Layers Revisited

J. Scott Marcus, Wissenschaftliches Institut für Kommunikationsdienste GmbH

Douglas C. Sicker, University of Colorado at Boulder

Abstract

As network technology continues to evolve, the traditional regulatory models that served us well in the United States in the past have largely broken down. New services such as VoIP reveal glaring challenges and asymmetries in U.S. law and regulation. This breakdown reflects technological change, changes in the structure of the industry and the market, and regulatory stagnation. How can we hope to fix these problems going forward?

We propose a modest, simplified interpretation of the Layered Model of telecommunications policy: Transmission, Applications, and Content. Within each layer, it is still necessary to articulate principles that will guide regulatory decisions, as the Layered Model does not in and of itself do so. We do so by borrowing heavily from concepts articulated in the European Framework for the regulation of electronic communications.

1. Introduction

The current breakdown in telecommunications law and regulation in the United States is widely recognized. Many or the leading regulatory and policy theorists have focused their efforts on a thought model known as the Layered Model, arguing that public policy should in general follow the same concept of layering as the Internet.[1]

Unfortunately, this thread of research is largely embroiled in unproductive controversy at present. One of the most widely read papers on the Layered Model is usually interpreted as advocating economic regulation of last mile facilities in order to address the market power of incumbents.[2] Predictably, this led to a firestorm of lobbying and of funded research both by those who possess last mile facilities, and by those who do not possess facilities but depend on access to them. The resulting debate has, in our view, shed far more heat than light.

Rather than wasting the reader's time with these now rather shopworn arguments, we choose instead in this paper to go back to first principles, and to try to understand what the Layered Model *can* achieve for the regulator, and what it *can not*.

The core of the Layered Model is that it can be useful to decompose the regulatory environment, and that regulatory problems need to be addressed in the milieu in which they are found. These modest, basic propositions seem to us to scarcely be arguable, provided that it is possible to reach an appropriate decomposition. It seems to us then that the real shortcomings of the Layered Model[3] are:

- That it fails to tell the regulator anything at all about how to regulate within each layer;
- That the decision models for how to regulate within a layer still need to be developed, since the Layered Model itself does not provide them; and
- That the most widely cited papers use a layer decomposition that seems to us to be inappropriate.

These concerns can be addressed. We see the application of the Layered Model, not as a panacea, but rather as a modest first step of regulatory analysis in an analytic process that necessarily involves at least one further step.

---

[1] The proponents include: Douglas C. Sicker and Joshua L. Mindel, *Refinements of a Layered Model for Telecommunications Policy*, 1 J. TELECOMM. & HIGH TECH. L. 69 (2002); Richard S. Whitt, *A Horizontal Leap Forward: Formulating a New Communications Public Policy Framework Based on the Network Layers Model*, 56 FED. COMM. L. J. 587 (2004); John T. Nakahata, *Regulating Information Platforms: The Challenge of Rewriting Communications Regulation from the Bottom Up*, 1 J. TELECOMM. & HIGH TECH. L. 95 (2002); Rob Frieden, *Adjusting the Horizontal and Vertical in Telecommunications Regulation: A Comparison of the Traditional and a New Layered Approach*, 55 FED. COMM. L.J. 207 (2003); Lawrence B. Solum & Minn Chung, *The Layers Principle: Internet Architecture and the Law*, (University of San Diego School of Law, Public Law and Legal Theory Research Paper No. 55) (2003), *available at* http://ssrn.com/abstract=416263; and Philip J. Weiser, *Law and Information Platforms*, 1 J. ON TELECOMM. & HIGH TECH. L. 1 (2002).

[2] Whitt, *A Horizontal Leap Forward*, op. cit. Whitt's treatment of last mile market power in the final paper as published is in fact more nuanced than most people realize. His concerns are conditioned on the presence of market power.

[3] This critique is presented at slightly greater length in the following sections. A more comprehensive treatment is to appear as Marcus, "Beyond Layers", in the *Journal on Telecommunications and High Technology Law* (JTHTL) in 2006.

As for the decomposition, in previous work the authors independently arrived at a simplified three layer model,[4] which can be summarized as *transmission*, *application*, *content*. Many of the widely cited papers on the Layered Model use a four layer model, in which they distinguish between physical and logical transmission, assigning each to its own layer.[5] In our view, this is an extremely bad idea – it carries forward one of the most problematic and unsustainable aspects of the current legal system, violates technological neutrality, prejudges ultimate technological and market outcomes, and positively invites regulatory arbitrage. But these defects are easily corrected by merging the two lower layers into one.

The distinction between transmission and application, on the other hand, is important and even fundamental. Historically, the network and the service were one. Today, the value chain has been broken and reconstructed. The service provider may not be the same as the network provider – in fact, the service provider may not be able to predict in advance which networks will be used to deliver the service.[6] Under these circumstances, obligations intended for the network operator cannot meaningfully be imposed on the service provider, nor can regulation relevant to the service provider meaningfully be imposed on the network operator, in general. Traditional regulation failed to make this distinction, but future regulation must.

Returning to the notion of a second step of analysis, embodying regulation within a Layer, what does this imply?

The regulatory issues at each level are similar, but they manifest themselves somewhat differently at each layer. In each layer, there are three core reasons to regulate: (1) dealing with such market power and barriers to competitive entry as may exist; (2) meeting any societal needs that might not otherwise be adequately addressed by market mechanisms; and (3) managing limited resources, such as spectrum or telephone numbers. In all three areas, regulators should rely on market forces as much as possible; however, there are many well documented instances in each of the three areas where regulation is appropriate and necessary. We see much value in borrowing from European practice here, as the underlying decision principles in each of these areas are articulated much more clearly in European law than in American at present.

