

# How Deployment of NGN is Changing Costing and Charging Principles

J. Scott Marcus: Department Manager, NGN and Internet Economics

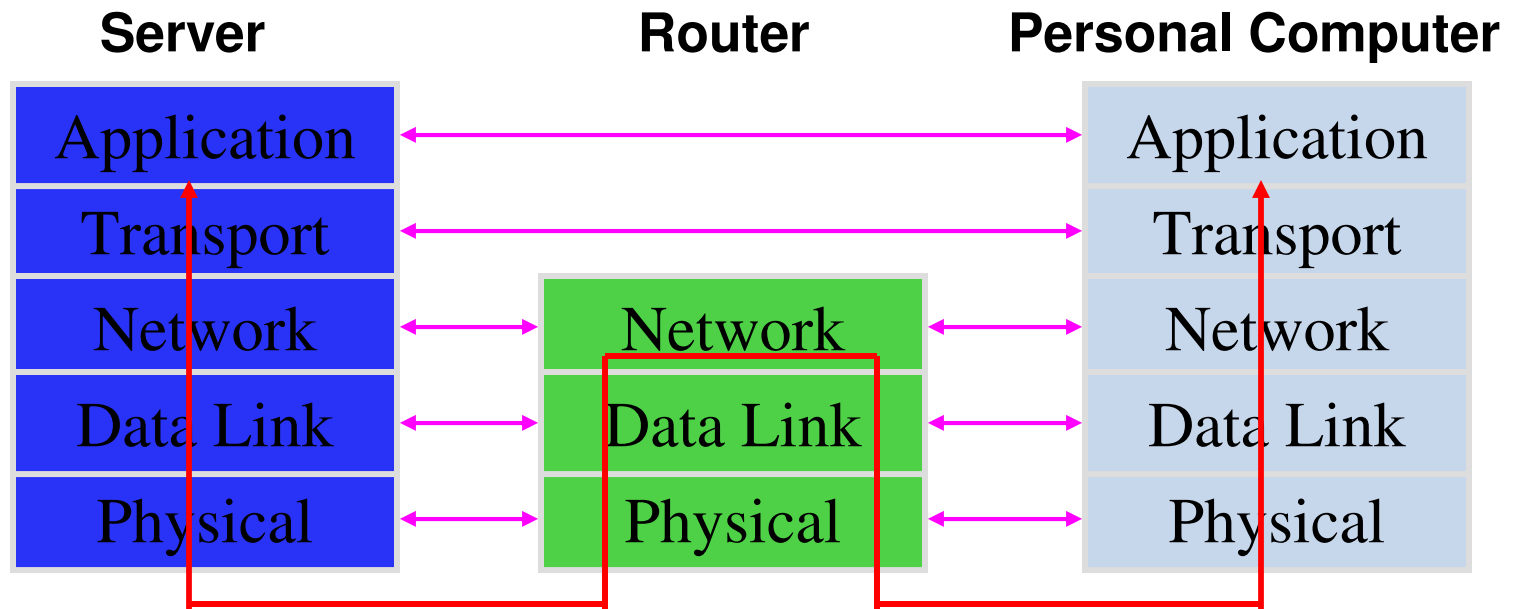
ITU, Geneva, 8 September 2008

# Implementing the NGN: Regulatory and Economic Aspects

- The evolution to an IP-based NGN
  - Key functional attributes of an NGN
  - From a policymaker's perspective, what changes with the migration to an IP-based NGN?
- Implications of the migration to NGN for interconnection
  - Wholesale costs and prices
  - Retail costs and prices
  - Application interconnection
  - QoS and network neutrality
- Conclusions

# The TCP/IP Reference Model

Layers interact with peer layers

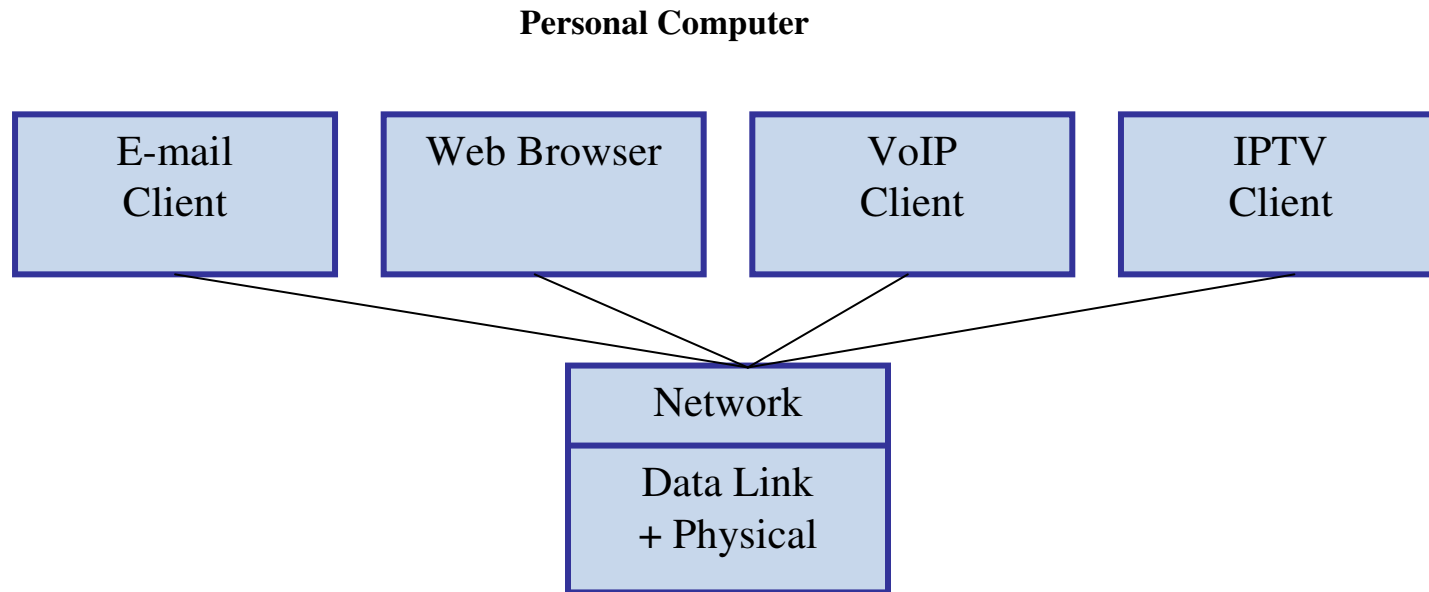


Layers derive services from successively lower layers

## The TCP/IP Reference Model

- Physical Layer – the transmission facilities.
- Data Link Layer – the logical management of physical transmission facilities.
- Network Layer – forwarding and routing (Internet Protocol, or IP).
- Transport Layer – provides applications with datagram (UDP) or virtual circuit (TCP) services, as needed.
- Application Layer – provides services to the user.
  - Web
  - Email
  - VoIP
  - IPTV
  - Peer-to-peer file sharing

# Implementing the NGN: Regulatory and Economic Aspects



Today, your broadband connection can support any combination of voice, video and data – provided that it is fast enough.

