-

ERG Common Statement on Regulatory Principles of IP-IC/NGN Core - A work program towards a Common Position

This **ERG Common Statement on Regulatory Principles of IP-IC/NGN Core - A work program towards a Common Position** (ERG (08) 26 final) sets out some regulatory principles focusing on the core network. It is based on the ERG report on IP interconnection (see ERG (07) 09, published in March 2007), tackling IP interconnection and its implications as one of the main challenges emerging out of the developments towards multi-service NGNs in the core network and also takes into account more recent developments.

A previous version of this document was published as ERG Consultation Document on Regulatory Principles of IP-IC/NGN Core (ERG (08) 26rev1).

The Consultation Report will be in a "Supplementary Document" (ERG (08) 26b final) which will also contain several additional parts with country case studies and other factual information referenced in the Common Statement. Responses not taken up here are dealt with in the Consultation Report more extensively as well as our evaluation to the answers to the questions.

ERG Common Statement on Regulatory Principles of IP-IC/NGN Core – A work program towards a Common Position

Table of content

List of Figures

List of Tables

Executive summary and conclusions

Next Generation Networks

The introduction of Next Generation Networks (NGN[1]), leading to a multi-service network for audio (including voice), video (including TV) and data, as well as new plans and investment in next generation access (NGA) sets the communications sector on the verge of a new era. These developments give rise to innovation opportunities at both the service and infrastructure level and may subsequently impact significantly on market structure. Additionally, due to the increased economies of scope of a multi-service network, cost savings are to be expected.

These developments can take different forms. Traditional PSTN operators including incumbents, plan to migrate towards NGNs, relying on the ITU-T[2] and ETSI[3] as relevant standardization bodies. The migration towards NGNs does not only relate to fixed network operators but also to mobile operators.[4] On the other hand, independent ISPs and ITPs (Internet Transport Providers) continue to develop their IP networks towards multi-service networks, relying more on the IETF as standardization body.

The term NGN was developed and defined by ITU-T[5]. It is, however, also used as a more general slogan for the use of IP technology when converting the telecommunications networks from traditional circuit-switched to packet-switched technology. An IP network that uses some deliberately chosen elements that are specified in NGN standards for improving its transmission performance may, therefore, also be referred to as NGN. Today the term NGN covers a broad performance spectrum from simple implementation of TCP/IP with low level best effort performance to intensive implementation of traffic management methodologies providing high level and stable transmission performance.

Since the liberalization of the telecommunications markets, network interconnection has been one of the basic requirements for enabling competition, because this is the only way a provider can make it possible for its own subscribers to communicate with subscribers of another network.

NGNs do not change the fundamental importance of interconnection to sustainable competition in both network infrastructure and of electronic communication services. However, taking into account the migration towards all-IP networks compared to arrangements in the PSTN,

---

1 A list of definitions and acronyms is presented in the Glossary.
2 See ITU-T information on NGNs in http://www.itu.int/osg/spu/ngn/.
3 See ETSI "TISPAN PUBLISHED NGN SPECIFICATIONS" in http://portal.etsi.org/docbox/TISPAN/Open/NGN_Published/PUBLISHED_NGN_ SPECIFICATIONS.doc.
4 Thus, the scope of this document covers fixed and mobile networks. When referring to "legacy networks" here, this generally relates to both types of networks. Encompassing mobile networks in this document is not to deny differences between fixed and mobile networks e.g. with regard to access issues.
5 NGN is defined by ITU-T (Rec. Y.2001) as follows: „A packet-based network able to provide Telecommunication Services to users and able to make use of multiple broadband, QoS-enabled transport technologies and in which service-related functions are independent of the underlying transport-related technologies...". In addition, certain criteria („fundamental aspects") are specified that must be met by an NGN.

changes, may be expected at the wholesale level, as the markets for network conveyance converge and service intelligence is decoupled from the (transport) network. With the gradual replacement of traditional interconnection by IP interconnection the question of possible future interconnection models becomes relevant.

Objectives

This "ERG Common Statement on Regulatory Principles of IP-Interconnection/ NGN Core" develops some general regulatory principles focusing on the core network in order to ensure a consistent application of the European regulatory framework. The provisions of the Regulatory Framework – such as the market analysis procedure - require that regulatory actions taken by NRAs are clearly evidence based. In case SMP has been found in a market susceptible to ex-ante regulation, regulatory obligations imposed must be appropriate and proportionate in relation to the nature of the problem identified.

The aim of this document is to present the current situation with regard to IP interconnection in Europe and to outline how the development towards NGNs may affect market regulation[6] in this matter. This document will analyse the effects this evolution might have on interconnection regimes and develops some general principles with regard to regulatory treatment of IP interconnection and interoperability issues reflecting the development towards multi-service NGNs. The ERG intends to increase regulatory certainty for all market players at this time when many stakeholders are considering or actively moving towards NGN implementation.

