organized into several layers, and these functions are independent of each other—that is, the functions (and thus the layers) are *separated.* In other words, the internal workings of the layers are hidden from each other because that is the only way layers can be independent of each other. In fact, encapsulation is a method of *information hiding* in software design.[96]

Although separation of layers may seem obvious (why divide function into different layers if not to separate them?) the concept makes clear or articulates a very important design choice implicit in the layers model—that the layers are separated for a sound reason of network engineering, and hence that functions should not cross layers unless there is an exceptional reason to do so. For example, the designers and programmers are not free to implement application layer functions in the IP layer. It is implicit in and fundamental to the layers model that the application layer is created and placed above the network layers because we want to put application functions there and not at the lower network layers.

Therefore, the end-to-end principle—which, as described by Lessig, "keep[s] intelligence in a network at the ends, or in the applications, leaving the network itself to be relatively simple"[97]—follows from (and is an articulation of) the implicit design principle inherent to the layers model of the TCP/IP protocol.[98] One might be tempted to argue that the TCP/IP's layered model is a result of the end-to-end

95 *See id.*

96 GRADY BOOCH, OBJECT-ORIENTED ANALYSIS AND DESIGN 49 (1994).

97 LESSIG, *supra* note 1, at 34.

98 Computer scientist Edward Felten, reacting to a prior version of this Article, questions this point, writing:

principle, not the other way around. That argument, however, would be erroneous historically, conceptually, and functionally. Historically, the end-to-end principle was first articulated in the early 1980s,[99] and the layered model of TCP/IP was developed in the mid-1970s.[100] Conceptually, the layer separation does not follow from the end-to-end principle, because the end-to-end principle doesn't tell us to separate the TCP, IP, and physical layers, whereas the end-to-end principle does follow from the separation of the application layer from the lower network layers. Functionally, satisfaction of the end-to-end principle is guaranteed by the observance of separation of layers, but the reverse is not the case. The end-to-end principle simply does not dictate a robustly specified functional design for the network.

Our claim—that for the purposes of policy analysis, layers analysis is in a sense more fundamental than the end-to-end principle—should not obscure the fact that the end-to-end principle is important to the issues of Internet architecture and regulation on the Internet. The end-to-end principle is a guiding normative principle that clari-

---

> [Solum and Chung] are on shak[y] ground, though, when they relate their layering principle to the end-to-end principle that Lessig has popularized in the legal/policy world. (The end-to-end principle says that most of the "brains" in the Internet should be at the endpoints, e.g. in end users' computers, rather than in the core of the network itself.) Solum and Chung say that end-to-end is a simple consequence of their layering principle. That's true, but only because the end-to-end principle is built in to their assumptions, in a subtle way, from the beginning. In their account, layering occurs only at the endpoints, and not in the network itself. While this is not entirely accurate, it's not far wrong, since the layering is much deeper at the endpoints than in the core of the Net. But the *reason* this is true is that the Net is designed on the end-to-end principle. There are alternative designs that use deep layering everywhere, but those were not chosen because they would have violated the end-to-end principle. End-to-end is not necessarily a consequence of layering; but end-to-end is, tautologically, a consequence of the kind of end-to-end style layering that Solum and Chung assume.

Edward W. Felten, *Layers*, FREEDOM TO TINKER, June 18, 2003, *at* http://www.freedom-to-tinker.com/archives/000410.html. Felten is approaching the question from the standpoint of an engineer designing the system, and his point is correct from that perspective. The abstract concept of layering does not dictate end-to-end. Our claim, which is consistent with Felten's analysis, is that the layers model of the TCP protocol yields end-to-end. Nonetheless, Felten's perception clarifies our analysis and yields an important corollary—not just any layered protocol would guarantee the transparency of the Internet. This clarification would become important in the event that a basic redesign of the Internet's communications protocol were on the table. In that situation, both layers analysis and the end-to-end principle would be required to ensure transparency in what would be, in an important sense, a *new* Internet.

99 Saltzer et al., *supra* note 80, at 287.

100 Clark, *supra* note 86, at 114.

fies, articulates, and illuminates the implicit design principle inherent to the layers model of the TCP/IP. In the words of the original inventors of the end-to-end idea, "[e]nd-to-end arguments may be viewed as part of a set of rational principles for organizing layered systems."[101]

b. Separation of Layers and Transparency

From our discussion above, it also follows that the transparency in Internet architecture is an inherent, built-in characteristic of the layered architecture of the TCP/IP protocol. If transparency is loosely described for the present purpose as a nondiscriminatory characteristic of the network to application data (a more precise definition will be given later), then such nondiscriminatory behavior is a direct result of the way the data from the upper layer is treated by the lower layer as a consequence of the principle of separation of layers.

Because of the requirement of the layer separation principle, a lower layer does not impute intelligence or information function to the payload data received from the upper layer. A lower layer is "stupid" with respect to the data from the upper layers because, by design, the lower layer treats or is supposed to treat the payload as an encapsulated stuff the content of which it does not know nor wants to know about. Thus, the lower layer, by design, cannot or is not supposed to discriminate the payload from the upper layer based on its content. Nor is the lower layer allowed to modify the content. Hence, the lower layer is transparent with respect to the upper layer.

Furthermore, the expectation of transparency is an essential and inherent part of the workings of the layered protocol architecture. The horizontal communication within a layer cannot take place reliably unless it is expected that the lower layers will not modify or discriminate—that is, horizontal communication with respect to any given layer requires that all of the layers that are lower in the vertical hierarchy are transparent to the higher layer. Horizontal communication requires vertical transparency.

c. Layers, End-to-End, Transparency, and Innovation on the Internet

Transparency is a consequence of the layers model. Transparency means that the Internet is a neutral platform. Anyone can develop network applications with or on top of the TCP/IP protocol, and they will run on all of the networked computers and across all of the routers in the world running the TCP/IP protocol. All of the hun-

101 Saltzer et al., *supra* note 80, at 287.

dreds of millions of computers connected by the global network—including the high-speed Internet backbones—a colossal computational resource that far surpasses even those owned by the largest nations in the world, in effect become a commons that is placed at the disposal of every end user and developer. No permission is necessary for anyone to develop applications on top of the TCP/IP layers, or for the application to run on the vast global network that is the Internet.

Furthermore, guided by the end-to-end principle, nearly all user functions are implemented at the upper, application layer. Users more or less know what functionality they want, and developers use the information at the application layer to provide the functionality that users want. Thus, innovation is decentralized and placed in the hands of individual innovators. The Internet has become an innovation commons that has enabled the most powerful and diverse spur to innovation in modern times.[102]

It should be noted that much more than the architecture of the Internet is responsible for the rate and significance of the innovations associated with the growth of cyberspace. Open access to protocol documentation, to the code that implements it, to development tools, and to computer operating systems that support development and implementation, were all key ingredients. These factors for the production of innovation were a set of commons that coordinated with the Internet. Moreover, the substantial market incentives for Internet innovations are likely to have played a significant role in determining the pace of innovation. However, if those factors had been in place, but the fundamental architecture of the Internet had not enabled innovation, then, it seems likely, indeed all but certain, that the rate of innovation would not have been as great. As we have already seen through the example of the development of the World Wide Web, without transparency the Internet would have been qualitatively and quantitatively different. The external social benefits of the Internet have been, in large part, a product of its architecture. It is, of course, difficult to estimate the magnitude of these social benefits. As explained by Ira Magaziner, President Clinton's senior policy advisor, "almost two-thirds of the real growth of the U.S. economy [during the mid- and late-1990s came] from the Internet economy."[103]

C. Communication System Layers

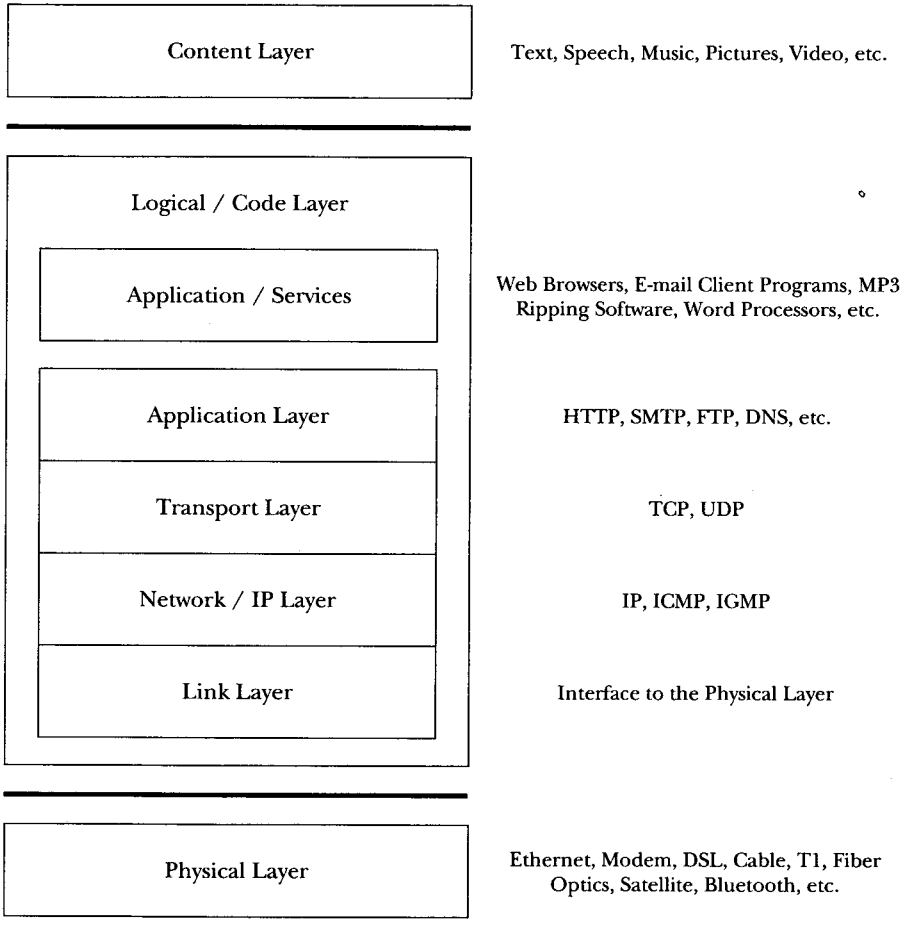
Another advantage of looking at Internet architecture in terms of layers is that the model fits more naturally with the layers framework

102 *See* LESSIG, *supra* note 1, at 5–23.

103 Magaziner, *supra* note 79, at 1166.

in a larger communication systems context—the layers framework proposed by Yochai Benkler. Benkler suggests that we understand a communication system by dividing it into three distinct layers. At the bottom is a physical layer, in the middle a logical layer, and at the top a content layer. The logical or the code layer, as Lessig calls it, is a software layer that includes the TCP/IP protocol layers, application software, and/or services.[104] This idea is illustrated by Figure 5:

FIGURE 5. TCP/IP LAYERS WITHIN COMMUNICATION SYSTEM LAYERS



Thus, our approach to layers analysis integrates the TCP/IP layers within a generalized communication systems layers framework. This integrated framework enables an analysis that uncovers important issues—issues that might have otherwise been overlooked—arising from the interaction of the TCP/IP layers with the content layer

104 *See* Benkler, *supra* note 13, at 562.

above and the physical layer below. We explore these implications in Part III, when we apply the layers principle to both the narrow conception of TCP/IP layers and the broader conception of layers in a communications system. At this point, we turn from engineering to policy and flesh out our development of the layers principle.

II. THE LAYERS PRINCIPLE: RESPECT THE INTEGRITY OF THE LAYERS

A. Explication of the Layers Principle

In this section, our aim is to provide both a general and abstract conception of the layers principle and a relatively particular and concrete set of guidelines of the implementation of that conception.

1. A First Approximation of the Layers Principle

We have already provided a preliminary statement of the layers principle: respect the integrity of the layers. We can now provide a fuller statement of the principle and a richer explanation of its meaning. The layers principle itself is general and abstract. It conveys a fundamental aim of Internet regulation. The two corollaries to the layers principle, layer separation and the minimization of layer-crossing regulation, operationalize this abstract goal, offering a concrete, action-guiding statement of its implications.

What normative ideal is expressed by the layers principle itself? The injunction to respect the integrity of the layers expresses a goal for Internet regulation. Regulation of the Internet should aim to avoid interference with the layered nature of Internet architecture. This aim is expressed as a negative—and not a positive—injunction.[105] That is, the layers principle tells regulators to avoid regula-

105 Of course, there is a positive flip-side of the negative injunction: the Internet should be layered, but this positive goal cannot (or should not) be pursued by Internet regulators. Take an extreme case. Imagine a statute that requires respect of the layers principle, and provides a private right of action for injured parties against the members of the Internet Engineer Task Force (or other body) that violated the principle. This would, quite obviously, be a terrible idea, and not just because it involves the courts. A similar provision enforced by a regulatory agency would be almost as bad—as would direct legislative mandates of layering (such as a statute that prohibited the use of a particular version of TCP/IP). Judges, regulators, and legislators lack the institutional capacity to make sound decisions regarding the implementation of the layers principle. *See generally* Cass R. Sustein & Adrian Vermeule, *Interpretation and Institutions*, 101 MICH. L. REV. 885 (2003). Similarly, adversary proceedings, notice and comment rulemaking, and interest group legislating are poor processes for deciding how to implement layering.

tions that would compromise the architecture of the Internet, but it does not tell regulators how they should achieve their policy goals.[106]

To whom is the layers principle addressed? The layers principle can be interpreted broadly, so as to apply to any actor with the power to affect Internet architecture, or narrowly, so as to apply to a more circumscribed set of Internet regulators. On the broad interpretation, the layers principle addresses software designers and network administrators—that is, it reaches what constitutional theorists might call private actors. On the narrow interpretation, the layers principle addresses governments or quasi-governmental entities, which we shall call *public Internet regulators.* This set includes international organizations with regulatory power, national and subnational legislatures, executive and administrative policymakers, as well as judges and other adjudicators—state actors in the parlance of constitutional theory. The Internet Corporation for Assigned Names and Numbers (ICANN) and related organizations[107] pose a special case. ICANN is a private, nonprofit corporation, organized under the laws of the State of California. ICANN acts in some ways as a regulator of the Internet, and it performs the so-called Internet Assigned Numbers Authority (IANA) functions, its coordination of the DNS, and the global Root. Likewise, the IETF and the Internet Architecture Board (IAB) play a role in coordinating the design of the TCP/IP protocol. We call ICANN and related organizations *transnational Internet regulators.*[108] For most of the remainder of this Article, we adopt the narrow interpretation of the layers principle, with the caveat that we understand the narrow interpretation to include transnational Internet regula-

106 Because the layers principle is negative, it does not provide a comprehensive framework for the evaluation of Internet policy. *See infra* Part II.A.3.

107 Other organizations related to ICANN for this purpose include the Internet Assigned Numbers Authority (IANA) and the IETF.

108 We realize that this terminology is potentially contentious. There are many different views about ICANN's nature. Here is a very brief survey of the possibilities: (1) ICANN is an agent of the U.S. government, operating on the basis of a contractual or quasi-contractual relationship with the Department of Commerce; (2) ICANN is simply a private entity that operates the IANA functions as a nonprofit private enterprise; (3) ICANN is a "club" or association of private and public Internet actors, including ISPs, backbone operators, Internet service providers, domain name registries and registrars, and other stakeholders; (4) ICANN is a nascent international organization (IO), evolving from one of the three models above, and evolving toward a formal, treaty based IO; or (5) ICANN represents a new paradigm, a transnational regulatory body, that exercises regulatory authority over the Internet without authorization from national governments.

We do not mean to take sides in the debate over ICANN's nature. We believe the phrase "transnational Internet regulator" is at least loosely descriptive of ICANN's nature, but we leave the question as to ICANN's ultimate status for another day.

tors. Our adoption of the narrow interpretation is motivated solely by the scope of our inquiry. We do not believe that private Internet regulators should ignore the layers principle; rather, our analysis for the most part does not extend to their actions.

2. Corollaries of the Layers Principle

The layers principle itself expresses a general and abstract goal for Internet regulation. We now explicate the content of the two corollaries to the layers principle.

a. The Principle of Layer Separation

The first corollary of the layers principle is the principle of layer separation. This principle directs public Internet regulators not to violate or compromise the separation between layers designed into the basic architecture of the Internet. The content of this principle can be expressed as a constraint on the relationship between information and the layers. The principle of layer separation proscribes any regulation that would require one layer of the Internet to differentiate the handling of data on the basis of information available only at another layer. This is the most general articulation of the principle of layer separation; it states the principle at the abstract level of information theory.

The most fundamental threat to the principles of layer separation would be a regulation that required the global ability to differentiate content at the transport layer and/or the IP layer on the basis of information from the content layer and/or the application layer. Imagine, for example, that regulators concerned with peer-to-peer file sharing decided to require Internet service providers or Internet backbone operators to proscribe the transmission of data packets that originated from a P2P application. The current architecture of the Internet makes this difficult or impossible; the data packets do not provide the information in readable form to the physical layer.[109] But, a fundamental change in the architecture of the Internet could make this information available. There could be a central registry for Internet applications, and the header for each packet of data could contain a registry number that identified the packet by application. Going down this road, however, would compromise the principle of layer separation, and hence compromise the transparency of the Internet.

109 So-called "packet sniffing" applications can intercept data packets and may be able to decode them. *See, e.g.*, Steve Gibson, *OptOut: How to Watch Spyware Watching You!*, *at* http://grc.com/oo/packetsniff.htm (last visited Feb. 24, 2004).

b. The Principle of Minimizing Layer Crossing

The second corollary to the layers principle is the principle of minimizing layer crossing. The second corollary directs Internet regulators to minimize the distance between the layer at which the law aims to produce an effect and the layer directly targeted by legal regulation. When we say "minimize the distance" we are employing a spatial metaphor for a logical relationship between the layers of the Internet. Recall that we identify six layers, arranged in a vertical hierarchy from top to bottom: (1) content, (2) application, (3) transport, (4) IP, (5) link, and (6) physical.[110] The closer the layers are in the vertical hierarchy, the smaller the distance between the layers. Adjacent layers are closest to one another. The maximum distance is between the content layer and the physical layer.

This concept can be stated more formally. We can assign each layer an ordinal value corresponding to the layer's position in the vertical hierarchy. The distance between a pair of layers is the absolute value of the difference between their ordinal values.[111] That is, the distance between the first layer, the content layer, and the second layer, the application layer, is one. Similarly, the distance between the third layer, the transport layer, and the sixth layer, the physical layer, is three.

---

110 *See supra* Part II.A.

111 The formulation in text can be notated as $D\,(L_xL_y) = \left|\, O\,(L_x) - O\,(L_y) \,\right|$, where $\|$ is the absolute value function, L_x is layer x and L_y is layer y, D is the distance function for any pair of layers x and y, and O is the ordinal ranking of a given layer, L, corresponding to the following ranking matrix:

Layer	Ordinal Rank
Content	1
Application	2
Transport	3
IP	4
Link	5
Physical	6

The principle of layer separation does not provide a formal mechanism for comparing within the class or regulations that violate the principle. Formally, we might say that layer separation is a binary function, yielding values of "violation" or "no violation." The principle of minimizing layer crossing, by contrast, does permit ranking within the class of layer-crossing regulations. The greater the number of layers crossed, the worse the regulation; the fewer the layers crossed, the better the regulation.

3. A Partial, Not Comprehensive, Set of Principles

Taken together, the principle of layer separation and the principle of minimizing layer crossing constitute a set of tools that yield robust evaluations of possible Internet regulations. Where the two principles do not provide clear guidance, recourse may be had to the more general conception of the layers principle. At this point, however, we should be clear that we have not provided a general theory of Internet regulatory policy.

Our analysis is aimed at a particular evil: violation of the integrity of the layers. What our analysis does *not* include is a framework for evaluating the various rationales that might be provided as justifications for a particular regulation. Of course, our theory does provide a goal for Internet regulators. In some cases, that goal might justify public Internet regulators acting to proscribe or discourage private action that threatens the integrity of the layers. Nonetheless, for the most part, our analysis provides only a negative program for Internet regulation. This suggests a natural question: is our theory incomplete or deficient without a more complete analysis on the positive side?

Ought we to develop a comprehensive theory of the various rationales for Internet regulation? Our answer to this question is no. We give this answer because the Internet is a pervasive medium of communication. The goals of Internet regulation run the gamut of substantive policy goals. These include protection of intellectual property, protection of consumers, safeguarding reputation, preserving privacy, discouraging the exploitation of women and children by pornography, encouraging free expression, and so forth. Because the Internet is a multipurpose, global communications medium, the positive program of theorizing the rationales for Internet regulation is identical with legal and political theory in general.

We avoid this larger enterprise for two reasons. First, and obviously, it is beyond the scope of this Article for practical reasons—the Article is long enough without presenting and defending a general theory of jurisprudence. Second, and more importantly, the ques-

tions of general legal and policy theory are perennially controversial. For example, welfarism contends with law as integrity and virtue jurisprudence. As we argue below, acceptance of the layers principle does not depend on the acceptance of any particular theory in general jurisprudence.[112] Rather, the layers principle can be supported by arguments generated from within a variety of jurisprudential perspectives.

4. The Constraining Force of the Layers Principle

The layers principle and its corollaries provide a powerful evaluative tool for legal analysts and policymakers. So far, however, we have not addressed a crucial question: what should policymakers do when faced with a conflict between the layers principle and some other policy goal? In this subsection, we provide an abstract answer to this question. That abstract answer will become more concrete in Part III below. Although our approach at this stage of our argument is very general, it is not indeterminate.[113] We argue that the layers principle should be entitled to substantial weight in Internet regulatory policy and that, absent compelling justification, marginal benefits do not justify violation of the layer separation principle, even if the apparent marginal benefits outweigh the apparent marginal costs. Where violation is justified, the principle of minimizing layer crossing suggests that regulators may not adopt a layer-crossing regulation if an alternative regulation that crosses fewer layers would achieve the regulator's goal at an acceptable level of cost.

We begin our analysis of the constraining force of the layers principle by examining *incrementalism* as an alternative approach to Internet regulation. The reasons for our rejection of incrementalism lead naturally to the approach we favor, based on the notion of a strong presumption or, metaphorically, a weighted scale. We then explain our reasons for rejecting an alternative approach, based on absolute prohibitions of regulations that violate the layers principle.

a. The Case Against Incrementalism

By incrementalism we mean an approach to policymaking that takes each decision on its own merits, without consideration of the cumulative impact of similar decisions.[114] Incrementalism is often the way to go. By making incremental decisions, we can focus on the *mar-*

112 *See infra* Part II.C.

113 *See* Lawrence B. Solum, *On the Indeterminacy Crisis: Critiquing Critical Dogma*, 54 U. CHI. L. REV. 462, 462, 473 (1987).

114 Incrementalism can be used in another sense to distinguish small-scale changes from large-scale changes. That use is related to ours, but it is not identical.

ginal costs and benefits of each decision—an approach that may offer the best chance of reaching the optimal outcome available in the decision space. In the context of Internet regulation, however, incrementalism is a poor institutional strategy for three reasons: (1) incrementalism leads to a scope of decision problem—the tyranny of small decisions; (2) incrementalism is ill suited to decisions in informational environments characterized by ignorance, that is in situations in which there is uncertainty that cannot be reduced to risk; and (3) incrementalism requires that low-level decisionmakers, legislators, judges, and administrators possess certain institutional capacities that they almost always lack. Each of these three problems is considered in turn.

i. Tyranny of Small Decisions

An incrementalist approach to the impact of Internet regulation on the layered architecture of the Internet might use the familiar decision procedure of ad hoc, case-by-case balancing.[115] For each Internet governance decision, the policymaker would weigh the benefits of the decision against all the costs, including the marginal effect of the decision on the transparency of the Internet. If marginal costs, including a loss of innovation resulting from the marginal reduction in transparency, outweighed the benefits, then a proposed regulation would be rejected; if the benefits exceeded the marginal costs, it would be adopted.

Incrementalism (or case-by-case balancing) assumes an answer to the scope of decision question: On what class of cases should our decisions operate? Should we take each case as it comes or should we adopt a general rule (or principle) that decides (or guides decision in) a class of cases? Incrementalism faces a well known scope of decision problem, sometimes called the tyranny of small decisions.[116] This problem can be illustrated with a classic objection to act-utilitarianism with respect to promise-keeping. Act-utilitarianism takes each decision whether to keep or break a promise as the relevant scope of decision. In each case, the act-utilitarian determines whether keeping or breaking the promise will produce more utility. If we consider only one decision, this procedure is unproblematic, assuming we are interested only in consequences. But if this procedure were adopted gen-

115 *See generally* Eugene Volokh, *The Mechanisms of the Slippery Slope*, 116 HARV. L. REV. 1026 (2003) (describing mechanisms by which a decision now (t^1) can constrain the feasible set of choices at some point in the future (t^2)).

116 Alfred E. Kahn, *The Tyranny of Small Decisions: Market Failures, Imperfections, and the Limits of Economics*, 19 KYLOS 23 *passim* (1966).

erally, it would erode the institution of promising. I could not rely on you keeping your promise, because you will break it for a marginal utility gain. The act-utilitarian might reply that the calculation of utilities should take into account the effect of promise-breaking on the institution of promising, but the critic will reply that any one broken promise will likely produce no marginal effect on the institution itself. Rule-utilitarianism solves this problem by adopting a different answer to the scope of decision question. We decide what general rule with respect to promise-keeping will maximize utility, and then act in accord with the general rule.

A case-by-case approach to Internet transparency would lead to a scope of decision problem. Suppose we are deciding whether to compromise the integrity of the layers in order to address a particular policy goal. Incrementalism would counsel us to weigh the benefits of the proposed layer-violating regulation against the costs, including the effect on transparency. But, for any particular regulation, the effect on transparency is not likely to be appreciable. Application programmers can cope with any particular single compromise in layer separation. So long as there is only one violation (or a very small number of violations), software and network design engineers can program around the violation. But, as the violations increase, the complexity of the programming task increases in a more or less logarithmic fashion. In this way, a series of seemingly cost-beneficial incremental decisions can lead to a pattern of decisionmaking that leads to an outcome in which costs outweigh benefits. In theory, a perfectly informed decisionmaker could avoid this problem; in practice, however, the slippery slope may be unavoidable, given incremental decisionmaking.[117]

ii. Ignorance

Internet regulation is fraught with problems of unforeseen and unintended consequences. Indeed, in a very real sense the central purpose of this Article is to provide a framework for Internet regulation that avoids unintended consequences. The problem of unintended consequences has more than one cause. In part, this problem results from the limited institutional capacities of courts, legislatures, and administrative agencies. We discuss this aspect of the problem below.[118] In part, the problem of unintended consequences is also a function of irreducible uncertainty. We discuss that aspect of the problem in this subsection.

---

117 *See* Volokh, *supra* note 115.

118 *See infra* Part II.A.4.a.iii.

One source of irreducible uncertainty stems from what we call the transparency thesis:[119] layer-violating regulations inherently damage the transparency of the Internet. As we illustrated with the example of Tim Berners-Lee and the World Wide Web,[120] damaging transparency increases the cost of innovation. But how are these costs to be weighed for case-by-case balancing? For any particular decision whether to realize some gain at the cost of damaging transparency, there are several sources of uncertainty as to the magnitude of the cost. First, we do not know what innovative applications may be affected by the transparency loss. Until Tim Berners-Lee conceived of the World Wide Web, there was no way to know that a transparent Internet would enable its invention. Second, we do not know how much cost a particular loss of transparency will impose on any particular innovation. This problem is compounded by the fact that transparency-damaging changes in Internet architecture will interact with each other in unexpected ways. Third, we do not know what value particular innovations will produce. The value created by the World Wide Web did not become apparent until after networking effects had passed a tipping point that resulted in the widespread use of web browsers and the creation of substantial content that could be accessed by users. Even after this point was reached, the magnitude of the value was extremely difficult to estimate—witness the wild swings in market capitalization of Internet based enterprises. The combination of these three sources of uncertainty means that the cost of damaging transparency cannot be reduced to an expected value—none of the values necessary to calculate the amount of the loss discounted by its probability are available to decisionmakers.

These particular uncertainties are familiar to economists in the form of Arrow's work on the valuation of innovation.[121] We cannot know the benefits of innovation until they are realized,[122] but we must value innovation to make investment decisions. In the case of Internet regulation, we cannot know the innovation costs of damaging

119 *See infra* Part II.D.1.