## Implementing the NGN: Regulatory and Economic Aspects

- How does the migration to NGN change things:
  - For the network operator?
  - In terms of public policy and regulation?

## Implementing the NGN: Regulatory and Economic Aspects

- There are three primary reasons for regulation of electronic communications, all related to market failure:
  - Addressing distortions of competition, especially those caused by some form of market power.
  - Addressing social needs that the free market might not, typically because the social value exceeds the private value to parties that might otherwise invest.
  - Allocating scarce resources that are unique to each country.
- The move to NGN raises issues in all three areas.

## Implementing the NGN: Regulatory and Economic Aspects: Key Changes

- In the value chain by which services are delivered to end-users.
- In the ability of different service providers (not all of whom are network operators) to compete with one another for the same services.  
Distinctions of cable versus telecommunications, fixed versus mobile, wired versus wireless all become less relevant.
- In the speed and the character of network access, and thus in the ability to offer bandwidth-hungry services (e.g. video).
- In the ease with which certain public needs can be satisfied.
  - Access to emergency services.
  - Lawful intercept.
- In the ability of end-users (or software developers on their behalf) to create new capabilities in the end-user's device (PC), often without the active involvement of the network operator.



## Implementing the NGN: Regulatory and Economic Aspects

- The use of Internet Protocol enables any application, or combination of applications, to be delivered over “any” transmission medium (or over a path containing any combination of transmission media).
  - Data services
  - Voice (VoIP)
  - Video (IPTV)
- This decouples, for the first time, the *service* from the *network*.
- The decoupling is not absolute. The transmission medium must provide suitable characteristics in terms of performance, loss, and other aspects of Quality of Service.
- Some transmission media are particularly appropriate for certain jobs.
  - Cable and PON are especially well suited to delivering the same video streams simultaneously to multiple destinations (point-to-multipoint).
  - Mobile solutions may be less suitable than wired for high bandwidth.

## Implementing the NGN: Regulatory and Economic Aspects

- *Bundling*: a single network operator / service provider can offer voice, video and data to their customers (*triple play*).
  - Economies of scope for the service provider.
  - These economies are typically reflected in discounts to the consumer.
  - Higher effective switching costs for the consumer.
- *Independent third party service providers*: service providers (e.g. SIPgate, Skype and Vonage for VoIP) offer a service over the customer's broadband service.
  - In the distant past, the telephone network operator provided the voice service.
  - Today, an independent service provider can offer voice or video services without operating a network of its own.

## Implementing the NGN: Regulatory and Economic Aspects

- This transition signals large changes in the character of competition for end-user customers going forward.
  - One model: telecommunication network operators compete with cable operators for the complete triple play business of the end-user. This could be said to be the primary U.S. model.
  - Another model: third party service providers (without their own networks) compete with network operators for perhaps one or two services by offering price-performance advantages or superior features or service.

## Implementing the NGN: Regulatory and Economic Aspects

- The ability of third party service providers to deliver services (e.g. voice, video) to the user is critically dependent on the underlying network. The service is decoupled from the network, but the markets are *vertically related*.
- In Europe, these markets are likely to be reasonably competitive. In some other parts of the world, last mile network operators often possess market power that is not fully addressed by regulation.
- Where a firm operates in two vertically related markets and has market power in only one of them, the firm may be motivated to project its market power into the market that would otherwise be competitive. Economists refer to this as *foreclosure*.
- It is concerns over possible foreclosure that motivate a debate in the U.S. over *network neutrality*.

## Implementing the NGN: Regulatory and Economic Aspects

- The introduction of NGN has complex implications for the development of competition. Will market power increase, remain constant, or decrease overall?
  - **Decrease:** IP-based NGNs will introduce new forms of competition, by enabling new last mile access options and decoupling network provision from competition for services.
  - **Remain constant:** Unfortunately, the move to IP does not necessarily eliminate traditional market power, including:
    - Last mile market power
    - Call termination monopoly
  - **Increase:** The migration might, moreover, enable the emergence of new competitive bottlenecks, especially at the upper layers of the network.
- Which effects will predominate?
- How should policymakers respond?

## Implementing the NGN: Regulatory and Economic Aspects

- The evolution to an IP-based NGN
  - Key functional attributes of an NGN
  - From a policymaker's perspective, what changes with the migration to an IP-based NGN?
- Implications of the migration to NGN for interconnection
  - Wholesale costs and prices
  - Retail costs and prices
  - Application interconnection
  - QoS and network neutrality
- Conclusions

## Implications for Interconnection

- Migration from switched fixed/mobile networks to NGN “marries” the traditional telephone network with the Internet. Which rules prevail?
- Switched fixed and mobile networks – regulated arrangements.
  - Regulation to address market power.
  - Termination fees in the absence of regulation will tend to be very high, for both large and small operators.
  - Lack of interconnection implies a connectivity breakdown.
- Internet – peering arrangements are typically “Coasean”.
  - Peering: two providers exchange traffic only for their respective customers, often (but not always) with no explicit charges.
  - Sharing of facilities costs for interconnection may be unequal.
  - In most countries, no regulation of peering.
  - Lack of interconnection usually does not imply a loss of connectivity, but may have implications for costs.