Scope and structure of the document

This document is based on the ERG report on IP interconnection (see ERG (07) 09) published in March 2007, tackling IP interconnection and the implications of a move to this, as one of the main challenges emerging out of the developments towards multi-service NGNs in the core network and also takes into account more recent developments. In addition, some work of the Regulatory Accounting PT has been annexed to this document.

The document was published for consultation on the ERG website (consultation period June 4 2008 – July 11 2008). 21 responses were received (one of these confidential). The document has been revised in the light of these responses.

A Consultation Report, summarizing the main arguments in responses received to each question is published as Part 1 in a separate Supplementary Document. Country studies, technical background information and implications on regulatory accounting constitute parts 2-4 of the Supplementary Document. The points which received particular emphasis in the responses, including separation of service and transport and charging mechanisms, are dealt with briefly in this document and will be considered further in the Consultation Report.

---

6 Issues regarding technical regulation e.g. network integrity or security are not addressed here.

Since the publication of the ERG IP-Interconnection report, a number of NRAs have continued to work on NGN interconnection, monitoring progress in industry forums, engaging with market players and developing policies and procedural provisions.[7] One NRA has decided to require SMP operators in (old) Markets 8-10 to provide IP interconnection. This obligation has not yet been fully implemented.[8] Another NRA in a draft decision for the market on fixed voice termination defined in a technologically neutral manner determined that SMP operators have to provide interconnection on a regional level at no more than 20 PoI.[9] No other NRA has yet taken a definitive decision on a native IP interconnection product.

In view of the multi-service nature of NGNs and the decoupling of service[10] and transport, the paper will look at IP interconnection in general, and is not confined to voice interconnection.

Issues relating specifically to voice interconnection in future networks will nevertheless be treated in some detail as traditional network operators are mainly preoccupied with migrating their voice service to NGN/IP networks. Furthermore VoIP is the service best known to telecommunications network operators, whereas other services like VoD, IP-TV or presence services are still in early stages of evolution, and are less well understood at the service level. However, VoIP traffic already constitutes a relatively small fraction of overall IP traffic, in the future it is to be expected this fraction will continue to be small in comparison to other services.[11]

Given that the migration towards NGNs is in different stages in different countries, migration issues are not dealt with in detail in this paper[12] and are identified as a topic meriting further study.

Furthermore, the paper will explicitly not look at access to electronic communication services. Access has been addressed by the CP NGN Access. The distinction between access and interconnection can best be explained[13] as follows. Access enables an operator to utilize the facilities of another operator in the furtherance of its own business and in service of its own customers, while interconnection enables an operator to establish and maintain communications with customers of another operator.

After providing an Introduction in Part A, which inter alia outlines the drivers for change towards all-IP networks, Part B briefly sets out some background and facts necessary to analyse the regulatory implications of developments towards NGN:

---

7 See e.g. NGNuk (2007), AKNN (2007), Bundesnetzagentur (2008).
8 For details, see Italian Country Case Study in Annex 2.
9 If a competitor grants access at less locations, the same interconnection tariffs applies as if KPN had offered access at the 20 regional PoI. See http://www.opta.nl/download/201452+ontwerpbesluit+marktanalyse+vaste+gespreksafgifte+%28telefonie%29.pdf.
10 In this document the term service will be used in a wide sense signifying services provided to the end-user. Therefore use of the term service is not confined to the service layer described in some NGN documents but includes parts of the control layer (relevant to services), the service and the application layer.
11 Cisco estimates global consumer Internet traffic until 2011, cf. WIk-Consult (2008), p. 42.
12 This was also pointed out in several responses, see Consultation Report.
13 See WIK report The future of IP Interconnection 29-01-2008 paragraph 1.2.1

- it provides a summary of country studies;
- it describes and compares existing interconnection arrangements in the PSTN and the IP worlds;
- it describes different charging mechanisms and derives their impact on the possibility to abuse market dominance;
- it sets out main features of network structure and topology, including number and hierarchy of network nodes and hence possible interconnection points;
- it explains how quality of service may impact on interconnection.

Part C deals with the regulatory implications and challenges resulting from the expected developments towards NGNs:

- it starts setting out the relevant sections of the current framework and looks at the implications of the changes foreseen in the Review;
- it analyses how the separation of transport and service, the possible introduction of network performance classes and charging mechanisms may impact on the definition of relevant markets;
- it identifies possible bottlenecks and SMP positions at the transport as well as the service level;
- it identifies regulatory measures regarding symmetric regulation, based on USO directive and SMP-remedies
- it identifies probable changes of costs and appropriate pricing mechanisms in case an SMP-position is determined and *ex-ante* regulation imposed;
- it analyses different charging mechanisms (CPNP/ Bill & Keep) as a core element of an interconnection regime for termination service and presents a work plan for further study.

