

# Prospects for Differentiated Quality of Service within and among IP-Based Next Generation Networks (NGNs)

J. Scott Marcus<sup>1</sup>

## Abstract

As networks evolve from the traditional into Next Generation Networks (NGNs) based on the Internet Protocol (IP), many network providers expect to differentiate their networks from the public Internet by offering lower network delay or more stringent limits on variability of delay; moreover, they expect in some way to improve their profitability by charging for better-than-best efforts quality of service. It has long been recognized that enhanced QoS could be beneficial to the performance of some applications, especially bidirectional real time voice and video, and the technology has been mature for a decade (and is widely used *within* networks); nonetheless, few if any networks have implemented differentiated QoS *between* networks. As an additional concern, experience from the United States calls into question the willingness of consumers to pay a significant premium for enhanced QoS, even though they may claim to value it. Finally, the question has become mired in the network neutrality debate in the United States – many experts are questioning whether differentiated QoS might do more harm than good, and in particular whether it might serve as an avenue for the exploitation of market power.

Taking into account the technology and the economics of network interconnection, and these unhappy experiences in the United States, what can we say about prospects for deployment and adoption of differentiated QoS between providers in Europe and the rest of the world?

---

<sup>1</sup> Author's current addresses: WIK-Consult GmbH, Bad Honnef, Germany, [scott@scottmarcus.com](mailto:scott@scottmarcus.com).

## Contents

Prospects for Differentiated Quality of Service within and among IP-Based Next Generation Networks (NGNs) .....	1
1. Introduction.....	3
2. Background.....	4
2.1 Migration to IP-based Next Generation Networks .....	4
2.2 Technical background.....	5
2.2.1 Applications delay requirements.....	5
2.2.2 Delay in packet networks.....	5
2.2.3 Technical solutions .....	7
2.2.4 Implications for billing and accounting systems .....	7
2.3 Economic background .....	9
2.3.1 General theory of telephone network interconnection.....	9
2.3.2 Economic theory of IP-based interconnection.....	10
3. Price discrimination: A good thing, or a bad thing? .....	11
4. Glacial adoption of differentiated QoS between networks .....	13
5. Public policy implications.....	15

## 1. Introduction

Many services are offered with varying levels of quality, and at different prices. For more than a decade, it has been understood that, while the Internet was ideally suited for delay-insensitive packetized traffic, that it was less than ideal for delay-sensitive traffic such as two way real time voice. It was quickly recognized that differentiated Quality of Service (QoS) was the appropriate technical solution.

Differentiated Quality of Service (QoS) has been widely implemented *within* IP-based networks for perhaps a decade. Nonetheless, implementations *between* IP-based networks have been rare, even though the technology is relatively straightforward.

As the Public Switched Telephone Network evolves in the direction of IP-based NGNs, there are obvious technical and functional advantages to be sought in migrating the interconnection of those networks to an IP basis – not just for delay-insensitive data applications such as e-mail, but also for delay-sensitive applications such as real time bidirectional voice.

Price and service differentiation are commonplace in many industries. When we purchase airplane or rail tickets, we do not consider it pernicious that we are offered a choice of first and second class tickets. Why has price differentiation come to be viewed so differently in the case of the Internet?

- Does differentiated QoS enhance consumer welfare, or detract from it?
- Where and how is differentiated QoS likely to appear as networks evolve into IP-based NGNs?
- Where is it unlikely to appear without “help”?
- What are the likely consequences if differentiated QoS is not available for IP interconnection?
- What public policy initiatives should be considered, if any?

## 2. Background

These questions have technical and economic aspects. In this section, we provide background on both.

Section 2.1 explains the key implications of the migration to NGN insofar as it relates to QoS take-up. Section 2.2 provides background on the technical underpinnings of QoS in IP-based networks, while section 2.3 provides background on the economics, especially as it relates to the interconnection of IP-based networks.

### 2.1 *Migration to IP-based Next Generation Networks*

There is a widespread move globally for networks to evolve to IP-based Next Generation Networks. Instead of a single-purpose network designed to deliver a single service (voice telephony), the telecommunications network is evolving into an integrated network that delivers multiple services (voice, video and data) over a single, versatile and highly capable transmission platform.