120 *See supra* note 75 and accompanying text.

121 *See* Arrow, *supra* note 78, at 609–26.

122 We believe that in the case of the Internet, the ignorance problem with respect to the benefits of innovation are acute. In other contexts, historical experience may permit policymakers to reduce the uncertainty problem to a problem of risk. For example, pharmaceutical companies may be able to reduce their decisions to invest in research to expected utility (profit) calculations. We assert, although we do not have a strong empirical foundation, that it is intuitively plausible to believe that we lack the basis for making expected utility calculations with respect to the benefits of Internet innovation that might be damaged by transparency-compromising Internet regulations.

the transparency of the Internet, but we must consider those costs when formulating Internet regulatory policy.

Rational choice theory (or decision theory) attempts to give a formal account of how rational beings can make decisions under conditions of uncertainty.[123] Let us distinguish between two different kinds of uncertainty. When we are uncertain about the consequences of our actions but are able to estimate the probabilities of various actions, let us call this kind of uncertainty *risk*. When we cannot estimate probabilities, let us call this kind of uncertainty *ignorance*.[124]

Incrementalism, or case-by-case balancing is an appropriate decision strategy for decisions involving uncertainty as risk. Under conditions of risk, rational choice theory suggests that we should select the alternative with the highest expected utility.[125] We calculate expected utilities by discounting the utility of each possible outcome of a choice by the probability of its occurrence.[126] But, as we have seen, decisions whether to violate the transparency of the Internet are not susceptible to expected utility calculations on a case-by-case basis. For any particular case, we lack sufficient information to estimate the costs and their probabilities. Decision theorists have proposed a variety of strategies for decisionmaking under conditions of ignorance, but for reasons that are beyond the scope of this Article, all of these strategies have significant problems.[127]

All is not lost, however. By changing the scope of decision, we can transform the information space in which the decisions—i.e., whether to compromise the transparency of the Internet—are made. Although in any particular case we have no reliable basis for estimating the magnitude or probability of innovation costs, we do have reason to believe that the cumulative impact of a general practice of sacrificing transparency for marginal benefits will be the imposition of very large costs. For this reason, the uncertainties associated with innovation losses suggest that incrementalism is a poor decision strategy as compared to an approach that works with a large-scale scope of

123 For an account of decision theory, see generally R. DUNCAN LUCE & HOWARD RAIFFA, GAMES AND DECISIONS (1957); and MICHAEL D. RESNICK, CHOICES (1987).

124 *See generally* Lawrence Solum & Stephen Marzen, *Truth and Uncertainty: Legal Control of the Destruction of Evidence*, 36 EMORY L.J. 1085, 1149–50 (1987) (delineating the difference between uncertainty based on risk as when based on ignorance); Lawrence B. Solum, *You Prove It! Why Should I?*, 17 HARV. J.L. & PUB. POL'Y 691, 696–98 (1994) (delineating the difference between uncertainty based on risk versus ignorance).

125 *See* Solum & Marzen, *supra* note 124, at 1151.

126 *See id.*

127 *See* Lawrence B. Solum, *To Our Children's Children's Children: The Problems of Intergenerational Ethics*, 35 LOY. L.A. L. REV. 163, 217–18 (2001).

decision. The point of the layers principle is simply to provide a decisional heuristic that widens the scope of decision.

iii. Institutional Capacity

In addition to the tyranny of small decisions and ignorance, there is a third reason to doubt the efficacy of incrementalism or case-by-case balancing as a decision strategy for legal regulation of the Internet; we call this third reason *institutional capacity*.[128] Our thesis is simple: the agents who make case-by-case decisions about Internet regulation will frequently lack the institutional capacities necessary for consideration of the impact of Internet regulation on transparency on a case-by-case basis. Judges, legislators, and regulators are relatively ill prepared to understand Internet architecture in a subtle, fine-grained way. For this reason, they are more likely to make good decisions by respecting the layers principle as a general rule than by attempting a case-by-case assessment of the effects of particular regulations on the transparency of the Internet.

We believe this point is supported by uncontroversial assumptions about the institutional capacities of legal decisionmakers. Judges, legislators, and regulators are not software engineers. Although these policymakers may possess a variety of virtues, including the ability to grasp new subject matter, the task of making case-by-case judgments about the impact of Internet regulation on transparency is a particularly difficult one. This is because assessing the impact of particular layer-violating regulations on transparency requires the ability to imagine the impact of the regulation in the concrete and particular context of trying to program a new application. Few legal actors have experience as programmers, making this act of imaginative reconstruction particularly difficult.[129] Moreover, this problem is compounded by the two other sources of difficulty for incrementalism: the tyranny of small decisions and the ignorance problem. The combination of all three difficulties suggests that entrusting case-by-

128 *Cf.* Cass R. Sunstein & Adrian Vermeule, *Interpretation and Institutions*, U. CHI. JOHN M. OLIN L. & ECON. WORKING PAPER NO. 156, PUB. L. & LEGAL THEORY WORKING PAPER NO. 28 (2002), *available at* http://papers.ssrn.com/sol3/papers/cfm?abstract_id=320245 (arguing for consideration of institutional capacities in the context of approaches to interpreting legal texts).

129 *See* HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT (Thomas Gilovich et al. eds., 2002); Cass R. Sunstein, *Hazardous Heuristics*, U. CHI. JOHN M. OLIN L. & ECON. WORKING PAPER NO. 165, PUB. L. & LEGAL THEORY WORKING PAPER NO. 33, at 5–9 (2002), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=344620.

case incremental decisionmaking to judges, legislators, and regulators is unlikely to produce reliable outcomes.

It might be argued that the problem of institutional capacity could be remedied by creating a specialized regulatory agency, with expert decisionmakers aided by a staff that included competent network software engineers. There may be particular contexts in which such a body would provide an appropriate policy formulation mechanism, but this strategy is not generally available for Internet regulation. The reason is obvious. The Internet is a general purpose medium of communication. The problems of Internet regulation run the whole gamut of issues faced by the law, from intellectual property and electronic commerce to freedom of speech and privacy. The difficulty with a specialized agency is that it would face a generalized docket of regulatory issues. This problem is compounded by the fact that Internet regulation is divided between local, state, national, and international agencies. Our hypothetical Internet regulatory agency would be faced with a dilemma. On the one hand, the agency could take a narrow view of its jurisdiction—leaving important Internet-affecting decisions to actors with the requisite institutional competence. On the other hand, the agency could take a broad view of jurisdiction—essentially becoming a super-legislature of the Internet. Neither solution is satisfactory. The narrow view leaves the problem unsolved, and the broad view is not feasible, as it would trample the constitutional authority of national and subnational legislatures.

b. A Constraining Framework for Internet Regulation

If incrementalism does not provide an adequate framework for the integration of the layers principle into the policy formulation process, what is the alternative? In this subsection, we argue that the layers principle and its corollaries should be viewed as establishing a framework for analysis of alternative policies for Internet regulation that incorporates a strong presumption against regulations that violate the layers principle. Policymakers who adopt this framework of analysis should consider two questions before adopting a layer-violating regulation: (1) Is the regulatory interest compelling?, and (2) Are layer-respecting regulatory alternatives available? We contend that this framework can provide an adequate basis for institutional utilization of the layers principle as a practical constraint on legal regulation of the Internet. We begin our exposition of these ideas by explaining our notion that this framework for analysis should utilize principles as the fundamental analytic unit.

i. The Status of a "Principle" in Law and Policy

As an expositional heuristic, legal norms can be sorted into three general classes—*rules*, *standards*, and *principles*—that vary in terms of the degree of constraint they impose upon decisions.[130] Because different legal theorists use these terms differently, we will stipulate the following meanings. Rules are the most constraining and rigid. If the layers principle were a rule, then it would strongly constrain decisions. For example, public Internet regulators could impose regulations that would require a lower layer to obtain information from a higher layer without *violating* the rule. Standards provide an intermediate level of constraint. Standards guide decisions but provide a greater range of choice or discretion—for example, a standard may provide a framework for balancing several factors. Principles are less constraining still. As a principle, the layers principle and its corollaries provide mandatory considerations for public Internet regulators. Whereas standards identify an exhaustive set of considerations for adjudication or policymaking, a principle identifies a nonexhaustive set, leaving open the possibility that other considerations may be relevant to the decision.

There is an obvious difficulty with our choice of a principle, rather than a rule or a standard. Because principles provide only weak constraints on decisionmaking, it might be argued that principles are simply not up to the task of providing the necessary guidance to public Internet regulators. In the subsections that follow, we explain how principles can provide meaningful constraints and why we reject the more constraining alternatives, i.e., a rule or standard based approach to the protection of layers.

ii. Compelling Regulatory Justification

The law is familiar with the problems of case-by-case balancing that we have examined.[131] One strategy for decisionmaking that ame-

130 *See* Barnabas Dickson, *The Precautionary Principle in CITES: A Critical Assessment*, 39 NAT. RESOURCES J. 211, 222–23 (1999); Kathleen M. Sullivan, *Foreword: The Justices of Rules and Standards*, 106 HARV. L. REV. 22, 56–59 (1992). Dworkin's important discussion can be found at RONALD DWORKIN, TAKING RIGHTS SERIOUSLY 22–28 (1977); *see also* Larry Alexander & Ken Kress, *Against Legal Principles*, 82 IOWA L. REV. 739, 740 (1997).

131 *See, e.g.*, 44 Liquormart, Inc. v. Rhode Island, 517 U.S. 484 (1996) (Thomas, J., concurring) (criticizing the *Central Hudson* test as "very difficult to apply" because of "the inherently nondeterminative nature of a case-by-case balancing 'test' unaccompanied by any categorical rules"). There is vast literature on this topic. *See, e.g.*, Kathleen M. Sullivan, *Cheap Spirits, Cigarettes, and Free Speech: The Implications of* 44 Liquormart, 1996 SUP. CT. REV. 123, 158.

liorates these problems is the notion of a presumption[132] or, metaphorically, a weighted scale. We propose that the layers principle and its corollaries be treated as presumptive rules of decision. This puts the burdens of production and persuasion[133] on the advocate of a layer-violating regulation. What kind of justification should be required to overcome the presumption? By analogy with constitutional law, we suggest that public internet regulators adopt the following decisional heuristic: before adopting a layer-violating regulation, a regulator must articulate a compelling regulatory justification.

What constitutes a compelling regulatory justification? Our first reaction to this question is that is that it is the wrong question, or that it is not a well framed question. No definite criteria for what constitutes a compelling regulatory justification can be set out in advance, although we wish it could. We can gloss or restate the notion of compelling regulatory justifications. By a compelling regulatory justification, we mean a justification that is truly substantial. The evil to be avoided or the good to be gained should be very large or of great import. One way to clarify what we mean is by analogy to well known standards in constitutional law. The most obvious example is the idea of a "compelling state interest," well known from equal protection doctrine. Another analogy is the "clear and present danger" formulation for freedom of speech. Merely glossing or restating the idea of a compelling regulatory justification can only take us so far—more clarification is needed.

Another strategy for clarifying the meaning of the compelling regulatory interest standard is to give examples of interests that count as compelling and those that do not. On the one hand, we take it that a serious threat to national defense would count as a compelling regulatory interest. On the other hand, we take it that the elimination of content that was merely offensive would not count as a compelling regulatory interest. We could continue the listing of interests that are (or are not) compelling indefinitely, but this procedure does not deal

132 "Presumption" is a slippery term in legal discourse. In the law of evidence, presumptions are sometimes considered "bursting bubbles." *See* R. Alexander Acosta & Eric J. von Vorys, *Bursting Bubbles and Burdens of Proof: Disagreements on the Summary Judgment Standard in Disparate Treatment Employment Discrimination Cases*, 2 TEX. REV. L. & POL. 207 *passim* (1998); Leo H. Whinery, *Presumptions and Their Effect*, 54 OKLA. L. REV. 553, 556 n.14 (2001). That is, we sometimes say that a presumption can be overcome by any evidence, no matter how slender. In other contexts, however, the law treats presumptions as having "weight" and requires that a presumption be overcome by a significant quantum of evidence, e.g., "clear and compelling evidence." We mean presumption in the latter, and not the former, sense.

133 *See* Solum, *supra* note 124, at 700–01.

with the kind of case that is likely to be problematic—one in which the interest is serious, but not clearly compelling.

Marginal cases are the kind in which guidance is needed, but it is precisely those cases where guidance is more difficult to provide. In marginal cases, we offer the following three rules of thumb or decisionmaking heuristics:

Rule One: It should count in favor of an interest counting as compelling that there is widespread social agreement on its importance; it should count against compelling status that there is substantial dissensus about the importance of the value;

Rule Two: It should count in favor of an interest counting as compelling that the interest has a very large economic value; it should count against compelling status that the economic value is small or highly uncertain; and

Rule Three: It should count in favor of an interest counting as compelling that the nature of the interest is qualitatively important (e.g., life, fundamental constitutional rights, and so forth); it should count against compelling status that the interest is qualitatively insubstantial (e.g., convenience, entertainment value, etc.).

We realize that our discussion will strike many readers as inadequate. Our position is that the compelling interest standard is an improvement over ad hoc unstructured case-by-case balancing. It is a reasonable compromise between unguided discretion and a hard and fast rule. It is far better than nothing. Moreover, as public Internet regulators gain experience in applying the standard, its abstract terms should begin to acquire more concrete meaning. In other words, the problem of underdeterminacy[134] should gradually ease.

iii. Layer-Respecting Alternatives

The power of the requirement that violations of the layers principle be justified by a compelling regulatory justification is enhanced when combined with a second requirement: mandatory consideration of layer-respecting alternatives. Thus, we suggest that the layers principle can be implemented at a practical level by adherence to the following decision procedure: before a layer-violating regulation is adopted, decisionmakers should be required to consider the availability of layer-respecting alternatives and alternatives that minimize layer crossing. The fact that a layer-respecting alternative is more costly or less effective should not, by itself, constitute sufficient justification for adopting the layer-violating regulation. Rather, the proponent of the

134 *See* Solum, *supra* note 113, at 473.

regulation should be required to show that considerations of cost or effectiveness are compelling regulatory justifications for rejection of the layer-respecting alternative.

iv. Institutional Difficulties with Principles

By itself, the layers principle is highly abstract and general; it is too soft to guide decisionmaking. Even after we add the two corollaries and the accompanying decision heuristics, we do not have hard-edged rules. Questions remain. What kind of a justification counts as compelling? What is adequate consideration of layer-respecting alternatives? Answering these questions requires the exercise of practical judgment that is sensitive to the regulatory context. More particularly, sound application of our principles requires an appreciation of the importance of the presumptions expressed by the layers principle and its corollaries. The fact that the notion of a compelling regulatory justification requires the sensitive exercise of practical judgment suggests an objection to our suggestion that the layers principle is superior to incrementalism (or case-by-case balancing). If the application of the layers principle requires sound practical judgment and a contextualized assessment of regulatory justifications, then won't the layers principle suffer from the same problems (tyranny of small decisions, ignorance, and institutional incapacity) that plague instrumentalism?

We want to frankly acknowledge these problems. Our formulations are necessarily vague or fuzzy at the edges. There will, of course, be cases in which the asserted regulatory justification is clearly not compelling, or where layer-respecting alternatives provide an obviously adequate alternative. But, there will also be cases in which the application of our preferred interpretation of the layers principle will call for the sensitive exercise of practical judgment by decisionmakers. Given the limited institutional capacities of Internet regulators, we should expect that judges, administrators, and legislators will sometimes err. Given that the best solution, fully-adequate institutions, is not available,[135] we argue that treating the layers principle and its corollaries as strong presumptions represents the second-best solution to the problem of institutional design. But, before we reach that conclusion, we need to consider an important alternative: the possibility that the layers principle could be treated as an absolute rule.

---

135 We put to the side the question whether the institutional capacities of public Internet regulators can be improved, for example, by educational programs, judicial colleges, and so forth.

c. The Rationale for Rejecting Absolutism

Why shouldn't the layers principle be treated as a rule rather than a presumption? Isn't this the natural solution to the tyranny of small decisions, the problem of ignorance, and the limited institutional capacity of Internet regulators? Our answer to these questions is "no." The absolutist approach is undesirable for several reasons. First, as a matter of theory, the values protected by the layers principle do not possess the necessary trumping force that would justify an absolute rule. What are those values? In some cases, rights are at stake. For example, Internet regulations may impinge on freedom of speech. In those cases, the layers principle can aid the analysis of the legal questions, but the trumping force of the rights should be considered on the merits of the particular case. In other cases, the interests at stake are properly assessed as consequences that can be subject to a process of comparative evaluation—even if the information necessary for formal cost-benefit analysis is unavailable.

Second, an absolutist approach is undesirable precisely because of the problem of uncertainty that we have discussed. We believe that the value of transparency is well established. The layers principle is supported by sound considerations of network engineering. But, there is no reason to believe that these principles of network design are written in stone for all time. As the Internet evolves, it is possible that superior architectures may be conceived. Moreover, just as the Internet changed the total electronic communications system, there may be similar revolutionary innovations in the future. An absolute rule (especially a constitutional rule) would be based on the assumption that the general facts on which the argument for the layers principle relies are eternal facts, but we have no reason to believe that this is the case.[136]

Third, an absolutist approach, even if it were justified as a matter of theory, would likely be unavailable as a practical matter. Internet regulators are unlikely to accept the layers principle as an absolute; they are more likely to accept the layers principle as a presumption. That is, the layers principle operates in the world of nonideal theory or the world of the second-best. There are theoretical mechanisms by which an absolutist regime could be institutionalized, ranging from a constitutional amendment to statutory mandates. If, as a practical matter, these options are outside the feasible choice set, the question

136 We acknowledge that rules can be changed as circumstances change, but once entrenched, hard-edged rules may be difficult to change. If the rules are entrenched constitutionally, then rule change is even more difficult. Having recognized these issues, we now set these questions aside.

becomes: What are the best available feasible alternatives? The argument of this Article is that the layers principle is feasible and better than the alternatives that we can envision.

5. A Restatement of the Layers Principle and Its Corollaries

We can restate the layers principle and its corollaries as follows:

The Layers Principle. Public internet regulators should not adopt legal regulations of the Internet (including statutes, regulations, common law rules, or interpretations of any of these) that violate the integrity of the layers absent a compelling regulatory interest and consideration of layer-respecting alternatives.

Corollary One: The Principle of Layer Separation. Public Internet regulators should not adopt any regulation that would require one layer of the Internet to differentiate the handling of data on the basis of information available only at another layer, absent a compelling regulatory interest.

Corollary Two: The Principle of Minimizing Layer Crossing. If compelling regulatory interests require a layer-crossing regulation, public internet regulators should adopt the feasible regulation that minimizes the distance between the layer at which the law aims to produce an effect and the layer directly targeted by legal regulation.

Whereas the preliminary formulations of the layers principle were too abstract to operate as practical principles to guide regulatory action, the restatement of the layers principle and its corollaries are formulated to act as practical principles of action.

B. The Role of the Layers Principle

So far, we have formulated the layers principle and its corollaries as principles for Internet regulators in general. In this section, we make the principles even more concrete by considering their application in a variety of institutional contexts.

1. In Regulation and Legislation

The role of the layers principle in legislation and regulation is relatively straightforward. As we illustrate in Part III, layer-violating regulations are tempting to legislators. Legislation is frequently aimed at problems that occur in the content layer. The layers principle can operate in legislative and regulatory processes in a variety of particular, concrete, institutional ways. Some obvious possibilities are that: (1) legislative and regulatory staff can incorporate the layers principle and its corollaries into their analysis of proposed legislation

or regulatory action; (2) participants in the legislative and regulatory process (e.g., those who provide comments in notice and comment rulemaking or those who testify in legislative committee hearings) can raise arguments for or against particular proposals on the basis of the layers principle; or (3) regulators or legislators who specialize in Internet regulation can become acquainted with the layers principle through staff briefings or other educative mechanisms.

We recognize that the potential for the layers principle to influence legislative and regulatory proceedings is limited by the likelihood that other factors may dominate these processes. Legislation and regulation processes are likely to be dominated by special interest politics on the one hand and ideology on the other. The layers principle is unlikely to prevail if it suggests results that are in direct competition with more powerful causal forces.

Nonetheless, we believe that the layers principle can play a valuable role in lawmaking and rulemaking. First, regulators and legislators are not impervious to rational argument. Second, the layers principle is grounded on arguments that transcend particular ideologies and interests. Internet transparency yields a great social good. Although there may be antitechnology ideologies (Luddites) that would actually approve of barriers to innovation, this is not likely to be a dominant political ideology. Moreover, a variety of different interest groups converge on support for a transparent, innovation enabling Internet.

More simply, our point is that the layers principle can be used in legislative and regulatory debate. Although our presentation of the layers principle is prolix and long-winded—after all, this is a law review article—the principle itself need not suffer from these defects. The layers principle itself can be explained and defended in a manner compatible with the norms of public political debate. The layers principle is a simple and powerful idea. The layers principle is compatible with the values of the mainstream of both the political right and left. The layers principle can appeal to a wide range of economic interest groups.

2. In Adjudication

The layers principle can play an important role in public policy debate, and it can also play a similar role in adjudication. Because legal processes are governed by norms distinct from those that govern legislation and regulation, the presentation of the layers principle to judges will differ from its presentation to legislators and regulators. How different? Our answer to this question is fine-grained. We will

briefly survey the ways in which the layers principle can come to bear in common law adjudication, statutory interpretation, and constitutional adjudication.

a. Common Law Adjudication

We begin with common law adjudication. At first blush, it might seem that the layers principle is an unlikely candidate as a principle or rule recognized by the common law. After all, the common law is based on an accumulation of precedents, and the layers principle is novel in two dimensions. First, the domain of application for the layers principle is the Internet, and the common law has very limited experience with the Internet. Second, the layers principle itself is novel. Common law time and Internet time are measured on scales that vary by orders of magnitude. If the layers principle is comparatively ancient as a principle of Internet design, it is a mere babe by the standards of the common law. It is not surprising then, that only a handful of common law decisions rendered in American jurisdictions have recognized the layered nature of the Internet.[137]

How then might the layers principle come to be recognized by the common law? Widely accepted understandings of common law adjudication suggest a variety of techniques by which the common law could absorb or incorporate the layers principle. Consider the following simplified model of the process by which this could take place:

Step One: A court must decide an issue of common law that in some way implicates the layered nature of the Internet, and recognizes that the integrity of the Internet is one factor that should be

137 *See* British Telecommunications PLC v. Prodigy Communications Corp., 217 F. Supp. 2d 399, 407 (S.D.N.Y. 2002) ("To establish a TCP/IP connection, the transport-layer protocol software initiates a request to connect to a special protocol port of the Web server."); Fed. Trade Comm'n v. Crescent Pub. Group, Inc., No. 00-CIV-6315, 2001 WL 128444 (S.D.N.Y. Feb. 16, 2001):

> "Internet" refers to the global information system that is logically linked together by a globally unique address space based on the Internet Protocol (IP) or its subsequent extensions/follow-ons; is able to support communications using the Transmission Control Protocol/Internet Protocol (TCP/IP) suite or its subsequent extensions/follow-ons, and/or other IP-compatible protocols; and provides, uses or makes accessible, either publicly or privately, high level services layered on the communications and related infrastructure described herein, including, but not limited to, the following forms of electronic communication: electronic mail and email mailing lists, the World Wide Web, USENET newsgroups, Internet Relay Chat, instant messaging, wireless data messaging, FTP/file transfer protocol, and remote computer access from anywhere in the world thereto.

Id. at *2.

considered in determining whether an existing line of cases should be extended to the new situation.

Step Two: A second court faces an identical, similar, or analogous situation, and actually relies upon the layers principle as a necessary or sufficient group for its decision on the facts of the particular case before it.

Step Three: A third or subsequent court faces an identical, similar, or analogous situation, and relies upon the prior court's reliance on the layers principle as evidence that there is a general common law rule or principle requiring respect for the layered nature of the Internet.

This familiar process can lead to the emergence of the layers principle and its corollaries as a principle—an articulated premise for reasoning—by common law courts. The process could be accelerated to the extent that other public Internet regulators recognize and articulate the layers principle.[138]

b. Statutory Interpretation

The layers principle could also play a role in statutory interpretation. Of course, there are many different theories of statutory interpretation. Some theories emphasize strict adherence to the plain meaning of the text; others give more weight to legislative intent or legislative history.[139] Some approaches to statutory interpretation give weight to the general aims or goals of legislation; others allow a variety of policy concerns into the mix of factors that should be considered.[140] We are agnostic about these debates. We concede that some

138 *See generally* GUIDO CALABRESI, A COMMON LAW FOR THE AGE OF STATUTES (1982) (suggesting that the common law can respond to statutory change).

139 *See* Stephen Breyer, *On the Uses of Legislative History in Interpreting Statutes*, 65 S. CAL. L. REV. 845 *passim* (1992) (arguing for the use of legislative history). *But see* Nicholas S. Zeppos, *The Use of Authority in Statutory Interpretation: An Empirical Analysis*, 70 TEX. L. REV. 1073, 1088 (1992) (noting the dearth of support for originalist statutory interpretation).

140 Even an overview of the literature is beyond the scope of this Article. *See, e.g.*, WILLIAM N. ESKRIDGE, JR., DYNAMIC STATUTORY INTERPRETATION (1994); T. Alexander Aleinikoff, *Updating Statutory Interpretation*, 87 MICH. L. REV. 20 (1988) (promoting consideration of legislative history); Frank H. Easterbrook, *Statutes' Domains*, 50 U. CHI. L. REV. 533 (1983) (including legislative history as a factor to consider); Daniel A. Farber, *Statutory Interpretation and Legislative Supremacy*, 78 GEO. L.J. 281 (1989) (describing a preference for adherence to legislative intent); Philip P. Frickey, *From the Big Sleep to the Big Heat: The Revival of Theory in Statutory Interpretation*, 77 MINN. L. REV. 241 (1992) (describing the movement away from sole reliance on legislative intent); Thomas W. Merrill, *Chief Justice Rehnquist, Pluralist Theory, and the Interpretation of Statutes*, 25 RUTGERS L.J. 621, 624 (1994) (emphasizing legislative intent and history

powerful theories of statutory interpretation would rule out use of the layers principle as an aid to construction. Judicial practice, however, is quite varied. Judges tend to be pragmatic in their approach to statutory interpretation, employing a variety of techniques, including straightforward policy analysis.

As a practical matter, then, the layers principle might come to be employed as an aid to statutory interpretation in a variety of circumstances. Here are a few:

Interpreting abstract and general statutory mandates. When a statute related to Internet governance contains broad language, e.g., "in the public interest," the layers principle can give more particular and concrete meaning to the ambiguous statutory command.

Gap filling and tension resolving. When a statutory scheme or a related set of statutes requires resolution of an issue, but fail (due to oversight or deliberate omission) to provide a directive, the layers principle can step in to fill the gap.

Narrowing and broadening. Frequently, statutory language can be read broadly or narrowly without going outside the limits of the text. The layers principle can aid courts in determining whether to give Internet-related legislation a broad or a narrow reading.

Avoiding absurd consequences. Courts frequently read statutes so as to avoid absurd consequences. A reading of a statute that would require a serious violation of layer separation, and hence that would endanger the viability of the Internet, should be considered absurd—unless the evil to be avoided is of the highest order or the statutory language and legislative history clearly indicates that the legislature knew what it was doing.

This is only an incomplete list of the many situations in which the layers principle might be brought to bear on an issue of statutory interpretation.

c. In Constitutional Adjudication

How might the layers principle play a role in constitutional interpretation? One thing is for sure: there is no Layers Clause. Nonethe-

---

over textualism); Richard A. Posner, *Legal Formalism, Legal Realism, and the Interpretation of Statutes and the Constitution*, 37 CASE W. RES. L. REV. 179 (1986) (suggesting that legislative history can be, but does not have to be, utilized); Frederick Schauer, *Statutory Construction and the Coordinating Function of Plain Meaning*, 1990 SUP. CT. REV. 231 (promoting the use of plain meaning); David L. Shapiro, *Continuity and Change in Statutory Interpretation*, 67 N.Y.U. L. REV. 921 (1992) (noting the Supreme Court's shift away from legislative history and toward plain meaning); Cass R. Sunstein, *Interpreting Statutes in the Regulatory State*, 103 HARV. L. REV. 405 (1989) (discussing purpose, intent, and history as tools of interpretation).

less, the layers principle might come to play a role in constitutional interpretation in a variety of contexts. Before considering the specifics, we pause for an obligatory caveat. There are a multitude of approaches to constitutional interpretation, both in the academy[141] and on the bench.[142] Some constitutional theories emphasize the plain meaning of the text; others focus on the original meaning of the text or the original intentions of the Framers. Other approaches emphasize the idea of a living constitution or the notion of contemporary ratification. Some constitutional theories would approve of the use of the layers principle in constitutional interpretation; others would disapprove. Of necessity, we must set these controversies aside. As in the case of statutory interpretation, we assume that most judges are not theoretical purists.