Furthermore, a Glossary is provided in Annex 1.

The starting point: Interconnection in legacy networks and in existing IP networks

Interconnection is the physical and logical linking of public communications networks used by the same or a different undertaking in order to allow the users of one undertaking to communicate with users of the same or another undertaking.[14] The terms interconnection or interconnection services used subsequently include any additional agreements related to interconnection such as supply, transit or termination, in addition to the actual traffic handover at the point of interconnection, because these aspects are dealt with together with network interconnection in terms of market regulation.

Currently interconnection in the PSTN and IP networks are characterized by a number of important differences. The use of different technologies led to different interconnection products. Different charging mechanisms and different regulatory regimes have developed.

PSTNs were designed to be capable of establishing connections (channels) with fixed transmission characteristics. Thus, they provide a fixed transmission performance since they are circuit-switched and use pre-defined PCM channels. When first established, PSTN networks had a clear focus on voice services (additional services such as fax and dial-up Internet access were introduced later). Interconnection between “traditional” circuit-switched networks reflects this focus on telephone services as transport and service are closely linked with each other.

An IP network is an all-purpose network based on packet switched technology using the Internet protocol. It provides a platform for the delivery of multimedia services. In principle, any service can be realised with a specified quality level, if the performance objectives of the service can be met by the network. Since packet switched networks are designed to “only” provide transport resources and to support any service, a separation between transport and service layer can be made. When discussing aspects of interconnection this has to be addressed.

The most common wholesale billing regime in the PSTN is Calling Party’s Network Pays (CPNP), usually covering transport and service. Under this system, the network of the party who placed the call (the originating network) makes a payment to the network of the party that received the call (terminating network). Thus, at the wholesale level the whole call is paid by the caller’s network. The terminating network provider has an incentive to charge high termination rates (sometimes this strategy is called raising rival’s cost), because it is not its own customer who finally has to pay them. The rationale of CPNP is based on the assumption that the costs are caused solely be the calling party’s network.

Interconnection arrangements in IP-based networks exist either in the form of transit, peering or Internet Exchanges (IX). The direction of traffic flows does not play a role for these ar-

---

14 Definition based on Art. 2 b Access Directive. In brief some of the most important terms used throughout the paper such as interconnection, quality of service, network performance, best effort or interoperability are given in the Introduction. For a fuller definition of the terms see Glossary.

rangements. Traffic flows in both directions are added in determining charges. Therefore, there is no need to distinguish between origination and termination for billing purposes. In transit agreements, the Internet/broadband access provider pays for connectivity to the upstream network for upstream and downstream transmission of traffic. In peering agreements, normally there are no payment flows, as long as traffic imbalances do not exceed a certain, specified limit. Those ISP who fulfil the requirements for peering can choose between peering and buying transit services. The market is generally taken to function more or less competitively as long as broadband access providers have a choice of transit providers.

The way transit and peering agreements work implies that the access provider is not entitled to any payment when taking over traffic at his agreed PoI and physically terminating a data flow, e.g. a VoIP call on its network. Such a wholesale regime, where each network bears the costs of terminating traffic coming from other carriers itself, is called Bill & Keep. The carrier will bill these termination cost on its network and any payments for upstream connectivity to its customer. As long as there is sufficient competition for broadband access at the retail level, the access provider has an incentive to keep transit cost low, since too high a cost, if passed on to the end-user, may induce the latter to change supplier.

Interconnection in the PSTN is subject to both symmetric regulation (obligation to interconnect for all market players) and asymmetric regulation (SMP parties mostly subject to *ex-ante* cost-based regulation). The case for regulation was considered strongest for call termination markets. Absent regulation the so called “termination monopoly” of the terminating network operator arises due to the interplay of several factors, namely the physical bottleneck position to terminate the call, coupled with the control of the E.164-number and the CPNP mechanism entitling the terminating network operator to a termination fee from the calling party’s network operator.

In IP networks interconnection developed without any regulatory intervention although the obligation to negotiate for interconnection applies to IP networks as well. These agreements have been largely outside the scope of activity of NRAs and there are no known filings of disputes.

The crucial question for NRAs will be how interconnection evolves in future networks with the different regimes coalescing. Future IP interconnection products in an “all IP” environment – at least for telephone services – could be governed either in a PSTN-like manner or in the varieties of IP interconnection as they are known today.