# Implications for Interconnection

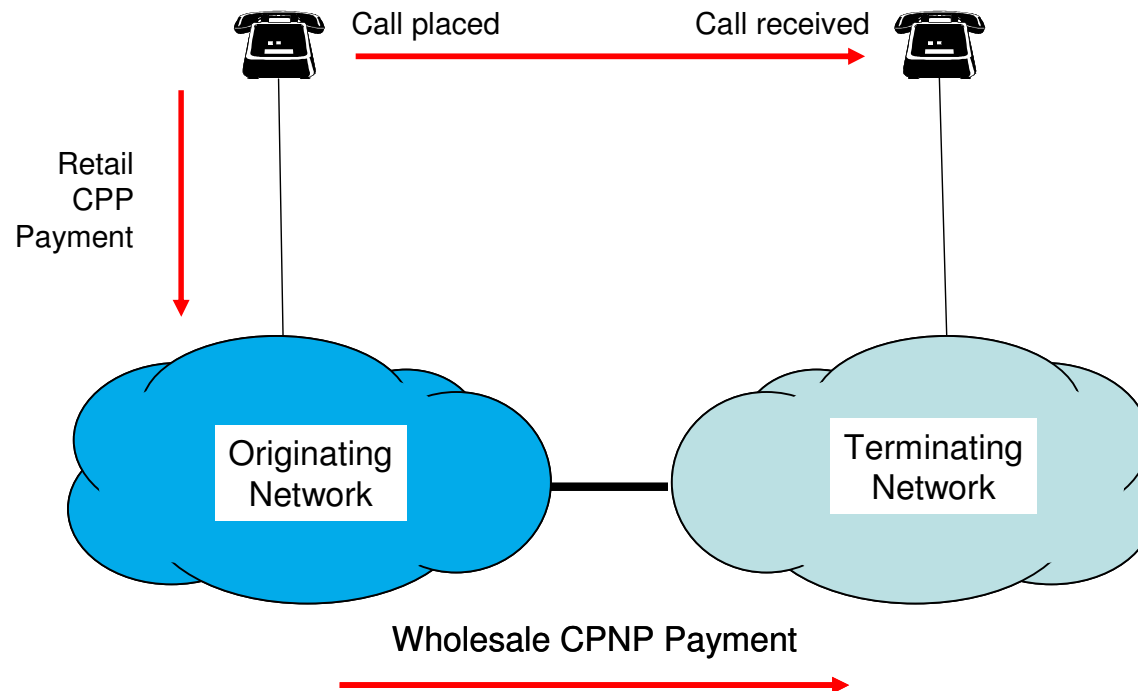
## Traditional Fixed and Mobile Interconnection Models

- Bulk of revenues comes from voice telephony
  - However, voice represents a sharply declining percentage of traffic
- Retail arrangements
  - Calling Party Pays (CPP)
    - Traditional arrangement based on presumed cost causality and presumed internalization of call externalities
  - Receiving (Mobile) Party Pays (RPP/MPP)
    - Shared utilities from calls, receiver sovereignty
    - True RPP systems are rare today.
  - Flat rates: Calls included in monthly fees (bandwidth)
  - Banded flat rates (buckets of minutes): “banded” flat rate



# Implications for Interconnection Traditional Fixed and Mobile Interconnection Models

- Calling Party's Network Pays (CPNP) wholesale arrangements



- An alternative (US and a few other countries) is to have negotiated arrangements, often resulting in no charges (Bill and Keep).

# Implications for Interconnection

## Internet Interconnection Models

- **Peering**

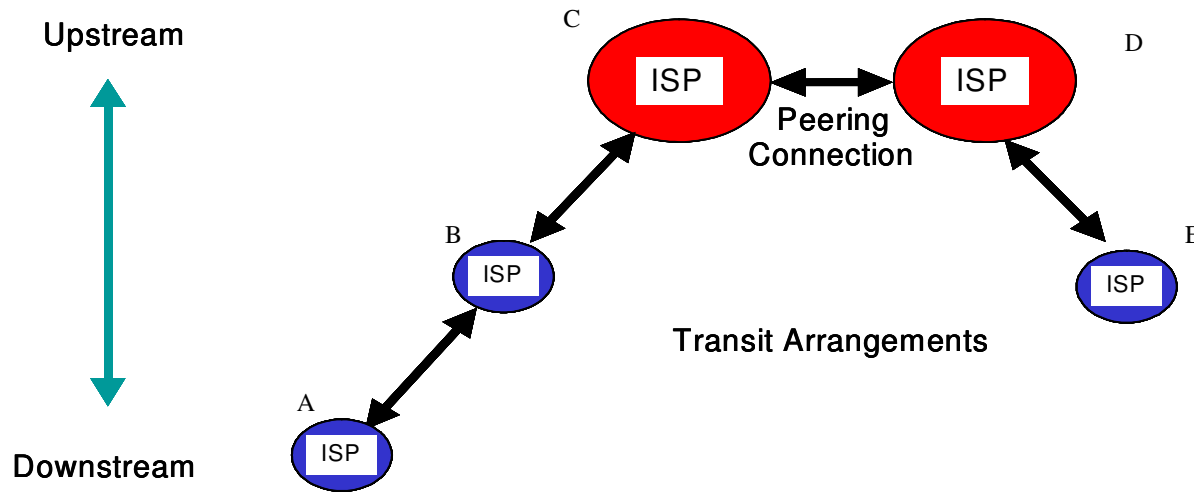
- Definition: a bilateral agreement between ISPs to carry traffic for each other and for their respective customers. Peering does not include the obligation to carry traffic to third parties.
- Symmetric
- Often “bill and keep”

- **Transit**

- Definition: The transit provider (the ISP) carries traffic for the transit customer, but the transit customer is not under any obligation to carry traffic for the transit provider. The transit customer pays the provider, not the other way around.
- Asymmetric
- Transit customer pays
- Typically connects to entire Internet

# Implications for Interconnection Internet Interconnection Models

Two peers and their respective transit customers



<i><u>Parties</u></i>	<i><u>Interconnection Arrangement</u></i>	<i><u>Typical Nature of Agreement</u></i>	<i><u>Typical Commercial Arrangements</u></i>
A – B	Transit	Bilateral	Payment reflects capacity, and may reflect volume of traffic or near-peak traffic level.
B – C	Transit	Bilateral	
E – D	Transit	Bilateral	
C – D	Peering	Bilateral	Often done without payment

## Implications for Interconnection Traditional Fixed and Mobile Interconnection Models

- Conditions for efficient (“Coasean”) negotiated arrangements for voice have included (1) reciprocity and (2) capped x2F rates.

Table 1: Coasean arrangements: Backbone peering versus U.S. Bill and Keep

	<u><i>U.S. Mobile Operators</i></u>	<u><i>Backbone ISPs</i></u>
<b>Obligation to Interconnect</b>	Applies to all carriers.	No regulatory obligations in most countries. Some backbone ISPs adhere voluntarily to guidelines or to principles of non-discrimination.
<b>Constraints on fees charged</b>	Must by law be reciprocal (equal in both directions).	Generally unconstrained.