The first and most obvious way in which the layers principle is relevant to constitutional interpretation focuses on the First Amendment freedom of speech.[143] One of the primary themes in free speech doctrine is overbreadth, the notion that a regulation that impacts more speech than is necessary for its regulatory goal is constitutionally suspect.[144] There is a close connection between First

141 Not even the longest, fattest footnote could do justice to this literature. *See generally* Akhil Reed Amar, *Intratextualism*, 112 HARV. L. REV. 747 *passim* (1999); Lino A. Graglia, *"Interpreting" the Constitution: Posner on Bork*, 44 STAN. L. REV. 1019, 1024 (1992); Richard S. Kay, *Adherence to the Original Intentions in Constitutional Adjudication: Three Objections and Responses*, 82 NW. U. L. REV. 226 *passim* (1988); Michael W. McConnell, *Textualism and the Dead Hand of the Past*, 66 GEO. WASH. L. REV. 1127 *passim* (1998); H. Jefferson Powell, *The Original Understanding of Original Intent*, 98 HARV. L. REV. 885, 887–88 (1985); David A. Strauss, *Common Law Constitutional Interpretation*, 63 U. CHI. L. REV. 877 *passim* (1996); Adrian Vermeule & Ernest A. Young, *Hercules, Herbert, and Amar: The Trouble with Intratextualism*, 113 HARV. L. REV. 730 *passim* (2000).

142 *See generally* Frank H. Easterbrook, *Textualism and the Dead Hand*, 66 GEO. WASH. L. REV. 1119, 1119 (1998); Antonin Scalia, *Originalism: The Lesser Evil*, 57 U. CIN. L. REV. 849, 854 (1989).

143 And, freedom of speech is itself the subject of vast literature and an enormous set of contending theories. *See, e.g.*, Lawrence Byard Solum, *Freedom of Communicative Action: A Theory of the First Amendment Freedom of Speech*, 83 NW. U. L. REV. 54 (1989) (proposing a meaning for the phrase "freedom of speech").

144 *See, e.g.*, Broadrick v. Oklahoma, 413 U.S. 601, 611–12 (1973).

> It has long been recognized that the First Amendment needs breathing space and that statutes attempting to restrict or burden the exercise of First Amendment rights must be narrowly drawn and represent a considered legislative judgment that a particular mode of expression has to give way to other compelling needs of society.

Id. For commentary, see Alan K. Chen, *Statutory Speech Bubbles, First Amendment Overbreadth, Improper Legislative Purpose*, 38 HARV. C.R.-C.L. L. REV. 31 *passim* (2003). For an example in the context of the Internet, see Michael Johns, *The First Amendment and Cyberspace: Trying to Teach Old Doctrines New Tricks*, 64 U. CIN. L. REV. 1383, 1418–20

Amendment overbreadth analysis and one of the key supporting arguments for the layer's principle—the fit thesis: "A given lower layer necessarily has substantial innocent use with respect to problems that originate at the upper layers."[145] The connection between the two ideas is intuitive and obvious. One way to regulate content (speech) on the Internet is to attack the source; for example, to require Internet service providers to block the IP addresses of websites that serve up the proscribed content. One of the implications of the fit thesis is that such layer-crossing regulations are inherently overbroad. The second corollary to the layers principle, minimize layer crossing, is a natural fit with First Amendment overbreadth doctrine.

Yet another theme in First Amendment jurisprudence is the idea that regulations of speech on the basis of content are suspect—and that courts may balance the benefits of such regulations against their costs.[146] The layers principle has nothing to say about content regulation. Indeed, there is a sense in which the layers principle runs against the grain of First Amendment doctrine, which generally favors content-neutral time, place, and manner restrictions over content-based regulations. Thus, it might be argued that free speech doctrine allows regulations that violate the integrity of the layers—so long as the regulation is not targeted at particular viewpoints or subject matters. This is true, as far as it goes.

However, when a content-based regulation also violates the layers principle, then layers analysis can play a role. Because content-based regulations trigger a balancing test, the layers principle can play a role by assisting the courts in the difficult task of making open-ended bal-

---

(1996). For a discussion of overbreadth and underinclusiveness, see Kenneth W. Simons, *Overinclusion and Underinclusion: A New Model*, 36 UCLA L. REV. 447, 451–55, 482–88 (1989).

145 *See infra* Part II.D.2.

146 Geoffrey R. Stone, *Content Regulation and the First Amendment*, 25 WM. & MARY L. REV. 189, 190 (1983) [hereinafter Stone, *Content Regulation*] ("The Supreme Court tests the constitutionality of content-neutral restrictions with an essentially open-ended form of balancing."). *See generally* Erwin Chemerinsky, *Content Neutrality as a Central Problem of Freedom of Speech: Problems in the Supreme Court's Application*, 74 S. CAL. L. REV. 49 (2000); Daniel A. Farber, *Content Regulation and the First Amendment: A Revisionist View*, 68 GEO. L.J. 727 (1980); Steven J. Heyman, *Spheres of Autonomy: Reforming the Content Neutrality Doctrine in First Amendment Jurisprudence*, 10 WM. & MARY BILL RTS. J. 647 (2002); Paul B. Stephan III, *The First Amendment and Content Discrimination*, 68 VA. L. REV. 203 (1982); Geoffrey R. Stone, *Content-Neutral Restrictions*, 54 U. CHI. L. REV. 46 (1987); Geoffrey R. Stone, *Restrictions of Speech Because of its Content: The Peculiar Case of Subject-Matter Restrictions*, 46 U. CHI. L. REV. 81 (1978); Susan H. Williams, *Content Discrimination and the First Amendment*, 139 U. PA. L. REV. 615 (1991); Note, *Content Regulation and the Dimensions of Free Expression*, 96 HARV. L. REV. 1854 (1983).

ancing determinate enough to guide decisions in individual cases. In particular, the layers principle suggests that layer-violating regulations that are content-based should only be upheld if there is a compelling regulatory justification[147] and no layer-respecting alternative is available.[148]

A second possible avenue of incorporation of the layers principle into constitutional doctrine is via the dormant Commerce Clause.[149] Dormant Commerce Clause doctrine has, for some time, been in a state of confusion, but one theme in the cases is the notion that the states are not permitted to regulate in ways that impose undue burdens on interstate commerce. In *American Libraries Ass'n v. Pataki*,[150] the Southern District of New York struck down a New York statute that prohibited the intentional use of any "computer communications system" to transmit to minors content that "depicts actual or simulated nudity, sexual conduct or sadomasochistic abuse, and which is harmful to minors."[151] The court found that the statute violated the dormant Commerce Clause in three ways:

> First, the practical impact of the New York Act results in the extraterritorial application of New York law to transactions involving citizens of other states and is therefore per se violative of the Commerce Clause. Second, the benefits derived from the Act are

147 *See supra* Part II.A.4.2.ii.

148 *See supra* Part II.A.4.2.iii.

149 *See* Goldsmith & Sykes, *supra* note 40, at 788–91; Ari Lanin, Note, *Who Controls the Internet? States' Rights and the Reawakening of the Dormant Commerce Clause*, 73 S. CAL. L. REV. 1423 *passim* (2000); *see also* Lemley & Lessig, *supra* note 32, at 10. Lemley and Lessig note that

> [t]he principle of End-to-End is not unique to computer networks. It has important analogs in American constitutional law and in other legal contexts. Vis-à-vis the states, for example, the dormant commerce clause imposes an End-to-End design on the flow of commerce: No state is to exercise a control over the flow of commerce between states; and the kind of control that a state may exercise over commerce flowing into that state is severely limited. The "network" of interstate commerce is to be influenced at its ends—by the consumer and producer—and not by intermediary actors (states) who might interfere with this flow for their own political purposes. Vis-à-vis transportation generally, End-to-End is also how the principle of common carriage works. The carrier is not to exercise power to discriminate in the carriage. So long as the toll is paid, it must accept the carriage that it is offered. In both contexts, the aim is to keep the transportation layer of intercourse simple, so as to enable the multiplication of applications at the end.

Id.

150 969 F. Supp. 160 (S.D.N.Y. 1997).

151 *Id.* at 163 (quoting N.Y. PENAL LAW § 235.21 (McKinney 1997)).

> inconsequential in relation to the severe burdens it imposes on interstate commerce. Finally, the unique nature of cyberspace necessitates uniform national treatment and bars the states from enacting inconsistent regulatory schemes. Because plaintiffs have demonstrated that they are likely to succeed on the merits of their claim under the Commerce Clause and that they face irreparable injury in the absence of an injunction, the motion for a preliminary injunction is granted.[152]

Layers analysis is relevant to the dormant Commerce Clause in at least two respects. First, layers analysis provides a framework for the identification of the burdens on commerce created by state laws that regulate the Internet. Any state regulation that compromises layer integrity would substantially affect the global Internet. Second, layers analysis can play a role in dormant Commerce Clause balancing that is similar to the role suggested with respect to First Amendment balancing, above.

The layers principle might also be relevant to equal protection analysis. Assuming the standard three tiers of equal protection scrutiny, we can reach the following tentative conclusions. First, if a regulation simultaneously discriminated against a suspect or quasi-suspect class *and* violated the layers principle, then layers analysis might be highly relevant to the question whether the regulation in question possessed the necessary degree of fit between ends and means. In particular, the fit thesis suggests that layer-crossing regulations are likely to be overinclusive and underinclusive.[153]

3. In the Design of Institutions for Internet Governance

It goes without saying that the layers principle is relevant to the decisions made by transnational Internet governance institutions, such as ICANN or the IETF. These institutions are populated by network engineers; they understand the importance of transparency and layer separation better than we do and far better than governmental Internet regulators ever will. The layers principle is already a norm within these institutions. We shall not preach to the converted.

We shall, however, step back and briefly consider the implications of the layers principle for the design of institutions for Internet governance. Transnational Internet governance institutions, such as the IETF and ICANN, should be organized so that their institutional design is consistent with their role as the guardians of a transparent Internet.

---

152 *Id.* at 183–84.

153 *See infra* Part II.D.2.

There are, however, reasons to believe that this role might be compromised. One threat comes from the movement to democratize Internet governance structures.[154] Undoubtedly, many of the advocates of Internet democracy favor a transparent Internet and would endorse the layers principle. There are, however, reasons to believe that Internet democratization might actually undermine, rather than enhance, the ability of Internet governance institutions to safeguard the integrity of the layers. There are a number of reasons for believing that this might be true.

Advocates of Internet democracy may support direct elections of the governing bodies of institutions such as ICANN. Such elections are subject to capture by various interest groups that may have goals that conflict with the layers principle. For example, national governments may favor national control over the Internet. Commercial interests, e.g., the intellectual property industry, wish to reconfigure the Internet's architecture so as to permit greater control of content by intermediaries, such as Internet service providers.

Even populist Internet democracy might threaten the layers principle. For example, popular majorities may strongly favor content regulations, including regulations on child pornography, other adult content, hate speech, radical speech, and so forth. If populist groups were able to control Internet governance institutions through democratic elections, they might deliberately compromise the principle of layer separation in order to facilitate content regulation.

Democratic elections are likely to bring a different set of institutional competences to the governing bodies of transnational Internet regulators such as ICANN. Whereas the current ICANN governance structure favors Internet insiders, democratic elections might result in the election of populists with little or no technical expertise. The elected leadership would be less likely to appreciate the importance of layer separation and the various ways in which it could be damaged.

In the end, democratization may or may not threaten the ability of transnational Internet governance institutions to safeguard the integrity of layers. Our argument is very modest. We believe that the layers principle ought to be a consideration of fundamental importance in designing Internet governance institutions. That is, we be-

154 *See* Internet Democracy Project, *Homepage*, *at* http://www.internetdemocracyproject.org (last visited Feb. 11, 2004) ("The Internet Democracy Project seeks to enhance the participation of Internet users worldwide in non-governmental bodies that are setting Internet policy and to advocate that these bodies adhere to principles of open participation, public accountability and human rights.").

lieve that Internet governance institutions should be biased in favor of maintaining the integrity of layers.

C. *The Foundations of the Layers Principle*

What are the deep foundations of the layers principle? That is, how is the layers principle grounded by general jurisprudence, political philosophy, and moral theory? These are questions that for the most part we wish to avoid and evade. In this section, we will lay out the reasons for our failure to deliver the theoretical goods.

Our fundamental reason for avoiding foundational questions is that such questions are endlessly controversial. In legal theory, consequential theories like welfarism[155] contend with fairness based approaches such as "law as integrity"[156] and aretaic theories such as "virtue jurisprudence."[157] In political philosophy, egalitarian theories like "justice as fairness"[158] do battle with communitarian theories,[159] Hobbesian approaches,[160] and libertarian theories.[161] In moral theory, virtue ethics[162] contends with deontology[163] and utilitarianism.[164] None of these great historical debates over normative theories of law, politics, or ethics is likely to be settled soon. Moreover, issues at each of the three levels (law, politics, and morality) can affect the resolution of controversies at other levels. We want to be clear. We are not endorsing relativism or taking a stand against normative objectivity. We are simply stating the obvious. Arguing for a normative theory is outside the scope of this Article.

Instead, we will suggest (but not mount a full argument) that the layers principle can be supported from within a variety of theoretical perspectives. In other words, the layers principle can be the subject of

---

155 *See generally* LOUIS KAPLOW & STEVEN SHAVELL, FAIRNESS VERSUS WELFARE (2002).

156 *See generally* RONALD DWORKIN, LAW'S EMPIRE (1986).

157 *See generally* Lawrence B. Solum, *Virtue Jurisprudence: A Virtue-Centered Theory of Judging*, 34 METAPHILOSOPHY 178 (2003).

158 *See* JOHN RAWLS, A THEORY OF JUSTICE 3–53 (1971).

159 *See* MICHAEL J. SANDEL, LIBERALISM AND THE LIMITS OF JUSTICE *passim* (2d ed. 1998).

160 *See* S.A. LLOYD, IDEALS AS INTERESTS IN HOBBES'S LEVIATHAN: THE POWER OF MIND OVER MATTER *passim* (1992).

161 *See* RANDY E. BARNETT, THE STRUCTURE OF LIBERTY: JUSTICE AND THE RULE OF LAW 63–83 (1998).

162 *See* ROSALIND HURSTHOUSE, ON VIRTUE ETHICS *passim* (1999).

163 *See* IMMANUEL KANT, GROUNDWORK OF THE METAPHYSICS OF MORALS 48 (Mary Gregor trans., 1998).

164 *See* JEREMY BENTHAM, AN INTRODUCTION TO THE PRINCIPLES OF MORALS AND LEGISLATION 1–7 (Athlone Press 1970) (1781).

what John Rawls calls an overlapping consensus[165] and Cass Sunstein calls an incompletely theorized agreement.[166] We assert that this is so with some confidence, because the case for the layers principle is quite powerful.[167] Consequences count for almost every plausible account of law, politics, or morality. Damaging the integrity of layers would have very bad consequences, and almost any sensible theoretical approach to Internet regulation must take that fact into account. Similarly, any sensible account of regulation will count alignment of means and ends as a good; the fit thesis demonstrates that layer-crossing regulations are unlikely to cure the evil they target and are likely to have unintended ill effects.

There is one final reason for our confidence that the layers principle can be supported by those who adhere to a wide variety of theoretical perspectives on moral, political, and legal theory. We have formulated our normative injunction as a principle, as opposed to a rule or a standard. This formulation creates a good deal of play in the normative joints. It allows those who might oppose the rigid enforcement of a layers rule to support a principle that can accommodate concerns outside the scope of the layers principle. Of course, some theoretical perspectives may find resources within their theories to support a more robust version of the layers principle. It may be that given a particular legal theory, the case can be made that public Internet regulators should never violate the integrity of the layers. We neither agree nor disagree with such views. Instead, we rest our case on the shallow arguments that we make in this Article. Our decision not to go deep is a deliberate one; hence there is no more for us to say about the ultimate moral and political foundations of the layers principle.

---

165 *See* JOHN RAWLS, POLITICAL LIBERALISM *passim* (1993).

166 Cass R. Sunstein, *Incompletely Theorized Agreements*, 108 HARV. L. REV. 1733, 1735–36 (1995).

167 We should note, however, that different theories of law, politics, and morality will assign different roles to the layers principle. Here is one example. We have argued that the layers principle might be relevant to statutory interpretation. But, some views of statutory interpretation assign priority to the plain meaning of statutory language or to the intentions of the legislators as revealed in legislative history. It could well be the case that the concerns identified by layers analysis, while relevant to the interpretation of statutory language, cannot easily be accommodated by plain meaning or legislative intent theories. This would move the layers principle from the courts to the legislature. Similar points could be made about the role of the layers principle in the common law or in constitutional interpretation.

D. A Summary of the Case for the Layers Principle

We are at a turning point in our argument. We can now summarize the case for the layers principle, picking out two theses, the fit thesis and the transparency thesis, as the central premises of the first half of our argument.

1. The Transparency Thesis

Our purpose here is to tie together the threads of our transparency argument. We begin by making clear and explicit precisely what we mean by transparency. Initially, the idea of transparency was discussed in the context of the IP layer (or network layer) of the TCP/IP protocol suite. In this setting, the "transparency" refers to the original concept of the function of the IP layer where the data packets flow from a source IP address to a destination essentially unaltered.[168] However, as we discussed above, transparency is a more general concept applicable to all TCP/IP layers. A generalized test of transparency may be given as follows:[169] (1) the lower layer does not access or otherwise analyze the content of the payload received from the upper layer; and (2) the lower layer does not process the data differently based on the upper layer information.

We have already demonstrated that layer-violating regulations inherently compromise transparency. We have also established that the transparency of the Internet produces a very great social good, an innovation commons that has served as an engine of economic growth and cultural evolution. And the layers principle, especially the first corollary, formulates the conclusion of this argument in the form of the injunction: do not regulate the Internet in a way that would violate the integrity of the layers.

The transparency thesis summarizes this argument. We can state it as follows: the fact that layer-violating regulations damage transparency, combined with the fact that Internet transparency lowers the cost of innovation, provides compelling support for the principle of layer separation; public Internet regulators should not violate or compromise the separation between layers designed into the basic architecture of the Internet.

The transparency thesis is one of the two pillars of our case for the layers principle. We now turn to the second.

---

168 Internet Transparency, *supra* note 90.

169 Hans Kruse et al., *The InterNAT: Policy Implications of the Internet Architecture Debate*, PROC. 28TH RES. CONF. ON COMMUNICATIONS, INFO. & INTERNET POL'Y 9 (2000), *available at* http://www.csm.ohiou.edu/kruse/publications/InterNAT_v4.pdf.

2. The Fit Thesis

Layer-crossing regulations—regulations at one layer that attack a problem at another layer—are undesirable for a second reason, independent of their effects on transparency. This second problem is a problem of fit between regulatory ends and regulatory means. It stems from the following fact: a given lower layer necessarily has substantial innocent use with respect to problems that originate at the upper layers.

This fact is a necessary result of the separation of layers and consequent transparency that are inherent in the layers model of Internet architecture. As we discussed above, a lower layer in the Internet layers model is required by architecture to be transparent to the upper layers. The lower layer does not know, or is not supposed to know, about the content of the payload data received from the upper layers. Consequently, the lower layer must necessarily have substantial innocent use with respect to the problems that originate at the upper layers.

Furthermore, because the lower layer is not supposed to, or expected to, know about the upper layer information, the upper layers, by design, do not pass down to the lower layer complete information necessary to perform effective discriminatory functions. This is just another aspect of the separation of layers. Thus, by architecture, the lower layers lack the information to effectively perform discriminatory functions with respect to the problems that arose at the upper layers. Therefore, for the same reason of information mismatch discussed above in the end-to-end argument section, layer-violating regulations—regulations targeted at a lower layer to discriminate against the problems in the upper layer—are inherently overinclusive and underinclusive.

We can summarize this argument as the fit thesis, which can be stated as follows: the fact that layer-crossing regulations result in inherent mismatch between the ends such regulations seek to promote and the means employed implies that layer-crossing regulations suffer from problems of overbreadth and underinclusion; avoidance of these problems requires Internet regulators to minimize the distance between the layer at which the law aims to produce an effect and the layer directly targeted by legal regulation.

Together, the transparency thesis and the fit thesis summarize our case for the layers principle. We now move from theory to practice.

III. APPLYING THE LAYERS PRINCIPLE

Does the layers principle yield attractive results when applied to particular cases? Does layers analysis illuminate problems of Internet regulation or does it instead obscure them? We answer these questions in the discussion that follows. We organize our discussion by categorizing the nature of the layer violation. At the macro level, we distinguish between violations that occur in the TCP/IP layers[170] and those that occur in the more generalized communication system layers.[171] This distinction is important. The TCP/IP layers express the architecture of the Internet. The communications system layers are analogous in some respects, but different in one crucial way: the communication systems layers are not the architecture of the Internet.

A. Application at the TCP/IP Layers

In the discussion that follows, we discuss particular regulations that fail to respect the integrity of the TCP/IP layers. We apply the layers principle and its corollaries to these regulations, and show that layers analysis is normative, attractive, and illuminating. We can begin with a quick overview of the purposes that would motivate regulators to target the TCP/IP layers.

1. Introduction: Purposes of Regulations Directed at the TCP/IP Layers

We introduce our discussion of examples with a preliminary question: Why are Internet regulators tempted to violate the layers principle? One kind of regulation that directly violates the architectural principles of the Internet is a regulation directed at a lower TCP/IP layer in order to address problems that originated at an upper layer—usually, although not necessarily, the content layer. This type of regulation is almost always motivated by the desire to aggressively deal with problems by targeting the technology that enables the undesirable conduct. It is a sort of "cut the problem down at its knees" approach.

Consider, for example, the ripping program that creates MP3 files from music CDs. Without the Internet, all one can do with the ripping software is download the MP3 files to a device like Rio or iPod (MP3 players for personal use), or users can mix and burn the music files (also by and large for personal use). But, the Internet changes this picture fundamentally by enabling mass distribution of MP3 files through peer-to-peer file sharing programs, such as the now defunct

170 *See infra* Part III.A.
171 *See infra* Part III.B.

Napster, KaZaA, and so forth. The net effect is that, because of the Internet, the ripping application is now a part of an extremely efficient global distribution system of digital music (although the ripping application is not technically a part of a network protocol in the sense that the HTTP or SMTP protocol is).[172] Looked at this way, it is not surprising at all that the stakeholders of vested interests, such as the music distribution industry, believe that the Internet represents a threat. When faced with a sudden, imminent threat, it is understandable that the reflexive reaction is to target the enabling technology, the Internet. In their aggressive "cut the problem down at its knees" approach, the industry might seek to persuade public Internet regulators to adopt legal norms that target one or more of the TCP/IP layers.

Another motivation for violating the integrity of the layers is the desire of regulators to reassert control over content. Because of the global reach of the Internet and its fundamental architecture of transparency, the Internet has become an extremely powerful medium to disseminate any type of speech across the globe. As Lessig insightfully pointed out, "[w]e have exported to the world, through the architecture of the Internet, a First Amendment *in code* more extreme than our own First Amendment *in law*."[173] As we know, many policymakers and stakeholders find such power of the Internet threatening, or at least objectionable. It is no surprise, then, that many of the reactions have been to target the networking (TCP/IP) layers to impair or destroy the transparency of the Internet. Although these reactions are understandable, it is quite another matter whether such regulations violate the architectural principles of the Internet are wise or desirable on balance for everyone concerned.

Yet another example of the desire to control content is provided by regulations that seek to block all data from a particular IP address due to the nature of the content coming from the site. Some filtering programs operate this way. The proxy servers and IP filters can be configured to do the same. These programs target or operate at the IP layer, and discriminate the data based on the nature of their content. Instead of treating the data received from the content layer as payload, the blocking programs either analyze the payload data or make assumptions about the nature of the content based on the IP address, then discriminate—i.e., allow or drop the data—on the basis of the nature of the content of the payload data.

---

172 LESSIG, *supra* note 1, at 123, 123–26.

173 LESSIG, *supra* note 15, at 167.

We now move to an even more concrete level, by discussing the application of the layers principle to a variety of specific problems in Internet regulation. All of the examples involve highly controversial regulations or disputes, the analysis of which is also disputed or controversial. For each of the examples, we shall demonstrate that layers analysis—i.e., analysis under the layers framework introduced above—clarifies the issues or brings out important issues that have been outside the reach of existing methods of analysis. Although our discussion is organized by the nature of the layer violation implicated by the regulation discussed, it may be helpful to the reader to present a brief descriptive preview. We will discuss the following examples:

(1) The Serbian Internet Interdiction Myth:[174] We discuss the implications of the urban legend that the Clinton administration attempted to interdict Serbian access to the Internet during the NATO campaign against Serbia.

(2) Myanmar's "Cut the Wire" Policy:[175] We examine Myanmar's policy of limiting Internet access by strictly controlling physical links to the Internet.

(3) China's Great Firewall:[176] We analyze China's strategies for blocking objectionable content.

(4) The French *Yahoo!* Case:[177] We investigate the French government's attempts to force Yahoo! to exclude French end-users from Nazi paraphernalia auctions.

(5) Cyberterrorism:[178] We speculate about the use of TCP/IP based regulations to control cyberterrorist activities.

(6) Pennsylvania's IP Address Blocking Child Pornography Statute:[179] We review Pennsylvania's statute that requires destination ISPs to block the access of Pennsylvania customers to the IP addresses of servers that provide child pornography.

(7) Port Blocking and Peer-to-Peer File Sharing:[180] We hypothesize on the use of port blocking as a strategy for the control of peer-to-peer file sharing of copyrighted content.

(8) Regulation of Streaming Video at the IP Layer:[181] We lay out the issues raised by regulations at the TCP/IP layer to control streaming video content.

---

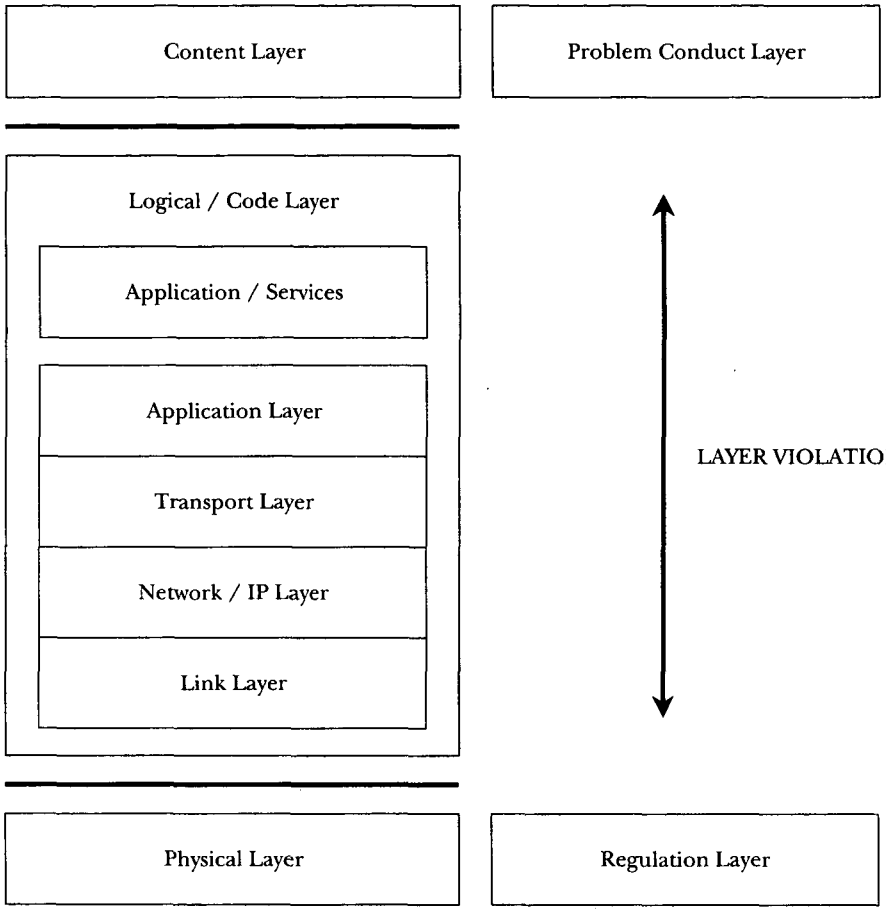
174 *See infra* Part III.A.2.a.
175 *See infra* Part III.A.2.b.
176 *See infra* Part III.A.3.b.
177 *See infra* Part III.A.3.c.
178 *See infra* Part III.A.3.d.
179 *See infra* Part III.A.3.e.
180 *See infra* Part III.A.4.a.
181 *See infra* Part III.A.5.

