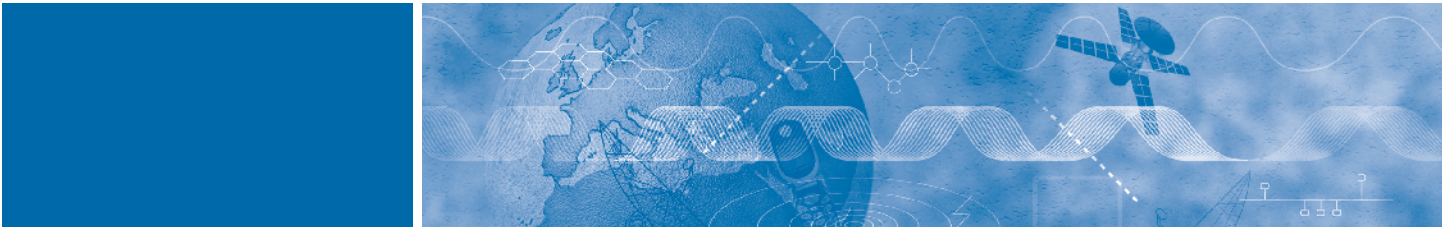


12 September 2007

Freedom, Regulation, and Net Neutrality

William E. Taylor



*Die ich rief, die Geister,
werd' ich nun nicht los!*¹

Introduction

The current debate over network neutrality—in all meanings of that phrase—embodies at its core a fundamental contradiction. Some parties (“Netheads²”) extol the *freedom* of the Internet: freedom from metering and from such government interference as taxation, commercial regulation, or political or moral censorship. For others (“Bellheads”), the Internet is a business, and their facilities and the services that ride upon them must be managed and priced to be as profitable as possible. The contradiction arises when Netheads advocate government regulation to protect themselves from potential anticompetitive actions or monopolistic exploitation by those Bellheads who supply last-mile broadband access to the network.

In this debate, there is plenty of room for factual disagreement regarding the market conditions that may or may not make such anticompetitive behavior likely or profitable. Similarly, well-meaning people can disagree on the costs and benefits of regulation: that is, on the relative incentives to innovate and invest in different forms of content and applications on the one hand, or in different facilities or network services on the other. However, disagreement between Netheads and Bellheads on the fundamental desirability of regulation is surprising because advocating regulation to preserve Internet freedoms is inherently inconsistent. Properly understood, “net neutrality” is reminiscent of the “Surprising Barbie!” marketing phrase in which a product is positively characterized by its most

¹ “From the spirits that I called, Master, deliver me.” Goethe, *The Sorcerer’s Apprentice*.

² Rob Frieden distinguishes a Nethead culture advocating a global connectivity through a seamless and unmetered voluntary interconnection of networks from a Bellhead culture that promotes managed traffic, metering, and network cost recovery: see “Revenge of the Bellheads: How the Netheads Lost Control of the Internet,” *Telecommunications Policy*, 26, No. 6 (Sept./Oct. 2002) at 125-144.

negative attribute. Whatever else it does, a doll never surprises. The Internet today is not and cannot ever be “neutral” across a vast array of applications, and advocating regulation to preserve neutrality in this context risks the outcome lamented by the sorcerer’s apprentice.

The Markets, the Players, and the Problem

Internet markets are complicated and identifying the players and their roles will help fix ideas. Think of “application providers” as firms at one end of a network supplying content, e.g. web pages, voice communications, search engines, massive multi-player games, etc. Some “content” also facilitates the sale of other products and services, e.g., Landsend.com and Amazon.com. “Customers” are end-users at another end of the network, who value and purchase content (and sometimes products and services) from application providers. Customers, or at least their eyeballs, also represent services sold in the form of advertising by some application providers (e.g., Yahoo!) in this classic two-sided marketplace.³ “Access providers” are the telecommunications carriers that supply last-mile, dedicated access to the Internet to application providers at one end and customers at another. “Carriers” include access providers at the ends of the network, as well as the suppliers of backbone network facilities that interconnect with one another ultimately to connect applications providers and customers.