The authors are of the opinion that, with these clarifications, the Layered Model can be a useful if modest first step in regulatory analysis.

Section 2 briefly summarizes possibly useful background, including the evolution of U.S. telecommunications law as it relates to these issues, and the TCP/IP reference model that underlies the regulatory Layered Model. The reader who is familiar with these topics can skip this discussion without loss of continuity. Section 3 provides a somewhat expanded explanation of why we prefer three layers to four. Section 4 seeks to develop decision rules within each layer. Section 5 provides a brief summary.

---

[4] See Marcus, "Beyond Layers", ibid., and an unpublished but widely distributed FCC OPP Working Paper by Douglas C. Sicker, Joshua Mindel and Cameron Cooper, *The Internet Connection Conundrum*, 1999.

[5] Cf. the Whitt and Werbach papers, op. cit.

[6] Consider, for example, a Voice over IP (VoIP) provider such as Vonage or Skype.

2. Background

This section provides a brief review of how the United States got into this mess, and of the TCP/IP reference model that underlies the Layered Model of regulatory policy. The reader who is familiar with these topics may skip this section without loss of continuity.

2.1 Unstructured communications law in a layered world

To understand these goals, one must first review the current regulatory structure and its origins. The current communications regulatory structure is often described as a "silo" model, with regulation of a service aligning with the type of underlying physical infrastructure on which a service is offered. (This is not to imply that the titles were written as a means of regulating the service or the specific infrastructure, but this has been the result. Indeed, Title II is not inherently tied to the copper/fiber infrastructure or even telephony, but rather to the concept of common carriage.) In the traditional regulatory model, there is a separate "silo" for each type of infrastructure or platform – wireline (twisted copper pair), cable (coaxial) or spectrum (wireless). For example, voice (telephony) service delivered over wireline (copper twisted pair), is regulated under Title II – Common Carrier, but voice service delivered over spectrum radio waves is regulated under Title III – Wireless. Thus the same service, voice, has come to be regulated differently according to the physical infrastructure over which it is delivered.

This Communications Act reflects the technological and market conditions that existed over the course of its evolution. In its beginning (and for many decades), telephony service was a monopoly industry provided largely by the Bell System. It was thought to be far too expensive, not profitable and not efficient for competing companies to redundantly lay tens of thousands of miles of wire and build the associated facilities to provide competing telephony service. Thus, it made sense to model telephony regulation after existing "common carrier" regulation (like train service or a utility). Radio developed, and the Radio Act of 1912 and the subsequent Radio Act of 1927 were enacted to regulate that industry. In 1934 all communications were brought under one regulatory construct by the 1934 Communications Act, with the Radio Act absorbed into the Communications Act largely intact under Title III. Later, broadcast television burst onto the scene, and was regulated under the Communications Act as it also operated over spectrum. Broadcast television outpaced regulatory efforts for some time. Later, cable television developed, but it was initially a small rural phenomenon that was not expected to overtake broadcast television. Congress enacted a number of laws over the years that were for the most part incorporated into Title VI of the Communications Act. Meanwhile, mobile telephony developed and was regulated under Title III, as it also utilized the spectrum; however, portions of Title II are also applicable. Many other bits and pieces were added to the Communications Act over the past twenty years in support of lawful intercept (CALEA), 911 emergency services, and spectrum auctions.

Thus, the Communications Act reflects an accretion over a period of more than seventy years. The risk in a process of accretion is that the end product loses coherence – and that certainly has been the case for the Communications Act. The

major operative components of the Act (Titles II, III and VI) interact in complex and difficult-to-predict ways, but there has never been a serious effort to harmonize the disparate pieces of the Act as a whole. The major revision in 1996 imperfectly codified a small portion of the Computer Inquiries (discussed in Section x), but for the most part otherwise increased the complexity and unwieldiness of the text rather than streamlining it. Moreover, the Communications Act interacts with other laws in complex and hard-to-predict ways.

The challenges to this system are generally characterized as *convergence*. Services that had once been uniquely associated with one kind of underlying network became available over another, thus jumbling the underlying legal and regulatory assumptions. Convergence is often thought of as an Internet phenomenon, but it has in fact been around for a long time. Regulatory disparities and asymmetries became an issue as cable television became a stiff competitor to broadcast television, and again as cellular (telephony) service began to replace wireline telephony. Thus, the same service, delivered over different infrastructure, fell under disparate sections of law and regulation. These disparities and asymmetries likely had a negative impact on innovation, competition and consumer benefits.

The advent of the Internet and the "digital revolution" accelerated these challenges. Services converged onto a single superstructure – the Internet Protocol (IP) environment – that could ride over all existing physical infrastructures. In this environment cable can deliver voice service and high-speed internet service, as well as television service; wireline telephony providers can deliver video service and high-speed internet service, in addition to voice service; wireless telephony providers can also deliver internet service, streaming video and other services in addition to voice; and voice service, audio or video broadcasts, streaming video, audio downloads and more services can be delivered over the Internet, provided by ISPs -unaffiliated or affiliated with a cable company (over cable modem), or a phone company (over dial-up or DSL), or by a competitive provider (over leased dial-up or DSL facilities of the incumbent phone company). As each new service evolves, there is a need to classify it in order to determine under which regulations it falls. However, regulating based on the underlying infrastructure no longer seems appropriate when each infrastructure is no longer tied to just one service, but when they can all deliver a multitude of competing services. Entities provide competing services to consumers who do not generally distinguish between the same service delivered over different infrastructure. However, this differing infrastructure causes these entities to operate under vastly disparate regulatory conditions.