2. Physical Layer Regulation Aimed at Content Layer Problems

The first category of layer-crossing regulations comprises regulatory actions that target the physical layer in order to solve a problem at the content layer. The most extreme example that falls within this category would be disabling or denying the physical connection—"cutting the wire"—due to content layer problems. Nations at war, for example, might resort to such extreme measures. A blockade against a nation could include denial of communication services under a severe embargo. Or, a politically threatened, unstable regime might cut off communication links to the outside world, including Internet connections. In a connected world, where the communication network is rapidly becoming an essential part of the everyday functioning of society, the network itself could come under attack as a result of very serious political conflicts.[182] Under the layers framework, these regulations directed at the physical layer due to problems at the content layer can be represented by Figure 6 below:

182 *See* William Yurcik & David Doss, *Internet Attacks: A Policy Framework for Rules of Engagement*, PROC. 29TH RES. CONF. ON COMMUNICATIONS, INFO. & INTERNET POL'Y 2 (2001), *available at* http://www.sosresearch.org/publications/tprc01.PDF; *see also* INST. FOR SECURITY TECH. STUD. AT DARTMOUTH C., CYBER ATTACKS DURING THE WAR ON TERRORISM: A PREDICTIVE ANALYSIS (2001), *available at* http://www.ists.dartmouth.edu/ISTS/counterterrorism/cyber_a1.pdf (using case studies to discuss attacks on the Internet).

FIGURE 6. REGULATION DIRECTED AT THE PHYSICAL LAYER DUE TO PROBLEMS AT THE CONTENT LAYER



We begin by discussing the possibility that physical links to the Internet could be cut during a war—using the NATO campaign against Serbia as the focus for discussion.[183] We then address Myanmar, a nation that drastically limits its physical links to the Internet.[184] We wrap up this discussion by generalizing from the two examples.[185]

a. Internet During Wartime: The Case of Serbia

During the NATO war against Serbia in 1999, the possibility of the United States cutting off Internet connection to Yugoslavia be-

183 *See infra* Part III.A.2.a.
184 *See infra* Part III.A.2.b.
185 *See infra* Part III.A.2.c.

came a highly visible, if short lived, issue among the Internet user community.[186] A Belgrade based ISP, BeoNET, claimed that Loral Orion—a United States based satellite communication service company that provides the bulk of Internet bandwidth to Yugoslavia—was about to cut off the satellite link to the country under the U.S. trade embargo against Yugoslavia.[187] The fears stemmed from Executive Order 13,121, which went into effect on May 1, 1999. The Executive Order prohibited trade with Serbia and Montenegro, including a broad prohibition of supply of any goods or services and any transaction in goods or services.[188]

Although some communication services such as telephone connections are usually exempt in trade embargoes, it was not clear from the order itself whether Internet access via a satellite link provided by Loral Orion was covered by the Executive Order. In addition, there were concerns about the cyberattacks launched by "hackers" in Serbia against U.S. and NATO sites after the bombing on Serbia began.[189]

Apparently, Loral Orion did discuss the cut-off issue with the U.S. government.[190] However, the Clinton administration decided against cutting off the satellite link and instead encouraged the Serbians to use the Internet to break the media clamp down by the Milošević government.[191] Such liberal Internet policy may have been influenced by the recognition that many of the Internet users in Yugoslavia were against the Milošević regime.[192] It is an open question whether the U.S. government would have taken such a liberal approach if the majority of the Internet community in Serbia had not been pro-West or anti-Milošević, or if the cyberattacks from Yugoslavian users were quite serious. In fact, during the earlier U.N. embargo against Yugoslavia in 1992, the BitNet link—the only available Internet connection to Ser-

186 Carlotta Gall, *Yugoslavians Fear Web Will Be Cut Off*, N.Y. TIMES ON THE WEB, May 15, 1999, *at* http://www.nytimes.com/library/world/europe/051599kosovo-internet.html. *See generally* STUART BIEGEL, BEYOND OUR CONTROL? 126 (2001).

187 Matthew Broersma, *Clinton Encourages Serbia Net Access*, ZDNET NEWS, May 13, 1999, *at* http://zdnet.com.com/2100-11-514639.html.

188 Exec. Order No. 13,121, 64 Fed. Reg. 24,021 (May 5, 1999).

189 Ellen Messmer, *Serb Supporters Sock It to NATO, U.S. Web Sites*, CNN.COM, Apr. 6, 1999, *at* http://europe.cnn.com/TECH/computing/9904/06/serbnato.idg.

190 Bob Woods, *White House Won't Impede Yugoslavia's Net Access*, NEWSBYTES, May 14, 1999, *at* http://www.alb-net.com/kcc/051599.htm#5.

191 Broersma, *supra* note 187.

192 Dražen Pantić, *Internet in Serbia: From Dark Side of the Moon to the Internet Revolution*, FIRST MONDAY, April 1997, *available at* http://www.firstmonday.dk/issues/issue2_4/pantic; *see also* PBS, *Bringing Down a Dictator: It's Time*, *at* http://www.pbs.org/weta/dictator/rock/itstime.html (last visited Feb. 24, 2004) (discussing Milošević's attacks on the media and the response by internet users in the former Yugoslavia).

bia and Montenegro at the time—was cut off.[193] The West presumably did not know or care about the pro-West tilt of the Yugoslavian Internet user community at the time.[194]

Although Serbian Internet access was not interdicted by severing the physical link, it could have been. How would this case be handled by layers analysis? Interdiction at the physical layer to solve a problem at the content layer is a layer-crossing regulation. Layers analysis suggests that, as a consequence, interdiction would be both overinclusive and underinclusive with respect to the regulatory goal—to disrupt communications over the Internet that would aid the Serbian regime. This suggestion is confirmed by the Serbian interdiction case. Had the link been severed, the regime in Serbia would not have been able to use the Internet to communicate as effectively with the outside world; interdiction would achieve some of the desired effect. But, interdiction would be overinclusive. Interdiction would have affected antiregime groups within Serbia by denying Serbians access to antiregime content. Interdiction would also have been underinclusive, because of the nature of the Internet's basic architecture. The Internet is designed to route around damaged physical links. Had one link gone down, communications would still have flowed via other physical links. Cutting off all the physical links would be difficult because of the wide variety of physical pipelines that can carry data. Satellite, fiber optic cable, copper wire, and microwave are all capable of performing the job. Total interdiction would have required the cooperation of all of Serbia's neighbor states.

Notice, however, that layers analysis does not necessarily lead to the conclusion that the wire should not be cut in wartime. Layers analysis requires a compelling regulatory justification, but disrupting enemy communications during wartime is surely the kind of interest that would count as compelling. Layers analysis requires the consideration of layer-respecting alternatives, but it seems quite possible that in some wartime situations, no such alternative will be adequate. In other words, layers analysis maps reasonable intuitions about how war-

193 Milan Serba, Internet Connectivity in Eastern Europe (Richard Budd ed., draft Sept. 1992), *at* http://www.eff.org/Infra/Foreign_and_local/Multinational/east_and_central_europe_net.paper; *see also* Eric Bachman, *Digital Communication via the Internet in a War Zone: Conflict Resolution and the Internet*, PROC. INET96, at H2 (1996), *available at* http://www.isoc.org/isoc/whatis/conferences/inet/96/proceedings/h2/h2_2.htm.

194 The reaction of the Serbian Internet community to the U.S. sanction in 1999 is quite understandable in light of their previous experience of total Internet connection cut-off in 1992.

time Internet interdiction should be analyzed by public Internet regulators.

b. Internet and Regimes in Crisis: Myanmar

In contrast to the externally imposed communication disruption by hostile nations at war, cutting off or severe restrictions of the physical communication link can be self-imposed by regimes in crisis. The current situation in Myanmar (formerly known as Burma) is illustrative for this type of extreme regulation at the physical layer.

Under a 1996 law, every telephone, fax machine, modem, or computer in Myanmar must be authorized by and registered with the government authorities. The law, passed in the midst of a major crackdown on popular political activities, imposes imprisonment of seven to fifteen years for unauthorized possession of a computer, modem, or fax machine.[195] The law is not a mere technicality or empty threat. In fact, in 1996, James Leander Nichols, Honorary Consul to Norway and a friend of the opposition leader,[196] was jailed for owning unauthorized fax machines. He died in prison several months later under questionable circumstances.[197]

The draconian Burmese telecommunications law has since been extended to specifically cover the Internet.[198] The law prohibits Internet access unless granted by the authorities, and requires prior government approval of every web page created in the country.[199] The government can effectively control all Internet access from Myanmar because the only ISP in the country is the government-run Myanmar Post and Telecommunications.[200] E-mail service is permitted to a lim-

195 Matthew Pennington, *Fearing Free Speech Pandora's Box, Myanmar Rulers Block Internet*, CNEWS, Apr. 18, 2000, *at* http://www.canoe.ca/TechNews0004/18_myanmar.html.

196 Aung San Suu Kyi is the leader of the National League for Democracy, who won by a landslide victory in a general election in 1990. Placed under house arrest by the military since her electoral victory, she was awarded the Nobel Peace Prize in 1991. *Id.*

197 AMNESTY INT'L, MEDICAL CONCERN: MYANMAR: DEATH IN CUSTODY OF LEO NICHOLS (July 16, 1996); *see also* U.N. GAOR, 51st Sess., Agenda Item 113(c), U.N. Doc. A/51/204 (1996), *available at* http://www.un.org/documents/ga/docs/51/plenary/a51-204.htm (expressing the concern of the European Union about the deteriorating political situation in Myanmar, and calling on the Myanmar authorities to respect human rights).

198 *Updates on Media Law Reforms*, 1 COMM. L. IN TRANSITION NEWSL., Feb. 12, 2000, *at* http://pcmlp.socleg.ox.ac.uk/transition/issue04/updates.htm.

199 *Id.*

200 Myanmar's Net Inc., *Internet Access in Myanmar*, *at* http://www.myanmars.net/myanmar/internet.htm (last visited Feb. 16, 2004).

ited number of foreign businesses and Burmese users, but access to the World Wide Web or any website outside the country is essentially prohibited, except for a few government officials for security monitoring purposes.[201]

In Myanmar, a politically unstable regime threatened by popular opposition imposed an extreme kind of regulation from the physical layer by strictly controlling, under threat of severe criminal punishment, communication lines, equipment, and network hardware. By doing so, the regime has ensured that it has the power, when necessary, to impose on the country a complete communication blackout, including "cutting the wire" to the Internet. Layers analysis of the Myanmar case is structurally similar to analysis of the Serbian interdiction case. But, of course, the two cases are quite different. Most policy analysts would agree that wartime Internet interdiction is justified under the right conditions; the very same analysts would likely condemn the Myanmar regime's "cut the wire" policy. This fact illustrates a conclusion that we have already reached—layers analysis provides a framework for the analysis of Internet regulations, *but* this framework is necessarily incomplete. The difference between wartime interdiction and the self-imposed isolation of a totalitarian regime is the following: some wars are justified, whereas no totalitarian regime should be supported. Layers analysis gives us an angle of attack on the "cut the wire" cases, but it does not do all the work.

c. Lessons from the "Cut the Wire" Example

In the Serbia and Myanmar examples above, the regulation is primarily or solely directed at the physical layer. However, the object of such regulation is to deny or restrict the data exchanged over the Internet—that is, to strike at the content. Also, the conflict arose at the content layer, as all political conflicts must, but the physical link was attacked as the most extreme form of blockade on the content.

The extreme nature of "cutting the wire" regulation is quite obvious. In virtually all circumstances, the physical link must have substantial innocent use, and "cutting the wire" regulation is extremely overinclusive. Clearly, "cutting the wire" is the ultimate destruction of transparency, as no content data whatsoever will get through. If imposed on an entire nation or region, transparency of the Internet over the whole area would be completely destroyed. Such extreme regula-

201 Sandy Barron, *Myanmar Works Hard to Keep the Internet Out*, N.Y. TIMES ON THE WEB, July 14, 2000, *at* http://www.nytimes.com/library/tech/00/07/cyber/articles/14myanmar.html; *see also* Myanmar's Net Inc., *supra* note 200 (reporting that the Internet is still operating in Myanmar).

tion can only be justified, if ever, by extremely compelling reasons, such as an all-out war of mass destruction.

When the problems at the content layer—the layer highest in the Internet layer hierarchy—are attacked at the physical layer—the lowest layer—we have the biggest layer violation that leads to the most severe and substantial innocent use problem, as well as the most complete destruction of transparency. Regulators must be intuitively aware of the extreme and obvious nature of the problem, as such regulation is deployed only under highly unusual circumstances. Nevertheless, the "cut the wire" example is instructive in that it shows a clear example of layer-violating regulation and problems with such regulation.

3. IP Layer Regulation Aimed at Content Layer Problems

Regulation at the physical layer represents the most extreme form of layer-crossing regulation. Moving one layer up, we now consider regulations that operate at the IP layer, but aim to solve a problem at the content layer. We begin with some introductory remarks, and then proceed to consider a series of examples.

a. Introduction: Definition and Comparisons

In one sense, this section represents the core of our Article. This is where the action is, as becomes apparent from the four examples we shall discuss—China's blocking of foreign sites,[202] the French *Yahoo!* case,[203] responses to cyberattacks or terrorism,[204] and Pennsylvania's child pornography statute[205]—and some additional examples that are important but left to a footnote.[206] The issues involved in these examples are diverse and complex in comparison to the issues involved in

202 *See infra* Part III.A.3.b.

203 *See infra* Part III.A.3.c.

204 *See infra* Part III.A.3.d.

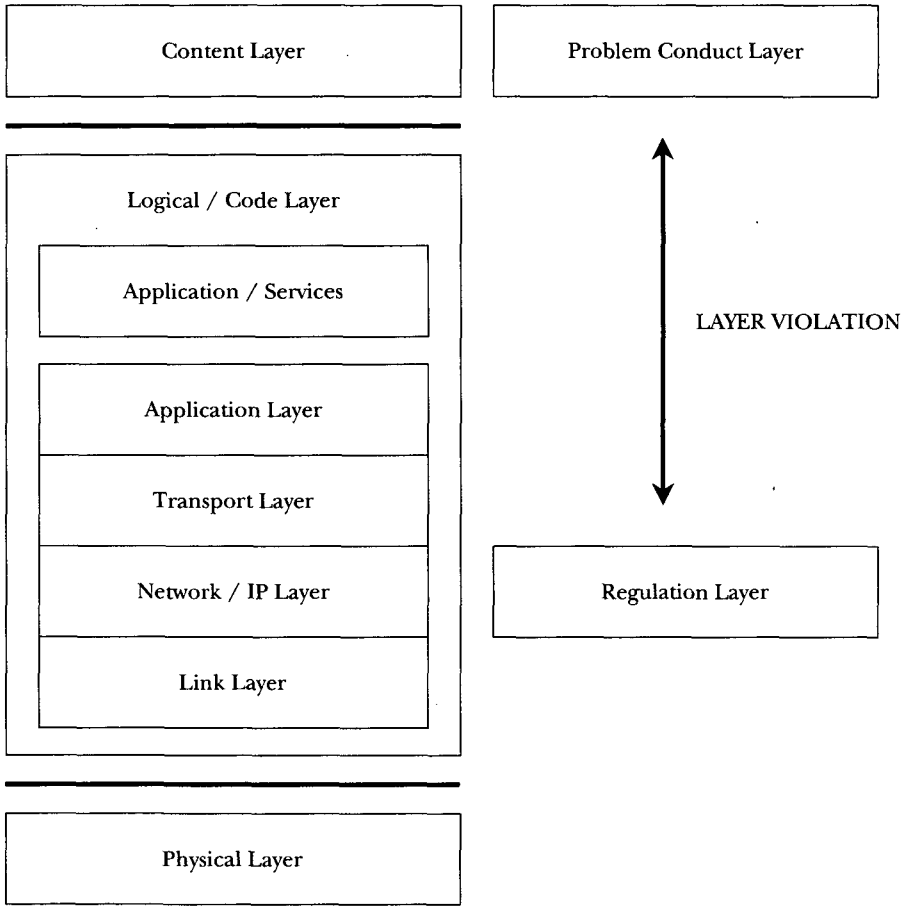
205 *See infra* Part III.A.3.e.

206 Our list of examples is not intended to be exhaustive. Jonathan Zittrain provides at least two more: (1) "A German court has held that approximately sixty destination ISPs in the state of North-Rhine Westphalia can lawfully be asked to block German customer access there to two U.S.-hosted Web sites determined by the German government to contain banned Nazi propaganda," Jonathan Zittrain, *Internet Points of Control*, 44 B.C. L. REV. 653, 683 (2003), and (2):

> [I]n a short-lived case that would have proved an interesting test of the Digital Millennium Copyright Act's provisions on injunctions, thirteen record companies filed a lawsuit in August 2002 to force five major domestic ISPs, in their role as destination ISPs and backbone providers within the Internet cloud, to block their customers' Web access to www.listen4ever.com, an al-

regulations at the physical layer. Furthermore, some of the concepts introduced may not be familiar to most of the readers. We begin with the task of defining what we mean by regulation at, or directed at, the IP layer. Then, an explanation of how it is different from regulation at other layers follows. The category of regulations at the IP layer that are addressed to problems at the content layer is illustrated by Figure 7:

FIGURE 7. REGULATION DIRECTED AT THE IP LAYER DUE TO PROBLEMS AT THE CONTENT LAYER



In the "cut the wire" example above, few will dispute that the regulation was directed at the physical layer. Although the regulation may ultimately seek to control the content, the objects of the regula-

legedly unauthorized China-based source of those companies' copyrighted music.

Id.

tion's control are the physical communication links. In an analogous way, regulation directed at the IP layer seeks to control a function or functions at the IP layer. As discussed above in the TCP/IP section, the IP layer is the layer above the transport layer in the Internet layer hierarchy.

Many regulations at the IP layer are not motivated by a desire to control content. For example, switching over to the next generation IP protocol—IPv6—will require global coordination of all Internet users, ISPs, and backbone operators for the process to be fully effective.[207] Although a successful upgrade to IPv6 would be an unprecedented, monumental worldwide cooperation with a staggering global reach, it has nothing to do with content. The object of global coordination is purely about the "plumbing" of the Internet.[208]

Some regulations at the IP layer, however, are motivated by problems at the content layer, or have the ultimate goal of controlling content. A regulation that requires blocking an entire site (i.e., a host or hosts from the site), due to concerns about the site's content, is one such example. The offending site or user is identified with the IP address of the host machine, and all data coming from, or going to, the IP address is blocked. In effect, the problem at the content layer is associated with the identifier at the IP layer—namely, the IP address. The regulation is directed at the IP layer because it seeks to control an IP layer function—that is, delivery of data from one host to another according to its IP addresses. The regulation says: "If the data are coming from or going to the IP address *xxx.yyy.zzz*, then don't deliver that data."

Technically, the IP address blocking can be implemented in several different ways at different levels or layers. The blocking can be implemented by an IP filter (typically, an IP level device driver), by an IP router that runs at the IP layer, or a proxy server (or firewall) that runs at the "application level."[209] However, the technical implementation methodology should not be confused with *the layer of the function* that is being sought to be controlled. Regardless of the implementa-

207 *See generally* ERIC CARMÈS, THE TRANSITION TO IPv6 (2002), *available at* http://www.isoc.org/briefings/006/isocbriefing06.pdf.

208 Our discussion is somewhat oversimplified. One of the benefits of IPv6 will be to vastly expand the supply of IP addresses. Currently, there is a shortage of IP addresses in some regions, due mostly to distribution rather than the absolute number of IP addresses. These distributional issues may have some impact on content and, therefore, there may be a connection between IPv6 and content problems. These issues are complex, and we set them aside for the purpose of this Article.

209 *See generally* WILLIAM R. CHESWICK & STEVEN M. BELLOVIN, FIREWALLS AND INTERNET SECURITY (1994); ARI LUOTONEN, WEB PROXY SERVERS (1998).

tion detail, the controlled function, or the function where the restriction is imposed—delivery of data from one IP address to another—is a function at the IP layer.

And, this is not a mere semantic distinction. The difference is a fundamental and architectural one. Given TCP/IP's layered architecture, the only place where a host machine is identified is at the IP layer. The DNS host name—e.g., "www.amazon.com"—is used purely for the convenience of human users. When the data travel from one machine to another over the Internet, the source and destination are identified by the IP addresses *only*. Thus, when an entire host (or entire site) is blocked, it is fundamentally an IP layer function that is being regulated.

A World Wide Web content filter at the application layer provides an illuminating example of the fundamental differences between regulations targeted at different layers. As discussed above, HTTP is the application layer protocol for the World Wide Web.[210] The fundamental basic building block of HTTP is a "document." When a user visits a website, every "page" presented to the user in the browser is a "document" that is identified by a uniform resource locator (URL). A URL might look like: "http://www.amazon.com/index.htm" if the DNS host name is used, or "http://157.242.144.232/index.htm" if the IP address is used directly.[211] It is crucial to understand that a URL does *not* identify a host machine; rather, it identifies *a document* at a host machine. The host is identified by an IP address at the IP layer, and the web document is identified by a URL at the application layer.

Within HTTP, there is no more basic unit of data than a document. As an application layer protocol, HTTP has no provisions for identifying a host machine, routing data packets, or delivering data bits. It relies on the lower layer protocols for those functions. Thus, within HTTP (the application layer protocol), it is *not* possible to block an entire site or host; it can only block an individual *document.* In the Amazon.com website example given above, blocking the home page of the web site—identified as "http://www.amazon.com/index.htm"—would not block other documents on the site. For example, "http://www.amazon.com/books.htm" would not be blocked, and users can get to the document directly by entering the URL directly in their browser's address field. Of course, the entire site would be blocked effectively if every document from the site is blocked. But, within the HTTP or at the application layer, the blocking will have to be done one document at a time. And this is precisely how the con-

210 *See supra* text accompanying note 60 (discussing HTTP).

211 This is not the actual IP address for www.amazon.com.

tent filtering at the application layer would work. For example, the World Wide Web Consortium's Platform for Internet Content Selection (PICS) specification works this way. Each document identified by a URL is labeled with a standardized content rating, and the content filter program allows or blocks each document according to the user setting and the PICS label.[212] Any other technology that performs selective filtering within the web protocol or any application layer protocol would have to work by blocking individual documents or URLs.[213]

If, on the other hand, a whole site or a host is blocked as one unit using a single identifier, it will have to be done with an IP address. A URL operates at the application layer, whereas an IP address operates at the IP layer. These are fundamentally distinct constructs that identify architecturally distinct objects: the documents and host machines, respectively. Therefore, regulation that requires blocking an entire site or a host is a regulation that is directed at the IP layer, regardless of how it is implemented.

It should be pointed out, however, that both types of blocking are layer-violating regulations. Blocking an entire site or a host due to concerns about its content is a regulation directed at the IP layer as a result of problems at the content layer. Blocking of URLs, on the other hand, is directed at the application layer in order to address problem conduct at the content layer.

All Internet filter programs employ one or more of these methods.[214] The problems of filtering or blocking programs are well known.[215] Among others, they are notoriously overinclusive and underinclusive. A detailed analysis of the problem of filtering programs will be given later in this Article. For now, it is illuminating to note that, between the two methods of blocking, there is a significant difference in the severity of overinclusiveness and underinclusiveness. When a whole site is blocked due to some offending content at the site, all data on the host machine is blocked. When, however, blocking is done at a document level, the effect of overinclusion or under-

212 World Wide Web Consortium, *PICS Frequently Asked Questions, at* http://www.w3.org/2000/03/PICS-FAQ (last visited Feb. 13, 2004).

213 Because document names frequently include IP addresses and/or domain names, it would be possible to block multiple URLs by searching for the IP address or domain name within the URL string.

214 Junichi P. Semitsu, *Burning Cyberbooks in Public Libraries: Internet Filtering Software vs. The First Amendment*, 52 STAN. L. REV. 509, 513–19 (2000).

215 Am. Library Ass'n v. United States, 201 F. Supp. 2d 401, 408 (E.D. Pa. 2002) (citing the expert report of Benjamin Edelman); *see also* Geoffrey Nunberg, *The Internet Filter Farce*, 12 AM. PROSPECT, Jan. 1, 2001, *available at* http://www.prospect.org/print/V12/1/nunberg-g.html (discussing some of the problems with internet filters).

inclusion is limited to the document. Therefore, blocking at the IP layer—blocking of an entire site or a host—is inherently more overinclusive and underinclusive than blocking at the application layer—blocking of individual documents.

And, the same goes for transparency. The blocking at the application layer is done at a finer subunit of data than the blocking at the IP layer—data in a document on a machine as opposed to all data on an entire machine. Thus, blocking at the IP layer inherently has greater impact on the transparency of the Internet than blocking at the application layer.

Now, we can point to an interesting and important pattern that emerges from layers analysis. The "cut the wire" regulation blocks all data from all machines in a country or geographical region; blocking an IP address blocks all data from a host machine; and blocking a URL blocks all data contained in the document. Thus, when addressing the offending conduct at the content layer, the substantial innocent use problems and transparency impairment are most severe when the regulation is directed at the physical layer, less severe when at the IP layer, and lesser still at the application layer. Therefore, the severity of the layer violation is greater when the regulation is attempted at a lower or deeper layer in the TCP/IP layer hierarchy in order to address the problems at an upper or higher layer.

Layer-violating regulation at the IP layer, then, is regulation directed at the IP layer due to problems at the content layer. As most regulators refrain from "cut the wire"-type regulation due to the intuitively obvious nature of the problem, layer-violating regulation at the IP layer is the most common form of severe layer-violating regulation. This type of regulation is frequently deployed in the context of one of the most interesting and controversial areas of Internet regulation—regulation across national boundaries. Such regulations are invariably motivated by the desire to control the extreme transparency of the Internet on the part of the national governments. Below, issues brought up by national regulation are considered by looking at three examples—China's regulation of the Internet, the French *Yahoo!* case, and post-September 11 responses to cyberattack or terrorism. A fourth example, a Pennsylvania child pornography IP address blocking statute, deals with the same issues at the state (subnational) level in a federal system.

Our discussion of the example of China's regulation of the Internet is important, in part, because it is illustrative of a general problem that the Internet poses to developing countries. Typically, such nations do not possess the political and cultural tradition of Western democracies, but want to develop toward advanced society, at least in

technological and economic aspects.[216] The French *Yahoo!* case represents the dispute among developed nations or Western democracies instigated by a transparent Internet. And, the concerns about cyberattacks or cyberterrorism must be considered, as a changed attitude among people and decisionmakers due to post-September 11 realities is likely to shape the discussion about the transparency of the Internet. The Pennsylvania child pornography case provides yet another example—this time in the context of a category of content that is almost universally condemned as evil, and as properly the subject of regulation. We now turn to these four examples—beginning with the case of China's attempts to block IP addresses.

b. China's Regulation of the Internet

China's regulation of the Internet presents a fascinating study of the intersection of law, social policy, and technology (i.e., architecture of the Internet) in a society in transition—a developing nation that aspires to be a world leader in commerce and technology while maintaining an authoritarian political system, or at least deferring possibly inevitable political reforms. Since the ascendance of Deng Xiaoping in 1978, the government of China (or rather the political leadership of the Chinese Communist Party) has pursued two fundamental goals: economic prosperity and political stability.[217] The Internet poses an unprecedented dilemma for the Chinese government in its dual policy of pursuing economic prosperity through economic liberalization while maintaining a single party authoritarian political system through pervasive political control—and the dichotomy or tension between the two objectives is especially apparent in the area of the Internet.[218]

Many in the Chinese leadership perceive the information and knowledge industry, including the Internet, to be the key to China's future prosperity. By taking advantage of its superior human capital, China hopes to rapidly catch up and become competitive with the

216 The Myanmar example, on the other hand, illustrated an example of countries that do not want to develop toward a Western style liberal democracy.