This paper examines the economics of one aspect of net neutrality, so-called “access tiering.” The question here is whether the current Internet standard of “best effort” carriage of data packets should be modified to permit voluntary prioritization of traffic: i.e. to allow carriers to provide different qualities of service in some dimensions for a price to different applications, different application providers, or different customers. Differential quality of service (QoS) arises in a packet-switched network when the network experiences congestion. At every switching point in a network, switches (called “routers”) receive traffic in the form of packets, examine the destinations of the packets, and attempt to send each packet down its proper route. When more packets arrive than can be sent out, routers respond by using alternative routes, by queuing packets for transmission and ultimately, if their buffers are full, by dropping packets. All of these responses can degrade the QoS experienced by the customer.⁴ Today, and from its inception, the Internet generally treats all forms of traffic indiscriminately with respect to the routing, queuing, or dropping of packets, ignoring the nature of the service or the identities of the application providers, the access providers, and the customers. Thus:

- packets that correspond to the download of a webpage are accorded the same treatment as packets comprising a VoIP conversation despite the fact that the resulting delay and “jitter” have vastly different effects on the customers’ experiences for these applications;
- packets corresponding to applications initiated by an AT&T DSL customer are delayed or dropped as frequently for AT&T applications as for Vonage, Google, or Yahoo applications initiated by that customer; and
- packets corresponding to some real-time medical procedure between a doctor and a distant hospital are subject to the same delay and jitter as packets carrying Yahoo instant messages between bored high school students in math class.

³ Think of customers as honeybees, websites as flowers, and pollination as advertising revenue.

⁴ Albeit in different ways for different services. Increased average delay may be unnoticeable for web-browsing but annoying for downloading large files. Increased maximum delay may degrade performance for VoIP conversations or streaming audio or video applications.



That all traffic is treated equally in this sense is probably the origin of the notion that the Internet is “neutral.” However, as these examples show, neutrality with respect to packets does not translate into neutrality with respect to applications. It is here that the net neutrality debate has teeth and where economics has the most to add to the discussion.

Opponents of access tiering argue that permitting access providers to charge application providers for higher QoS would degrade the performance of other applications, reduce innovation and investment in content applications, and allow carriers to leverage their alleged bottlenecks in access into anticompetitive advantages in the various markets for applications.

On the other hand, proponents of voluntary access tiering respond that: a market for priority might change the current distribution of the effects of congestion across applications, but that change would make customers vastly better off; rather than reducing the incentive to innovate and invest in content, access tiering would make it possible to develop applications more highly-valued by customers that are currently prohibitively expensive to implement due to congestion from lower-valued traffic; and the limited or non-existent ability and incentive of carriers that are vertically integrated into applications to profitably discriminate against their competitors’ applications are insufficient to warrant *ex ante* regulation and should be controlled—as they are elsewhere in the economy—by the stringent application of antitrust or competition law.

Degraded Performance

Properly construed, access tiering has no effect unless network facilities are congested. Unfortunately, the tragedy of the commons ensures that, over time, the demands of applications will expand to fill the capacity available. In the long run, bandwidth, reliability, and low latency service will be rationed by some mechanism. Some advocates of net neutrality legislation insist that *price* not be used as the rationing mechanism, at least in part because such an allocation might degrade the quality of their favorite applications. For example,

The current legislation, backed by companies such as AT&T, Verizon and Comcast, would allow the firms to create different tiers of online service. They would be able to sell access to the express lane to deep-pocketed corporations and relegate everyone else to the digital equivalent of a winding dirt road.⁵

Instead, packets would be routed without regard for cost or for the customer’s willingness to pay for additional QoS and irrespective of the application or the identity of the customer, the application provider, or the carrier(s).

When facilities are congested, some traffic must inevitably be delayed or degraded. Changing the assignment of priority—from the current first-in first-out priority—to something else could degrade the current quality of some applications relative to others when facilities are congested.

But using any mechanism to assign priority other than one that reflects cost and consumers’ willingness to pay for priority can impose massive welfare losses on society.⁶

⁵ Lawrence Lessig and Robert W. McChesney, “No Tolls on The Internet,” Washington Post, 8 June 2006; page A23.

⁶ For example, the welfare-maximizing rule would price the different services at differential markups over the different incremental costs, where the markups were inversely proportional to the price elasticity of demand.



In a simple example, Litan and Singer estimate the consumer surplus from one QoS-needy application—online massive multiplayer gaming⁷—at between \$700 million and \$1.5 billion per year in 2009, surplus that would be lost or reduced if net neutrality regulation were effective in equalizing priority.⁸ Add to that figure consumer surpluses from other high-bandwidth, low latency services such as delivery of IP high-definition video.⁹ That these applications (and others) would be priced out of the market in an unmanaged, net-neutral network is clear from some estimates of Richard Clarke. According to Clarke, the monthly cost to provide capacity for current subscribers in an unmanaged network is about \$47 per subscriber. To add sufficient capacity to provide two standard-definition video channels would triple that cost. To add two high-definition video channels would increase the cost by a factor of 10.¹⁰ While cost estimates are uncertain, it is evident that the simple expedient of allowing backbone capacity to increase over time in an unmanaged network would render many high-bandwidth services unmarketable at a massive sacrifice of economic welfare.