The FCC tried to deal with regulation of nascent computer networks in Computer Inquiry I,[7] II[8] and III,[9] creating *basic* and *enhanced* service classifications. Basic services –"the common carrier offering of transmission capacity for the movement of information"[10] fell under common-carrier regulation and enhanced services remained

---

[7] Regulatory & Policy Problems Presented by the Interdependence of Computer and Communication Services & Facilities, *Notice of Inquiry*, 7 F.C.C.2d 11, ¶. 1, (1966) [hereinafter *Computer I*].

[8] Amendment of Section 64.702 of the Comm'n's Rules and Regulations. (Second Computer Inquiry), *Final Decision*, 77 F.C.C.2d 384, (1980) [hereinafter *Computer II*]

[9] Amendment of Section 64.702 of the Comm'n's Rules and Regulations. (Third Computer Inquiry), *Report & Order*, 104 F.C.C.2d 958 (1986) [hereinafter *Computer III*].

[10] *Computer II*, *supra* note 20 at ¶. 93.

unregulated.[11] Sicker has observed that this can be viewed an early example of a Layered Model approach, as it reflects treatment of the transport network that is distinct from the treatment of the services that ride on that network.[12] The Telecommunications Act of 1996 carried these regulatory concepts over into law by defining a *telecommunications service* as "the offering of telecommunications for a fee directly to the public, or to such classes of users as to be effectively available directly to the public, regardless of the facilities used"[13] and an *information service* as "the offering of a capability for generating, acquiring, storing, transforming, processing, retrieving, utilizing, or making available information via telecommunications."[14]

These minor regulatory overlays were not sufficient to close the gap between the inherited regulatory structure and the vastly changed communications environment of today. Most experts would agree that the Communications Act as it stands no longer fits existing conditions. Many have examined the Layered Model as a possible framework with which to examine policy issues going forward.

2.2 The TCP/IP Reference Model

A computer network can be conceived of as the interconnections of computers that allow communication. The content, scope, size, speed and reliability of the network vary depending on its protocols and implementation. *Protocols* are pre-established rules or means of communication. They are simply a set of valid messages, rules and formats that govern the communication among communicating peers.[15] Protocol layering is a common technique to simplify networking designs by dividing them into functional layers, and assigning protocols to perform each layer's task. Protocol layering produces a number of sub-functions, each with a few well-defined tasks. The concept of layering relies on breaking a complex task into smaller subsets, each of which addresses a specific issue. Each layer provides a well-defined set of services to the layers above it and depends services provided by lower layers for its own operation. This segregation into well defined functions creates modularity.[16]

The Internet protocols are arranged in essentially independent layers with the Internet Protocol (IP) itself at the middle or "waist" of the stack. The protocol stack broadens above the waist to support a wide range of transport and application layers including email, the Worldwide Web, and file transfer protocols. The protocol stack broadens below the waist to ride on a wide range of underlying networks using a variety of technologies including frame relay, ATM, ADSL, fiber optic systems, and so on.

---

[11] *Id*. at ¶¶. 100–01.

[12] *See* Douglas C. Sicker, *Further Defining a Layered Model for Telecommunications Policy*, (Paper Presented at the Telecommunications Policy Research Conference (TPRC) 2002), p. 5. *Available at* http://intel.si.umich.edu/tprc/papers/2002/95/LayeredTelecomPolicy.pdf.

[13] 47 U.S.C. 153, § 3 (46) (2005).

[14] 47 U.S.C. 153, § 3 (20) (2005).

[15] For a detailed explanation of protocol layering, *see* Srinivasan Keshav, An Enginerring Approach to Computer Networking: ATM Networks, the Internet, and the Telephone Network, ISBN 0201634422] Yes, use the book reference.

[16] *See generally*, *Protocol Layering*, *at* http://www.freesoft.org/CIE/Course/Section1/4.htm (visited 08/10/2003).

Modularity of the protocol layers promotes an environment wherein providers compete with products that will interoperate. The modularity, coupled with well-understood specifications, facilitates the introduction of new technologies and new applications, thereby stimulating growth.

Physical and Data Link Layers: The Physical Layer covers the network hardware, physical cabling or a wireless electromagnetic connection. It also deals with electrical specifications, collision control and other low-level functions. The Data Link Layer attempts to make the physical link reliable and provides the means to activate, maintain and deactivate the link. These layers collectively contain the protocols used to deliver data to the other computers and devices that are attached to the network. TCP/IP was designed to be independent of the network access or transport platform. In this way, TCP/IP can be used to connect differing network technologies such as Ethernet, ATM or Frame Relay. Independence from any specific network technology gives TCP/IP the ability to be adapted to new technologies.

Internet Layer: This layer is responsible for routing messages through networks.

Transport Layer: The protocol layer just above the Internet layer is the transport layer. It is responsible for the reliability and integrity of the end-to-end communications.