217 Ken Davies, China's International Economic Policy, Lecture at the University of Hong Kong (Oct. 19, 2000), *at* http://members.tripod.com/~Ken_Davies/hku1.html; *see also* Pieter Bottelier, China's Economic Transition and the Significance of WTO Membership, Huang Lian Memorial Lecture at Stanford University (Nov. 17, 1999), *at* http://credpr.stanford.edu/pdf/lienlecture.pdf. *See generally* NICHOLAS R. LARDY, CHINA'S UNFINISHED ECONOMIC REVOLUTION (1998) (discussing China's economic reform policies); JUN MA, THE CHINESE ECONOMY IN THE 1990s (1997) (discussing the changes in the Chinese economy during the 1990s).

218 Philip Sohmen, *Taming the Dragon: China's Efforts to Regulate the Internet*, 1 STAN. J. E. ASIAN AFF. 17, 17 (2001).

West.[219] On the other hand, because of its global reach and transparent nature, the Internet also opens up floodgates to ideas and influences that the Chinese government wants to keep out, including unsanctioned political activities that could threaten China's monopoly political system. Thus, the People's Republic of China is faced with the problem of allowing enough of the Internet necessary to take advantage of its benefits, but stopping enough of it sufficient to neutralize potential threats. Examining the methods by which China is attempting to achieve these conflicting goals offers a revealing glimpse of the interaction of law and policy, and how such regulatory attempts are manifested due to the inherent layered architecture of the Internet.

i. The Great Wall: Controlling Access at the Physical Layer

China's Internet regulation has a physical access control component strikingly similar to that of Myanmar.[220] Under the law, the Chinese government has a monopoly over all Internet connections going into and out of the country.[221] The regulations designate the Ministry of Information Industry (MII) as the gatekeeper to the Internet, and access to the global Internet by networks (ISPs) is restricted exclusively to a handful of channels—the national backbones—provided or sanctioned by the MII.[222] The ISPs are required to apply for a license for connection to the MII backbone, and in so doing, must provide information on the location of their host computers as well as the nature and scope of their networks. The individual users of the Internet who sign up with ISPs are also required to register with the local Public Security Bureau (PSB). The registration process provides the PSB with information on the location of each computer connected to the Internet, as well as its owner, for surveillance and monitoring purposes. Failure to register is punishable with fines or a prison sentence.[223]

Unlike Myanmar, however, many of the draconian regulations at the physical layer are not zealously enforced, especially against individual users.[224] If having a computer and Internet connection comes with constant surveillance and an ever-present threat of imprison-

219 *Id.*; *see also* William Foster & Seymour E. Goodman, *The Diffusion of the Internet in China*, *at* http://iis-db.stanford.edu/pubs/20022/chinainternet.pdf (last modified Sept. 12, 2000).

220 *See supra* Part III.A.2.b.

221 Sohmen, *supra* note 218, at 20.

222 *Id.*; *see also* Foster & Goodman, *supra* note 219, at 35.

223 Sohmen, *supra* note 218, at 20, 23.

224 Foster & Goodman, *supra* note 219, at 35.

ment, such regulation would inhibit the growth of the Internet industry in China by scaring the Chinese population away from the Internet; China would not reap the economic benefits that the Internet promises. Thus, although the drastic physical access control regulations similar to those of Myanmar are still technically on the books in China, they are not enforced due to this overriding policy objective. Nevertheless, in order to control potentially threatening activities, the Chinese government resorts to a less severe form of Internet regulation—i.e., regulation at the IP layer.

ii. The Great Firewall: Blocking Content at the IP Layer

Under the "Computer Information Network and Internet Security, Protection and Management Regulations," access to certain objectionable materials over the Internet is prohibited in China.[225] The list of prohibited materials includes those: subversive of state power or the socialist system; damaging to national unity; inciting discrimination between nationalities; disturbing to social order; propagating feudal superstition; pornography, gambling or violence; insulting or libelous; or violating the constitution or other laws.[226]

Words such as "subversive," "damaging to national unity," "disturbing social order," and "feudal superstition" are not defined, and obviously are vague. However, foreign websites such as those of the *New York Times*, the BBC, CNN, and Stanford University are routinely blocked because these sites somehow run afoul of Chinese law.[227] In the Fall of 2002, the Chinese government blocked access to Google for about three weeks.[228]

Blocking of sites is delegated to the ISPs, with the PSB sending out a list of websites to be blocked from users in China.[229] The ISPs are required by law to follow the directions from the PSB, and essentially operate as agents for the government.[230] Technically, the ISPs in China are required to block all traffic from or to specific IP addresses handed down by the government. As the Chinese government has a monopoly over the physical connection to the global Internet, and all

225 *See generally* Jonathan Zittrain & Benjamin Edelman, *Empirical Analysis of Internet Filtering in China, at* http://cyber.law.harvard.edu/filtering/china (last modified Mar. 20, 2003).

226 *See id.*

227 U.S. Embassy Beijing, U.S. Dep't of State, *China's Internet "Information Skirmish"* (Jan. 2000), *at* http://www.usembassy-china.org.cn/english/sandt/webwar.htm.

228 Joseph Kahn, *China Toughens Obstacles to Internet Searches*, N.Y. TIMES, Sept. 12, 2002, at A3.

229 Sohmen, *supra* note 218, at 20.

230 Foster & Goodman, *supra* note 219, at 35.

ISPs that connect to the government sanctioned backbones are required to be licensed—and thus are required to follow government orders—the government can completely block all objectionable content from all of China, at least theoretically. It would be like building an electronic Great Wall around China—hence the Great Firewall.[231]

In reality, however, voluntary compliance by Chinese ISPs is uneven.[232] Also, uses of circumventing technologies—such as the peer-to-peer applications connecting to anonymizing proxy servers outside China—make it difficult to block access based on the IP address of the true or ultimate destination.[233] Nevertheless, in contrast to physical access control, China is dead serious about enforcement of its IP layer regulation. First, rather than entirely relying on the ISPs to faithfully block the banned sites, the Chinese authorities monitor and filter all Internet traffic going through China's eight primary gateways to the global Internet.[234] Presumably, the packets from the banned IP addresses are dropped at these backbone gateways. At the other end of the ISPs, installation of site blocking software is required on all end user computers with public access, such as the PCs in Internet cafes. Recently, Shanghai police closed down almost 200 Internet cafes in the city during a weeklong sweep for not blocking the sites as required under the law.[235] Last year, Chinese authorities reportedly shut down 17,000 Internet bars that failed to install the site blocking software.[236] Although closing down the service and confiscating the computers is directed at the physical establishment, it should properly be regarded as a part of IP layer regulation, as the nature of action is enforcement of the regulations at the IP layer—i.e., blocking access to sites at specific IP addresses. Of course, the ultimate goal of the regulation is to control social or political unrest by restricting the flow of information in and out of China—all problems at the content layer. China's regulation of the Internet is, as explained above, a clear example of layer-

231 A. LIN NEUMANN, COMM. TO PROTECT JOUNALISTS, THE GREAT FIREWALL (2001), *available at* http://www.cpj.org/Briefings/2001/China_jan01/Great_Firewall.pdf.

232 *See* Jasper Becker, *China Wrestles an Online Dragon*, CHRISTIAN SCI. MONITOR, June 19, 2002, at 6, 6.

233 Doug Nairne, *China Tightens Web Control*, S. CHINA MORNING POST, Aug. 14, 2002, *available at* http://asia.cnet.com/newstech/industry/0,39001143,39071903,00.htm.

234 Dion Wiggins & Louisa Liu, *China's Internet Strategy: Struggling to Maintain the "Great Firewall,"* GARTNER NEWS, Apr. 1, 2002, *at* http://security1.gartner.com/story.php.id.177.s.1.jsp.

235 *Shanghai Cracks Down on Net Cafes*, REUTERS, May 6, 2002, *available at* http://www.wired.com/news/politics/0,1283,52330,00.html.

236 *Id.*

violating regulation targeted at the IP layer due to problems arising at the content layer.

iii. Application of Layers Analysis to China's Regulation of the Internet

China's regulation of the Internet is highly controversial, and has been the object of vigorous protest or criticism from inside and outside of China. The criticisms can be thought of as belonging to one of three broad categories of arguments: (1) China's policy of censorship is objectionable because it is politically or morally wrong (i.e., violates human rights, free speech rights, etc.), whether or not the censorship takes place on the Internet; (2) regardless of whether China should exercise censorship within its borders, it should not do the same on the Internet because the Internet is different; and (3) even if China were to engage in content control on the Internet, it should not do what is currently done because the approach is flawed. The first position is outside the scope of our analysis in this Article.[237] The next two arguments are examined with an emphasis on what layers analysis can tell us about the issues. First, we present an analysis of the general layer-violating characteristics of China's Internet regulation. Then, issues of censorship are analyzed under the layers analysis approach.

a) General Layers Analysis of China's Internet Regulation

China's regulation of the Internet is perhaps a severe example of layer-violating regulation targeted at the IP layer due to problems arising at the content layer. The innocent use problem is substantial, as most of the materials on the blocked websites would be beneficial or desirable to millions of users within China, and are matters of little or no concern even to the Chinese government. The regulation clearly impairs the overall transparency of the Internet by preventing access to popular Internet services from a very large region for a very large number of people.[238]

---

237 Of course there are a variety of arguments against China's position. For a survey of free speech theory, see Lawrence B. Solum, *Freedom of Communicative Action: A Theory of the First Amendment Freedom of Speech*, 83 NW. U. L. REV. 54 (1989). There is a further question whether China's policy on free speech is an internal matter or is properly within the scope of concern of foreign governments and international organizations. *See generally* JOHN RAWLS, THE LAW OF PEOPLES (2000).

238 At the end of 2001, there were over 33 million Internet users in China. *See Internet Surfers in China Hit 33.7 Million*, PEOPLE'S DAILY, Jan. 16, 2002, *available at* http://www.china.org.cn/english/2002/Jan/25362.htm.

However, the problems are certainly not as severe as would result from "cutting the wire" regulation at the physical layer, as the dramatic differences between the Internet uses in Myanmar and China clearly show. It is very interesting to note that, although China has on the books exactly the same type of drastic regulations at the physical layer as Myanmar, the Chinese regulators consciously chose to regulate at the IP layer to impair the functioning of the Internet—transparency and substantial innocent use—to a lesser degree. In their attempt to achieve the conflicting goals of economic liberalization and political control, the regulators in China are attempting to strike a right balance by fine tuning the grades of transparency of the Internet—that is, transparent enough to take advantage of its benefits, but opaque enough to keep out potential threats. And, the mechanism they chose to adjust the transparency is regulation at different Internet layers—moving from the physical layer to the IP layer. Whether or not the Chinese regulators are aware of the nature of their acts, they are in fact traversing the scale of severity in layer-violating regulations. Destruction or impairment of transparency is most severe when the problems at the content layer are regulated at the physical layer, which is at the bottom of the Internet layer hierarchy. It is less severe when done at the IP layer, a layer above the physical layer. China's attempt to achieve a balance of conflicting demands of economic liberalization and political control is manifested in the Internet regulation as traversing the scale of severity in layer-violating regulations, reflecting the inherent layered architecture of the Internet and transparency flowing from it.

b) The Internet as *Res Publicae* in Cyberspace

Despite the harmful effects of its severe layer-violating regulation, China may stand on a firm ground when it argues that China, as a sovereign nation, has the right to control what happens on the computer networks located within its borders in a way that it sees fit. As pointed out above, many object to this position on political or moral grounds. Others argue that, regardless of what China does with other media or methods of communication, China should not censor the Internet, because the Internet is different. In fact, some argue that no nation, including China, should restrict activities that take place on the Internet. But, these arguments do not clash with the contention that as a sovereign, China has the right to autonomously determine its own Internet regulatory policies.

There are, however, arguments that do question the right of sovereign national governments to regulate cyberspace. Perhaps the

most extreme position against censorship on the Internet is that cyberspace is its own autonomous space that is not subject to the jurisdiction of nations or governments in real space.[239] Although there are good reasons to doubt this claim, consideration of this position is not analyzed in this Article in part because there is already substantial literature on the topic.[240] Instead, we focus on a somewhat narrower position that argues against censorship on the Internet, even if cyberspace is not completely outside the reach of government regulations. This position—sometimes called the global free speech argument—states that the Internet should be off limits to any government censorship, with a possible exception for the materials that all or most nations find objectionable or dangerous.[241] The argument is that, due to the global and borderless nature of the Internet, national or local laws restricting freedom of expression on the Internet violate the rights of Internet users around the world.[242] For example, if citizens of one country are prohibited from discussing political issues critically online, then not only are their rights infringed upon, but so are the rights of others around the world to seek and receive that information. Similarly, a country's efforts to block certain content from outside its borders implicates the right of those in other countries to impart information.[243]

Of course, a sovereign nation has the legal right to regulate unwanted information within its borders. And, the authors of the global free speech argument need not disagree with this point about sovereign legal authority. Nevertheless, the proponents of the global free speech movement contend that sovereign nations may have good reasons to favor free speech on the Internet, even if they restrict free speech through other media. With traditional media dissemination of information, such as through books and magazines, the information is embodied in tangible physical objects that can be subjected to national or local regulations within geographically defined jurisdictional boundaries. However, on the Internet, information is not tied to physical objects located within geographical boundaries. Furthermore, there are no recognizable national borders in cyberspace. Therefore, the argument goes, it is difficult to enforce national or lo-

239 John Perry Barlow, *Declaration of the Independence of Cyberspace* (Feb. 8, 1996), *at* http://www.eff.org/~barlow/Declaration-Final.html.

240 *See, e.g.*, BIEGEL, *supra* note 186, at 25–49.

241 *See* GLOBAL INTERNET LIBERTY CAMPAIGN, "REGARDLESS OF FRONTIERS": PROTECTING THE HUMAN RIGHT TO FREEDOM OF EXPRESSION ON THE GLOBAL INTERNET (1998), *available at* http://www.cdt.org/gilc/Regardless_of_Frontiers.pdf.

242 BIEGEL, *supra* note 186, at 28.

243 *Id.* at 29.

cal regulation of speech on the Internet without creating substantial problems of overinclusion and underinclusion.[244] Content censorship by one nation affects all Internet users around the world, and regulations that aim to fence particular content out of geographic national boundaries are doomed to be partially effective at best.

The global free speech argument can perhaps be better appreciated and sustained when considering the nature of the Internet as a public conduit for the flow of ideas and information. Carol Rose has argued that the Internet, along with the rest of the modern communication systems, can be considered to be analogous to public roads, bridges, and waterways in real space—*res publicae* under Roman law.[245] *Res publicae* in real space were the conduits of commerce, while the Internet of today is the backbone of the exchange of ideas and information, including the commercial activities that form the "information economy."

Public channels of commerce in real space were and still are subject to formal governance mainly due to the problems of physical congestion and overuse of resources. But, even when there is little or no formal regulation, the users of public channels were nevertheless expected to observe customary rules of behavior, as exemplified by American case law dealing with citizen behavior on the informal roads of the nineteenth century American countryside.[246]

The expectation of orderly conduct had largely to do with the purpose and nature of the public roads. Public channels of commerce are by nature highly efficient means to promote social welfare, because, as lanes of trade are more widely used, the benefits from trades grow exponentially (due to the network effect), while the cost of congestion and overuse increase only arithmetically.[247] This increased efficiency engendered by smooth flowing corridors of exchange was the reason why the roads and bridges were built despite significant initial investment. Thus, maintaining order and minimizing obstructions that disrupt the smooth flow of exchange was fundamental to the existence of the public channels of commerce in real space. Consequently, the users on the public channels of trade were required, or expected, to behave in an orderly fashion to avoid or minimize obstruction to the free flow of traffic.

---

244 *Id.* at 28.

245 Carol M. Rose, *Romans, Roads, and Romantic Creators: Traditions of Public Property in the Information Age*, YALE L. SCH. PUB. L. & LEGAL THEORY RES. PAPER SERIES, Feb. 2002, at 12, *available at* http://papers.ssrn.com/abstract=293142.

246 *Id.* at 10–11.

247 *Id.* at 9.

An analogous claim can be advanced with respect to the Internet. While traditional *res publicae* allowed relatively free flow of tangible physical goods, the Internet has enabled nearly frictionless flow of ideas and information around the world. And, as we have witnessed for the past several years, the resulting diverse cross-fertilization of ideas has brought about the most significant bursting of creativity and innovation in generations. Therefore, just as free flow of commercial traffic was fundamental to the public roads in real space, unfettered free flow of information and ideas is critical to the very existence of the Internet.

One important difference, however, is that problems of congestion and overuse of resources are relatively insignificant on the Internet compared to traditional *res publicae* in real space, due to the tremendous information carrying capacity of the Internet's communication infrastructure and the relative ease of adding more capacity.[248] Thus, unlike the traditional *res publicae*, conduct related to congestion or overuse is not the critical disruptive behavior that the Internet's code of conduct needs to address. Rather, the main obstruction to free flow of ideas and information on the Internet comes from direct censorship of ideas themselves. Therefore, on the Internet, the implicit code of conduct mandated by the very purpose and nature of the *res publicae* is not to engage in, or at least to minimize, censorship of ideas and information.

The public is expected to conform to both legal regulation and informal social norms in *res publicae*: do not obstruct traffic while on public roads; do not censor ideas while on the Internet. This is conduct expected of the public as a condition to its use of public resources. People who disregard these rules not only violate the rights of the parties who are directly affected, but also infringe upon the rights of all others—the public—by impairing the effective functioning of the *res publicae*. Thus, there appears to be some merit to the global free speech argument that says censorship by one nation affects all Internet users around the world. The proponents of global free speech are in essence calling for the recognition of a *norm*, an implied or customary code of conduct on the Internet as *res publicae* for the flow of ideas and information when claiming that no nation should engage in censorship on the Internet unless the material is objectiona-

248 This is not to say that the problems of congestion do not exist on the Internet or that such problems are not important. The point is that those problems on the Internet are far less significant and much easier to overcome than the similar problems on the public channels of commerce in real space. A very interesting point to note is that the relative ease of adding more communication capacity on the Internet comes from one of the fundamental architectural features of the Internet.

ble or dangerous to all or most nations in the world. Regardless of the validity or practicality of such an argument, this much is clear: given the nature of the Internet as the public conduit with global reach for flow of ideas and information—*res publicae* in cyberspace—and given that the Internet is a global network with no built-in national or regional boundaries, national regulation of ideas and information on the Internet is a highly problematic undertaking, *even if* the regulation purports to target activities within the national borders.

c) The Great Firewall in Public Space

The implied code of conduct in *res publicae* starts with general rules, but allows for exceptions or accommodations depending on the circumstances. For example, although one is expected to avoid obstructing traffic on a public road, sometimes obstruction is unavoidable because the user's car malfunctions. In such situations, the user is expected to minimize obstruction by moving the car to the side of the road.

Similarly, although governments of Western democratic nations have attempted to exercise censorship over some material on the Internet, such regulations are exceptions rather than the rule. In fact, debates over regulations are all about how exceptional circumstances or compelling reasons justify exceptional censorship.[249] By way of contrast, in China censorship is the general rule and access is the exception. This conclusion is apparent from the permission based regulation of Internet access in combination with extensive control over the physical links. Thus, China's content censorship on the Internet is contrary to the implied code of conduct expected of all users on the Internet—a shared public space for free flow of ideas and information. In effect, China is seeking benefits from being on the Internet without respecting the standard of behavior expected of everyone who is in the public space of the Internet. Thus, one might argue that China has an obligation of fair play to reciprocate and adhere to the norm of open access to information that prevails on the global Internet.[250]

On the other hand, China could argue that it has the right to regulate activities that take place on the networks located within its borders. After all, many corporations or organizations in the West

249 *See, e.g.*, Reno v. ACLU, 521 U.S. 844, 875 (1997) (recognizing that the government interest in protecting children from obscene material should not suppress adult free speech rights to the extent of allowing only what is suitable for children).

250 *Cf.* John Rawls, *Legal Obligation and the Duty of Fair Play*, *in* LAW AND PHILOSOPHY 3 (S. Hook ed., 1964).

have firewalls, and few challenge their control of activities on their internal network. In fact, China's initial approach to the Internet was building a "national intranet"—a sort of giant private network with tightly controlled access to the outside world.[251] Within the U.S. legal system, an easy answer would be that the actions of government are treated differently from those of private parties. But, the issue is about the standard of conduct on *global res publicae* that applies to end users of the Internet throughout the world—be it government body or an individual—using the public resource. Thus, China's potential counterargument cannot be simply dismissed by invoking a U.S. constitutional principle. To the extent that much of the subglobal network of networks within China may be distinct *res publicae*, China can claim the right to have control over them as a sovereign nation.

Although China's initial approach to building a giant intranet has been abandoned, the Chinese government still requires the ISPs to block listed sites and filters all Internet traffic in and out of the country at its eight gateways to the global Internet. In effect, China is acting as if it is a super-giant corporation that connects to the Internet over a gigantic network of firewalls or proxy servers—at least, that seems to be the desired Internet model being sought by the current government. Many in the West would find this model quite unsettling. Some would point out that China is not a corporation, but the world's most populous nation. But, that does not readily explain why China should not be able to do what corporations can do on the networks over which they have legitimate control—be it private ownership or state sovereignty.

It is important to remember that our analysis has a limited scope. Two important questions have not been analyzed here. First, we have not addressed the question whether China should, as a matter of political morality, recognize strong rights of freedom of expression. There may well be a compelling case for such recognition, but we do not address that question here. Second, we have not explored the implications of international human rights law and theory on China's Internet policy. There may be a case that China is obligated by international law to ease its restrictions of Internet access. Moreover, there may be a case that foreign nations or international organizations may legitimately pressure China to recognize more robust rights of freedom of expression. Instead, our focus has been on arguments that assume, arguendo, that China may legitimately control domestic

251 Nick Wingfield & Courtney Macavinta, *China's National Intranet*, CNET NEWS.COM, *at* http://news.com.com/2100-1023-262013.html (last modified Jan. 15, 1997).

communications. We claim that, nonetheless, China is obligated to respect the open nature of the Internet. Because our claim relies on weak assumptions,[252] it provides reasons that can be accepted by the Chinese leadership, even if the leadership rejects strong theories about freedom of speech in particular, or human rights in general.

d) Layers Analysis of China's Censorship on the Internet

Our discussion above illustrates a type of problem that often arises in debates over Internet regulation. Although a general principle, such as the end-to-end principle or the implied code of conduct on the Internet as *res publicae* in cyberspace, clarifies or illuminates hitherto unconsidered issue, analysis under it sometimes fails to gain traction due to the very generality of the principles. When this happens, layers analysis frequently can provide an effective way out of the impasse. By identifying the layer of the problem conduct, as well as the layer where the regulation operates, layers analysis provides more concrete and textured analysis of the issues by placing the disputed function at a proper layer and providing a correct focus on the relevant operation of the Internet. That is, a layers approach provides traction at the point where theory meets practice.

As applied to the "Internet as public conduit" argument, an important starting point is to note that the public conduit function of the Internet operates mainly at the IP layer and at the physical layer below it. It can certainly be argued that upper layer spaces such as the application layer are also public spaces. But the point is that the IP layer is the greatest common denominator of the publicly shared resources on the Internet. The public networks—WANs, regional networks, and the backbones—all operate at the IP layer and below. However, it is the function of the IP layer (or the IP protocol) to fuse a multitude of disparate networks into a single unified, seamless network.[253] This is why the proper technical name for the IP layer is the network or Internet layer.[254] In fact, the original meaning of "Internet"—which is still used in much technical literature—was a public

252 By "weak assumptions," we mean that our assumptions are undemanding and uncontroversial. "Strong assumptions," by way of contrast, are demanding and controversial.

253 Internet Transparency, *supra* note 90 ("The key to global connectivity is the inter-networking layer. [IP protocol] allows for uniform and relatively seamless operations in a competitive, multi-vendor, multi-provider public network."); *see also* David S. Isenberg, *The Dawn of the Stupid Network*, ACM NETWORKER 2.1, Feb./Mar. 1998, at 24–31, *available at* http://www.isen.com/papers/Dawnstupid.html.

254 1 STEVENS, *supra* note 18, at 2.

network of IP routers.[255] Thus, from the point of view of applications, it is the IP network that is the global public conduit for the free flow of information.

The second crucial point is about the nature of IP routing—i.e., the nature of packet switching in the public IP networks. The routers in the public IP networks are supposed to perform simple, "dumb," or "stupid" network functions: just forward the packets to the desired destination. However, there is no predestined path of travel from the origin to the destination point. IP routing is done through a "hop-by-hop" basis—that is, the packet is routed to any router that is willing to pass the packet to another router from which the destination is reachable.[256] Thus, the function of the "Internet" (public IP network) does not depend on the operation of any particular router or component network. One router is like any other. The uniform IP protocol, along with the "stupid" hop-by-hop routing design, makes the Internet a single, unified, seamless global network, where networks or routers can be added as indistinguishable "commodities."[257] This aspect of Internet design is largely responsible for the legendary scalability, efficiency, and survivability of the spectacularly successful global Internet.

For our purpose, the key point to note is that the design of IP routing, or the process of IP packet switching, relies on the availability of all routers reachable between the source and destination points. Because no geographical distinction is made between the routers, all of the public routers around the world are part of the seamless global public Internet, and everyone on the Internet essentially relies on all of those routers in order for the Internet to be the medium of seamless global communication. In fact, on the global Internet, it is not unusual for packets to be routed through several countries en route to a destination.

This fundamental property of the Internet as a public IP network is inconsistent with China's Internet policy. For example, if China were to implement a gigantic national intranet, the routers inside the private intranet by definition do not participate in the public IP network. Therefore, the public outside China would not be able to rely on the thousands of routers located inside China as part of a public IP network. However, by seeking benefits from the Internet, China takes advantage of the millions of routers in other countries made available

255 INFORMATION SCIENCES INST., UNIV. S. CAL., INTERNET PROTOCOL: DARPA INTERNET PROGRAM SPECIFICATION 1 (Jon Postel ed., 1981), *available at* http://www.ietf.org/rfc/rfc0791.txt ("The Internet Protocol is designed for use in interconnected systems of packet-switched computer communication networks.").

256 1 STEVENS, *supra* note 18, at 38.

257 Isenberg, *supra* note 253, at 30.

for the global public Internet. In effect, China's tens of millions of users would burden the world's public IP networks without contributing its share by allowing its routers to be used by the global IP network. This point provides a distinction between private corporations and China. Running the Internet as an intranet may be acceptable or tolerable if done by an organization with thousands of users, but, if practiced by a nation with tens of millions of users, the result would be very disruptive to the operation of the global Internet as a public conduit.

China gave up on building a national intranet fairly early on.[258] Rather, it blocks or filters traffic at the backbone gateways to the global Internet, and requires blocking on the ISPs that connect to the backbone. Although this practice does not make the routers inside China unavailable to the global IP network per se, it interferes with the proper functioning of the seamless, transparent IP network by complicating the "stupid" routing function. For example, any packet that is routed through the sites or networks blocked by Chinese authorities could get dropped at the gateways or blocking points, even if the final destination of the packet is not located inside China. This is because the routing is done on a hop-by-hop basis—i.e., for any router, the only thing it knows about a received packet is its destination address and the address of the immediately preceding router. A router typically is not aware of the address of the origination point. Since it is difficult to tell the geographical location of the destination on the basis of the IP address, the blocking might be simply based on the source IP address, in which case, packets routed through the blocked sites or networks will get dropped, even if the final destination of the packets is not located inside China.

In effect, the routers in China would be committing a sort of routing fraud, by agreeing to route the packet, then dropping it in the middle. Although the routers themselves may not have been programmed *intentionally* to mislead, the resulting effect would be the same and this effect is a foreseeable consequence of the policy China would adopt. To the extent that China claims a sovereign right to control routers and networks located within its borders, China cannot avoid its culpability for the results of its specific blocking policy on the global Internet. This is an unavoidable consequence of China's layer-

258 On the other hand, it is interesting to note that Myanmar and Vietnam still maintain the "intranet" model. Myanmar was discussed above. For the case of Vietnam, see Minutes of the U.N.D.P. "Partnership to Fight Poverty" Meeting, *Internet Players in Vietnam* (Oct. 31, 2001), *at* http://www.isoc-vn.org/www/archive/011013-Minutes.html.

violating regulation at the IP layer that is enforced to control activities at the content layer. Implementing content controlling functions deep at the IP layer, especially at the backbones or national gateways, is fundamentally inconsistent with the workings of the Internet as a global public conduit—the *res publicae* for flow of ideas and information. Therefore, layers analysis proves at least a narrower version of the global free speech argument—that is, due to the global, public nature of the Internet (global public network of networks of IP routers), censorship of content deep in the IP layer infringes upon the rights of everyone on the Internet by placing an unnecessary and unexpected obstacle on the global public conduit shared by everyone.