Among other potential issues in voice (telephony) services, the interplay of the following factors in an NGN world should be carefully analysed:

- Physical bottleneck for termination

  Depending on the degree of competition at the access level, the broadband access provider may have the option of barring traffic to and from particular servers. Thus, this form of market power refers to the ability to deliver any service (not only voice);

- Control of the E-164 number (the subscriber's telephone number)

  It confines market power to the service provider as long as it is not possible for more than one provider to terminate calls to a specific number;

- Role of the charging mechanism (e.g. CPNP)

  The extent to which the market power enabled by the physical bottleneck for termination can be abused depends on the charging mechanism and the direction of payment flows.

Drivers for change

Future electronic communications networks will be packet-switched, mostly or completely based on the IP. They will be multi-service networks, rather than service-specific networks, for audio (including voice), video (including TV services) and data rather than service-specific networks, allowing a decoupling of service and transport provision.

Developments towards all-IP networks can take different forms. On the one hand, traditional PSTN network operators including incumbents plan to migrate towards NGNs, relying on the ITU[15] and ETSI[16] as relevant standardization bodies and on the other hand, independent ISPs and ITPs continue to develop their IP networks towards multi-service networks, relying more on the IETF as standardization body.[17]

Important features of NGN are:

- separation of transport and service;
- assurance of Quality of service;
- cost savings;
- network topology.

Separation of transport and service

A core feature of NGN architecture is the separation of the main functional levels, i.e., generally a distinction can be made between transport and service. This separation of transport and service constitutes an essential feature of NGN specifications in both the specifications on NGN (ITU-T) and NGI (IETF).[18] A crucial point is the adoption of open and standardised

---

15 See ITU-T information on NGNs in http://www.itu.int/osg/spu/ngn/.
16 See ETSI "TISPAN PUBLISHED NGN SPECIFICATIONS" in http://portal.etsi.org/docbox/TISPAN/Open/NGN_Published/PUBLISHED_NGN_ SPECIFICATIONS.doc.
17 See IETF working groups "Session PEERing for Multimedia INTerconnect –speermint" (http://www.ietf.org/html.charters/speermint-charter.html).
18 See e.g. the description of ITU architecture in ERG (2007), Ch. A1.1 (p. 39). For a comparison of NGN and NGI see Hackbarth/Kulenkampff (2006), Ch. 3.4.

interfaces between each functional level in order to allow third parties to develop and create services independent of the network. More specifically ITU-T Rec Y.2001 (2004) defines a number of necessary criteria for an NGN among them “separation of control functions among bearer capabilities, call/session, and application/service” and “unrestricted access by users to different service providers”:

In the present Internet this allowed easier service creation (as it lowered entry barriers) which contributed to the success of the Internet and TCP/IP protocol suite.

The architecture of NGNs can include the principle of separation of transport and service that has so far been characteristic of IP networks. This principle could also support service provision by third parties by providing open interfaces, even though but transport could become service-specific to reflect the different network performance needs of different services (real time etc.).

This separation of transport and services is also expected to be reflected in the respective interconnections services, i.e. service interconnection and transport interconnection.[19]

- Transport interconnection includes the physical and logical linking of networks based on simple IP connectivity irrespective of the levels of interoperability. It is characterised by the absence of the service-related signalling, implying that there is no end-to-end service awareness. Consequently, service-specific quality of service and security requirements are not necessarily assured. At the transport level, only objectives for those parameters[20] are negotiated that affect the transmission performance at the point of interconnection (e.g. availability) and the IP packet transmission performance via interconnected networks.

  This definition does not exclude that some services may provide a defined level of interoperability or the establishment of transport classes to guarantee specific quality parameter at this level.

- Service interconnection in this paper is understood as including solely service-specific aspects.[21,22] It consists of logical linking of network domains, having access and control of resources including the control of signalling (i.e. session based service-related signalling[23]). Depending on the kind of service, different aspects must be considered. For example, in the voice service, the call server interconnection is required for call setup and disconnection.

---

19 See Ch. 1.3.1 (p. 5) stating that with IC taking place at separate levels, this may require defining different IC services.

20 For e.g. bit rate, delay and packet loss ratio.

21 This differs from ETSI/TISPAN’s definition of ‘service oriented interconnection’ also including transport related information. See.below Section 3.2 in Annex 3.

22 This may have an important implication. Given this understanding, service interconnection may not fulfil the definition of Art 2 (a) Access Directive (see Sec. C.1.1) because it does not necessarily include the physical linking of NGN domains (see Sec. C.3.2).

23 Information that allows identification of the end-to-end service that has been requested.

Interconnection between services from different operators requires a minimum set of technical (e.g. defined by a SLA) and commercial conditions to be fulfilled by both operators.

These conditions may include *inter alia*:

(i) mutual policies for exchange of data (including transcoding the information mapping of quality of service information (if applicable), service control information and network protocols);

(ii) agreement of charges;

(iii) agreement on performance and reliability levels.

Other aspects such as security for example, may need to be considered.