Source: wik-Consult

## Implications for Interconnection

### How things work today: Wholesale and retail

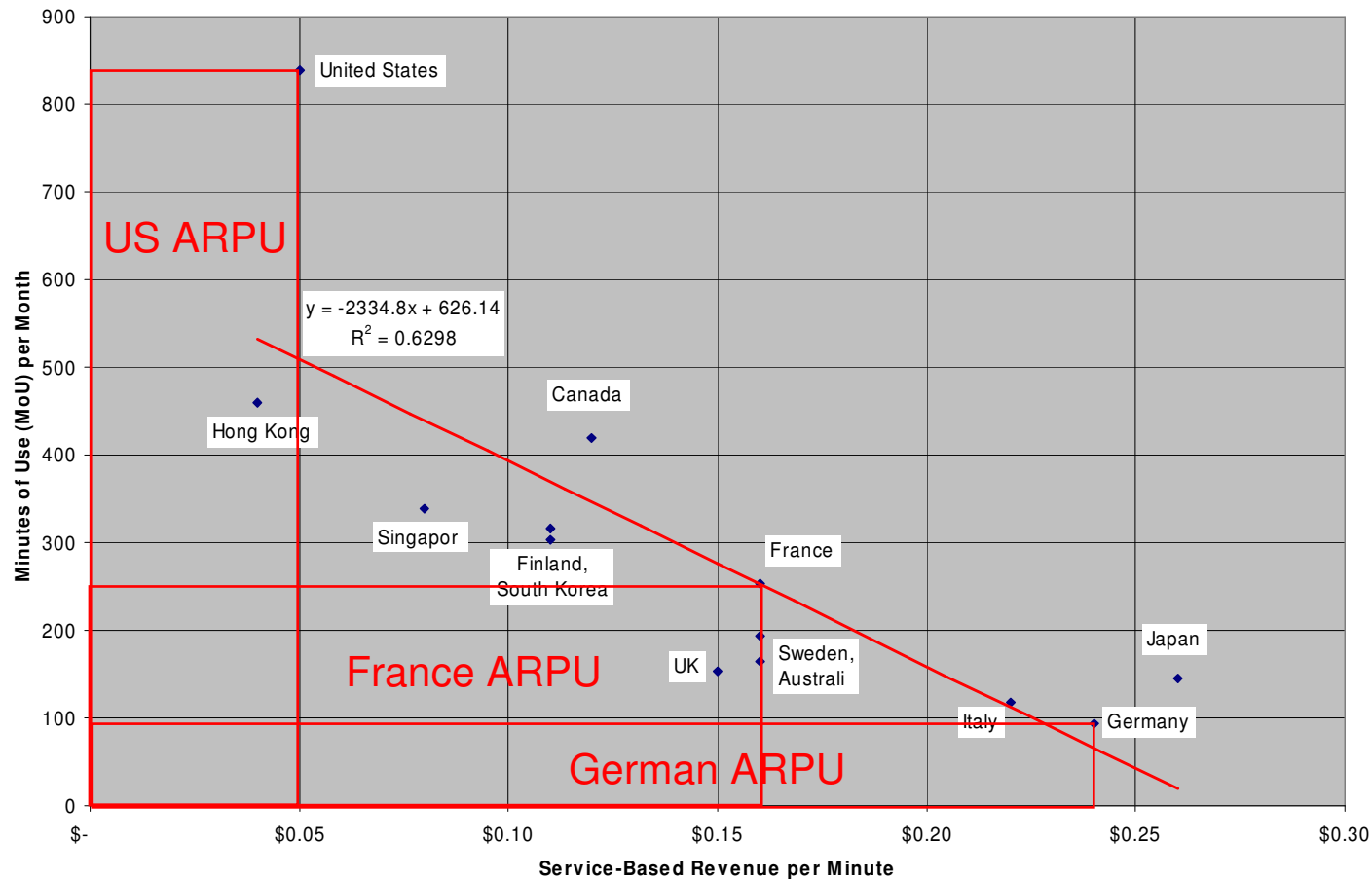
- In an unregulated CPNP system, carriers will tend to establish very high termination charge levels (the *termination monopoly*).
  - Smaller operators will be motivated to set termination fees even higher than large operators.
  - The problem is addressed in the EU by regulating all rates.
- Several factors contribute to the termination monopoly.
  - Since the charges are ultimately borne by *another operator's customers*, normal market forces do not adequately constrain them.
  - Customers have little visibility into termination fees.
- Termination charges at the wholesale level have some interaction with retail pricing arrangements.
  - The termination fee generally sets a floor on the retail price.
  - Where termination fees are high, they generally limit the applicability of flat rate or “buckets of minutes” plans.

# Implications for Interconnection

## How things work today: Wholesale and retail

### Mobile Usage Responds to Unit Price

Merrill-Lynch Data 4Q2006, as quoted by the U.S. FCC 12th CMRS Report



Note: ML does not count on-net mobile in CPNP countries thus under-counting about 20%.

## Implications for Interconnection

### How things work today: Wholesale and retail

- CPNP with high mobile termination rates tends to lead to:
  - Subsidies for mobile adoption, and thus rapid penetration.
  - High retail prices.
  - Exclusion of calls with high termination from flat rate plans.
  - Low usage.
- Rapid penetration is beneficial (but whether it needs stimulus depends on national circumstances); the other aspects are harmful.
- There is only limited interaction between wholesale and retail arrangements.
  - CPNP with high termination rates practically mandates CPP retail arrangements.
  - Low or zero termination rates, however, place no constraints on retail arrangements.

## Implications for Interconnection

### Appropriate termination rates

- Optimal rates could depend in complicated ways on
  - Costs
  - Network externalities
  - Call externalities
  - Calling habits (induced calling)
  - Institutional factors (CPP, RPP traditions; reciprocity)
  - Customer heterogeneity: Targeting of customers (e.g., ISPs)
  - Very hard to determine and will depend on circumstances
  - Very likely range:  $0 \leq p^* \ll \text{LRAIC}$
- Prof. Littlechild describes determination of an optimal termination rate as a Sisyphean task.
- Even if one cannot say that a particular rate is *right*, perhaps it is possible to identify rates that are *wrong*?