On the other hand, it is unlikely that non-layer-violating or less restrictive solutions would work for China's current policy objective. Given the types of content it wants to filter out, e.g., the *New York Times* and CNN, no content level labeling schemes such as PICS would work for China. Policing high at the content layer, i.e., policing the users in real space, is always an option. In fact, the authorities in China have recently launched a broad censorship and monitoring program called "Golden Shield," which involves, among other measures, deploying a gigantic network of digital surveillance cameras to monitor the users and PCs in China's 200,000 or more Internet cafes.[259] The system is reported to "incorporate speech and face recognition, closed-circuit television, smart cards, credit records and other surveillance technologies."[260] It is highly doubtful, however, such a system would be effective in the long run, due to its highly complex and resource intensive character, as well as the political cost and social burdens that come from the extraordinarily intrusive nature of the system. Furthermore, such an Orwellian censorship system is unlikely to be conducive to fostering an atmosphere of intellectual creativity needed for China to take advantage of the benefits of the Internet.

It is interesting to note that the Golden Shield project was instigated largely due to the difficulties or ineffectiveness of the Great Firewall model.[261] Such difficulty seems to bear out the prediction of the fit thesis[262]—that China's severe layer-violating regulation has fundamental architectural problems that would prevent China from becoming a full-fledged participant in the global Internet. Attempting to impose similar restrictions at the content layer, however, demands ex-

259 Nairne, *supra* note 233.

260 Wiggins & Liu, *supra* note 234.

261 *Id.*

262 *See supra* Part II.D.2.

traordinarily complex and intrusive measures that may have costly political consequences. It is quite possible that China's current policy regarding the Internet is a fundamentally unstable approach that is doomed to fail. Perhaps a political system that requires blocking of the contents of the *New York Times* or CNN is fundamentally at odds with the country's full participation in the global Internet.

c. The French *Yahoo!* Case

Our next application is the French *Yahoo!* case, in which the government of France sought to require Yahoo! to block access to particular content based on IP addresses as a proxy for national identity.[263] Although our discussion focuses on *Yahoo!*, there are other similar cases. One such example is Google's blocking of more than one hundred Internet sites from the French and German versions of its web based search engine.[264]

Yahoo! is an Internet portal. The name stands for "Yet another hierarchically organized outline," reflecting Yahoo!'s origins as a collection of web links organized hierarchically in outline form. Yahoo! is now a general purpose portal site, incorporating the original outline and adding a wide range of web based services, such as yellow pages, mapping, movie guides, a search engine, and so forth. Yahoo! originally had a North American focus, reflecting its California origins and the fact that Internet use was heavily concentrated in North America at the time Yahoo! came into being.[265] Today, Yahoo! has a global reach, with specialized Yahoo! portals for a variety of nations,[266] including France.[267] Among Yahoo!'s offerings is an auction site, which is accessible from its French portal; and among the items offered for sale on Yahoo!'s auction site are war memorabilia, including items connected to the Nazi party. The French government at-

263 *See generally* Current Development, *Enforcement of French Court Order Barring Yahoo! Internet Auction Sale of Nazi Material Would Violate First Amendment*, 19 COMPUTER & INTERNET L. 24 (2002); Elissa A. Okoniewski, Comment, Yahoo!, Inc. v. LICRA: *The French Challenge to Free Expression on the Internet*, 18 AM. U. INT'L L. REV. 295 (2002); Pamela G. Smith, Comment, *Free Speech on the World Wide Web: A Comparison Between French and United States Policy with a Focus on* UEJF v. Yahoo! Inc., 21 PENN ST. INT'L L. REV. 319 (2003).

264 *See* John Schwartz, *Study Tallies Sites Blocked by Google*, N.Y. TIMES, Oct. 25, 2002, at C8.

265 *See* Yahoo!, *The History of Yahoo!—How It All Started . . .*, *at* http://docs.yahoo.com/info/misc/history.html (last visited Feb. 13, 2004).

266 *See* Yahoo!, *More Yahoo!*, *at* http://docs.yahoo.com/docs/family/more.html (last visited Feb. 13, 2004) (listing twenty-five regional Yahoo!s).

267 *See* Yahoo! France, *Homepage*, *at* http://fr.yahoo.com (last visited Feb. 13, 2004).

tempted to prevent Yahoo! from auctioning Nazi items to or from France.

The French *Yahoo!* case illustrates that the problems created by the Internet's global transparency are not limited to China or non-Western nations. France and the United States are two of the oldest and most advanced modern democracies in the world. However, even among such nations, the transparency of the Internet can give rise to conflicts as a result of the clash of differing cultural, political, or legal norms or values. The French *Yahoo!* case is the case in point.

The French *Yahoo!* case raises important questions regarding the interaction of national sovereignty, the nature of the Internet or cyberspace, and the nature and limitations of regulation of the Internet. First, the case makes clear that the problems of clashing cultural and political values that the Internet's transparency brings cannot be simply brushed off by denouncing the perceived backwardness of developing countries or non-Western societies. France and the United States are not sworn enemies who subscribe to diametrically opposed worldviews. Quite the contrary, they are among the closest allies whose political and cultural values are similar in many important ways—perhaps indistinguishable from the perspective of some non-Western cultures. Yet, there are significant enough differences between them that result in exactly the same kind of problems that the transparent Internet brings to China or non-Western nations.

The case also serves as an interesting illustration of an important type of layer-violating regulation at the IP layer—nationality or territory based policing of the Internet using IP addresses. Under this type of regulation, restrictions on the contents are imposed along the national or territorial boundaries, using IP addresses to identify the national or territorial location of the users. Below, implications and problems of such regulation are analyzed under the layers analysis framework. Then, possible non-layer-violating alternatives are considered. Finally, a non-layer-violating solution in the code layer, use of digital certificates, is proposed as a much more effective approach with far fewer problems than the layer-violating regulation at the IP layer.

i. The French Ruling

In early 2000, two French organizations, La Ligue Contre Le Racisme Et L'Antisémitisme (League Against Racism and Antisemitism—LICRA) and L'Union Des Étudiants Juifs De France (French Union of Jewish Students) brought a legal action against Yahoo! in

the Tribunal de Grande Instance de Paris (the "French court").[268] The French petitioners claimed that the sale of Nazi and Third Reich-related goods through Yahoo!'s auction site that is accessible in France violated section R645-1 of the French Criminal Code, which prohibits exhibition of Nazi propaganda and artifacts for sale.[269] Yahoo! is never a party in the transaction, as any buying or selling takes place directly between the individual users. The transactions are automated through Yahoo!'s auction server, which is located in the United States.[270] Nevertheless, because any French citizen is able to access the Nazi-related materials on the website, the French court found that Yahoo!'s auction site violates French law, and entered an order requiring Yahoo! to block French citizen's access to the areas of Yahoo!'s auction site that offer for sale any Nazi-related materials.[271]

Yahoo! asked the French court to reconsider its order, claiming that it was technologically impossible to selectively block French citizens from its auction site.[272] Relying on the testimonies from experts, however, the French court found that blocking access from French citizens was technologically feasible.[273] The panel of experts, including Vinton Cerf, one of the original designers of the Internet, testified that some seventy percent of the IP addresses of French users or users in French territory could be correctly identified to be located within the French territory.[274] By combining the IP address method with voluntary identification or registration, the experts opined that Yahoo! could filter French users with a success rate approaching ninety percent.[275] The French court thus denied Yahoo!'s request and reaffirmed its earlier order.[276]

ii. The U.S. Court's Response

The final ruling of the French court appears to be based on the technical feasibility of blocking users on the Internet based on their geographical location. However, a U.S. federal district court, a paral-

268 Yahoo!, Inc. v. La Ligue Contre Le Racisme et L'Antisemitisme, 169 F. Supp. 2d 1181, 1184 (N.D. Cal. 2001).

269 C. PÉN. R645-1.

270 *Yahoo!, Inc.*, at 1183.

271 *Id.* at 1188.

272 *Id.* at 1184.

273 *Id.* at 1185.

274 *Experts Testify in French Yahoo! Case Over Nazi Memorabilia*, CNN.COM, Nov. 6, 2000, *at* http://www.cnn.com/2000/TECH/computing/11/06/france.yahoo.trial.ap [hereinafter *Experts*].

275 *Yahoo!, Inc.*, 169 F. Supp. 2d at 1185.

276 *Id.* at 1194

lel decisionmaking body in the United States, did not rest on such grounds when rejecting enforcement of the French order in this country. Rather, the U.S. court concluded that the French order could not be enforced in the United States because it would violate the First Amendment of the U.S. Constitution, *regardless* of the technical feasibility of the selective nationality based blocking.[277]

In response to the French court's ruling, Yahoo! filed a complaint against the French organizations in the U.S. District Court for the Northern District of California, seeking a declaratory judgment that the French court's orders were neither cognizable nor enforceable under the laws of the United States.[278] The defendants immediately moved to dismiss on the basis that the district court lacked personal jurisdiction over them.[279] That motion was denied.[280] Yahoo! then moved for summary judgment on its declaratory judgment claim.[281]

When considering the case on its merits, the court first noted that the case was not about the right of France or any other nation to determine its own law and social policies.[282] "A basic function of a sovereign state is to determine by law what forms of speech and conduct are acceptable within its borders," said the court.[283] The issue was, according to the court, "whether it is consistent with the Constitution and laws of the United States for another nation to regulate speech by a United States resident within the United States on the basis that such speech can be accessed by Internet users in that nation."[284]

Although there is potentially a substantial choice of law issue, the court adopted the U.S. laws by simply stating that "it must and will decide this case in accordance with the Constitution and laws of the United States" without engaging in international choice of law analysis.[285] Having decided that U.S. laws applied, the court found that the French order was a viewpoint based regulation that was impermissible under the First Amendment of the U.S. Constitution. The court said that "[a]lthough France has the sovereign right to regulate what speech is permissible in France, this court may not enforce a foreign order that violates the protections of the United States Constitution by

277 *Id.*

278 *Id.* at 1186.

279 *Id.*

280 *Id.*

281 *Id.* at 1185.

282 *Id.* at 1186.

283 *Id.* Hence, at least one court in United States should have no problem with China's regulation of the Internet.

284 *Id.*

285 *Id.* at 1187.

chilling protected speech that occurs . . . within our borders."[286] Accordingly, Yahoo!'s motion for summary judgment was granted.

iii. Application of Layers Analysis to the French *Yahoo!* Case

The French *Yahoo!* case raises important questions regarding the interaction of national sovereignty, the nature of the Internet or of cyberspace, and the nature and limitations of regulation of the Internet. Many of the questions raised do not find ready answers under existing legal frameworks, be they local, national, or international. And, proper analysis of all of the issues raised by the French *Yahoo!* case would be more than enough to consume several full length articles.

For example, neither the French nor U.S. court's analysis of personal jurisdiction and choice of law issues is illuminating or satisfying. Does Yahoo!'s website, located within the United States, violate French law? Is it subject to the jurisdiction of the French courts? If so, which law applies? The French court completely sidestepped the issues by simply declaring that Yahoo!'s U.S. website violated French law because the French users can access the site, without engaging in even a cursory analysis of why such a position is justified and what consequences such a position entails. The U.S. district court did somewhat better, although the court's analysis still leaves many issues unresolved. However, discussion of the complex issues of personal jurisdiction and conflict of laws is not taken up in this Article, as a proper treatment would require a full length article on this topic alone.

Instead, this Article will focus on the feasibility and desirability of the regulation at the IP layer. Assuming, arguendo, that a national sovereign can legitimately regulate content on the Internet along national territorial boundaries, is it possible to do so based on IP addresses? Even if possible, is it wise to do so? What are the alternate approaches? Are there better solutions that are more effective?

a) General Layers Analysis of the French *Yahoo!* Case

Even within these narrow confines, the French and U.S. courts' analyses are rather incomplete or outright unsatisfactory. The French court, for example, simply accepted the report from the experts on the technological feasibility of identifying territorial identity of users based on their IP addresses. It then turned the expert opinion on the narrow question of technical feasibility into a legal mandate without

286 *Id.* at 1192.

any analysis or critical assessment of the consequences of such order. The absence of analysis seems especially inadequate when considering the fact that the main expert witness, Vinton Cerf, publicly expressed his reservation on the wisdom of the blocking measure before the French court's final ruling.[287]

The U.S. district court, on the other hand, sidestepped the issue of blocking by concluding that, *regardless* of the technical feasibility of the selective nationality based blocking, the French order could not be enforced in the United States because it would violate the First Amendment of the U.S. Constitution by "chilling protected speech that occurs . . . within our borders."[288] However, the court did not explain how blocking French citizens (and/or users from French territories) chills speech in the United States. Did the court mean that chilling occurs in the United States because Yahoo!'s servers are located in this country? If the only people affected by the blocking measure are all physically located outside the United States, how does any chilling of speech occur in this country? The court provided no analysis of these fairly obvious questions.

In this regard, the U.S. court would have benefited from analysis of the "technological feasibility," as it would have shed light on the issue of chilling effect in a very concrete way. Among the methods of identifying the nationality of the users, voluntary registration is most likely to be an ineffective method. Only the IP address based method would have any significant basis for success. Taking the experts' opinion at its face value, it is possible to identify the nationality of the location of the computers on the Internet on the basis of their IP addresses with about seventy percent accuracy. Thus, the blocking ordered by the French court would block about seventy percent of the users from France, as well as many outside the country. At least theoretically, it is possible that some of the thirty percent of the users who are erroneously blocked reside in the United States. Therefore, it is at least plausible that there will be some chilling effect in the United States.

The real problem, however, is what the seventy percent figure represents. First of all, the experts themselves emphasized that "there

287 *See Experts*, *supra* note 274. Cerf, a founding father of the Internet, expressed doubts about whether such attempts were worthwhile. "There are 100 million Internet sites in the world," he said, "[i]n five years, there will be a billion. Even if we only block some of them, the list is long. And if we block too many of them, we risk blocking the whole system." *Id.*

288 *Yahoo!, Inc.*, 169 F. Supp. 2d at 1192.

is no evidence to suggest that the same will apply in the future."[289] That is, the figure has no predictive value. Seventy percent is a statistical figure that is estimated from the accidental history of how IP addresses were allocated around the world.[290] It is not based on architecture or design constraints of the Internet. And there is nothing in the architecture of the Internet to prevent reassignment or reallocation of blocks of IP addresses to a different region or country. Thus, the seventy percent figure given by the experts is not a stable figure that can be a basis for lasting regulation or policy. Therefore, any territorial blocking based on IP addresses is inherently overinclusive and underinclusive. How severe the problem is may be anyone's guess at a particular point in time.

The second problem is that IP address based blocking can be easily circumvented by readily available technologies such as the anonymizers. With such technology, users connect to a site through another server that hides the true origin of the user, and determination of the geographic location of the user based on IP address is consequently made impossible.[291] Thus, anonymizers can render the regulation ordered by the French court entirely underinclusive where it really counts. That is, those users in France who insist on purchasing the Nazi-related materials on the Yahoo! site will easily circumvent the IP based blocking, and the regulation would have very little impact.

Granted, there are technologies that can be used to "dynamically discover" the geographic location of the hosts based on their IP addresses.[292] But, these technologies involve some sort of guessing game—e.g., guessing physical distance by network response delay—that cannot guarantee a high rate of success.[293] In fact, under the current Internet architecture, any effort to map IP addresses to geographic location is bound to be a guessing game. This is because allocation and distribution of IP addresses are fundamentally activities

289 La Ligue Contre Le Racisme et L'Antisémitisme v. La Société Yahoo! Inc., T.G.I. Paris (Nov. 20, 2000), *available at* http://www.cdt.org/speech/international/001120yahoofrance.pdf.

290 Memorandum from K. Hubbard et al., Network Working Group, to the Internet Community (Nov. 1996), *at* http://www.ietf.org/rfc/rfc2050.txt.

291 *La Ligue Contre Le Racisme et L'Antisémitisme*, *supra* note 289.

292 *See, e.g.*, Venkata N. Padmanabhan & Lakshminarayanan Subramanian, *An Investigation of Geographic Mapping Techniques for Internet Hosts*, 2001 PROC. ACM SIGCOMM 01, at 173, *available at* http://www.acm.org/sigs/sigcomm/sigcomm2001/p14-pabmanabhan.pdf (evaluating the performance of various techniques for determining the geographic location of Internet hosts, and identifying challenges in deducing geographic location from an IP address).

293 *Id.* at 178.

above the TCP/IP layers—addresses are assigned without any *architectural* constraints or mandates—and information about them is simply not communicated to the IP layer, as required by the principle of separation of layers. Thus, all efforts to map IP addresses to geographic location are in effect layer-violating actions, and as such, mapping has the inherent problems of being overinclusive and underinclusive—problems which we summarize as the fit thesis.[294]

Therefore, the U.S. district court's conclusion that enforcement of the French court's order would have a chilling effect on protected speech in this country may have been correct after all. It is important to note, however, that the crucial supporting reason for the chilling effect conclusion can only be identified by analyzing technological feasibility, using the tools provided by the layers model. This leads to an observation that has substantial significance, both in the French *Yahoo!* case and in other contexts: whenever technological feasibility is mentioned in the context of Internet regulation, what is really involved most of the time is the interaction between the architecture of the Internet and the law. Recall that the code thesis told us that the nature of Internet or cyberspace is determined by the code that implements it. Code is the prime regulator on the Internet. Thus, in Internet regulation, what is really discussed under technological feasibility is the possibility of change in code (or architecture) as the prime regulator of the Internet. The legal regulation will be only as good as is permitted by the architecture of the Internet, and the nature and limitations of the legal regulation will be determined by the nature of the code being implemented. Therefore, whenever an analyst or policymaker is evaluating a given Internet regulation, analysis of a proposed change in code and its interaction with legal regulation should be included as part of a comprehensive analysis. One of the goals of this Article is to show that layers analysis in conjunction with end-to-end analysis provides an effective approach to the analysis of regulation by law in the context of the architecture that determines the effect the law will have. The analysis of the layer-violating regulation ordered by the French court in this subsection, as well as the analysis of the Internet as *res publicae* in cyberspace in the China section above, illustrates this point.

b) Non-Layer-Violating Alternatives Analysis

If we were to take the French court's order to require blocking of French users based on their IP addresses, that would be a regulation

294 *See supra* Part II.D.2.

directed at the IP layer, due to problems at the content layer. In contrast to China's blocking of entire websites and services, however, the French court's order mandates only a selective blocking of certain materials on Yahoo!'s site that are found to be objectionable in France. In addition, the blocking is directed at the IP addresses of the users accessing the site, not the site itself.

On the other hand, the blocking would affect a large number of people—i.e., most users in France and many outside the country. Thus, it would significantly impact overall Internet transparency by affecting the transparency for a large number of people. In addition, the regulation is likely to suffer from rather substantial innocent use problems, as discussed above. A consequence of the substantial innocent use problem is a chilling effect on protected speech in the United States.

What is France to do, then? Are there non-layer-violating solutions that work just as well or better? What if there is no non-layer-violating solution? Given that controlling importation of Nazi-related material is justifiably important for France, should we accept the layer-violating solution because no viable non-layer-violating solution exists, and the regulation is justified by a compelling reason?

Consider non-layer-violating alternatives. The most obvious alternatives are traditional models of regulation in real space at the content layer. They are: traditional real space regulation within national borders, international agreements, and supranational regulation frameworks such as that provided by the World Trade Organization (WTO).

Under traditional national real space regulation, importation or transaction of offending materials can be regulated at or within the nation's borders. For example, if a French citizen were to purchase a Nazi item on Yahoo!'s auction site from a foreign seller, France can stop the material from coming into the country by enforcing the ban on Nazi materials at customs entry points. If the transaction is between French citizens, the task of control is not any different from general policing of those materials within France.

The problem with this approach, however, is speech and digital content. If what is prohibited is speech—as is the case in France where *display* of the Nazi-related item itself is banned whether or not any sale takes place—then controlling physical goods at or within France's borders has no effect on banned speech on the Internet posted by sites outside the national borders. Digital content presents even greater problems. Digital materials can be downloaded, bought, and sold directly over the Internet. Such material could be legal in one country but illegal in another. For example, Hitler's *Mein Kampf*

may be available for sale in eBook format to buyers in France over the Internet from a seller in the United States, where such material is not illegal. Although, in theory, nations can regulate possession, buying, and selling of illegal materials within their borders, national regulation of digital content transmitted over the Internet would be nearly impossible unless the activities on the Internet can be effectively monitored and selectively blocked.

Alternatively, nations can agree on a uniform standard of conduct on the Internet through some international cooperation mechanism.[295] Then, the policing can presumably be done by each nation within its borders, which may not be any different from regulation of any other problem. However, a problem common to all international agreement approaches is that the process of negotiation and settlement is extremely complicated, laborious, and slow.[296] Thus, most international agreement models may not be suitable approaches for regulation on the Internet, due to dizzying complexity and vast diversity of the issues involved, as well as the lightning speed of the change of technology and patterns of conduct on the Internet.

Another problem characteristic of all existing models of international regulation is that they are primarily designed for negotiation and transaction among relatively sophisticated parties, such as governments, public organizations, or businesses.[297] They are not designed to deal with millions of individual citizens around the world. As an illustration of this point, consider the problem of tariffs in international e-commerce transactions. Enforcement of tariffs at nations' borders involves rather complicated questions of classification, valuation, and origin.[298] The amount of tariff imposed depends upon which category the product belongs under the nation's tariff schedule, how much the good is worth, and from what country the product

295 For an overview of various international regulatory approaches to the Internet, see BIEGEL, *supra* note 186, at 157–86.

296 *See* Helfer & Dinwoodie, *supra* note 44, at 145–46.

> International dispute settlement mechanisms do not spring up overnight. Instead, they are carefully built, often with agonizing slowness, through a process of give-and-take among negotiators wrangling over the subject matter of disputes, the procedures for adjudicating them, the identity of parties who can bring claims, and the authority of the decision makers who will rule on those claims.

Id.

297 LESSIG, *supra* note 15, at 197. ("[I]nternational agreements for the most part are agreements between sophisticated actors. Before cyberspace, ordinary consumers were not international actors.").

298 *See* DAVID SERKO, IMPORT PRACTICE: CUSTOMS AND INTERNATIONAL TRADE LAW 69–133 (2d ed. 1991).

is deemed to originate. Although standards for these issues have been made uniform under the WTO for its member countries, these questions frequently involve complex legal issues that are beyond the reach of most ordinary citizens. When citizens of different countries buy or sell digital products through the Yahoo! auction site, the issues of tariff determination and enforcement present great difficulties to the traditional model. Neither the individuals nor Yahoo! can be expected to know how to determine the tariffs. And Yahoo! cannot be expected to enforce tariff imposition and payments to proper government bodies. Enforcement of quotas has the same difficulties. At present, the WTO has a temporary moratorium on customs duties for digital transactions over the Internet, in part due to these difficulties.[299]

The WTO framework is by far the strongest and most effective international agreement mechanism with real enforcement powers.[300] But, the current WTO system, even limited to the questions of international trade, does not seem to be an effective approach to international Internet regulation. Moreover, the WTO is limited in scope to trade issues, and for that reason is not an appropriate vehicle for addressing many problems raised by digital content. In sum, for resolving much more complex problems like the clash of cultural, political, or legal values concerning speech, as represented by the French *Yahoo!* case, none of the existing models of real space regulation at the content layer is likely to be effective.

It seems, then, that regulation at the code layers is inevitable.[301] If we accept the proposition that regulation of Nazi materials is important for the French government, then it may be necessary to institute some change in code—i.e., changes to the architecture of the Internet that are justified by a compelling purpose. When considering appropriate code changes, however, the architecture of the Internet must be taken into account. The principle of minimizing layer crossing tells us that the closer the regulation layer is to the content layer, the less the impairment of transparency and the impact on substantial innocent use. Minimizing layer crossing tells us to place the solution as close as possible to the content layer where the function is ultimately

299 *See* WORLD TRADE ORGANIZATION, THE GENEVA MINISTERIAL DECLARATION ON GLOBAL ELECTRONIC COMMERCE (1998), *available at* http://www.wto.org/english/tratop_e/ecom_e/mindec1_e.htm.

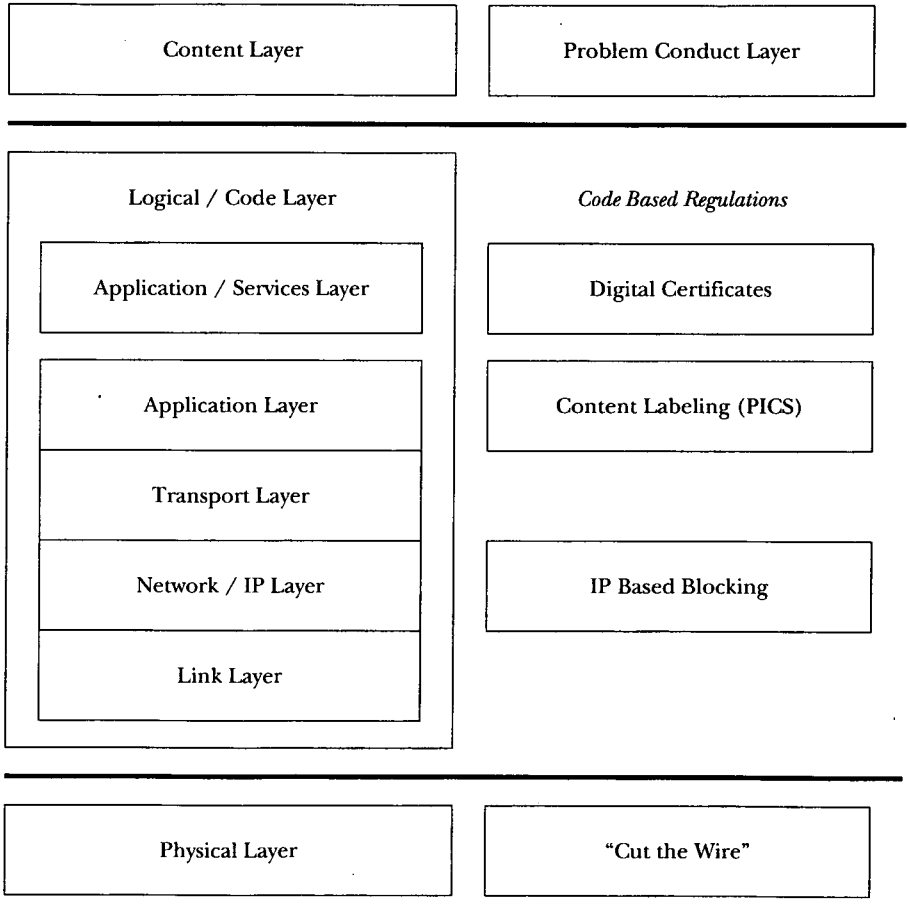
300 Joost Pauwelyn, Note and Comment, *Enforcement and Countermeasures in the WTO: Rules are Rules—Toward a More Collective Approach*, 94 AM. J. INT'L L. 335, 339 (2000).

301 It is taken as obvious that France will not resort to "cut the wire" regulation at the physical layer.

consumed. The principle says that the closest-possible-to-the-content-layer solution will be more effective because more complete information is available to implement the function properly.

The code change ordered by the French court is blocking based on IP addresses—i.e., code change at the IP layer. Other choices are content labeling such as the PICS at the application (protocol) layer, and digital certificates at the application/services layer.[302] These options are illustrated by Figure 8 below.

FIGURE 8. CODE BASED REGULATION CHOICES



Among the choices, the layer principle and end-to-end argument tell us to choose the digital certificate solution. The principles predict

302 Recall that application protocols are for Internet communications at the application level, such as HTTP, while applications/services are application programs themselves, such as web browsers.

that the digital certificate approach will be, in effect, the least restrictive or discriminating means to achieve the compelling end of the French government. And this is, in fact, the case. Digital certificates are electronic files that can serve to identify the users or computers, including their nationality or territorial location.[303] The certificates are issued by a trusted third party and are encrypted with strong encryption technology so that they cannot be forged or tampered with. Verification of identity can be done securely, privately, and anonymously by using automated authentication servers.[304] Under this approach, the French government can enforce installation of digital certificates with national identification on the computers and web browser programs located within France's territory. Then, Yahoo!'s auction server can receive the digital certificates from the users or computers that access the site, and the server can refuse access to the areas that contain Nazi-related items, if the users or computers are found to be from French territory—based on the authentication result of their digital certificates. This approach would have a far higher success rate than a complicated guessing game of unstable mapping of IP addresses to geographic locations.