In Europe, the best known example is British Telecom (BT), which has committed to migrate at least half of its network to an NGN by 2009, and which intends to decommission its traditional switches as quickly as possible once NGN infrastructure is in place. As I write this paper in November of 2006, BT has just gone live with real customers on 21CN.<sup>2</sup>

In North America, one rarely speaks of NGNs, but roughly the same migration to IP-based infrastructure is anticipated. Meanwhile, cable television networks are undergoing a similar evolution over time.<sup>3</sup>

In an IP-based world, applications operate on an *end-to-end* basis. In other words, they are present in the end user's system, and in the system with which the user is interacting (typically either a server or another end user's system), but not in any of the systems in between. This elegant design represents one of the key strengths of the Internet, in that it means that the transmission system can be a relatively simple transmission plane.<sup>4</sup>

In the Internet, interconnection constitutes the simple transmission of IP datagrams. Forwarding engines ("routers") need not understand the applications whose traffic they carry. Again, the decoupling of the applications from the underlying transmission simplifies the design, and also greatly facilitates the evolution of all aspects of the system over time.

---

<sup>2</sup> BT, "First customers go live on BT's 21st century network", at <http://www.btplc.com/News/Articles/ShowArticle.cfm?ArticleID=90018709-f48e-409a-a1e4-9b49fd423335>.

<sup>3</sup> In the case of cable networks, however – and also in the case of telecommunications networks based on Passive Optical Network (PON) technology – there may be a tendency to continue to deliver video using traditional means for some time. These networks are already exceptionally well suited to video.

<sup>4</sup> Cf. Isenberg *The Stupid Network*.

By design, then, the routers know next to nothing about the applications whose traffic they carry; however, the application has the ability to signal the routers as regards any special requirements that its traffic might have. It is then up to the routers to accommodate these needs as best they can, *to the extent that the organizations that operate them wish them to do so*. In an IP-based network, performance is assured on a statistical basis – it can be met on average, but cannot in general be guaranteed.

## **2.2 Technical background**

In the following sections, we discuss the nature of application requirements, the basic technical characteristics of packet transmission in the Internet as they relate to variable delay, and the implications of differentiated QoS for billing, accounting, and Operational Support systems.

### **2.2.1 Applications delay requirements**

What sort of performance does a user need from the network? This is a function of what the user is attempting to do.

For typical data applications – email, for example – the user may need the ability to send a fair amount of data, but a great deal of packet delay is acceptable. Even if packets were to take several seconds to get through the network – which is an extremely long time by the standards of today’s networks – the user would receive his or her email and would likely be satisfied with the service.

For streaming video – one way transmission akin to watching television – the network needs to transmit a great deal of information, but can still tolerate some variation in delay, as long as the user will accept a second or two of delay when the transmission begins. This is so because the receiving system can *buffer* the video data. As long as individual packets are not delayed more than the original start-up delay, the buffer corrects for occasional slow-downs. Even occasional loss of data will not necessarily result in performance unacceptable to the user, as long as the receiving system is designed to try to smooth over the gap.