If transport interconnection and service interconnection were bundled, as was the case in the PSTN, we may talk about transport+service interconnection (corresponding to the concept of Service oriented Interconnection according to ETSI/TISPAN).

In most NGNs planned by incumbents, services tend to be provided using centralized platforms (Media Gateway, Softswitch). Operators who have market power may not have an incentive to open their networks to competition at the service level and may want to limit use of these capabilities.[24] This impacts on the ability of independent service providers to integrate their services into the NGN platform.

Furthermore, such a configuration of services and the centralization of the control function have implications for the locations at which traffic can be handed over to other networks or received from other networks. This feature will therefore be crucial for the possibility of defining interconnection points.

Therefore, NRAs may have to ensure that interconnection is possible at specific functional levels in a reasonable manner. This may prove to be particularly challenging since a number of “telco” network operators are of the view that a horizontal separation of transport, service (and control) levels is neither appropriate nor in their interest, particularly if they want to guarantee quality of service (see IMS[25], adaptation of IMS to fixed networks, etc.). Most respondents to the Consultation corroborated that their understanding of NGN implies a continuation of vertically integrated provision of transport and services as has been the case in legacy “telco” networks. If network operators continue applying a vertically integrated approach, implying a simple transfer of the PSTN world to the NGN world, then the same kind of regulation may also need to be transferred respectively, if the underlying economic problems have not changed.

---

24 See Katz/Shapiro (1985)
25 See Ericsson’s White Paper, “IMS – IP Multimedia Subsystem” (2004).

Assurance of quality of service

Another important distinctive feature of NGN relative to current IP networks is the expected ability to provide a range of service quality levels. However, the implementation of NGN has yet to prove to what extent this is technically and commercially feasible.

This ability could be key in enabling a wide range of services to be carried using a single network infrastructure and hence realizing the economies of scope that could be achieved. Real-time services including conversational and interactive services (for example videoconference calls and on-line gaming) tend to convey ever-richer media content and therefore require ever higher bit-rates. Furthermore, multiple services can be combined into a single user's session. A network operator has to take into account the performance objectives of these services and plan and operate its network accordingly. For this purpose, it may want to implement QoS mechanisms.

The term quality of service (QoS) is a broad concept (see Glossary) as it covers all aspects influencing the user's perception of the quality of the service; related to the network performance (such as reliability or availability) as well as to other factors, such as terminal equipment[26], codecs or customer support. However, the term QoS is sometimes used when referring to part of these aspects and often misunderstood as a synonym for network performance, which is defined and observed as performance of a telecommunication network (or sections of a network) by using objective performance parameters.

In the current practice of interconnection of IP networks, even though there are QoS mechanisms implemented for internal purposes, there is no technical implementation of QoS mechanisms across interconnected networks resulting in a "best effort" transmission performance, i.e. no service prioritisation. This "QoS model" is characteristic of today's IP network but does not necessarily result in a low transmission performance and thus a low quality of service.

PSTN interconnection did not entail any formal QoS guarantee either, but transmission plans were established over many years to provide a comparatively high assurance of voice service quality at the expense of greater complexity.

If a guaranteed QoS over IP-based networks with a PSTN-comparable level of performance is desired, one has to modify and adapt the IP transport model in a way that connections with reliable and fixed transmission characteristics (transport classes) are possible. Moreover, the ability to provide services with guaranteed QoS may not only be applied to a single network but must be maintained over the whole chain of interconnected networks involved in the provisioning of a specific end-to-end service.

It is not clear yet which strategies will be chosen by the different operators. Depending on the strategy (stringent systems versus systems guaranteeing QoS on a statistical basis only),

---

26 For instance by selecting and configuring the terminal equipment the user can substantially affect the resulting end-to-end quality.

specific additional agreements and mechanisms may have to be implemented. Interoperability, i.e. the ability of systems and services to exchange information and mutually use the information that has been exchanged[27], is necessary to make interconnection work in practice.

The implications of different transport classes will need to be analysed in some detail, making NGN interconnection potentially more complex. QoS is therefore potentially a new dimension in the interconnection of NGN, and could be an important focus for NRA because it could enable new forms of discrimination between a larger operator's services and those provided by interconnecting competitors.

Cost savings in NGNs

The development of NGN core networks offers opportunities for increasing efficiency and for innovation. Economies of scope can be expected to result from the convergence of services onto a single network infrastructure which will convey them potentially leading to a reduction of total costs. At the same time, the capacity and processing power of modern network equipment can result in increased economies of scale and in innovation in services and in packaged retail propositions. The greater economies of scale in NGNs are likely to increase the challenges in promoting effective competition in the provision of services solely reliant on core infrastructure.