Application Layer: The application layer is the highest layer of the TCP/IP protocol stack. It provides services to human or automated users.[17]

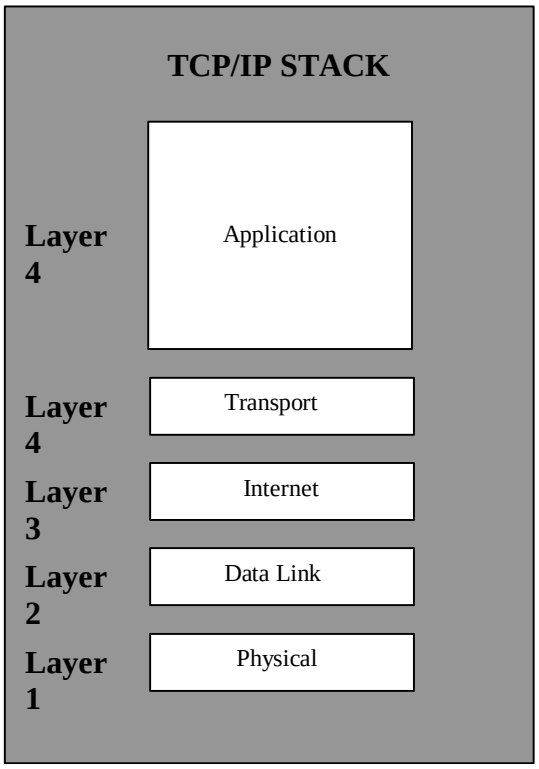

Figure 1. The TCP/IP reference model.

[17] *See Understanding TCP/IP (*Cisco Documentation), *at* http://www.cisco.com/univercd/cc/td/doc/product/iaabu/centri4/user/scf4ap1.pdf

3. How many layers?

We noted in the introduction that we respectfully take issue with the particular decomposition of regulatory functions into layers taken in two of the most widely cited papers on the Layered Model. The Whitt and Werbach[18] papers both base their analysis on a four layer model, where physical transmission is the lowest layer and logical transmission is the next-to-lowest. In our view, this decomposition violates the principle of technological neutrality, thereby carrying forward one of the most problematic and unsustainable aspects of the current legal system. It also prejudges ultimate technological and market outcomes, and positively invites regulatory arbitrage.

3.1 Technological Neutrality

One of the most useful principles to emerge from European regulatory practice is the notion of *technological neutrality*. In an era of technological and market convergence, services should be regulated as much as possible based on the economic characteristics as the service as perceived by the user, not on basis of the underlying technology. To do otherwise is to invite irrational results and regulatory arbitrage.[19]

To be sure, this principle must be applied with some sensitivity. Some regulatory obligations may have already been put in place because they were easy to implement under the technology of the day; the new service might perhaps more cost-effectively achieve the necessary functional capabilities in a different way. Thus, prematurely and reflexively imposing identical obligations on a new service may simply result in a failure to deploy the new service. But this is the exception that proves the rule: the principle of technological neutrality should be adhered to in the absence of a compelling need to do otherwise.

Paradoxically, the "classical" four layer Layered Model could be said to violate technological neutrality. It implies different treatment for the logical transmission compared to the physical transmission layer. Depending on how one draws the lines, this might imply, for example, different treatment for Voice over IP (which rides on top of the logical transmission layer) as compared to traditional PSTN telephony (which could be said to ride on top of the physical transmission layer).

The four layer model is a natural extension of current U.S. law – it is a natural way to enable regulation of traditional facilities, while prohibiting regulation of transmission over the Internet. Unfortunately, this means that it exemplifies the problems and the sloppy analysis that characterize current U.S. law. From the time of the Computer Inquiries, we have treated transmission over the Internet as being fundamentally different from transmission over the traditional network, even though they are functionally equivalent.

---

[18] Both works op. cit.

[19] This is generally consistent with principles advanced by the eminent economist Hayek, who advocated the use of market mechanisms to ensure survival of the fittest among competing solutions. See Nicholas Garnham's keynote address at EuroCPR, Barcelona, 2004.

There is no basis for this disparate treatment – neither in logic nor in the Communications Act. The Stevens Report sought to justify the regulatory status of data transmission over the Internet by a rather tortured argument[20] that Internet access was inextricably intertwined with the e-mail service offered by the ISP[21], and with the ever-popular network news application (nntp).

There was, to be sure, a reasonable public policy rationale for avoiding premature regulation of data transmission over the Internet. First, there was the desire to avoid stifling innovation by premature regulation. Second, there was the concern that proper regulation of Internet services was difficult, and that premature efforts to do so would be counterproductive.

All of that may have held in 1998, but it is not clear that it still holds in 2005. The systematic unregulation of the Internet, solely because it is the Internet, was a great idea whose time has come. And gone. As increasingly critical services migrate to the Internet, it is inevitable and proper that certain elements of PSTN-like regulation will follow.[22]

3.2 Prejudging technological outcomes

The division of the Transmission Layer into physical and logical layers implicitly presupposes that such a division either is or soon will be relevant to all services. If some legacy services were to remain indefinitely on the PSTN under traditional models, it is difficult to imagine how those traditional services would be regulated under a four layer model.

It is fashionable to assume that all services will rapidly migrate to the Internet. In reality, we do not know that this will happen. We do not know how quickly it will happen, if it happens. And we can not be assured that there might not be a rapid migration afterwards to something else.

Uncertainty is always a risk for the regulator. The problem in this instance is that the distinction is not inherent to the service being provided. Rather, the dividing line is an artifact, not of the service being delivered, but rather of the manner in which it is delivered. The latter is much more malleable, much more readily manipulated, than the former.

3.3 Regulatory arbitrage

Any system where regulatory obligations flow from the manner in which a service is delivered, rather than the more fundamental nature of the service itself, will inevitably encourage service providers to deliver the service in the manner that provides for the most favorable regulatory treatment.