It is important to note that, although the digital certificate approach is a code layer solution, it does not target networking layers—i.e., one of the TCP/IP layers. Thus, it is not a layer-violating regulation—this concept is limited to regulations crossing into the TCP/IP layers. As a result, it has no effect on the networking function of the global Internet. And, its impact on the transparency of the Internet, as well as the problems of substantial innocent use, should be minimal, if any. The effects of adopting digital certificates are limited to the application/service space—it is like adopting any other program or software, such as choosing a web browser or media player. In fact, digital certificates are already in wide use without any impact on the network function of the Internet. Whenever we go online with a "secure site," such as an online purchase at Amazon.com, we are using the server's digital certificate to encrypt the data so that the transaction is "secure," even though the transaction data is sent over the public networks that make up the global Internet. There is little evidence that such wide use of digital certificates on the Internet has any deleterious impact on the workings of the Internet. Such a change with

303 *See* Netscape, *Tech Briefs*, at http://wp.netscape.com/security/techbriefs/index.html (last visited Feb. 20, 2004).

304 *See, e.g.*, VeriSign Inc., *Authentication Service Bureau*, *at* http://www.verisign.com/products/asb/index.html (last visited Feb. 20, 2004) (detailing the company's product offerings that enable an enterprise to authenticate the identities of customers, medical professionals, and business entities in online transactions).

minimal impact would certainly be justified by the French government's compelling purpose to control the dissemination of Nazi materials.

Our purpose is not to advocate (or oppose) a digital certificate approach. It may well be that the disadvantages of digital certificates outweigh their advantages. It is possible that the French government simply cannot achieve its goal of effectively regulating Nazi paraphernalia. Our point is that analysis of the French *Yahoo!* case is illuminated and clarified by layers analysis. Without layers analysis, the French government is essentially shooting in the dark without night vision equipment. With layers analysis, the French government has the tools to assess the effectiveness and costs of various policy alternatives.

d. Cyberterrorism

Although there appears to be a non-layer-violating solution for the French *Yahoo!* case, it is still an open question to ask whether the Internet's global transparency should be maintained at all cost. The area where this question is most relevant and pressing is the issue of security—i.e., concerns about cyberattacks or cyberterrorism over the Internet—especially in light of the post-September 11 realities. The security concerns in this regard are a clear example of a compelling purpose.

Security issues that arise because of the transparent global Internet, or "cybersecurity," are identified as one of the key areas of security concerns by the current U.S. administration.[305] The new Homeland Security Administration includes an Information Analysis and Infrastructure Protection Division. The new department would comprehensively assess the vulnerability of America's key assets and critical infrastructures, including the information and telecommunications systems. Most of the groups or agencies within the federal government that deal with computer or cybersecurity are set to be transferred to this department.[306] In addition, the President's budget for 2003, the federal government's first post-September 11 budget, specifically requested significantly increased funding for "cyberspace security, an essential new mission for the 21st century given our grow-

305 See Press Release, U.S. Dep't of Homeland Security, Remarks by Secretary Tom Ridge to the Council for Exellence in Government (Sept. 16, 2003), *available at* http://www.dhs.gov/dhspublic/display?content=1597.

306 Homeland Security Act of 2002, Pub. L. No. 107-296, 116 Stat. 2135 (2002).

ing dependence on critical information infrastructure, most importantly the Internet."[307] The President's budget emphasized:

> The information technology revolution has changed the way business is transacted, government operates and national defense is conducted. These three functions are now fueled by an interdependent network of critical information infrastructures of which the Internet is key. America must do more to strengthen security on the Internet to protect our critical infrastructure.[308]

The budget called for creation of a Cyberspace Warning Intelligence Network (CWIN) that would link the major players in government and the private sector to manage future cyberspace crises, in addition to requesting money for a feasibility study for a secure "GovNet."[309]

A recent cybersecurity report by Congress identifies critical infrastructures as: information and communications; electric power generation, transmission, and distribution; oil and gas storage and distribution; banking and finance; transportation; water supply; and emergency assistance.[310] The report further acknowledges the challenges "arising from an increased dependence on information systems and networks to operate critical infrastructures."[311] That is, all areas of critical infrastructures may be vulnerable or be affected by attacks or terrorism perpetrated via cyberspace or over the Internet.

Already, we have seen a "hacker" from Sweden disable portions of the emergency 911 system in southern Florida, and a Massachusetts teenager disable communications to an aviation control tower.[312] What would happen if members of the al Qaeda network break into the computer system that controls the operation of a nuclear power plant? A national power grid? The Hoover Dam? The results could be as devastating as (or far worse than) the September 11 attacks on the World Trade Center and the Pentagon.

There are now calls to build a "national firewall"—China's Great Firewall does not seem so bad now—and quarantine or block Internet traffic from certain nations.[313] Suddenly, the transparency of the In-

---

307 GEORGE W. BUSH, SECURING THE HOMELAND, STRENGTHENING THE NATION 19 (2002), *available at* http://www.whitehouse.gov/homeland/homeland_security_book.pdf.

308 *Id.*

309 *Id.* at 22.

310 JOINT ECONOMIC COMM., 107TH CONG., SECURITY IN THE INFORMATION AGE: NEW CHALLENGES, NEW STRATEGIES 12 (Comm. Print 2002).

311 *Id.* at 13.

312 *Id.* at 21.

313 Simson Garfinkel, *Leaky Cyber Borders*, TECH. REV., June 2002, at 31.

ternet may be regarded as much more of a liability than an asset. When there are real concerns about the consequences of porous physical borders, arguments for having no borders—transparency—in cyberspace may not fall on sympathetic ears.

The concerns over cyberterrorism are compelling reasons for adopting regulations, although compelling motivations are never the reason to adopt misguided or inappropriate means without proper analysis. Nevertheless, if the West or the United States adopts regulations that severely impair the transparency of the Internet, it is most likely to be a result of hurried or reflexive reactions to September 11, or subsequent threats of terrorism and a serious erosion of everyday safety. And, once adopted in the context of security concerns, the layer-violating regulations will also be used in other contexts.

On the other hand, we also need to keep in mind that transparency across national borders—global transparency—was really not the purpose or original design of the Internet. For the original Internet, transparency meant within the United States or, at most, transparency among the Western nations. As we discussed, the Internet was justified, if not created, in some part, by the Cold War military project to build a communication network that would survive catastrophic failures.[314] It was designed to be used initially by the U.S. military, government agencies, universities, and researchers working on the project, who were by and large trusted parties. Thus, the advantages of transparency clearly outweighed any possible problems it might cause.

That certainly is not the situation on today's global Internet. Given the post-September 11 realities, perhaps the world is not yet ready for the utopian total transparency of the original Internet design. Certain transparency losses—"tuning" of transparency across national borders—may be necessary considering today's global political reality. Even if this is the case, however, there still is the question of what the appropriate mechanism is to control or tune transparency across national borders.

There is a need to consider, for example, whether IP based blocking really gives us better security. As evidenced in the French *Yahoo!* case, the problems of overblocking and underblocking resulting from the inherent violation of the Internet design will plague effectiveness as well. The IP address was never designed to be the basis of geographic or national identification. As we saw, solutions implemented at the IP layer to address the problems at the content layer inherently

314 *See* KATIE HAFNER & MATTHEW LYON, WHERE WIZARDS STAY UP LATE: THE ORIGINS OF THE INTERNET 54–81 (1996).

cannot work effectively because it is very difficult, if not impossible, to convey content layer information—geographic allocation and assignment of IP addresses—to the IP layer. Furthermore, as analyzed in the China section, national firewalls operating at the IP layer have serious destructive consequences to the proper functioning of the global IP networks. If national IP firewalls are deployed by the United States and a few other major Western nations, it would essentially mean the end of Internet as we know it—at the seamless global network.

As was the case in the French *Yahoo!* case, use of a digital certificate with encoded national ID would be both more effective and less damaging to the transparency of the Internet. Granted there may be problems with misappropriation or counterfeiting of digital certificates, but such a problem isn't any worse than the problem of controlling passports in real space. National firewalls—really application proxy servers in this case—that control traffic based on digital certificates would work much more effectively than IP based firewalls, with far less harmful effect on the workings of the global Internet.

e. IP Address Blocking and Child Pornography

Yet another example of an IP layer regulation targeted at a content layer problem is a Pennsylvania statute that requires ISPs to block access by Pennsylvanians to IP addresses associated with servers that provide access to child pornography.[315] The statute provides in relevant part:

> An internet service provider shall remove or disable access to child pornography items residing on or accessible through its service in a manner accessible to persons located within this Commonwealth within five business days of when the Internet service provider is notified by the Attorney General pursuant to section 7628 (relating to notification procedure) that child pornography items reside on or are accessible through its service.[316]

Jonathan Zittrain, in his important article, *Internet Points of Control*,[317] notes that the Pennsylvania statute represents an important new category of Internet regulation—regulations targeted at the "destination ISP."[318]

This form of regulation has strong attractions for public Internet regulators, for a variety of reasons. First and foremost, destination

315 18 PA. CONS. STAT. ANN. § 7622 (West Supp. 2003).

316 *See id.*

317 Zittrain, *supra* note 206.

318 *Id.* at 672.

ISPs fit easily into traditional paradigms of geographically based regulatory authority. Commercial ISPs enter into contracts with customers, and rely on local physical infrastructure (wireline telephony, cable, or satellite dish) to bring the Internet to their customers. Thus, an ISP "does business" in, and has "contacts" with, the nation, state, or locality that might seek to regulate the ISP.[319] Such contacts provide a basis for the two crucial prerequisites of effective regulatory power: personal (or territorial) jurisdiction and choice of law.

Second, destination ISPs do not have a substantial economic interest in contesting IP address blocking. From the customer's point of view, the impact is opaque and minimal—opaque because the customer may never learn that she could not connect with a server because her ISP was blocking the address; minimal because only a tiny fraction of all IP addresses will be blocked. Because customers are unlikely to switch ISPs because of compliance with an address blocking regulation, the ISP is not placed at a competitive disadvantage. Of course, address blocking imposes some costs on the ISP, but those costs are not likely to be so substantial as to justify resistance by litigation or lobbying.

Pennsylvania has employed this statute to force WorldCom to block the ISP addresses of at least two content providers.[320] WorldCom's objections to this action are illuminated in juxtaposition to layers analysis. WorldCom argued that it was unable to limit IP address blocking to Pennsylvania customers; it would have to block all WorldCom subscribers, irrespective of geography.[321] This argument is perfectly congruent with the fit thesis, because the regulation uses an IP layer solution to a content layer problem, it is inherently overinclusive. As Zittrain notes,[322] a similar problem has led one district court to invalidate an Internet regulation on the basis of the dormant Commerce Clause.[323] Moreover, IP address blocking is at best a crude tool, because IP addresses change over time or, in the case of an individual using a peer-to-peer file sharing program, may be assigned dy-

319 *Id.* at 672–73 ("Destination ISPs are by their nature local, easing jurisdictional concerns since ISPs will have equipment and assets within the reach of the interested jurisdiction.").

320 *See id.* at 674–75 (describing proceedings in the Pennsylvania trial court).

321 Letter from Craig Silliman, Director of Technology and Network Legal, WorldCom, to John J. Burfete, Jr., Chief Deputy Attorney General, Office of the Attorney General of Pennsylvania (September 23, 2002) (on file with Jonathan Zittrain), *cited in* Zittrain, *supra* note 206, at 675.

322 *See* Zittrain, *supra* note 206, at 676.

323 *See* Am. Libraries Ass'n v. Pataki, 969 F. Supp. 160, 183–84 (S.D.N.Y. 1997).

namically. The IP addresses that are blocked may well have innocent uses the day, week, or year after the blocking order is made.[324]

Of course, child pornography is a great evil, and therefore, provides a compelling regulatory justification. Layers analysis suggests that a second question ought then to be asked: Is a layer-respecting alternative available? One alternative is takedown by the host ISP. In fact, WorldCom notified the firms that hosted the offending content and, in most cases, this resulted in the content being removed by the host.[325] Of course, the firms providing hosting service for the content (which might be online service providers such as Yahoo! or ISPs in the case of peer-to-peer file sharing) may be located outside of Pennsylvania. That fact may increase the cost of enforcement. Assuming that Pennsylvania had personal or territorial jurisdiction, it could obtain an injunction, which then could be recognized and enforced by another state or nation. If jurisdiction is unavailable or judgments will not be recognized, a treaty or other international solution may be available—although again at increased cost. Layers analysis suggests that the costs of extraterritorial enforcement alone do not justify preference for the layer-violating regulation. Unless those costs are so burdensome as to make layer-respecting enforcement measures infeasible, the layers principle suggests that Pennsylvania's approach is not justified.

4. Transport Layer Regulations Aimed at Content Layer Problems

Yet another category of layer-violating regulation includes regulations targeting the transport or TCP layer. The core idea of such regulations is to block access to content (i.e., target a content layer problem) by utilizing a transport layer regulation (e.g., block IP addresses that are associated with the targeted content). We analyze a mostly hypothetical example of this kind of regulation—the possibility that copyright owners would seek to compel ISPs to block the ports used by peer-to-peer file sharing programs.[326] This category of layer-crossing regulation is illustrated by Figure 9 below.

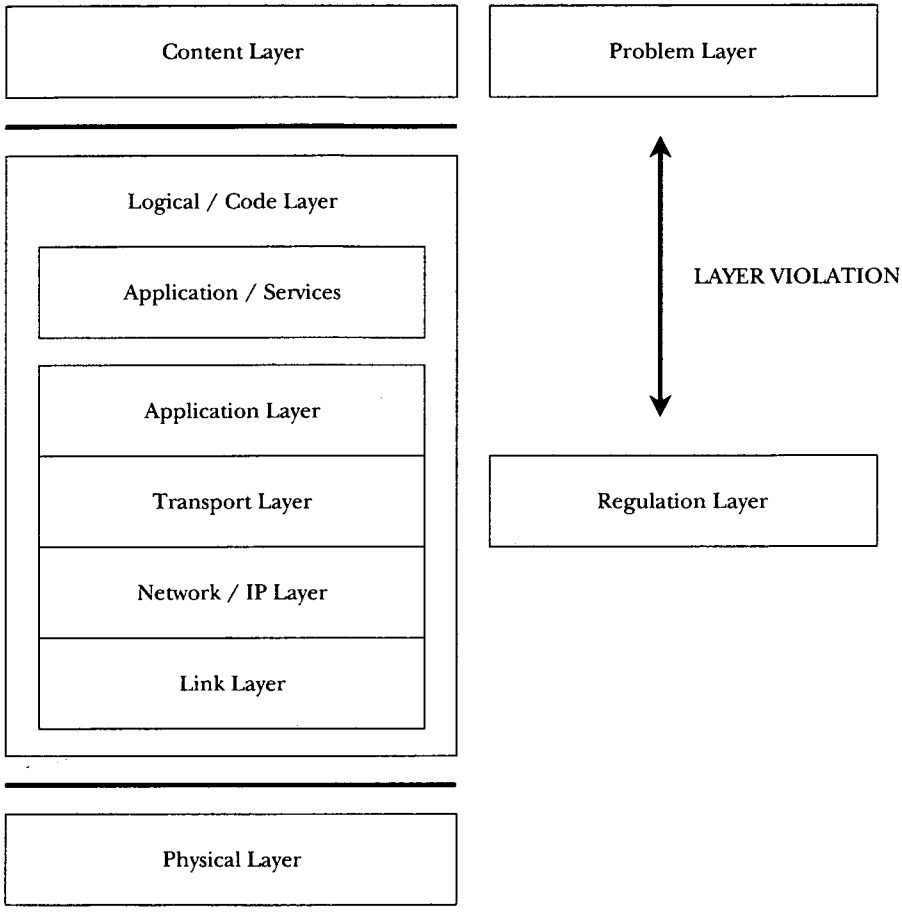
---

324 *Id.* at 679–80.

325 *Id.* at 675 (citing Letter from Craig Silliman, *supra* note 321).

326 *See infra* Part III.A.4.a.

FIGURE 9. REGULATION DIRECTED AT THE TRANSPORT LAYER DUE TO PROBLEMS AT THE CONTENT LAYER



a. Port Blocking and Peer-to-Peer File Sharing

This type of regulation could come up in the context of current P2P litigation.[327] The Digital Millennium Copyright Act (DMCA) requires ISPs to comply with the takedown provisions of the Act in order to avoid liability.[328] Due to DMCA concerns, many universities and private organizations have implemented the blocking of Gnutella or other peer-to-peer protocol traffic by blocking the TCP port used by the protocols. A public Internet regulator, i.e. a court, could order ISPs to block the TCP ports used by the P2P applications. No such

327 *See* Metro-Goldwyn-Mayer Studios, Inc. v. Grokster, Ltd., 259 F. Supp. 2d 1029 (C.D. Cal. 2003).

328 *See* 17 U.S.C. § 1201(a) (2000).

order has been issued so far, perhaps because the lawyers and judges are not aware of such possibilities yet.

Technically, a TCP port number is just a parameter used in the TCP header of the data packet. But every Internet application or application protocol must have a unique port number assigned to it in order to properly achieve its communication functions over the Internet. Using the port numbers encoded in the data packets, the TCP layer puts together the multitude of data packets received from the IP layer into coherent data streams intelligible to the applications.[329] For example, the port number for the web application protocol, HTTP, is 80. Among the myriad packets it receives, the TCP layer puts together the packets with the same port number 80 to reconstruct the communication data stream for HTTP, i.e., the web application.

The TCP ports are classified into three categories: well known ports, registered ports, and dynamic/private ports.[330] The well known port numbers are from 0 to 1023, and are assigned and managed by the Internet Assigned Numbers Authority (IANA), which is now a part of ICANN.[331] The registered ports range from 1024 to 49,151.[332] According to IANA, anyone can use or register a port number in this range, but IANA registers and lists the uses of these ports as a convenience to the community.[333] The dynamic/private ports are those from 49,152 through 65,535.[334]

All of the P2P applications use TCP port numbers above 1023—i.e., either registered or private ports. For example, Gnutella uses port 6346, and the FastTrack technology based applications, Grokster, KaZaA, and Morpheus, all use the port 1214.[335] However, many of the P2P applications can change port numbers on the fly in order to circumvent port blocking by universities or private organizations. In response, most organizations with firewalls block all ports higher than 1023, except for those specifically allowed by the organization.

---

329 1 STEVENS, *supra* note 18, at 226.

330 Internet Assigned Numbers Auth., *Port Numbers*, *at* http://www.iana.org/assignments/port-numbers (last modified Oct. 23, 2003).

331 *Id.*

332 *Id.*

333 *Id.*

334 *Id.*

335 *See Gnutella & Firewalls*, GNUTELLA NEWS, *at* http://www.gnutellanews.com/information/firewalls.html (last visited Feb. 13, 2004); Grokster Ltd., *Help and Frequently Asked Questions*, *at* http://www.grokster.com/helpfaq.html#FAQ7 (last visited Feb. 13, 2004).

Despite the understandable security and legal concerns on the part of the universities or private organizations, the wholesale blocking of the entire range of TCP ports raises the issue of its impact on overall Internet transparency. Although localized port blocking by private parties or schools would not impair the overall transparency significantly by itself, the cumulative effect could amount to a serious overall impairment if the practice becomes prevalent around the world. Furthermore, the TCP port blocking, as currently practiced, has serious substantial innocent use problems, because most ports above 1023 have innocent, legitimate, or noncontroversial uses.

b. Port Blocking and the Layers Principle

There was an important reason why the original design of the TCP/IP provided as many as 65,536 ports—far more than what was needed in the foreseeable future. They were laid down as a part of the initial Internet design to give users and developers the freedom to innovate without artificial scarcity of port numbers. For this reason, the decision to disable some ninety-eight percent of the ports without considering the impact on the functionality of the Internet is, at the very least, open to question.

The issues of TCP port use and assignment are not likely to go away. Eventually, the stakeholders may seek legal regulation directed at the TCP layer, either through the courts or the legislature. In addition to a port blocking approach, they may seek to deauthorize or deregister the ports through ICANN/IANA. At the moment, regulation over the "ownership" of the TCP port numbers is quite minimal. For the assigned and registered ports, one may apply for the port through IANA's website at www.iana.org. The dynamic/private ports are completely "free" or unregulated. And, there is no dispute resolution process for port assignment, in contrast to the ICANN process available for domain name assignments. Traditionally, TCP port assignment has not been a source of dispute because most of the popular Internet applications are client-server applications using well known assigned TCP ports. However, because of the explosive popularity of the highly controversial peer-to-peer applications using "registered" or private ports, the TCP port assignment or registration could become a new arena for legal challenges.

5. IP Layer Regulations Aimed at Transport or Application Layer Problems

Not all problem behaviors of layer-violating regulations arise at the content layer. Although most disputes arise at the content layer, it

is possible for conflict to rise at the level of one of the TCP/IP networking layers in a protocol war of a technological nature. An example of this type of dispute has come up in the hotly debated area of regulation in the broadband market—specifically, regulation of streaming video over high-speed cable Internet service.

a. Streaming Video over the Internet: A Brief Introduction

There are at least three different methods of delivering video or audio materials over the Internet: "downloading," "progressive downloading," and "continuous play."[336] First is downloading, where the entire file for the video program is first downloaded to the user's computer and then played after the transfer is completed.[337] An obvious disadvantage of this method is that the user must wait until the entire material has been downloaded before the user can start playing the material, even if the user is interested only in the beginning portion of the material. A method called progressive downloading addresses this shortcoming.[338] With this method, the playing of the material starts after a portion of the file has been downloaded, but the transfer of the yet-to-be-played portions of the file continues in the background so that the user has the perception of continuous play. Or the user might terminate the data transfer after a partial play if the material is found to be uninteresting.

The delivery methods based on downloading, however, present a very serious problem for the content providers. That is, a perfect copy of the program remains on the user's computer as a digital file, which in turn can be copied perfectly an unlimited number of times without authorization from the content owners. One way to deal with this problem is Digital Rights Management technology, where the number of times material can be played or copied is controlled by the content owners.[339] Another approach is to not leave a complete digital file on the user's computer in the first place. Under this approach, the video or audio program is played "live" directly into the user's computer over the Internet without leaving a permanent file on the user's com-

336 *See* KEXP 90.3 FM, *Streaming Media Frequently Asked Questions*, *at* http://www.kexp.org/listen/faq.htm (last visited Feb. 13, 2004).

337 *Id.*

338 *Id.*

339 Robert McGarvey, *Digital Rights Management: 10 Emerging Technologies That Will Change the World*, TECH. REV., Jan.–Feb. 2001, at 102, 103; *see also* Cross-Industry Working Team, *Managing Access To Digital Information*, *at* http://www.xiwt.org/documents/ManagAccess.html (last modified July 12, 1999) (addressing various issues surrounding the management of rights and permissions in the digital environment).

puter.[340] Most webcasts and Internet radio or television programs use this method.

Various methods of delivering video or audio programs "live," or in "real time" over the Internet are called streaming technologies.[341] Such technologies are needed because delivering video or audio program is a time sensitive or time critical task. To be effective, a video or audio program must be played at a correct speed or rate. Otherwise, Pavarotti's aria might sound like a munchkin's chant, or a fast-paced action movie might look like a slow-motion replay. Thus, the data packets that make up the movie or song must be delivered continuously in correct sequence at a correct rate—a process called "streaming."[342] Interactive "live" communication over the Internet, such as Internet phone or video conferencing, must also employ streaming technologies regardless of copyright concerns due to the "live" or "real time" nature of the communication.

b. Broadband Internet Service over Cable: A Very Brief Overview

Until recently, streaming video over the Internet has not been a source of widespread controversy, because the telephone modems most people utilized to get on the Internet were not fast enough to provide the data transfer rate required for live playback of any decent quality video material of any nontrivial length. The dispute arose as a result of the convergence of two separate lines of development occurring during the late 1990s: rapid growth of the broadband Internet industry, and vigorous expansion of traditional media companies into the Internet service market.[343]

In the context of Internet access speed, the term broadband refers to high-speed network data transmission capabilities ranging from more than a million bits per second (megabits per second) to billions of bits per second (gigabits per second).[344] In contrast, typical "56K" modems over traditional telephone connection can transmit at about

340 *See* KEXP 90.3 FM, *supra* note 336.

341 Jian Lu, *Signal Processing for Internet Video Streaming: A Review*, PROC. SPIE IMAGE & VIDEO COMM. & PROCESSING, Jan. 2000, at 1, *available at* http://streamingmedialand.com/sp4streaming2.pdf.

342 *Id.*

343 François Bar et al., *The Open Access Principle: Cable Access as a Case Study for the Next Generation Internet*, *in* THE ECONOMICS OF QUALITY SERVICE IN NETWORKED MARKETS (W. McKnight & John Wroclawski eds., forthcoming 2004), *available at* http://www.stanford.edu/~fbar/Drafts/OpenAccess-MITPress.pdf.

344 TOM SHELDON, ENCYCLOPEDIA OF NETWORKING 112 (1998).

45,000 bits per second.[345] Although high-speed transmission technologies have been in use for more than a decade by large organizations, ISPs, and Internet backbone operators, access to high-speed connections has become available to individual consumers or home users relatively recently. Broadband technologies currently available to home users are: digital subscriber line (DSL) service over telephone lines, digital communication over the cable TV wires, satellite communication, terrestrial or fixed wireless systems, and fiber optic lines (called fiber to the home—FTTH).[346]

Among these technologies, DSL and cable services are the most widely adopted broadband solutions in this country, in part because they utilize existing physical links already connected to most homes—i.e., phone lines and cable TV wires. Of the two, cable installations have about a two to one lead over the DSL installation base. According to a survey conducted by the Federal Communications Commission (FCC) of 4.3 million high-speed lines in service as of June 30, 2000, about one million were DSL service lines and about 2.2 million were cable connections.[347]

And, cable's lead over DSL is likely to continue for a few years.[348] A major constraint that limits the spread of DSL technology is that the user's home must be within about 12,000 to 18,000 feet (about three miles) from a telephone switching station equipped to handle DSL traffic.[349] Although certain equipment updates to the cable TV networks are necessary in order to support high-speed digital communication over the cable wires, the resultant effect of constraints are not as severe. At the end of 2000, only twenty-three percent of U.S. households were within DSL serviceable areas, while fifty-two percent could access high-speed digital cable services.[350]

Thus, cable technology is likely to be the most dominant method of high-speed Internet access for some time to come. However, the cable companies are also media companies whose traditional business

345 *See* Laurent Belsie, *Hurry up and Wait to Buy Newer, Faster 56K Modem*, CHRISTIAN SCI. MONITOR, Mar. 3, 1998, at 13, 13.

346 LENNARD G. KRUGER & ANGELE A. GILROY, CONG. RES. SERV., ISSUE BRIEF FOR CONGRESS: BROADBAND INTERNET ACCESS: BACKGROUND AND ISSUES 2–4 (2002), *available at* http://cnie.org/NLE/CRSreports/Science/st-49.cfm.

347 FED. COMMUNICATIONS COMM'N, HIGH-SPEED SERVICES FOR INTERNET ACCESS: SUBSCRIBERSHIP AS OF JUNE 30, 2000 (2000), *available at* http://www.fcc.gov/Bureaus/Common_Carrier/Reports/FCC-State_Link/IAD/hspd1000.pdf.

348 Tiffany Kary, *Cable Will Rule Broadband, Report Says*, CNET NEWS.COM, May 7, 2002, *at* http://news.com.com/2100-1033-901501.html?tag=prntfr.

349 This is a limitation on ADSL, a slower form of DSL technology. For faster DSL services, the distance needs to be even shorter. *See* SHELDON, *supra* note 344, at 312.