Therefore, NGN technology is expected to save costs (cheaper equipment, economies of scale and scope). The cost structure of NGN core networks is likely to contain a relatively high proportion of costs common to all services carried by the network, and a relatively low proportion of costs directly driven by the volume of each service (see Sec C.4 and Annex 4).

Network topology

It is as yet unclear exactly how NGN implementations will occur and they will differ in Member States by investment cycles, network size and other factors. Operators will "pick and choose" from the different architectures/features standardized in ITU and ETSI or other bodies.

This relates for instance to the implementation of transport classes relying on prioritisation versus best effort management with adjusted network dimensioning.

Furthermore, it is not yet clear as of now to what extent features such as the separation of service and transport will remain possible in the architectures finally implemented.

A number of operators intend to implement their NGN using centralized platforms for service provision, affecting the ability of independent service-providers to integrate their services into the NGN platform. Whether independent service providers will be able to do so also depends

---

27 Definition based on ITU-T Rec. Y.101

on the availability of open and standardised interfaces. Furthermore, such a configuration of services and the centralization of the control function have implications for the locations at which traffic can be handed over to other networks or received from other networks.

The separation of transport and services will be crucial for the definition of the interconnection points[28] and (number of) nodes. In any case, an immediate manifestation of the increased capacity of network equipment is likely to be a reduction in the number of nodes in which routing of traffic will occur in NGNs when compared with the number of nodes where traffic is switched today by traditional networks.[29]

However, currently operators consider to continue the implementation of interconnection for voice services in a bundled manner – combining transport and service - during the transitional period towards NGN.

It may have to be taken into account that short-term determinations that are only related to voice services do not reflect the multi-service nature of NGNs and therefore risk hampering or delaying introduction of the NGN principle as an overall concept and not fully exploiting the competitive potential offered by NGN in spreading innovative services rapidly.

Core elements of interconnection

An interconnection regime contains at least the following core elements:

- number and geographic location of interconnection points as well as functionality and hierarchy of these interconnection points. They are determined by the network architecture (for the network topology and functional layers in NGNs see Sec. B.4.4);
- the network performance needs to be specified (e.g. best effort and/or some additional transport classes specifying the relevant network performance parameters like jitter, delay and packet loss (see Sec. B.4.5);
- furthermore, an interconnection regime is characterized by the charging mechanism used, i.e. who pays for which part of the value chain. This has an impact on the market power that can be exerted by different market parties at different levels of the value chain (see Sec. B.3);
- in case interconnection is regulated according to *ex-ante* cost based regulation, costing and pricing principles have to be discussed (structure of tariffs according to hierarchy, accounting units such as minutes, bandwidth, etc.) (see Sec. C.5).

---

28 ERG (2007), p. II
29 ERG (2007), Ch. 1.3 (p. 6) and Ch. 3.3 (p. 16)

Conclusions

Multiservice networks[30]

Electronic communications networks will become packet switched, mostly or completely based on the IP. They will be multi-service networks, rather than service specific networks, for audio (including voice), video (including TV-services) and data networks, allowing a de-coupling of service and transport provision.

Developments towards NGNs give rise to innovation opportunities at both the service and infrastructure level and may subsequently impact significantly on market structure. Additionally, due to the increased economies of scope of a multi-service network, cost savings are to be expected.

These technical changes result in new and various possibilities for service provision on several network layers for both access providers and pure service providers. As such, NGN holds an important competitive potential that should be optimally exploited for spreading innovative services based upon IP networks.

Given that dynamic market processes often lead to unforeseeable innovations, it is import not to foreclose certain options and potentials.

More generally, NGNs could allow competition to develop in two mutually compatible ways. The first is based on investment in infrastructure, in which competing market players can seek to exploit both local loop unbundling and the efficiency of underlying technologies of NGN to build more efficient networks and aggregate traffic from a combination of services to help achieve economies of scale. The second is based on innovation in services, in which competing players control services transported on an SMP operator's network, thereby both lowering barriers to entry of new players and allowing competitors with more substantial investments in infrastructure to increase scale by using the SMP operator's network to extend the market addressable by their services.

It should be taken into account that short-term decisions related solely to voice services do not reflect the multi-service nature of NGNs, and therefore risk hampering or delaying the introduction of the NGN principle as an overall concept, and this will not allow the full exploitation of the competitive potential offered by NGN in rapidly spreading innovative services.

The ERG is therefore committed to creating a regulatory environment based on the tools of the ECNS framework in which the chances and innovative potentials of NGN services can flourish and can be passed through to consumers and business customers, so that they can use innovative services. The ERG is aware that the design of interconnection arrangements, plays a major role for exploiting these potentials.