---

[20] Cite to Stevens Report.

[21] This was prior to the existence of services such as hotmail.

[22] Consider, for example, the FCC's recent imposition of E-911 emergency services obligations and lawful intercept (CALEA) obligations on VoIP providers who interconnect to the conventional PSTN, and in the case of CALEA also on facilities-based broadband Internet access providers.

A regulatory system that favors Logical (i.e. Internet) Layer services over Physical Layer services leads ineluctably to market distortions. It necessarily violates the Hayekian ideal of survival of the fittest in the marketplace. The regulator has chosen the winner, not the marketplace.

3.4 A three layer model

For all of the reasons noted above, we think that a four layer model that distinguishes between physical transmission and logical transmission is problematic.

A three layer model, on the other hand, seems to be workable. The distinction between the application (service) and the transmission capabilities used to deliver it (network) is precisely what has changed. It is what motivates a change in regulatory model in the first place. And the distinction between the content is already well established in law and regulation.

The three layer model is simply:

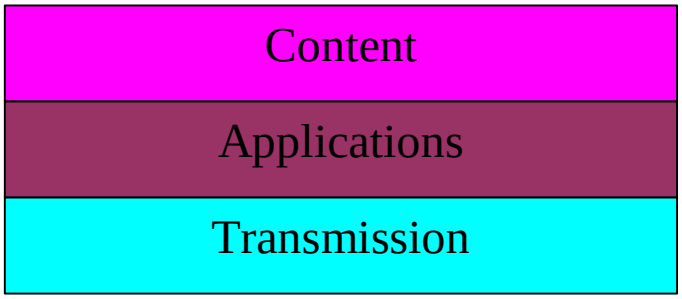

Figure 3. A simplified three layer model of regulation.

In a sense, this is not a new version of the Layered Model. It was there from the beginning, as one of several legitimate variations on a theme.[23]

In Sicker et.al. (1999), one of the earliest papers on the Layered Model, the authors sought to start with the model that technologists use to conceptualize the hardware and software associated with a network (i.e., protocol layers), and to use it as a lens through which to view long term policy decisions in general, taking service provider interconnection as a case in point.[24] It quickly became clear to the authors that the TCP/IP protocol suite and the OSI reference did not directly relate to policy issues or market reality. A simplified and abstracted model better captured the policy and market concerns that were of interest to the regulator.[25] For the purposes of that paper, the authors found it necessary to separate transport and applications into discrete layers, but did not find it necessary to address further layers.

Services and service providers are the focus of this model. The service layers distinguish between 1) physical services (e.g., access, transport), 2) application

---

[23] Cf. Whitt, pages 621-622.

[24] *See* Sicker et. al., *supra* note 4 at p. 27.

[25] *See* Sicker et. al., *supra* note 24 at 10.

services (e.g., directories, caching, voice, electronic mail), and 3) content (e.g., music, video programming). More specifically, the authors defined the layers as:

- Physical services: Providers of 1) Access and 2) Transport Services; including both best-effort and QoS services.
- Applications services: Providers of application services that rely on underlying access and transport services can be further subdivided into three subcategories: 1) directory service providers (e.g., DNS and other naming/numbering functions); 2) intermediate or middle service providers (e.g., multicasting and caching); and 3) end user service providers (e.g., voice, email, and hosting). One might argue that each of these three subcategories are distinct and should be treated as such, but the authors did not find it necessary to do so. It is sufficient to distinguish between the data delivery service and the applications that use or support the data delivery service.
- Content: Content providers that rely on underlying transport, access, Application directory, and Application-intermediate services. Examples of content include video, music, and telephony services.

These three layers correspond directly to the Transmission, Application and Content Layers that we put forward in this paper.

This early vision of the Layered Model did not seek to create an entire new model for the regulation of networks; rather, it sought to develop a tool for looking at networks in a more technically neutral and consistent manner. The model proved particularly useful in clarifying the importance of innovation and openness at the application layer.

4. Regulation within each layer

One point of commonality among the many papers that deal with the Layered Model is the recognition that regulatory problems must be addressed in the layer in which they are found. Thus, it may be unproductive or even counterproductive to try to remedy an Application Layer problem by applying a regulatory fix to Transmission providers, and vice versa.

Well and good. But what are the key regulatory problems in each layer, and what decision principles should most appropriately be applied to their solution?

4.1 The Transmission Layer

Much of regulation today is aimed at the Transmission Layer. The challenges of regulation in this layer are perhaps better understood than in the others.

There are a great many regulations that are relevant to this layer, but they can all be viewed as variations on a small number of themes. We would contend that there are only three real reasons to regulate at the Transmission Layer:

- to ensure consumer benefits by enabling competitive market entry, and by protecting the public from exploitation of competitive bottlenecks;
- to ensure that services necessary to public welfare are provided where economic incentives alone might not; and
- to manage limited resources, such as spectrum.[26]

What decision principles are appropriate to each of these three regulatory areas? What constitutes regulatory best practice, based on examples both in the United States and around the world?

Before proceeding to that discussion, we need to briefly digress to deal with a threshold question.

4.1.1 Which layer?

How can we properly identify the layer to which a particular service belongs?