## Implications for Interconnection

### Appropriate termination rates

- Mobile termination rates vary from one European Member State to another, but the European average of € 0.0967 per minute is likely well in excess of real average incremental cost.
  - Large differences between on-net and off-net pricing suggest that operators view their costs as being less than the termination fee.
  - Service-based revenue per MoU in the U.S. is about € 0.03 per minute. Cost is presumably less than price.
  - Cyprus prices termination at some € 0.02 per minute.
  - A WIK study found average incremental cost in Australia to be between € 0.03 and € 0.04 per minute, depending on share.
- A few years ago, India implemented termination rates for both fixed and mobile of about \$0.005 (half of a U.S. cent). They seem to have achieved excellent results, with low retail prices, high usage, and rapid penetration.

## Implications for Interconnection Challenges implied by NGN

- Network costs are driven by capacity requirements. In future integrated IP-based networks, where voice may represent only a small fraction of the traffic, total costs might have little to do with minutes of voice use.
- Traditional interconnection arrangements historically represented an attempt to use *wholesale* payments (between network operators) to correct for imbalanced *retail* payments (between service providers).
- To the extent that the network and service providers are different firms, and to the extent that voice is only a small fraction of the cost of the network, this system makes even less sense going forward than it did in the past.

## Implementing the NGN: Regulatory and Economic Aspects

- The evolution to an IP-based NGN
  - Key functional attributes of an NGN
  - From a policymaker's perspective, what changes with the migration to an IP-based NGN?
- Implications of the migration to NGN for interconnection
  - Wholesale costs and prices
  - Retail costs and prices
  - Application interconnection
  - QoS and network neutrality
- Conclusions

## Implications for Application Interconnection Voice Interconnection

- Does IP interconnection equate to voice interconnection?
- Different answers are emerging for IP-based service providers, fixed incumbents, and mobile operators.
  - Independent VoIP service providers, cable operators who offer VoIP: Interconnection arrangements based on VoIP peering are emerging. This requires the ability to determine which service provider serves which telephone numbers (e.g. by means of carrier ENUM).
  - Fixed PSTN: As networks migrate to IP-based NGNs, IP peering does not automatically imply the ability to use VoIP to connect to the fixed incumbent's voice services.
  - Mobile PLMN: The GSM-A already provides an IP backbone for interconnecting mobile operators, the GRX/IPX. All indications are that connectivity to the mobile operator's inherent voice service will be available only to other mobile operators, and possibly to other service providers who agree to equivalent conditions.

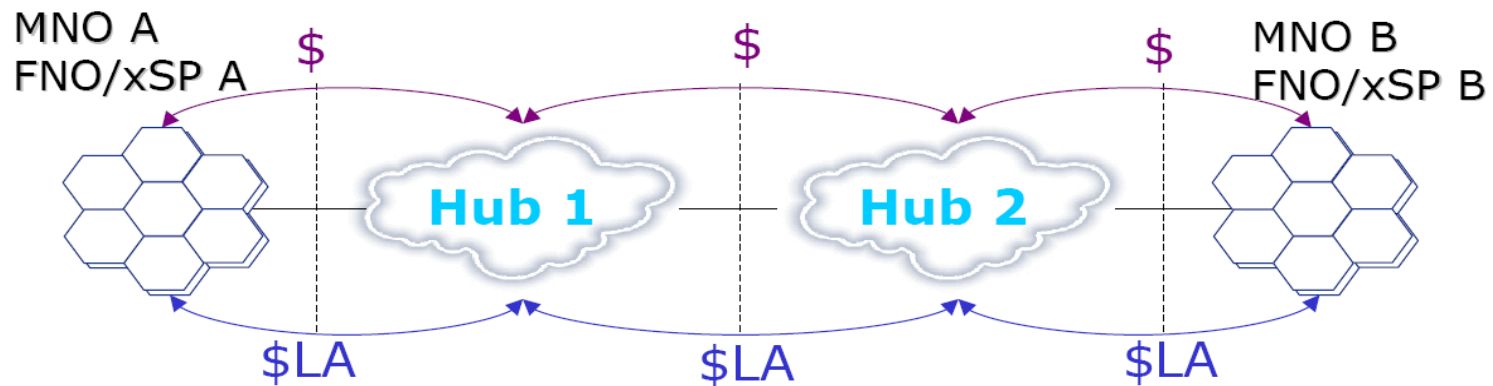
## Implications for Application Interconnection Voice Interconnection

- Evolution of GSM-A interconnection architecture from GRX to IPX
  - Migration from basic peering model to QoS-oriented approach
  - IPX internals will be on “public” IP addresses, but will not be externally accessible.
  - Key voice infrastructure, including carrier ENUM and SIP servers, will thus be accessible only to MNOs or (possibly) to firms that accept MNO pricing and quality arrangements.
- Advantage: “Hubbed” architecture potentially overcomes transaction cost problems in getting QoS-capable services launched. (Although the value of differentiated QoS is not proven.)
- Concern: “Hubbed” wholesale pricing could serve in effect to reinforce current arrangements not only for wholesale pricing, but also for retail pricing.

# Implications for Application Interconnection Voice Interconnection

## Cascading QoS

- E2E SLA & Operational KPI's
  - Money Flows Along the Value Chain (Event Charging)
    - No free peering
    - Empowering all parties to secure SLA's



Copyright GSM Association

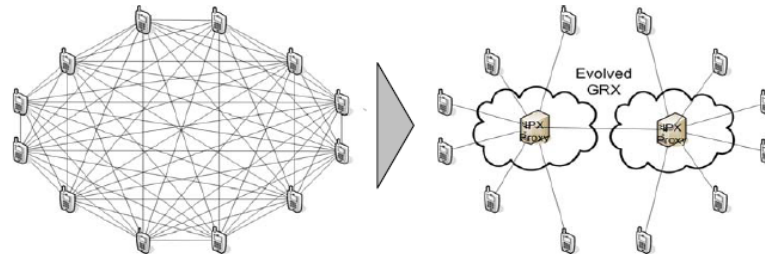
# Implications for Application Interconnection Voice Interconnection

## IPX Interconnect Models

IPX supports 3 different types of interconnect model

Interconnect Relationship	IPX Interworking Service	End to End QoS	Cascade Billing	Single Contract/ Connection
None - Internet		X		
Bilateral	Transport Only	✓	Termination billing takes place directly between Service Providers	Single contract with IPX provider but multiple contracts with connecting Service Providers
Bilateral	Service Transit	✓	✓	Single contract with IPX provider but multiple contracts with connecting Service Providers
Multilateral	Service Hub	✓	✓	✓

Multilateral interworking, or hubbing, is the principle benefit of the IPX. One contract many partners.