Network pricing that accounts for costs and customers' willingness-to-pay generally increases economic welfare relative to uniform pricing, particularly where costs are largely fixed, uniform marginal cost pricing does not recover the total cost of the network, and differential pricing would expand total network demand. With a single level of quality stemming from best-effort switching, under congestion, customers would be willing to pay more than the average price for QoS-needy applications but are prevented from doing so by regulation. At the same time, QoS-irrelevant applications are supplied at a higher level of quality than customers would be willing to pay for. Both errors entail welfare losses.

An implied but unstated concern of some proponents of net neutrality regulation is that consumers' demand for video services, if unchecked, would absorb all available capacity and thus degrade the Internet experience of other users.¹¹ While I might share Professor Lessig's apparent taste for video services, I know of no better mechanism to allocate scarce capacity than market prices. If providing high-definition streaming video is costly and customers' demand for the service is insatiable, then the response time for applications I prefer may suffer, but economic welfare as a whole may be enhanced.

⁷ Ask your children.

⁸ R.E. Litan and H.J. Singer, "Slouching Towards Mediocrity: Unintended Consequences of Net Neutrality Regulation," *Journal on Telecommunications and High Technology Law*, 2007, downloaded from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=942043.

⁹ And if the inefficient undersupply of these particular services leaves you unmoved, consider why such frivolous uses of scarce capacity are not charged the opportunity costs they impose on the socially more valuable uses you favor.

¹⁰ Clarke, Richard N, "Costs of Neutral/Unmanaged IP Networks" (May 2006) at 20. Available at SSRN: <http://ssrn.com/abstract=903433>.

¹¹ For example: "The incentives in a world of access-tiering would be to auction to the highest bidders the quality of service necessary to support video service, and leave to the rest insufficient bandwidth to compete." Testimony of Lawrence Lessig Before the Senate Committee On Commerce, Science and Transportation Hearing on Network Neutrality, 7 February 2006.

Innovation and Investment

The Internet has arguably been responsible for the greatest transformation of industry and commerce since the Industrial Revolution. It has spawned new and innovative services and generated billions of dollars of investment, both in applications and by carriers in new equipment and facilities. Some proponents of network neutrality regulation observe the so-called “end-to-end” design of the network and infer that this particular architecture—intelligence and control at the network edges and transparency at the center—is optimal for innovation. For example,

One consequence of this design is that early network providers couldn’t easily control the application innovation that happened upon their networks. That in turn meant that innovation for these network (*sic*) could come from many who had no real connection to the owners of the physical network itself...

This diversity of innovators is no accident. By minimizing the control by the network itself, the “end-to-end” design maximizes the range of competitors who can innovate for the network. Rather than concentrating the right to innovate in a few network owners, the right to innovate is open to anyone, anywhere.¹²

While admiration for the innovation and investment associated with the Internet is understandable, the implication that the end-to-end architecture is a necessary cause of that growth is not. With transparent best-effort switching, a wide variety of largely QoS-independent applications were devised and marketed successfully as the capacity of the Internet and penetration of broadband access increased. In a different architecture—say, one that permitted voluntary priority pricing—a different mix of applications would emerge, but there is no theory or evidence to suggest that the amount of innovation (somehow quantified) or, particularly, the value of those innovations to consumers would have been less. Why would the ability of an application provider to choose among a range of QoS standards for the network reduce its incentive to invest and innovate? Indeed, if priority prices reflected the costs of priority as well as consumers’ valuations of the applications that depend on priority, one would expect more valuable innovation in a market-determined network architecture rather than less.

Anticompetitive Discrimination

Of course, the assumption that priority prices would be set in effectively competitive markets is not one that network neutrality proponents would accept. Rather, Netheads point to concentration in the broadband access markets and infer that access and backbone facility providers would have the incentive and ability to exercise market power and to leverage it from access into the provision of applications. The economics of this claim are dubious.

First, the carriers’ ability to exercise or leverage market power from access is constrained by the presence of multiple broadband access providers, including telephone, cable, and wireless carriers. An access provider that degraded competitors’ applications would be handicapped in the competition in the access market.

¹² Testimony of Lawrence Lessig Before the Senate Committee On Commerce, Science and Transportation Hearing on “Network Neutrality,” 7 February 2006.

Second, the incentive to use a putative access bottleneck to derive a competitive advantage in an applications market is also questionable. Basic economics shows that—with some exceptions—monopoly profit (e.g., in access) can only be earned once in a vertically-integrated firm, and that additional profit cannot—in general—be earned by subsidizing competition in a downstream applications market from an upstream access monopoly.¹³

Third, despite the above, if priority pricing is thought to increase the likelihood of anticompetitive pricing (or other behavior) on the part of access providers, the efficient solution is the rigorous application, *ex post*, of the antitrust or competition laws. Generally speaking, *ex ante* economic regulation only makes sense in cases of natural monopoly, where an efficient industry structure precludes entry and inefficient monopoly pricing would be the natural consequence. In markets opened to competition, *ex ante* regulation is too blunt an instrument to distinguish between vigorous competition and anticompetitive acts. As a result, a reasonable answer to Netheads who fear anticompetitive leveraging of an access monopoly into markets for applications is to enforce the antitrust laws, rather than restricting *ex ante* the pricing and provisioning of services for which customers are willing to pay.¹⁴