The answer is not always obvious. Some earlier papers distinguished between phone numbers, which were ostensibly an artifact of the Physical Layer, and DNS names, which were an artifact of the Logical Layer.[27]

We would argue that this classification itself reflects a carry-over of old and outdated thinking. A long time ago, a telephone number was closely associated with a physical line. That day is long gone. An 800 number stopped being associated with a single physical line decades ago – it was, in effect, a *name* rather than an *address* (in the

[26] Justus Haucap and J. Scott Marcus, "Why Regulate? Lessons from New Zealand", to appear in *IEEE Communications Magazine*.

[27] Cf. Werbach, op. cit.

engineering sense), just as a DNS name is a name that is mapped to an underlying network address.

Today, the Vonage Voice over IP (VoIP) service enables Marcus's friends to call him in Bonn, Germany, using a phone number with a 202 (Washington, DC) area code. The service could also provide alternative phone numbers in Canada or the UK, all logically associated with Scott's phone in Bonn. But the direct association between phone number and line broke down a long time ago. For x11 numbers, toll free numbers (and other non-geographic numbers), for number portability, and for mobile phones, the number is more appropriately viewed as a name than an address.

We would contend that the telephone number is more closely associated at present with the telephone service, not with the underlying network. Thus, numbering is an artifact of the Application Layer, not of the Transmission Layer.

Once we recognize that a telephone number is most appropriately viewed as an Application Layer artifact, it becomes much easier to see how to treat the use of telephone numbers for VoIP and for PSTN telephony in a regulatorily symmetric fashion.

4.1.2 Competitive bottlenecks, competitive entry

Wherever possible, regulators should step back in order to enable competitive markets to function, impeded by unnecessary regulation.

Unfortunately, not every market is competitive. There are many aspects of telecommunications markets that may lead to weak or imperfect competition, including (1) low marginal cost and high fixed cost; (2) significant economies of scale; and (3) historically concentrated markets. Where competition is ineffective, some form of government intervention is required.

Most countries attempt to deal with this through a combination of (1) competition law (antitrust), which applies sanctions after the fact (*ex post*) in response to anticompetitive acts; and (2) regulation specific to the telecommunications sector, where remedies may be applied in advance (*ex ante*), even in the absence of documented anticompetitive behavior, in order to enable competitors to achieve market entry.

It bears noting that competition law alone does not appear to be sufficient. Reformist New Zealand attempted for years to operate with little or no *ex ante* regulation. They ultimately gave it up as a bad job, and implemented regulatory structures roughly comparable to those of other developed economies. In the absence of ex ante regulation, prospective new market entrants stalled – it was simply too hard and too time-consuming to achieve entry. The most daunting challenges to new entrants involved network interconnection.[28]

Happily, this one very important area of regulation can for the most part be viewed as a solved problem. The regulatory framework that the European Union put in place in July, 2003 is clearly working. It continues to face challenges, to be sure, but it is increasingly clear that it represents a very substantial improvement over everything

[28] Haucap and Marcus, op. cit.

that came before. It provides a unified and generally well-conceived framework for addressing the barriers to competition associated with market power.

The European approach is to (1) define a set of regulatory obligations that can appropriately be used to address market power where it may exist; (2) define an general set of markets to which *ex ante* regulation may be appropriate; (3) task national regulators with providing detailed market definitions suitable to their respective national circumstances, and identifying any service providers who may have dominance (*Significant Market Power,* or *SMP*) or those markets; (4) imposing a minimally adequate (*proportionate*) set of *ex ante* remedies on those providers, and only on those providers, that possess SMP on the market in question; and (5) lifting any corresponding obligations from service providers that do *not* have SMP on the market in question.[29]

The European approach requires serious economic analysis at a fairly detailed level. Doing so requires a significant investment of resources, primarily skilled analysis by economists. It also requires institutions that enable objective and impartial analysis. And it takes time to do the analysis correctly. Nonetheless, preliminary results are that it reaches appropriate conclusions most of the time, and in particular that it generally addresses technological and market convergence appropriately.[30]

From roughly 1970 to 2000, U.S. law and regulatory policy reached conclusions very similar to those of the new European system.[31] The need to address such market power as might exit was well understood. Only in recent years has there been substantial divergence in this country.

A series of FCC proceedings known as the Computer Inquiries[32] allowed telecommunications carriers to offer so-called *enhanced services* (services representing a mix of telecommunications and computing), and declined to regulate them because they felt that the market for enhanced services would be competitive. At the same time, the Computer Inquiries ensured that the market would indeed be competitive by requiring carriers to make any underlying telecommunications services available to third parties on nondiscriminatory and reasonable terms, which is to say on terms no less favorable than the terms on which their own enhanced services

---

[29] The European regulatory framework is discussed at length in earlier work. See Marcus, Federal Communications Commission (FCC) Office of Strategic Planning and Policy Analysis (OSP) Working Paper 36, "The Potential Relevance to the United States of the European Union's Newly Adopted Regulatory Framework for Telecommunications," July 2002, available at http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-224213A2.pdf. The article and derivative works also appear in: *Rethinking Rights and Regulations: Institutional Responses to New Communications Technologies*, Ed. Lorrie Faith Cranor and Steven S. Wildman, MIT Press, 2003; in the *Journal on Telecommunications and High Technology Law* 111 (2003); and in the *2004 Annual Review of the European Competitive Telecommunications Association* (ECTA). Citations to relevant European documents can be found there.

[30] See Marcus, "Europe's New Regulatory Framework for Electronic Communications in Action", presented at the 4th ZEW Conference on the Economics of Information and Communication Technologies, Mannheim, Germany, July 2004. Available at:

ftp://ftp.zew.de/pub/zew-docs/div/IKT04/Paper_Marcus_Invited.pdf.