Copyright GSM Association

## Implications for Application Interconnection Voice Interconnection

- Fixed incumbents, and even more so mobile operators, will likely resist external IP interconnection to their inherent voice service, except where charging arrangements mirror current arrangements. Mobile operators in particular have considerable means to resist.
- Tentative conclusion: IP Interconnection  $\neq$  VoIP Interconnection
- The same might conceivably prove to be true for some applications other than voice.

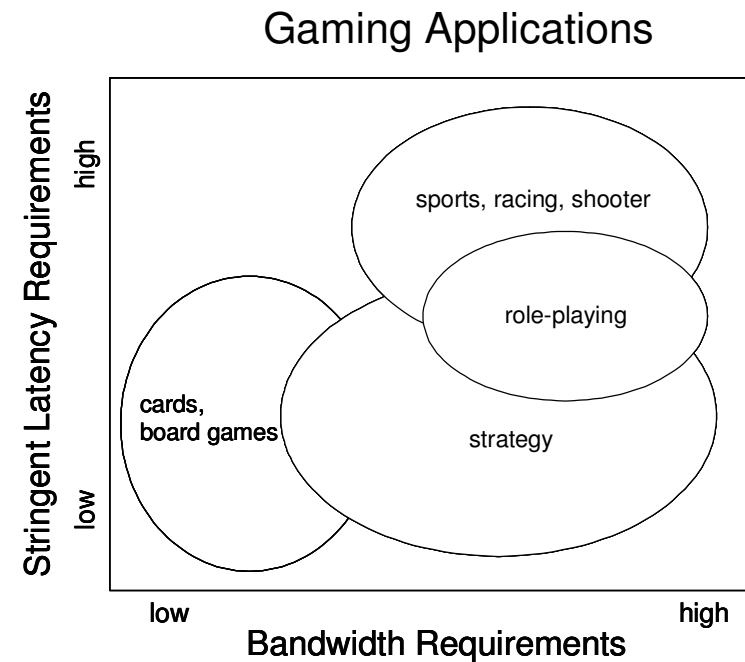
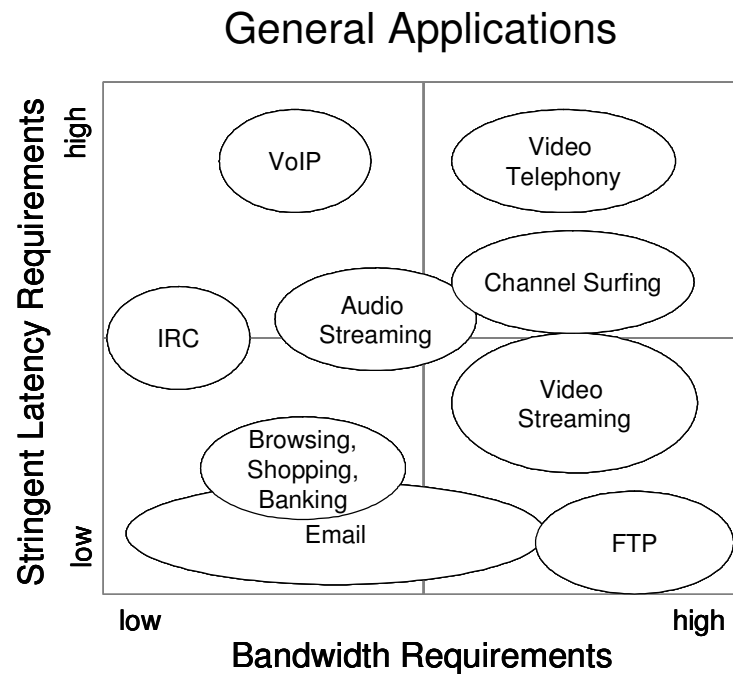


## Implementing the NGN: Regulatory and Economic Aspects

- The evolution to an IP-based NGN
  - Key functional attributes of an NGN
  - From a policymaker's perspective, what changes with the migration to an IP-based NGN?
- Implications of the migration to NGN for interconnection
  - Wholesale costs and prices
  - Retail costs and prices
  - Application interconnection
  - QoS and network neutrality
- Conclusions

# Implications for Interconnection

## Differentiated Quality of Service (QoS): Application Needs



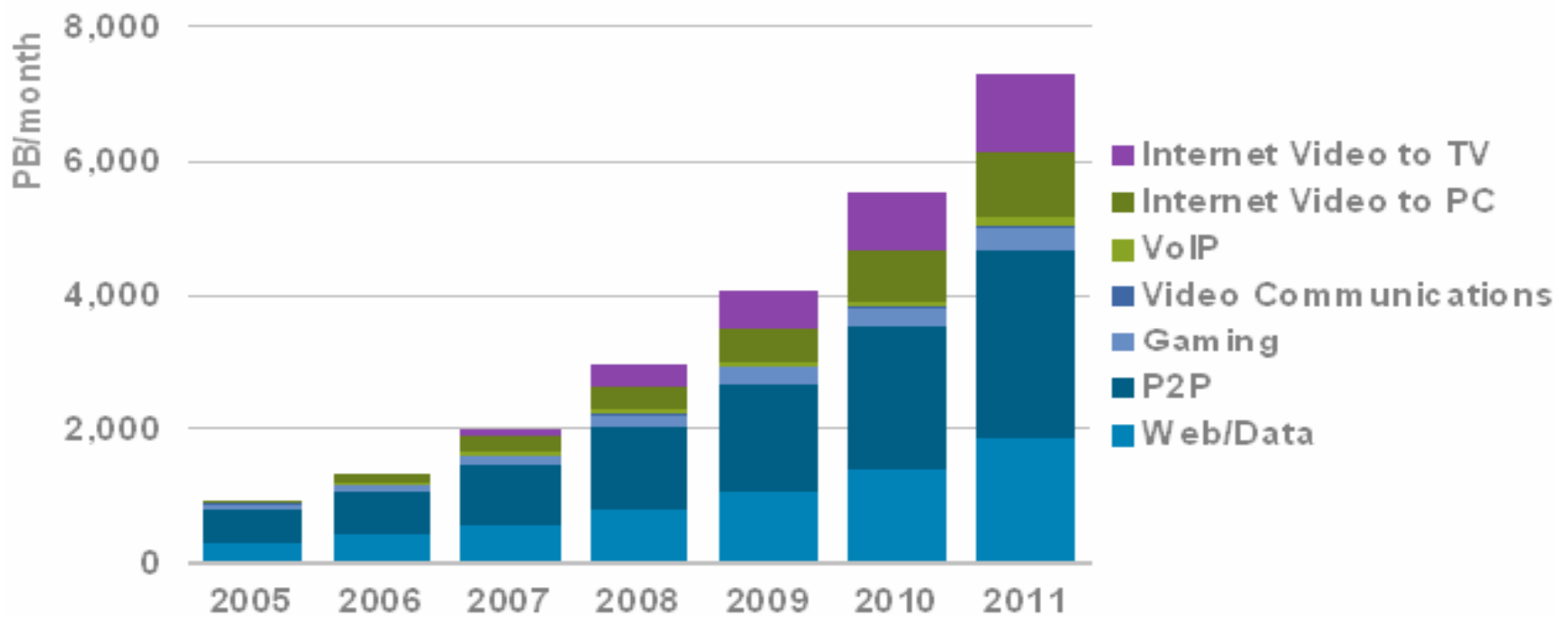
- Real time bidirectional audio: stringent requirements
- Email: liberal requirements
- Streamed audio and video: fairly liberal requirements. (Channel surfing?)