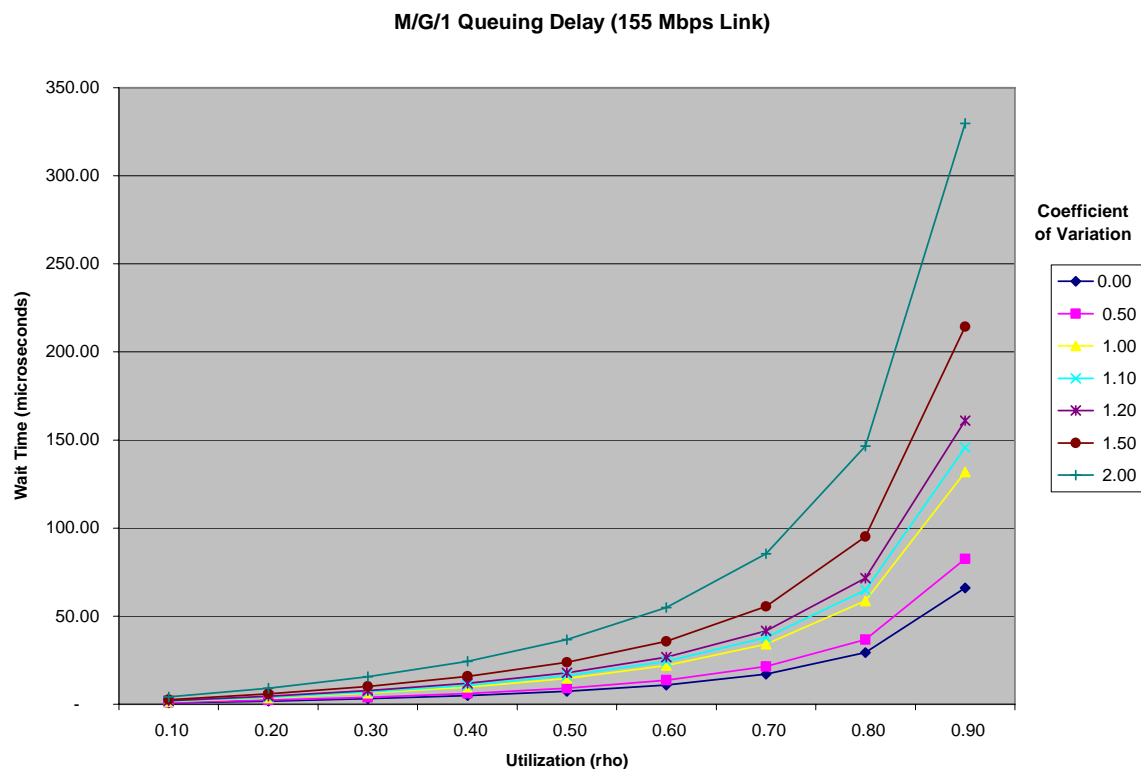
Two way real time voice is, however, much more demanding. Many tests over the years have found that, where delay exceeds about 150 milliseconds, people on both sides of the conversation are likely to start speaking at once (because neither knows that the other has already begun to speak). This phenomenon is familiar to those of us who have been around long enough to have used satellite circuits for telephone calls across the Atlantic. One can still conduct a conversation, to be sure, but nobody would prefer such a conversation to one conducted with low delay.

### **2.2.2 Delay in packet networks**

It is natural to begin by asking the degree to which normal Internet traffic would meet demands for delay-sensitive traffic. The answer is, quite simply, that normal Internet traffic performs well enough nearly all of the time. This is the reason why services such as Skype and Vonage have customers – they work well enough today over the public Internet, with no special provisions taken to ensure quality of service.

This follows from basic queuing theory. Queuing theory is the branch of mathematics that deals with waiting lines – in this case, the waiting line to place a packet onto a high speed link between two routers in the core of the Internet.

The standard formulae for variable delay depend on how busy the transmission link is, how fast it is, how big the packets are and how variable in size. They result in a family of curves, as shown in the figure below, where delay increases as the link becomes increasingly busy (moving toward the right of the graph). These curves reflect a 155 Mbps link, which is the *slowest* link that one is likely to find in the core of a high speed NGN. Variability in the length of the packets is reflected in the coefficient of variation of the packet length (and thus the service time), where a coefficient of variation of 0.0 denotes no variability at all, and 1.0 reflects a nominal degree of variability corresponding to exponentially distributed packet lengths. Some years back, when I designed networks for a living, I found a coefficient of variation of 1.2 to be typical for the Internet traffic of the day. It is thus immediately obvious that, even at exceedingly high loads of 90%, expected variable delay is less than 150 *microseconds*. Given that our “budget” is in the range of 150 *milliseconds*, it is no surprise that IP traffic performs adequately most of the time – we could afford about 1,000 router hops, even under heavy load (as long as no link is truly saturated with traffic).



This does not mean that measures to better manage traffic are irrelevant. First, there is the risk that some link is completely saturated with traffic – a risk that cannot be completely avoided in light of the bursty nature of Internet traffic.<sup>5</sup> Second, there are

<sup>5</sup> There has been a long and tiresome debate in the literature about the degree to which Internet traffic is *self-similar*. Suffice it to say that traffic patterns are not perfectly random, and that in consequence the risk of an occasionally saturated link is a bit higher than the standard equations would lead one to

the slower circuits at the edge of the network, especially shared media such as broadband based on cable television. And third, there is the matter of operation when the network is operation in a partially degraded mode due to the failure of some but not all components.

So the ability to provide QoS in a network has its place in a modern IP-based network. Still, the willingness of customers to pay a premium for a service that, most of the time and under most circumstances, provides no customer-visible benefits will surely be limited. This limited willingness to pay on the part of the customer naturally leads to limited willingness on the part of network operators to invest in making the capabilities available.