[31] See Marcus, "The Potential Relevance …", op. cit.

[32] Op. cit.

obtained them. In other words, the Computer Inquiries addressed the underlying market power.

The codification of the Computer Inquiries into the Telecommunications Act of 1996 captured much of the verbiage used to describe enhanced services in the Computer Inquiries;[33] however, it failed to capture the underlying rationale. Thus, when the FCC in recent years reclassified cable modem service, and also DSL bundled with Internet access, as information services not subject in general to regulation, they were under no obligation to address such market power as may exist. Indeed, they were under no statutory obligation to even determine whether market power on this market might exist at all. The question of market power for broadband is a complicated one, but it is inconceivable that there are not some geographic areas where market power is problematic. In fact, in the only known instance where the FCC inadvertently released a cursory economic analysis of the wired broadband marketplace, they found the market to be highly concentrated overall.[34] There are real questions about the effectiveness of competition on this market.

Conversely, the Layered Model is usually understood to imply that the last mile market power that necessarily resides in the physical transmission layer must be addressed.[35] This would be equally broken, but in the opposite direction. There are assuredly some geographic areas where market power exists on the wholesale and retail markets for broadband access; there are probably substantial geographic areas where broadband services could be considered to be reasonably competitive.

Market power should neither be presumed to exist, nor presumed not to exist. It should be analyzed objectively and quantitatively.

Competition / antitrust law and economics have been around for more than a century.[36] Antitrust cases are inherently high stakes games; consequently, a great deal of economic analysis has been funded over the years. The economists know quite a bit about how to identify market power.

Again, from a public policy perspective, the answer is clear: Market power should neither be assumed to exist, nor assumed not to exist, on a blanket basis. It is necessary to analyze the markets in question, at an appropriately granular level, using quantitative tools as much as possible.[37] The regulator needs to be willing to do this analysis, needs to have the time (and to *take* the time) to do the analysis, needs to approach the subject fairly and objectively, and needs to be funded to have the capabilities (in particular, highly trained economists) to do the analysis properly.[38]

---

[33] 47 U.S.C. section 153.

[34] In the MMDS/ITFS NPRM, the FCC found the wired broadband market to have high and worrisome HHI.

[35] As previously noted, the Whitt paper is often (but inaccurately) interpreted as saying this.

[36] Cite to Sherman Antitrust Act, circa 1890, and Clayton Act, around 1914.

[37] The DC Circuit has said much the same in …

[38] It bears mentioning that the FCC as presently constituted is not well suited to this task.

4.1.3 Ensuring that socially important services are provided

A nation's communications system is widely understood to represent vitally important national infrastructure. Market forces can and do serve to meet a wide range of societal needs; nonetheless, there may be gaps.

At the Transmission Layer, one of the most conspicuous gaps relates to the availability of communications services at reasonable price to remote areas, to areas of low teledensity, and to the poor. Commercial incentives will not necessarily make it cost-effective for any provider to offer service at a reasonable price under these conditions.

This shortfall is generally addressed through some form of *universal service*. Areas of high teledensity subsidize areas of low teledensity. Economists tend to be uncomfortable with universal service, because it inevitably introduces economic distortions. At the same time, most governments consider universal service to be necessary in order to ensure that no citizens are excluded from the at least the basic benefits of communications.

Here, there is not a clear cut best practice – universal service is under great stress everywhere. Most economists would recommend that the costs of universal service be distributed over as wide a contribution base as possible, in order to minimize distortions. Ideally, a government would pay for universal service out of general revenues.

This leaves an enormous question: In today's world, what services should be subsidized? Traditionally, the answer has been basic telephony. In the world of the future, will broadband Internet access replace basic telephone service? And does this imply that VoIP service becomes part of the package?

The provision of over-the-air broadcast television is an aspect of universal service that is often overlooked. Government tends to play an active role, effectively subsidizing over-the-air broadcast with cheap or free spectrum rights, as a means of ensuring that those of limited means can receive basic video broadcasting services even though they cannot afford to pay for cable television or satellite.

4.1.4 Managing limited resources

For the Transmission Layer, the conspicuous resource is spectrum. One might perhaps include Internet addresses at this level, since they are true addresses. As noted previously, phone numbers, which are often thought of a being tied to transmission, these days are more appropriately viewed as Application Layer resources.

As regards spectrum management policy, the leaders in terms of best practice are widely acknowledged to include the United States, the United Kingdom, Australia, New Zealand, and Canada. In all of these countries, there are strong movements afoot to reduce (but not eliminate) the role of traditional regulation, and instead to increase the use of market mechanisms.

This trend has many aspects. On the one hand, it implies a more property-like handling of spectrum: clearly defined rights, assignment through auctions, flexible use, and secondary markets. On the other, there is increased emphasis on collective use in commons arrangements (unlicensed spectrum).

Control of scarce resources administered has elements in common with addressing market power. Scarcity of spectrum can, indeed, create market power. But there are also differences, inasmuch as this kind of regulation deals with resources that are societally managed in the first place. These two forms or regulation interact, but they are not the same.

4.2 The Applications Layer

The key regulatory challenges at the Applications Layer can again be characterized as comprising (1) ensuring effective competition, (2) ensuring that societal goals are met, and (3) managing scarce resources. But each of these manifests itself in discernibly different ways at the Applications Layer than they do at the Transmission Layer.

4.2.1 Ensuring effective competition

Telecommunications regulation to deal with competitive bottlenecks has historically been focused on the Transmission Layer; however, that does not necessarily preclude the possibility of bottlenecks at the Applications Layer.