## Implications for Interconnection Differentiated Quality of Service (QoS): Application Needs

- For voice, if delay exceeds about 150 milliseconds, both sides may begin to speak at once.
- Not all video is delay-sensitive.
  - For real-time videoconferencing, similar considerations apply to delay; however, bandwidth requirements are far greater.
  - For streamed video, if it is permissible to wait a second or two at the outset, a jitter buffer can accommodate typical delays.
  - Interconnection is not relevant to all video. Much video is originated close to the end-user (within the end-user's own network, in order to save transmission costs).
- Certain interactive games may be highly delay-sensitive.
- Data applications tend to be less sensitive, but some (e.g. web-browsing) are more delay-sensitive than others (e.g. e-mail).

# Implications for Interconnection

## Differentiated Quality of Service (QoS): Application Needs



“The Internet is not collapsing under the weight of streaming video...”

Source: Cisco (2007)

## Implications for Interconnection Differentiated Quality of Service (QoS): Application Needs

- “In spite of the widespread claims of continuing and even accelerating growth rates, Internet traffic growth appears to be decelerating. In the United States, there was a brief period of ‘Internet traffic doubling every 100 days’ back in 1995-96, but already by 1997 growth subsided towards an approximate doubling every year ..., and more recently even that growth rate has declined towards 50-60% per year. ...Traffic growth rates of 50% per year appear to only about offset technology advances, as transmission capacity available for a given price steadily increases.”
- “[A]lthough service providers are pushing to throttle customer traffic, an argument can be made that they should instead be encouraging more traffic and new applications, to fill the growing capacity of transmission links.”

Source: “Minnesota Internet Traffic Studies (MINTS)” at  
<http://www.dtc.umn.edu/mints/home.html>.

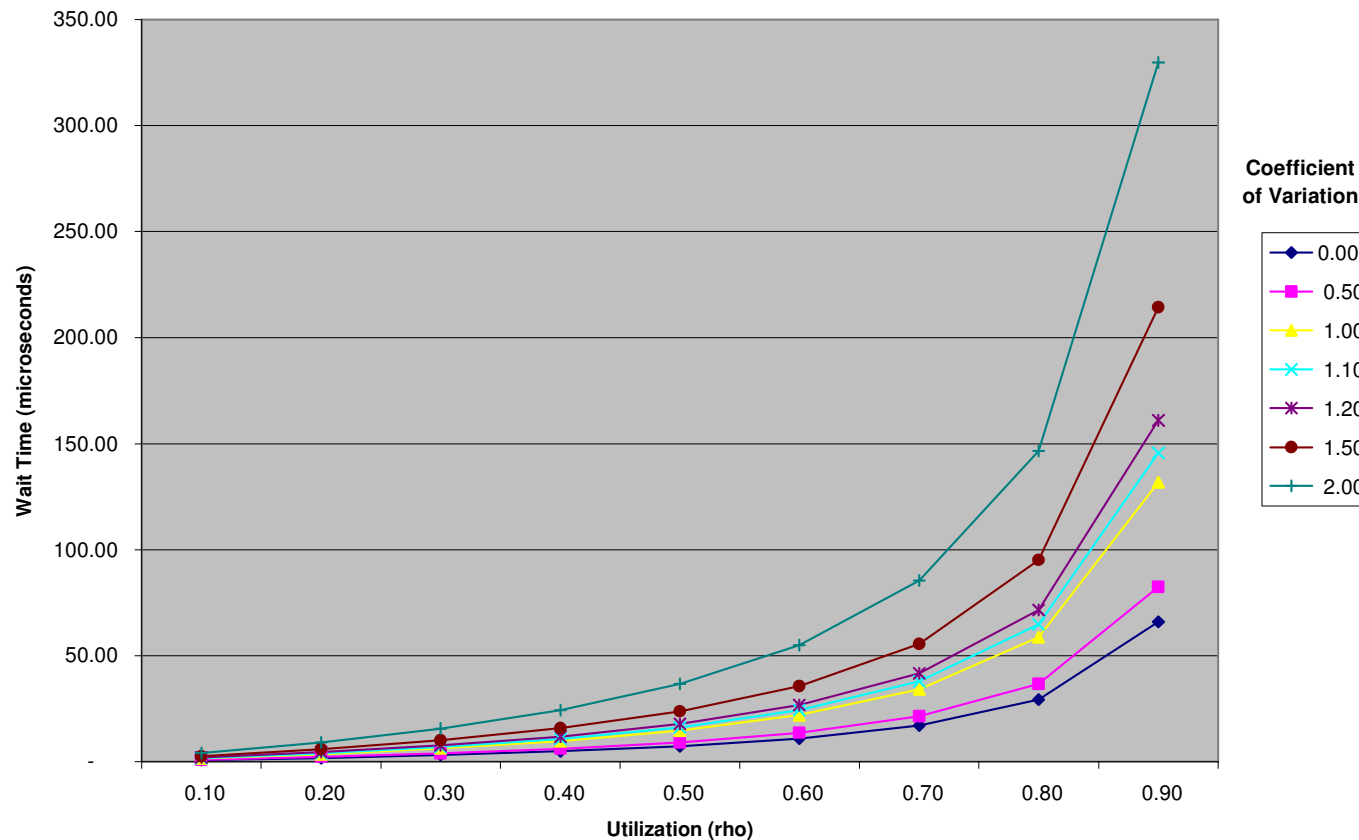
## Implications for Interconnection Differentiated Quality of Service (QoS)