### 2.2.3 Technical solutions

In the early Nineties, the IETF was active in evolving a series of relatively complex solutions under the rubric of the *Integrated Services Architecture*, as exemplified by the RSVP protocol. The common wisdom has been that these protocols were hopelessly complex.

In fact, my company (BBN) had a working network based on RSVP. It was delivering services to real customers. It was a technical success but a commercial failure. It was eventually shut down, not due to technical problems, but rather because the company never found enough customers who were willing to pay much of a premium to use the RSVP-capable network.

Be that as it may, the IETF subsequently evolved a much simpler set of communications protocols in conjunction with *DiffServ* (for *Differentiated Services*). DiffServ enables hop-by-hop traffic management, where selected packets are marked as having requirements other than best efforts. It is up to each router, then, to do what it can to implement the desired transmission characteristics (or to decline to do so). Various techniques can then be used to ensure hop-by-hop performance, with Multi-Protocol Label Switching (MPLS) being perhaps the most common.

DiffServ provides a more limited service than RSVP in the sense that it assures performance only on a hop-by-hop basis, rather than end-to-end. Still, it can provide adequate overall assurance at a statistical level.

DiffServ and MPLS are trivial to implement within a network, and are in use in many large networks today. Nonetheless, there is no significant use *between* networks. My contention is that the technology has long been basically mature, and that the lack of deployment reflects economic and business factors rather than technical ones.

### 2.2.4 Implications for billing and accounting systems

At the same time, the implications for Operational Support Systems (OSS) in support of differentiated QoS tend to be overlooked in most discussions. Technologists tend to focus more on the problem of getting the bits to move as they are supposed to move, and less on the problem of how to ensure that someone pays for those movements.

---

expect. The network design consequently needs to leave slightly more “headroom” than the basic mathematics would otherwise imply.

It has generally been assumed that a network operator would be willing to provide better-than-best-efforts quality only to the extent that either the end user or another network operator was willing to pay them a premium to do so. To the extent that this implies the need to account for QoS-capable traffic, it implies surprising complexity.

First, a pair of network operators would need to agree on how much QoS-relevant traffic each delivered to the other. Second, they would need to verify that each actually delivered the quality that it had committed to the other. Finally, each would need some tools to deter fraudulent use or “gaming” of the system. The first is trivial, the second and third are difficult. Finally, there would be the need to reconcile statistics, and to deal with discrepancies between the measurements of the parties.

Measuring traffic across a link would seem to be straightforward, and distinguishing among different marked classes of traffic is no harder. Capturing first-order statistics on traffic sent between the parties is straightforward. Even here, some prior agreements will be needed as regards what is being measured, and when – otherwise, there is the risk that network A has a slightly different view of the traffic delivered on the link from A to B than does network B, even though both are measuring (different ends of) the same link using substantially similar tools. And sampling intervals need to be mutually agreed, otherwise any measures of variability (quantiles, standard deviation) are likely to reach different conclusions due to the perverse effects of the Central Limit Theorem (if two sensors sample the same distribution, the one that is sampling at more frequent intervals will tend to see an apparently more lumpy distribution).

Reconciling data would be challenging. There is an old Dutch proverb: “Never go to sea with two compasses. Take one or three.” If the providers do not agree, whose statistics should govern? Is there scope here for a trusted intermediary, and if so who might that trusted third party be?

The challenges in verifying that the service was actually delivered are much more profound. In this case, network A needs to ensure that network B delivered the committed performance, and neither will want to rely on measurements provided by their respective end users. Network A thus needs performance statistics about network B’s network. At the same time, these networks are likely to be direct competitors for the same end users – network B is not about to let network A place sensors in its network. Both networks are likely to be skittish about providing internal performance data to one another.

It might be far simpler to bill, not for the use of the network, but rather for the services that benefit from differentiated QoS. Here, too, there are challenges – in an IP-based NGN, the service provide might not be the network provider. Moreover, it is quite possible, for reasons noted earlier, that services without QoS will compete successfully with services that are supported by QoS. It is not clear that network operators would be able to extract enough revenue from independent service providers to enable them to fund the differentiated services.



## 2.3 *Economic background*

The economic theory of network interconnection has been extensively analyzed by Armstrong (1998) and in a series of papers by Laffont, Rey and Tirole (1998). For a comprehensive treatment, see Laffont and Tirole (2000).<sup>6</sup> We provide only the briefest summary here, since the theory is well established.<sup>7</sup> Section 2.3.1 deals with economic analysis of the traditional telephone network, while section 2.3.2 deals with the equivalent theory for the Internet.