At the same time, there is much less regulatory experience in addressing barriers to competition at this level.

For example, a significant concern has arisen in terms of network externality effects, primarily in the context of mergers or proposed mergers of Internet backbones. Regulators on both sides of the Atlantic have recognized the potential competitive harms associated with permitting a firm to merge to the point where it controlled access to too large a number of end users, most notably in conjunction with the WorldCom/MCI merger and the proposed WorldCom/Sprint merger.

In the European Union, Article 5 of the Access and Interconnection Directive provides national regulatory authorities with substantial ability to impose "… to the extent that is necessary to ensure end-to-end connectivity, obligations on undertakings that control access to end-users, including in justified cases the obligation to interconnect their networks where this is not already the case…" In the context of European regulation, this is an extraordinary power, inasmuch as it can be applied even in the absence of Significant Market Power. At the same time, the intent is that it should be used very sparingly, precisely because it is otherwise unconstrained.

4.2.2 Ensuring that societal goals are met

For reasons previously noted, we consider it appropriate to treat telephony service as as Application Layer construct rather than a Transmission Layer construct. Thus, the Voice over IP (VoIP) service resides at this layer, even though the network on which the service rides is a Transmission Layer capability.

A great many regulatory obligations on voice telephony are most appropriately viewed as Application Layer instances of ensuring that societally vital services are provided, despite a lack of sufficient market incentives to achieve them solely through market mechanisms. For example, citizens recognize the value to society of

the ability to impose a wiretap for good cause; at the same time, no individual user would choose to pay a premium for the knowledge that his own phone could be tapped if need be. Similarly, we all recognize the value of being able to call for emergency services, but the value to society as a whole is probably much greater than the premium that most individuals would be willing to pay for the service.

European sector-specific telecommunications regulation draws a clear distinction between the service and the network; unfortunately, the *Universal Service Directive* (which specifies obligations relevant to retail services provided to consumers, including voice telephony, and not just in the context of universal service as the term is defined in the United States) fails to follow through with a clean allocation of obligations either to the service or the network.

The USD contains literally dozens of potential regulatory obligations on service providers. Among them are:

- Obligations to provide universal service – not just access to public telephone network, but also operator services, directory enquiry services, and pay phones.
- Obligations of transparency, nondiscrimination, carrier preselection, and retail price controls on firms that have market power at the level of retail services;
- Publication of information about commercial terms and conditions, as well as service quality;
- Access to emergency services; and
- Number portability.

There is no simple rule to magically tell the regulatory which potential Application Layer obligations are likely to be warranted, and which are not. Nonetheless, it is possible to state a few general guidelines. Marcus suggested in prior work that policymakers might choose to be guided in such cases by two criteria:

> **Balance**: Government should recognize both the risks of action and those of inaction, and make cautious and deliberate choices.
>
> **Minimalism**: Government should choose to err in general on the side of less regulation rather than more. Do not attempt a massive intervention where a less intrusive intervention might suffice. Do not intervene at all unless markets have shown themselves to be unable to deliver a socially important outcome.[39]

This argues for a sober and dispassionate cost-benefits analysis, considering also the likelihood that the capability in question would deploy without government intervention solely as a result of competitive market forces.

[39] "Evolving Core Capabilities of the Internet", *Journal on Telecommunications and High Technology Law*, 2004.

4.2.3 Managing scarce resources

Telephone numbers are a limited resource – not, perhaps, in the same sense as spectrum, but limited nonetheless.

Again, European regulation provides some useful insights as to possible policy desiderata. Article 10 of the Framework Directive directs national regulatory authorities to ensure reasonable access, prevent discriminatory practices, and assure open and transparent process.

4.3 The Content Layer

Application of the model to the Content Layer clearly raises issues that are not present in the Transmission and Applications Layers.

Ensuring effective competition is again relevant; again, it manifests itself differently in this layer. Traditional competition law concerns are only part of the story; the other part relates to the number of independent voices, and the degree to which local preferences are reflected in available content. In terms of economic analysis, media ownership rules are quite distinct from classic competition analysis.

Meeting societal goals where the market otherwise would fail to do so is also relevant at this Layer. For example, making educational programming available to children could be viewed in this light.

Different communications media have historically interacted very differently with content, largely as a function of different regulatory treatment. For telephony, the Transmission and Application Layer provider has generally been forbidden from interfering with or interacting with the user's content. For broadcasting, by contrast, the Transmission and Application Layer provider (for example, the cable operator) selects content in an effort to maximize its profitability. This selectivity could, in the absence of regulation, lead to monopsony effects (i.e. market power due to a limited number of wholesale purchasers of content).

Regulatory principles at this layer do not seem to lend themselves to tidy quantification and categorization to the same degree as other layers. Issues at this layer include freedom of speech, and indecency and obscenity.

5. Summary

Layers represent an interesting lens through which to view regulatory issues; however, the Layered Model does not inherently tell us what the regulator should be doing within each layer.

We find it necessary to return to the notion that we regulate for three reasons: (1) to ensure effective competition; (2) to achieve necessary social objectives where the market alone would not; and (3) to manage scarce resources.

What we observe is, first, that these three premises drive necessary regulation; and second, that they manifest themselves somewhat different within the different layers of the Simplified Layers Model.

In none of these instances does the Layered Model necessarily tell us to regulate them differently than they were regulated in the past; rather, it may help us to understand why we are doing what we are doing, and thereby to clarify what we need to do going forward.