---

30 See Section C

Separation of service and transport[31]

A core feature of IP networks is the separation of the main functional levels, i.e., generally, a distinction can be made between transport and service. This distinction potentially allows competition along the value chain more easily than in the PSTN world. A crucial point is the adoption of open and standardised interfaces between each functional level in order to allow third parties to develop and create services independent of the network. More specifically ITU-T Rec Y.2001 (2004) defines a number of necessary criteria for an NGN among them "separation of control funtions from bearer capabilities, call/session, and application/service" and "unrestricted access by users to different service providers":

This business model has led to the success of IP networks. In the long run, it must be expected that the separation of the functional levels is also reflected in the respective interconnections services.

The architecture of NGNs allows continuing the principle of separation of transport and service that has governed IP networks. It also is intended to allow for third party service provision by providing open interfaces. Another important distinctive feature of NGN relative to current IP networks is the expected ability to provide a range of service quality levels. This ability could be key in enabling a wide range of services to be carried using a single network infrastructure. However, the implementation of NGN has yet to prove to what extent this is technically and commercially feasible.

Therefore, NRAs may have to ensure that interconnection is possible at specific functional levels in a reasonable manner. This separation of transport and services is also expected to be reflected in the respective interconnections services, i.e. service interconnection and transport interconnection.

This may prove to be particularly challenging since a number of operators intend to implement their NGN using (centralized) platforms for service provision, affecting the ability of independent service-providers to integrate their services into the NGN-platform. Furthermore, such a configuration of services and the centralization of the control function has implications for the locations at which traffic can be handed over to other networks or received from other networks. Traditional telecom operators consider providing the implementation of interconnection in a bundled manner – combining transport and service –- during the transitional period towards NGN regarding voice services.

If network operators continue applying a vertically integrated approach, implying a simple transfer of the PSTN world to the NGN world, then the same kind of regulation may also need to be transferred respectively. In this case, the potential deregulatory opportunities which could arise from the separation of transport and service may not materialise. If in this case the underlying competition problems remain and/or potential new ones arise, regulation will continue to be required.

---

31 See Section A.5.1

The ERG supports the separation as expressed in the NGN concepts of ITU-T and ETSI-TISPAN. The ERG clarifies that this separation does not imply that both levels are completely independent. Even with a logical separation between transport and service, there is co-ordination and interaction between them. What is key for competition, however, is that the separation should allow transport and service to be provided by different parties. Service provision by independent third parties becomes possible, independent of transport technology and type of network access. This approach requires open interfaces for third parties.

The ERG is convinced that such a separation between transport and service would contribute to and promote the development of new and innovative services.

Topology[32]

The separation of transport and services will be crucial for the definition of the interconnection points. Transport and service interconnection might occur at different nodes and hierarchy levels. Considering the distinction between transport and service, transport interconnection could take place at a greater number of locations than service interconnection.

The currently classically used hierarchy concept of three physical levels, “local, regional, national”, may not be applicable in an IP network. The amount of physical levels in a NGN network may change from formerly 3 in a PSTN network to only two physical levels.

The number of network nodes/PoI at each hierarchy level for NGN is not yet decided upon (or relevant information is not available) in most countries. The empirical basis is not broad enough to derive substantial conclusions as the sample is too small to derive stable relationships between the number of nodes at different hierarchy level. This situation highlights the need for more transparency on SMP operators’ NGN plans.

In any case, an immediate manifestation of the increased capacity of network equipment is likely to be a reduction in the number of nodes in which routing of traffic will occur in NGNs when compared with the number of nodes where traffic is switched today by traditional PSTN networks. As the Consultation has confirmed one may therefore assume that interconnection will be driven towards higher levels of the network.[33] Whether/to what extent this will be the case may depend on the specifics of each network and may therefore differ among Member States.

The highest level in a NGN network could logically serve both double tandem and single tandem termination when the highest level is composed of more than 1 location and these locations each serve a specific region of the country under which a certain amount of metro nodes are implemented. The lowest interconnection level possible will be the metro node level in most countries, but this will depend on the design of the network in question.

---

32 See Section B.3.3

33 Such a reduction could lead to stranded investments; these are not restricted to either incumbents or competitors.

ERG considers that the maximum efficient number of PoI offered in NGNs should be used for applying the lowest interconnection rate accordingly, even in case where not all of these points are physically offered for interconnection.

Quality of Service[34]

Quality of service (QoS) is potentially gaining importance in the interconnection of IP / NGN.

IP networks are using packet-switched technology and provide simply transport capabilities irrespective of the services that are using the network while the intelligence and complexity that is necessary for the provision of services is relying on the end-devices.

Thus QoS issues are more complex than in legacy telephone networks where fixed resources are reserved to each call.

The term QoS is a broad concept as it covers all aspects influencing the user's perception of the quality of the service; related to the network performance (parameters like jitter, delay or packet loss) as well as to other factors (e.g. device, codec, help-desk). However the term is sometimes used when referring to part of these aspects and often loosely used as a synonym for network performance.