350 Bar et al., *supra* note 343, at 7.

is selling video programs. As such, broadband Internet over cable represents media companies' expansion into the rapidly growing (and potentially highly lucrative) Internet service market. A situation is created where the same company sells products at the content layer as a media company, owns the cable wires at the physical layer as a cable company, and has the ability to impose controls at the code layers as an ISP. Such vertical integration of functions across the layers may raise anticompetitive or antitrust concerns, especially when considering the cable companies' local regional monopolies in the high-speed Internet service market—perhaps the most important segment of the market because that is where the future lies.[351]

Most of these issues, however, are beyond the scope of this Article. Nonetheless, as the issues are of substantial importance for Internet regulation, we give a brief summary of the arguments made by Mark Lemley and Lawrence Lessig.[352] First, Lemley and Lessig observe that the cable companies' practice of bundling ISP service with access and prohibiting users from choosing another ISP removes ISP competition within the residential broadband cable market. This is a marked shift from the situation in highly competitive narrowband ISP markets. The authors argue that it is important that the ISP market remain competitive. The nature of ISP service is not inherently fixed, and they can and do provide a wide range of diverse services, including audio and video content. In short, the independent ISPs are engines for innovation in markets we do not yet imagine. Thus, vertical integration of ISP and access services by the cable companies threatens the future of innovation on the Internet by locking out independent ISPs from the broadband residential consumers. More generally, it is important to allow effective vertical competition. The Internet market generally has been characterized by massive shifts in the competitive center. Hardware companies (e.g., IBM) have been displaced by operating system companies (e.g., Microsoft); operating system companies have been threatened by browser corporations (e.g., Netscape) and by open platform "meta-operating" systems (e.g., Sun's Java). "Far and away the most important [factor in market structure] is that competition came . . . from another . . . layer."[353] Cable companies' vertical integration of content, ISP, and access market threatens

351 Mark A. Lemley & Lawrence Lessig, *The End of End-to-End: Preserving the Architecture of the Internet in the Broadband Era*, 48 UCLA L. REV. 925, 927 (2001).

352 *See id.*

353 *See* Timothy F. Bresnahan, *New Modes of Competition: Implications for the Future Structure of the Computer Industry*, *in* COMPETITION, INNOVATION AND THE MICROSOFT MONOPOLY: ANTITRUST IN THE DIGITAL MARKETPLACE 155, 161 (Jeffrey A. Eisenach & Thomas M. Lenard eds., 1999).

to stifle the innovative future of the Internet by eliminating this strategic competitive factor in the critically important residential broadband market.

c. Streaming Video over Cable Internet Service: The Controversy

An example of the possible effects of vertical integration is the controversial issue of cable companies' control of streaming video over their cable lines. Because cable companies' traditional business is the sale of video programs, there may be no incentive for them to allow streaming video content delivery from their potential competitors over their high-speed cable Internet services. In fact, when the cable companies started offering high-speed Internet services over cable, most of them did not allow full length streaming video on their cable lines.[354] For example, @Home, an ISP providing service over the AT&T cable lines, prohibited its users from downloading more than ten minutes of streaming video over the Internet.[355] The CEO of AT&T Broadband and Internet Services made clear that he "will not allow others to freely transmit movies and TV shows via his company's high-speed Internet connections."[356] He was reported to say: "AT&T didn't spend $56 billion to get into the cable business to have the blood sucked out of our vein."[357]

Such an aggressive posture and inflammatory rhetoric might have been influenced by the "Internet gold rush" high-stakes business environment of the late 1990s. The cable industry's attitude seems to have toned down quite a bit since then. To begin with, the limit on video delivery generated so much controversy that it drew the concerns of the chairman of the FCC, as well as some members of Congress.[358] Then, the limitation came up as a prominent issue during the closely watched and highly publicized FCC regulatory application process in connection with the AOL Time Warner merger. The Memorandum of Understanding (MOU) filed with the FCC by AOL Time Warner contained eleven points, one of which specifically addressed the video streaming issue: "AOL Time Warner will allow ISPs to provide video

354 Jerome H. Saltzer, *"Open Access" is Just the Tip of the Iceberg* (1999), *at* http://web.mit.edu/Saltzer/www/publications/openaccess.html; *see also* Lemley & Lessig, *supra* note 351, at 934–95.

355 *Excite@Home Keeps a 'Video Collar,'* MSNBC, Oct. 31, 1999, *at* http://zdnet.com.com/2102-11-501520.html.

356 David Lieberman, *Media Giants Net Change, Major Companies Establish Strong Foothold Online*, USA TODAY, Dec. 14, 1999, at B3.

357 *Id.*.

358 *See Excite@Home Keeps a 'Video Collar,' supra* note 355.

streaming. AOL Time Warner recognizes that some consumers desire video streaming, and AOL Time Warner will not block or limit it."[359]

The FCC's written question to the MOU raised the issue again, to which AOL Time Warner replied, reaffirming its commitment to allowing streaming video without restriction.[360] In a separate exchange, AOL Time Warner also had to answer the FCC's questions regarding its relationship with AT&T, which had an interest in Time Warner Entertainment through the MediaOne cable company.[361]

Since the AOL Time Warner merger, none of the major cable owner ISPs seems to have placed an explicit ban on streaming video traffic in its user agreement or policy documents. And @Home is no longer in existence after a highly publicized bankruptcy at the end of 2001—one of a series of the "dot com crashes" in that year.[362] Nevertheless, some cable owners, including AT&T, are still being accused of requiring independent subscriber ISPs not to sell streaming video on their own.[363] For example, Comcast requires their users to not engage in activities that would result in performance degradation of their networks.[364] At least in principle, receiving high-quality streaming video for an extended period of time can be viewed by the cable companies as prohibited activity that leads to performance degradation. Thus, although a specific ban on streaming video for individual users is no longer explicitly imposed, the cable companies' regulation of streaming video over the Internet may still be alive and well.

d. Streaming Video and Layer-Violating Regulation

The Internet service providers' concern about overuse or abuse of shared network capacity (the "bandwidth") cannot be limited or made specific to video traffic, as any type of data can cause network

359 *See* Memorandum of Understanding, between Time Warner Inc. & American Online, Inc. (Feb. 29, 2000), *available at* http://www.fcc.gov/mb/aoltw/mou.doc.

360 *See* Responses to Written FCC Questions, at 11–12 (Aug. 25, 2000), *available at* http://www.fcc.gov/mb/aoltw/techresp.doc.

361 Letter from Peter D. Ross, Counsel for America Online, Inc. & Arthur H. Harding, Counsel for Time Warner Inc., to Deborah Lathen, Chief, Cable Services Bureau, Fed. Communications Comm'n (Oct. 25, 2000), *available at* http://www.fcc.gov/mb/aoltw/attexp.pdf.

362 Consumer Affairs, *High-Speed Internet Just A Memory For Many*, CONSUMER NEWS, Dec. 3, 2001, *at* http://www.consumeraffairs.com/news/excite.html.

363 Consumer Federation of America et al., *Statement on the AT&T-Comcast Merger Submitted to the Subcommittee on Antitrust, Business Rights And Competition, Senate Judiciary Committee* (Apr. 23, 2002), *available at* http://www.consumerfed.org/CFA_et_al_ATT-Comcast_testimony.pdf.

364 *See, e.g.*, Comcast, *Terms of Service, Acceptable Use Policy* § viii (Nov. 11, 2003), *available at* http://http://www.comcast.net/terms/use.jsp.

congestion. However, the technologies developed to control network congestion can be employed to single out streaming video and treat it differently in a way that is not reasonably related to network congestion control.

One such example is limiting video traffic *from* a specific IP address. In fact, technical literature for Cisco's broadband router product gives this as a specific example for its use:

> [T]he network operator may specify policies to be executed for traffic, which either conforms to or exceeds a specified rate limit. The following actions may be executed in either case: Transmit (switch the packet) . . . . Drop (*discard the packet*) . . . . Example packet classification policies include . . . *video traffic* from *a specified IP address* is classified as medium priority.[365]

Limiting video traffic may be a valid exercise of network congestion control when, for instance, nearly all "bandwidth hogging" traffic is streaming video. In such a case, however, the limitation would need to apply equally to all video traffic to be effective, not just to those from a specific source. A more plausible explanation for limiting video traffic from a specified IP address would be that the ISP is attempting to limit or block video content delivery from a specific source.

If cable companies were to take such measures, however, the nature of their objection would not be against content per se. Although the ultimate concern of the cable owners is that of business interest—i.e., unwelcome competition against their own video content business—their regulating behavior is most directly aimed at protocols, not content. That is, the nature of concerns by the cable companies is not about the content itself, but the *nature* of streaming video traffic.

The streaming video protocols guarantee delivery of image frames at a steady video rate so that the series of images delivered appears to be the conventional video playback. The same image frames delivered by the ordinary web protocol, HTTP, would be ineffective as video material due to unsteady or erratic rate of delivery. In fact, none of the cable companies ever had restrictions banning delivery of individual images frame by frame by HTTP or FTP protocols. What the cable companies objected to was the *streaming nature* of the images delivered—i.e., the streaming protocols themselves.[366]

---

365 Cisco Systems, Inc., *Committed Access Rate*, *at* http://www.cisco.com/warp/public/732/Tech/car (last visited Feb. 20, 2004) (emphasis added).

366 On the other hand, blocking streaming video protocols would operate as effectively as the blocking of video content over the Internet, because the content provid-

Streaming video can be implemented at the application protocol layer using the Real Time Streaming Protocol (RTSP),[367] or at the transport layer using the Real Time Transport Protocol (RTP).[368] And restricting data transmission from an IP address is regulation at the IP layer because it targets an IP layer function. Thus, limiting streaming video from a specified IP address is a regulation at the IP layer because of an objection to upper layer protocols—RTP at the transport layer, or RTSP at the application layer. The restriction will typically be implemented in a router, as described in Cisco's technical literature quoted above, where packets from the specified IP address are looked at, then dropped or delayed if the protocol field of the IP packet indicates RTP or RTSP protocol. Thus, an application layer protocol—RTSP—or a transport layer protocol—RTP—is discriminated by an IP layer function. It is a layer-violating regulation, but the immediate objectionable conduct is not at the content layer. Rather, what are directly discriminated are upper level protocols.

Of course, the ultimate goal of the cable companies may be to prevent delivery of video content over the Internet. And video rate is arguably an integral and necessary part of video content. Our point, however, is that the video content is denied *indirectly* by directly blocking the protocols at the application or transport layer. Thus, the regulation can be viewed as one involving three or four layers, as illustrated in Figure 10 below.

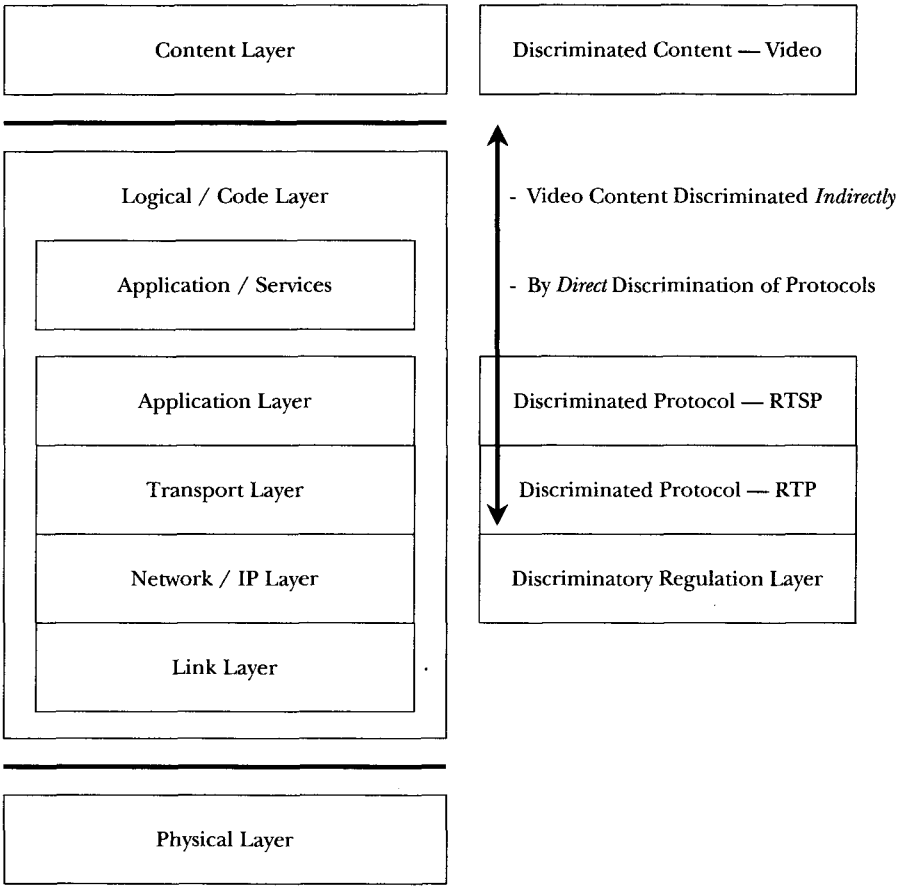
---

ers would not want the video files to be downloaded over FTP or HTTP due to unauthorized copying concerns, as discussed above.

367 Memorandum from H. Schulzrinne et al., Network Working Group, to the Internet Community (Apr. 1998), *at* http://www.ietf.org/rfc/rfc2326.txt.

368 Memorandum from H. Schulzrinne et al., Network Working Group, to the Internet Community (Jan. 1996), *at* http://www.ietf.org/rfc/rfc1889.txt.

FIGURE 10. LAYER-VIOLATING STRUCTURE OF STREAMING VIDEO REGULATION



e. Issues Raised by Streaming Video Regulation

Regulation of streaming video by cable companies provides an interesting example of layer-violating behavior *within* the TCP/IP layers. The example shows how complex regulating behaviors can become within the TCP/IP layers, although the ultimate goal may be control at the content layer. Indeed, this type of control may become more prevalent in the future, as the players become more familiar with the technologies. For this type of behavior, the end-to-end analysis approach focusing on the application ends is inherently incomplete. End-to-end analysis will completely miss the actual discriminatory functions operating inside the TCP/IP layers. And, the resulting effects at the content layer may not be conclusive of the discriminatory intent or effect. As discussed above, a cable company may

claim network congestion as the justification for blocking streaming video protocols. Such concern is certainly a compelling reason for an ISP or access provider, and cannot be dismissed without sufficient showing of facts to the contrary. Layers analysis, on the other hand, provides an analytic framework that is much more comprehensive as well as incisive by considering the entire spectrum of layers from the content to the physical layer, while placing the "core" TCP/IP layers at the center of the analysis.

Although a restriction on streaming video traffic from certain IP addresses is a layer-violating regulation within the TCP/IP layers, it nevertheless has problems similar to other layer-violating regulations. If these practices were adopted by many or all of the largest ISPs—which also happened to be cable media companies, such as AOL Time Warner or AT&T—the overall transparency of the Internet would be seriously impaired, as Internet transmissions using legitimately recognized protocols will not be able to reach a large number of users. Furthermore, the protocol blocking regulation would have a serious substantial innocent use problem. The discriminated streaming protocols have substantial innocent use because there is nothing inherently objectionable or illegitimate about the streaming protocols. In fact, the ISPs who are not in the video content business—e.g., noncable broadband providers such as DSL or T1 providers—do not prohibit streaming protocols. On the contrary, they encourage such usage in order to sell higher bandwidth services.

More specifically, there is nothing inherently objectionable about streaming protocols running over cable lines. Streaming traffic does not conflict with or impair any existing cable technology or equipment. If the cable company ISPs were not vertically integrated with the video content business, they would have no reason to object to streaming protocols. Thus, the cable lines or the IP routers employed by the cable company ISPs have substantial innocent use with respect to the streaming protocols at the upper layers.

As for ISPs' legitimate concerns over managing network bandwidth and controlling congestion—referred to as Quality of Service (QoS) issues in the network industry—there are non-layer-violating solutions such as Differentiated Services (DiffServ) or Resource Reservation Protocol (RSVP).[369] It should also be pointed out, how-

369 *See* Memorandum from S. Blake et al., Network Working Group, to the Internet Community (Sept. 1998), *at* http://www.ietf.org/rfc/rfc2475.txt; Memorandum from R. Braden et al., Network Working Group, to the Internet Community (Sept. 1999), *at* http://www.ietf.org/rfc/rfc2205.txt.

ever, that there is growing appreciation of the shortcomings of the traditional QoS approaches.[370]

6. The Case Against Layer-Violating Regulations Revisited

Having considered various examples of layer-violating behavior, we now revisit the abstract argument of Part II in light of our concrete discussion of particular examples. Layer-violating regulations on the Internet should be given very serious considerations before they are adopted because of (1) the transparency violations that undermine or chip away at the Internet's status as the innovation commons,[371] and (2) the substantial problems of fit inherent in such regulations.[372] What then is a proper approach to such regulations in light of our discussion of a variety of concrete problems (real and hypothetical) faced by public Internet regulators?

We have not proposed an absolute ban[373] on layer-violating or layer-crossing regulations because there may be cases where compromising the integrity of the layers is justified on the basis of a compelling regulatory interest.[374] The discussion of examples reveals a variety of reasons (some good, some bad)[375] that public Internet regulators might regard as compelling. Our discussion of cyberterrorism,[376] for example, suggests a compelling interest—national security. The on-balance justification will be especially strong if the localized loss of transparency—caused by the violation of the layers principle—would not affect the overall transparency of the Internet so as to threaten its function as innovation commons.

On the other hand, a permissive approach—allowing layer-violating regulations in general unless there is a good reason to prohibit them—may also not work, because we may totally miss the cumulative effect of many localized transparency violations by using an ad hoc,

370 Andy Oram, *A Nice Way to Get Network Quality of Service?*, O'REILLY NETWORK, Jun. 11, 2002, *at* http://www.oreillynet.com/pub/a/network/2002/06/11/platform.html.

371 *See supra* Part II.D.1.

372 *See supra* Part II.D.2.

373 *See supra* Part II.A.4.c.

374 *See supra* Part II.A.4.b.

375 The layers principle cannot guarantee that public Internet regulators will use good judgment or well grounded values in identifying compelling regulatory justifications. Thus, the regime in Myanmar may view its own survival as a compelling regulatory justification. Of necessity, the layers principle takes public Internet regulators as it finds them. Some are better; some are worse; some are terrible. Our attempt has been to articulate and shape the principle and its corollaries in a way that achieves the greatest practical good, given the world as it is.

376 *See supra* Part III.A.3.d.

case-by-case analysis.[377] By the time the cumulative effect is noticeable—the tyranny of small decisions—it might be too late to preserve the transparency of the Internet. The damage will have already been done, and the Internet would no longer be the innovation commons it once was. Our discussion of layer-compromising solutions to the threat to copyright interests posed by peer-to-peer filing is suggestive of a scenario whereby Internet regulators might slide down a slippery slope, ending with serious damage to the transparency of the Internet. Each individual act of IP address blocking[378] or port closing[379] would not be serious by itself, but the cumulative effect of many such actions would be an opaque Internet.

We have argued that the best approach is a principle raising a strong presumption against layer-violating regulations.[380] Our discussion of several concrete examples enables us to provide further guidelines for application of the layers principle at this point. We have seen that the presumption should be especially strong if one or more of the following factors are present: (1) when the layer-violating regulation affects or has potential to affect a large number of users—such as the regulations affecting the nation's largest ISPs or backbone operators, the regulations that affect the entire nation or nations, or the regulations that block most of the available TCP ports; or (2) when the layer-violating regulation is directed at a lower networking layer, such as the TCP, IP, or physical layer, due to problems at an upper end layer, such as the content or application layer.

When the layer-violating regulation affects a large number of users, the regulation's impact on overall Internet transparency would be especially destructive. Thus, policymakers need to be especially cautious when adopting or mandating such regulations. When the layer-violating regulation is directed at a lower end layer to counter the objectionable conduct at an upper end layer, the problem of substantial innocent use would tend to be especially serious—e.g., the "cutting the wire" regulation discussed above. Thus, a higher level of caution should be exercised when considering a regulation that targets the TCP, IP, or physical layers.

B. Application at the Communication System Levels

So far, we have analyzed the layer-violating regulations that target the networking layers—i.e., the TCP/IP layers or the physical layer in

377 *See supra* Part II.A.4.a.
378 *See supra* Part III.A.3.e.
379 *See supra* Part III.A.4.a.
380 *See supra* Part II.A.4.b.

the Internet networking context. Within the framework of communication system layers—the more generalized conception of layers articulated by Benkler—there is a class of regulations that have a structural similarity to the layer-violating regulations, but are different in such significant ways that they need to be analyzed separately. These regulations also target a lower layer in order to address problems that arose at an upper layer, but they do not directly target the networking layers. We shall call this type of regulation a *layer-crossing regulation in the communication system layers.*

1. Layer-Crossing Regulations in the Communication System Layers

A layer-crossing regulation may target, for example, a code layer in order to counter the problems that arose at the content layer. In fact, the anticircumvention provisions of the DMCA are an example of such a layer-crossing regulation. The DMCA prohibits manufacture or distribution of any technology, product, service, or device that circumvents copy protection technology.[381] Thus, the DMCA targets the circumventing programs (the codes) that are not a part of a network system—i.e., do not belong to the networking layers—due to concerns about issues at the content layer—i.e., copyright infringement.

Layer-crossing regulations in the communication system layers are structurally similar to the layer-violating regulations in the Internet layers in that they both target a lower layer to address problems from an upper layer. Typically, both types are motivated by similar desires—i.e., cut the problem behavior down at its knees by targeting the enabling technologies.

However, they are different in significant ways as a consequence of a critical distinction between the TCP/IP layers and the communication system layers. The TCP/IP layers represent the implementation of the design principles of the Internet—layer separation, transparency, and end-to-end. The communication system layers, on the other hand, are descriptive categories to conveniently organize the existing concepts in communication systems. Thus, there is no general transparency requirement or expectation across the communication system layers. The end-to-end argument in its general form, however, is still applicable in the communication system layers as a normative principle that argues against placing a function away from the level where it is ultimately consumed or used.

381 17 U.S.C. § 1201(a)(1)-(2), (b)(1) (2000).

2. A Layers Approach to the Substantial Noninfringing Use Doctrine

In the seminal 1984 *Sony* case, the owners of copyrights on television programs attempted to counter alleged infringing activities by targeting VCR technology.[382] Rather than going after the consumers who were actually taping their programs—the direct infringers—the television industry chose to attack VCR technology itself by suing the manufacturers of the video recording machines for contributory infringement.[383] The Supreme Court, however, refused to hold the manufacturers generally liable because the VCR technology was "capable of substantial non-infringing uses."[384]

Within the framework of communication system layers, the regulation that the television industry sought to implement in *Sony* can be viewed as a layer-crossing regulation that targets the physical layer—the VCR machines—in order to counter the objectionable conduct at the content layer—the unauthorized copying of the programs on TV. It is significant, then, to note that the Supreme Court ruled against a layer-crossing regulation when it found substantial innocent use at the lower layer with respect to the problems at the upper layer. When faced with a structurally similar problem, the law came up with a solution—the *Sony* substantial noninfringing use doctrine—that is consistent with arguments against layer-violating regulations presented in this Article based on the fit thesis.

Despite the analogy, there are fundamental differences between the communication system layers and TCP/IP layers. First, there is no transparency requirement across the communication system layers, whereas layer transparency is a fundamental design principle inherent in the TCP/IP layers architecture. Second, not all lower layer devices or technologies in a communication system layer have substantial innocent use. It is conceivable that, for certain devices or technologies, all or nearly all uses of the device are unlawful. In such cases, there is identity or near identity of the problem conduct and the technology, and there is no substantial innocent use of the technology or device.

In contrast, within the TCP/IP layers, substantial innocent use of lower layers with respect to problems at an upper layer is a necessary consequence of the architectural properties inherent in the TCP/IP layers architecture. Thus, the argument against the layer-violating regulations in the Internet layers is completely consistent with the solution provided by the Supreme Court in *Sony*. In fact, it is an a fortiori argument, because, in addition to the necessary substantial

382 Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417, 420 (1984).
383 *Id.* at 420.
384 *Id.* at 442, 456.

innocent use problem, the layer-violating regulations threaten to destroy or impair the Internet's tremendous value generating function by destroying or impairing its transparency.

It should also be pointed out that the *Sony* doctrine can be thought of as a kind of end-to-end principle applied to the field of legal regulation. Within the communication system layers, the *Sony* doctrine can be interpreted as suggesting a proposition that says: unless there is near identity of the problem and the technology, the law should not target the lower technology layer in order to regulate the problems at the upper layer. Where possible, a regulation should be directed at, or match, the layer where the problematic conduct arose.

3. The Implications of the Substantial Noninfringing Use Doctrine for the Layers Principle as a Legal Doctrine

This leads us to a crucial point about the application of layers analysis to layer-crossing regulations in the communication system layers. The Supreme Court's decision in *Sony* is based, at rock bottom, on a rationale that is structurally homologous to layers analysis. Although the *Sony* court did not discuss layers—it would have been a miracle of prescience if it had—the reasoning in *Sony* is fundamentally based on what we have called the fit thesis, the inherent overinclusiveness of regulating a communications technology with substantial innocent uses in order to regulate content. As we have demonstrated, the case for the layers principle is even stronger in the context of the TCP/IP layers than it is for the communications systems layers. The problems of fit created by layer-crossing regulations at the TCP/IP layer are even more profound than the problems of fit that undergird the *Sony* rule. To these problems of fit are added the even more compelling problems identified by the transparency thesis. In other words, the underlying rationale of *Sony* applies more strongly to layer-violating regulations in the TCP/IP layers than it does on the facts of *Sony* itself.

The implication of these facts for legal doctrine is profound. The layers principle rests on a foundation of both facts and norms. *Sony* shows that the layers principle is already embedded in the woof and warp of the law, at least in the United States.[385] The normative basis

385 We recognize that our discussion of the legal status of the layers is incomplete, because it is not comparative. A more thorough analysis would consider the question whether the *Sony* doctrine has siblings or cousins in other legal systems. Even this U.S.-centric analysis of the legal implications of the layers principle is long and complex, and the costs to the authors of a comprehensive extension of our analysis and research of other jurisdictions would be substantial.

for the layers principle is already anchored in the deep structure of American law. The factual basis for the layers principle rests on the well established findings of network engineering. In other words, the layers principle is, in a jurisprudentially significant sense, already there—in the constitution and case law—waiting to be discovered.

CONCLUSION

The layers model is the key to understanding the fundamental architectural principles of the Internet. These principles—layer separation, transparency, and end-to-end—are the cornerstones of low cost innovation. Without this transparency-preserving architecture, there would not have been a World Wide Web, peer-to-peer, or e-mail. For a variety of reasons (some good, some bad), public Internet regulators have begun to act in ways that compromise the Internet's basic architecture—violating layer separation and targeting regulations at one layer to solve a problem that arises at a different layer. Public Internet regulators are subject to strong temptations to act in ways that could compromise the integrity of the layers.

As a remedy, we propose that public Internet regulators should adopt the layers principle and its corollaries as a framework for the evaluation of any proposed Internet regulation. In general, a regulation should be directed at, or match, the layer where the problematic conduct arises. Absent compelling justification, the law should not target a lower layer in order to regulate the problems that occur at an upper layer. There should be a strong presumption against regulations targeting one or more of the lower networking layers (the TCP, IP, or physical layer). The presumption should be especially strong when the layer-violating regulation affects or has the potential to affect a large number of users, or when the target layer and the problem layer are far apart in the layers hierarchy. If a layer-violating regulation is justified, the regulation should be implemented at the layer closest to the problem layer. Care should be taken to implement the regulation narrowly so as not to create overbroad regulations, thereby creating or destroying important rights inadvertently and without due consideration. In every case, public Internet regulators should consider the availability of layer-respecting alternatives *before* violating the layers principle.

In the real world, public officials tend to focus on problems and solutions. Of necessity, public Internet regulators, especially in legislatures and administrative agencies, are concerned with incremental costs and benefits. It goes without saying that legislators and regulators are subject to political pressures. Moreover, most public Internet

regulators (courts, regulatory agencies, executive departments, and legislative bodies) are ill informed about network engineering. To them, the architecture of the Internet is little more than a mysterious black box. But, for Internet regulation to succeed, public Internet regulators must have a basis for understanding how proposed regulatory action will interact with the architecture of the Internet. For Internet regulators to avoid unintended consequences, they must gain an appreciation of the layers and the impact of layer-violating regulations on the transparency (and hence the functional value) of the Internet. The layers principle and its corollaries distill the complexities of network engineering into guidelines for regulation. They transform information about Internet architecture into norms for Internet regulation.

The future of innovation and creativity on the Internet hangs in the balance. It would be especially tragic if the most vibrant blossomings of innovation and creativity in modern times were extinguished unintentionally because we did not understand the nature of this accidental gift of history. This potential tragedy can be avoided if public Internet regulators follow a simple injunction: *respect the integrity of the layers.*