Unintended Consequences of the Regulation of Access Tiering

Ex ante economic regulation, particularly in markets opened to competition, necessarily imposes costs on society. These costs include the inherent market, technology, and investment distortions that stem unavoidably from economic regulation of any service—retail or wholesale. These distortions are particularly acute in current telecommunications markets where retail markets are subject to effective competition, where regulatory authority differs across competing platforms, and where the markets are characterized by rapid technological change and competing platforms or technologies subject to lock-in or path dependence.¹⁵ Such regulation does not merely transfer welfare among suppliers but inevitably distorts technical choices which can have large and irreversible welfare effects on consumers, reducing economic efficiency and productivity by distorting the competitive market outcome, and driving the market to an inefficient platform or technology.

In particular, detecting anticompetitive discrimination in priority would be difficult. Consider a carrier whose network exhibits higher jitter than other networks so that the quality of VoIP applications for its customers is degraded.

¹³ See, e.g., Robert H. Bork, *The Antitrust Paradox: A Policy at War With Itself* at 372-75 (2d ed. 1993) or Richard Posner, *Antitrust Law: An Economic Perspective* 171-74 (1976). The recognized exceptions to this basic theory include avoiding regulation in the upstream market, the presence of variable proportions in the downstream market, and an enhanced ability to price discriminate in the upstream market. See D.W. Carlton and J.M Perloff, *Modern Industrial Organization* (Fourth edition) Boston: Addison Wesley, (2005), Chapter 12.

¹⁴ In fact, each of the few examples of anticompetitive conduct by access providers cited by net neutrality proponents were quickly identified and corrected by the relevant enforcement agency.

¹⁵ For example, consider last-mile broadband access service, which can be provided by wireline, power line, or wireless carriers.



The most challenging possibility, for a policy standpoint, is that [the carrier] didn't take any obvious steps to cause the problem but is happy that it exists, and is subtly managing its network in a way that fosters jitter. Network management is complicated, and many management decisions could impact jitter one way or the other. A network provider who wants to cause high jitter can do so, and might have pretextual excuses for all of the steps that it takes. Can regulators distinguish this kind of stratagem from the case of fair and justified engineering decisions that happen to cause a little temporary jitter?¹⁶

Futile and counterproductive attempts to regulate the details of service quality are familiar from the worst days of airline regulation, including prescribing the maximum amount of leg-room, requiring that meals be limited to sandwiches, and establishing uniform additional prices for in-flight entertainment. While these examples appear silly, their consequences are not. Service quality and price are two sides of the same coin and, together with competition, such regulation reverses the competitive process, causing costs to move towards a regulated price rather than the other way around.¹⁷



In airlines, quality regulation gave rise to the sumptuous three-course sandwich. In telecommunications, if any lesson can be learned from the growth and diversity of the Internet, it is that networks and the applications that ride on them adapt extremely rapidly and unpredictably to changes in their environments. In the Internet setting, if some sorcerer's apprentice invoked a government agency to regulate *ex ante* the terms on which application providers, carriers, and customers were permitted to exchange vast amounts of traffic in real time, the consequences would be unimaginable.

¹⁶ E.W. Felton, "Nuts and Bolts of Network Neutrality," August 2006 (<http://itpolicy.princeton.edu/pub/neutrality.pdf>).

¹⁷ See A.E. Kahn, *The Economics of Regulation*, Vol II Boston: The MIT Press (1988) at 209-220.



About NERA

NERA Economic Consulting is an international firm of economists who understand how markets work. We provide economic analysis and advice to corporations, governments, law firms, regulatory agencies, trade associations, and international agencies. Our global team of more than 600 professionals operates in over 20 offices across North and South America, Europe, and Asia Pacific.

NERA provides practical economic advice related to highly complex business and legal issues arising from competition, regulation, public policy, strategy, finance, and litigation. Our more than 45 years of experience creating strategies, studies, reports, expert testimony, and policy recommendations reflects our specialization in industrial and financial economics. Because of our commitment to deliver unbiased findings, we are widely recognized for our independence. Our clients come to us expecting integrity and the unvarnished truth.

NERA Economic Consulting (www.nera.com), founded in 1961 as National Economic Research Associates, is a unit of the Oliver Wyman Group, an MMC company.

Contact

For further information contact:

William E. Taylor

Senior Vice President and Global Communications Practice Chair

Boston, MA

+1 617 621 2615

william.taylor@nera.com