### 2.3.1 General theory of telephone network interconnection

In the conventional telephone world, *retail* plans have typically been based on a system known as *calling party pays* (CPP), in which the party that places or originates or places a call pays for the call based on the number of minutes of use, while the party that receives the call generally pays nothing. This model reflects the tacit assumption that the party that placed the call is in some sense the *cost causer*. In recent years, there have been revisionist challenges to this view – if the party receiving the call did not perceive value, he or she would simply hang up (the principle of receiver sovereignty).<sup>8</sup> So the newer view argues that both parties benefit, that “... it takes two to tango”.

At *wholesale* level, *calling party's network pays* (CPNP) is the usual counterpart to CPP. Since there is no retail payment from the party receiving the call, the receiving party's network should be compensated by the calling party's network. Thus, a wholesale payment flows from the originating party's network to the terminating party's network.

What is known about these wholesale payments is that they tend to be set at rates much higher than would be the case under full competition. Once someone subscribes to a network, that network effectively derives market power (the *termination monopoly*) over the termination of calls, because the call originator has no ability to choose the terminating network. In Europe, these rates are generally regulated and are floating downward today, but are still quite high.

In North America, a completely different system (“bill and keep”) is in place. At the wholesale level, many network operators are under no obligation to make wholesale payments to one another; however, any mutually agreed payment rate must be symmetric. Most network operators set the rate at zero (i.e. they waive payments). Under bill and keep, network operators have great flexibility about how to set retail rates, but competitive forces have motivated most operators to adopt plans that are flat rate (or that are flat rate up to some maximum number of minutes – a “buckets of minutes” plan).

---

<sup>6</sup> *Competition in Telecommunications*, MIT Press.

<sup>7</sup> For a somewhat more comprehensive overview, see Marcus, “Interconnection in an NGN Environment”, a background paper commissioned for the ITU New Initiatives Programme workshop on “What rules for IP-enabled Next Generation Networks?” held on 23-24 March 2006 at ITU Headquarters, Geneva. Available at: <http://www.itu.int/osg/spu/ngn/documents/Papers/Marcus-060323-Fin-v2.1.pdf>. Also available as WIK Discussion Paper 274 (see [http://www.wik.org/content\\_e/diskus/274.htm](http://www.wik.org/content_e/diskus/274.htm)).

<sup>8</sup> Jeon, Laffont and Tirole (2004). See also Hermalin and Katz ().

If traffic and rates are both symmetric, one might imagine that network operators would not care about them since they net to zero. As Laffont and Tirole perceptively observed, this is not quite right – the rate matters because it is perceived as part of the wholesale marginal cost of making a call, and therefore tends to be reflected in retail rates. It is for this reason that flat rate plans in European countries usually exclude calls to mobile phones (where termination rates tend to be high).

Bill and keep countries tend to experience somewhat slower rates of take-up of mobile phones than CPNP countries, but much higher usage of the phones once they have been taken up. I would argue that bill and keep achieves clearly superior results for developed countries that have already achieved widespread adoption of phones.

### 2.3.2 Economic theory of IP-based interconnection

An economic literature on the economic theory of Internet interconnection has emerged in the last few years, largely inspired by a number of large merger cases in the late Nineties.

In Laffont, Marcus, Rey and Tirole (2003),<sup>9</sup> we analyzed the interconnection of two backbones subject to an access charge. We were specifically interested in understanding the possible implications of perturbing the system to incorporate multiple classes of service. The model led to straightforward results – access fees could be higher for classes of service where the end user perceived a greater economic surplus. Today’s Internet is a special case of this model, where access fees happen to be zero.

The level of the access charge is largely neutral overall; however, a low rate tends to favor content providers (e.g. websites), while a high rate might favor consumers. It is not immediately clear which outcome is socially preferable, inasmuch as the Internet can be viewed as a two-sided market<sup>10</sup> where both of the market must be present. The proliferation of websites has arguably made it attractive for more consumers to acquire Internet access, thus sustaining a virtuous cycle of increasing content and increasing subscribership.

There has been a tendency to think of Internet interconnection as being totally distinct from interconnection in the Public Switched Telephone Network, but in fact the two are closely linked. Tirole has observed that the differences between the two economic models flow from the “missing price”, that is, from the fact that the recipient of a telephone call (under CPP) is not charged.