- At a technical level, QoS is not fundamentally hard.
  - DiffServ is technically trivial.
  - MPLS in a single network is technically trivial.
  - Cross-provider MPLS is only marginally harder.
  - Even RSVP is not that hard. My former company, BBN, had working production RSVP-compliant networks in 1995!
- In terms of the basic economics, QoS is not fundamentally hard.
- Differentiated QoS *within* a network is, in fact, commonplace.
- Nonetheless, there is no significant roll-out of differentiated QoS *between* networks.

### ➤ **WHY NOT?**

# Implications for Interconnection Differentiated Quality of Service (QoS)

M/G/1 Queuing Delay (155 Mbps Link)



M/G/1 queueing analysis of the performance *of a single link*

(with clocking delay of 50  $\mu$ secs (284 byte packets) and a 155 Mbps link)

## Implications for Interconnection Differentiated Quality of Service (QoS)

- As we have seen, per-hop delay, even in a network with 90% load, is about 1,000 times less than the 150 millisecond delay “budget” for real-time bidirectional voice.
- IMPLICATION: Most of the time, and under normal conditions, **variable delay in the core of the network is unlikely to be perceptible to the users of VoIP or other delay-sensitive applications.**
- FURTHER IMPLICATION: **Consumers will not willingly pay a large premium for a performance difference that they cannot perceive.**
- Packet delay is more likely to be an issue:
  - For slower circuits at the edge of the network
  - For shared circuits (e.g. cable modem services)
  - When one or more circuits are **saturated**
  - When one or more components have failed
  - When a *force majeure* incident has occurred



## Implications for Interconnection Differentiated Quality of Service (QoS)

- In the US, and to a lesser degree in Europe, there has been an intense debate over *network neutrality*.
- Network neutrality has many different meanings. All relate to some form of traffic or pricing discrimination that is felt to be unwarranted.
- The debate is closely linked to notions of differentiated Quality of Service.

“The chief executive of AT&T, Edward Whitacre, told *Business Week* last year that his company (then called SBC Communications) wanted some way to charge major Internet concerns like Google and Vonage for the bandwidth they use. "What they would like to do is use my pipes free, but I ain't going to let them do that because we have spent this capital and we have to have a return on it," he said.”

## Implications for Interconnection Differentiated Quality of Service (QoS)

- Network neutrality will manifest differently in Europe than in the U.S, because:
  - The European broadband market has a richer competitive structure than that of the U.S.
  - The European regulatory framework, in conjunction with European competition law, provides much more comprehensive mechanisms for dealing with potential harms.
- The issue bears watching, but a major intervention does not appear to be warranted in Europe, and may not be warranted outside of the United States.

## Concluding Remarks

- The migration to NGN raises challenges to costing and pricing due to:
  - The ability of any transmission medium to carry any form of traffic.
  - The evolution of the telecommunications network from a voice-only network to a multi-service network where voice likely represents only a small fraction of the traffic.
  - The emergence of service providers who do not even have a network.
  - The changing cost structure of the network.
  - The understandable desire of existing operators to maintain their revenue streams.
- Wholesale interconnection obligations will continue to be required for the foreseeable future.
- Mobile termination rates are inefficiently high in many countries. Regulators should consider lowering termination rates.

## References

Marcus, Dieter Elixmann, Kenneth R. Carter, and numerous senior experts, *The Future of IP Interconnection: Technical, Economic, and Public Policy Aspects*, a study prepared for the European Commission, available at:

[http://ec.europa.eu/information\\_society/policy/ecomm/doc/library/ext\\_studies/future\\_ip\\_intercon/ip\\_intercon\\_study\\_final.pdf](http://ec.europa.eu/information_society/policy/ecomm/doc/library/ext_studies/future_ip_intercon/ip_intercon_study_final.pdf).

Marcus, “Network Neutrality: The Roots of the Debate in the United States”, *Intereconomics*, Volume 43, Number 1, January 2008. See:

<http://www.springerlink.com/content/g37k162urx11/?p=1a363b658dfb4d95accaecba21b38d5f&pi=0>.

Marcus, Dieter Elixmann, “Regulatory Approaches to Next Generation Networks (NGNs): An International Comparison”, *Communications and Strategies* number 69, first quarter 2008. Preliminary versions are available on the web sites of TPRC 2007, and of the Center for Communication Law & Policy (CCLP) at USC, at

<http://web.si.umich.edu/tprc/papers/2007/800/Regulatory%20Approaches%20to%20NGNs%20v1.2%20FINAL.pdf> and [http://cclp.usc.edu/assets/docs/RegulatoryApproachestoNGNs\\_AnInternationalComparison.pdf](http://cclp.usc.edu/assets/docs/RegulatoryApproachestoNGNs_AnInternationalComparison.pdf), respectively.

Marcus, Dieter Elixmann, Antonio Portilla Figueras, Klaus Hackbarth, Péter Nagy, Zoltán Pápai, and Mark Scanlan, *The Regulation of Next Generation Networks (NGN)*, 10 May 2007, a study for the Hungarian NHH, available at: <http://www.nhh.hu/dokumentum.php?cid=15910>.

Marcus, Dieter Elixmann, Christian Wernick, and the support of Cullen International, *The Regulation of Voice over IP (VoIP) in Europe*, a study prepared for the European Commission, 19 March 2008, available at:

[http://ec.europa.eu/information\\_society/policy/ecomm/doc/library/ext\\_studies/voip\\_f\\_f\\_master\\_19mar08\\_fin\\_vers.pdf](http://ec.europa.eu/information_society/policy/ecomm/doc/library/ext_studies/voip_f_f_master_19mar08_fin_vers.pdf).



**wik-Consult GmbH**  
Postfach 2000  
53588 Bad Honnef  
Tel 02224-9225-0  
Fax 02224-9225-68  
eMail [info@wik-consult.com](mailto:info@wik-consult.com)  
[www.wik-consult.com](http://www.wik-consult.com)