---

<sup>9</sup> With Jean-Jacques Laffont, Patrick Rey, and Jean Tirole, IDE-I, Toulouse, “Internet interconnection and the off-net-cost pricing principle”, *RAND Journal of Economics*, Vol. 34, No. 2, Summer 2003, available at <http://www.rje.org/abstracts/abstracts/2003/rje.sum03.Laffont.pdf>. An earlier version of the paper appeared as “Internet Peering”, *American Economics Review*, Volume 91, Number 2, May 2001.

<sup>10</sup> Cf. Rochet and Tirole (2004).

### 3. Price discrimination: A good thing, or a bad thing?

It has long been recognized that providers of goods or services could potentially achieve some pricing power and profitability by distinguishing their goods and services, and by offering different qualities at different prices to different groups of customers.<sup>11</sup> When we buy a ticket for a train or an airplane, we take it for granted that we may be offered first class and second class tickets, with a higher price for the former.

In some cases, price discrimination may be linked solely to the willingness of the customer to pay, and largely unrelated to underlying costs. When an airline offers cheaper tickets to passengers who are willing to stay overnight on Saturday, it has nothing to do with their costs; rather, it reflects the greater willingness to pay (lower elasticity of demand) of business travelers. Business travelers are able to pay more, but are in most cases unwilling to stay overnight outside of the Monday to Friday time frame.

In the absence of market power, this kind of price discrimination tends to enhance consumer welfare. Deregulation of the airline industry, and emergence of price discrimination, is generally acknowledged as having resulted in lower prices for consumers.<sup>12</sup>

In the United States, a recent debate has emerged over *network neutrality*. The arguments on both sides of this complex debate have been somewhat misplaced, in my judgment, but it is worth noting that a number of experts have implicitly objected to price discrimination and to the use of technology to support the excludability that would make price discrimination effective.

I have argued elsewhere<sup>13</sup> that the network neutrality debate emerged in the United States due to a “perfect storm” of three simultaneous market and regulatory changes:

1. The collapse of the wholesale market for broadband Internet access;
2. A series of mergers (Cingular/AT&T Wireless, SBC/AT&T, Verizon/MCI, and now AT&T/BellSouth) with insufficient conditions imposed; and
3. The overly hasty and ill-considered withdrawal of procompetitive regulation.

With that said, my feeling is that these concerns about price discrimination reflect excessive concentration in the U.S. market – regulatory experts are objecting to many practices that, in a healthy market, would be welfare-enhancing. In the U.S. context, these concerns are real; moreover, they cannot easily be fixed through regulation. The problems are too complex. The FCC has already demonstrated (first in the Madison

---

<sup>11</sup> See Hotelling (1929).

<sup>12</sup> To be sure, it also led to many unanticipated effects and to some unanticipated problems.

<sup>13</sup> “Is the U.S. Dancing to a Different Drummer?”, Communications & Strategies, no. 60, 4th quarter 2005. Available at: [http://www.idate.fr/fic/revue\\_telech/132/CS60%20MARCUS.pdf](http://www.idate.fr/fic/revue_telech/132/CS60%20MARCUS.pdf). Also available in intermedia (the journal of the International Institute of Communications), vol. 34, no.3, July/August 2006.

River proceeding, and again in the “Broadband Policy Statement”) that they lack the necessary expertise to distinguish between welfare-enhancing service discrimination versus harmful anticompetitive acts. In any case, once markets have been allowed to deteriorate to this degree, no regulatory fix is likely to be satisfactory. The fox is already in the chicken house, the horse has already left the barn.

In Europe, by contrast, the underlying markets are much more competitive; moreover, the regulatory system in Europe is likely to ensure that they remain competitive. Even in relatively concentrated markets such as Germany, most consumers can choose among multiple broadband service providers (many of them service-based rather than facilities-based). For the most part, the network neutrality debate has not emerged in Europe, and it is unlikely to emerge in the same form in which it has in the United States.

There are, of course, possible risks going forward. For example, if incumbent operators were to use differentiated Quality of Service to block independent offers of VoIP services, it seems to me that such conduct would raise serious concerns. But it is not clear that such a strategy would be effective, and European regulators probably have sufficient tools to deal with that kind of abuse were it to emerge.

In Europe, as long as regulators continue to ensure competitive underlying markets, offers of different quality of service at different prices is likely to enhance consumer welfare rather than to detract from it.

## 4. Glacial adoption of differentiated QoS between networks

With the foregoing background, we can now return to a key question: Given that the technology of differentiated QoS is not particularly challenging, and given its widespread use *within* networks, why has it been so slow to achieve deployment *between and among* networks?

From an economic perspective, the basic answer is obvious: Had the benefits of deployment clearly exceeded the costs, it would have deployed. Thus, one might infer that either the perceived costs are too high, or the perceived benefits too low, or perhaps both.

For reasons noted in section 2 of this paper, there are indeed questions as to whether the perceived benefits are too low. In addition, a series of challenges related to *network externalities* and to *transaction costs* have inhibited deployment.

Many industries experience network externalities. A service may be most useful when a great many people use it (and not just because of economies of scale). This is true of telephone service, and also of the Internet. My telephone is worth more if there are a great many people whom I can call. My Internet connection is worth more when there are a great many people to whom I can send an email, and a great many websites to which I can connect.

Getting a new service launched in a sector dominated by network externalities can be challenging. In effect, the externalities of the old service keep pulling you back. It is difficult to get past the *initial adoption hump* in order to achieve critical mass. The economist Geoff Rohlfs<sup>14</sup> has explained that different services got past the initial adoption hump in different ways. VCRs were initially purchased for time-shifting of television programs; only when enough consumers had purchased VCRs did a rental business emerge. CDs were successful because Matsushita and Phillips had commercial interests in both CD players and studios, and were thus motivated to ensure that both players and content were available.

Differentiated QoS *between and among networks* is subject to these network effects. The service has some value within a network. It might have great value if it were available to every destination on the Internet. If it were available to only two or three networks, then it is of limited value. Thus, the value of deployment might be significant to those networks that implement it later, but the initial benefit to the first two or three networks to deploy it is minimal.

At the same time, extending differentiated QoS to each additional network implies *transaction costs*. Agreements, monitoring tools, and coordination in general would need to be put in place. These costs might be roughly linear in the number of networks with which one network has agreements in place.

Thus, it is hard to get the process started, and it would be hard to get it to completion once it had been launched.

These concerns are not unique to differentiated QoS. A number of Internet capabilities are faced with similar economic challenges. The adoption of Internet Version 6 (IPv6, a new version of the Internet Protocol with a greatly expanded

---

<sup>14</sup> See *Bandwagon Effects in Technology Enterprise* (2003).

address range) and of DNSSEC (a security enhancement to the Domain Name System) have arguably been impacted by similar considerations.<sup>15</sup>

---

<sup>15</sup> See Marcus, “Evolving Core Capabilities of the Internet”, *Journal on Telecommunications and High Technology Law*, 2004, available at: [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=921903](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=921903).

## 5. Public policy implications

All of this raises thorny questions for policymakers. Will differentiated QoS emerge on its own, without “help”? Suppose the answer were no – is that a basis for government response, or is it merely a message from the marketplace that this service is not really needed?

To be sure, it raises complex questions for network operators as well. Is it better to wait until some combination of firms demonstrates a viable and working business model for differentiated QoS, or is it better to bull ahead, implement, try to make things work, and seek first mover advantages?

For public policymakers, it seems to me that intervention is inappropriate, or at least premature. The migration to NGN is still in its infancy – it is too soon to say what new market forces it might unleash. There is a good basis for skepticism, but it is too soon to firmly judge whether differentiated QoS will deploy.

Second, it is not clear that this is the kind of capability for which intervention is appropriate. We normally look to the market to signal which services are supported by adequate demand, and which are not. It is not clear that a failure to deploy would reflect a market failure. It might just be that demand for the capability is slack because it fails to meet any otherwise unfulfilled, compelling need.

Intervention is always risky. Returning to equine analogies, there is the risk of “betting on the wrong horse” – policymakers might inadvertently force adoption of a standard that is worthless or even harmful.

With all of that said, one action that might conceivably smooth the road to adoption would be the creation of some entity that could serve as a trusted third party to account for traffic (at various levels of QoS requested) between networks, including traffic statistics if desired. This is perhaps a candidate for an industry initiative rather than a government initiative.

The one clear recommendation to European policymakers that emerges at this time is to continue to endeavor to ensure that European markets remain competitive, especially the market for residential broadband. The kind of concentration that we have seen in the U.S. market, and the resultant increase in concerns about service quality differentiation, can be and should be avoided here in Europe.