If a guaranteed network performance over IP-based networks with a PSTN-comparable level of performance is desired one has to modify the IP transport technology in a way that connections with reliable and fixed transmission characteristics (transport classes) are possible.

This could lead to different qualities for electronic communication services. Operators are free to develop this as competitive markets are often built on quality differentiation, which can generally be considered to be welfare-enhancing. Nevertheless it covers a potential for anti-competitive behaviour.

This relates to the fact that there might only be a willingness to pay for a premium transport class in case the best effort class quality is "bad enough".

NRAs should prevent any anticompetitive behaviour from SMP operators that might intentionally degrade quality of the interconnection with some specific networks to benefit their own quality service. In order to avoid this possible deviation, NRAs can use existing tools to impose non-discrimination obligations on SMP carriers in case markets have been defined accordingly.

Therefore, it could be an important focus for NRAs because it could enable new forms of discrimination between a larger operator's services and those provided by interconnecting competitors.

---

34 See Section B.3.5

Non SMP carriers can be obliged to negotiate interconnection (Art. 5 AD). This instrument of symmetric regulation has been rarely used by NRAs but may under certain circumstances be an appropriate tool.

NRAs should have the possibility to recommend or even set minimum levels of quality of service if this is unavoidable to achieve sufficient end user service quality. However, NRAs should not prescribe the concrete mechanism to fulfill this minimum level. This should be left to network operators. ERG therefore welcomes the proposed provision in Art. 22 para 3 UD, but considers that the power to set minimum quality of service requirements should be entrusted directly to NRAs.

Evaluation of Review proposals[35]

In the Framework Directive, an additional sentence has been inserted in Art. 5 para 1 FD according to which undertakings providing ECNS can be required to submit information concerning future network, or service developments that could have an impact on the wholesale services made available to competitors. Given the importance of transparency on incumbent operator's NGN plan, the current more general wording in Art. 5 FD, the FD and other specific directives have not proven to be sufficient to justify information requests by NRAs directed to incumbent operators.

The provisions in the new section on security and integrity of networks and services (Art. 13a and 13b FD) are of relevance as the provisions on network integrity for fixed telephony service providers in Art. 23 UD are now extended to all kinds of ECNS including mobile and IP networks

Art. 5 para 4 AD which stipulated Member States have to empower NRAs to intervene at their own initiative where justified has been deleted. This deletion is likely to have a disadvantageous impact on NRA' s efforts to ensure interoperability of services in an NGN environment which will become more important. The ERG holds the view that the power of NRAs to act on their own initiative to ensure end-to-end connectivity / interoperability should be maintained in Art. 5 para 4 AD.

ERG welcomes the proposed new para 3 in Art. 22 UD allowing the Commission to adopt technical implementing measures concerning minimum quality of service requirements to be set by the NRA on undertakings providing public communications networks. However, the ERG considers that the power to set minimum quality of service requirements should be entrusted directly to NRAs.[36] As minimum quality of service requires measures on both the end-user and the network level, it should be clarified that NRAs can require minimum quality of service on the network level as well.

---

35 See Sections C.1.2 and C.4

36 The EP adopted an amendment to that effect in its vote on September 24.

In case it is not possible to do so in the UD, a second best option would be to empower NRAs in Art. 5 AD to set on their own initiative minimum quality of service requirements on operators of public communications networks.

Relevant markets[37]

Concerning relevant markets IP interconnection will influence market definitions.

Today the abuse of the physical termination bottleneck can be considered the main regulatory problem concerning interconnection. Each market for call termination on an individual fixed/mobile network is a monopolistic market with no tendency towards effective competition (1 network – 1 market). The reason for these monopolies is not only control over access lines and necessary routing information behind E.164 numbers. The CPNP charging principle allows to exploit the physical termination bottleneck. Therefore under CPNP a method of cost-based regulation is necessary.

By applying Bill & Keep, a billing regime without payment flows at the wholesale level, the abuse of the physical bottleneck for termination could be avoided if there is sufficient competition at the retail level.

When reviewing markets 2, 3 and 7 of the Recommendation (call origination / call termination) NRAs will have to analyse the impact of IP interconnection on relevant markets.

In a first step, NRAs will have to examine if these markets that had been traditionally reduced to interconnection for narrowband telephone services contain interconnection for IP based telephone services. Such markets would still be service-specific.

The division between transport and service may lead to other market definitions like markets for transport interconnection (without relation to specific services) and additional interconnection markets on the service level though it is open if such markets would be susceptible to *ex ante* regulation.

SMP bottlenecks

Interoperability[38]

Barring the access to functions like home Subscriber Server, user Profile / User Identity, location Information, call Session Control Function, charging Collection Function / Online Charging System, policy Decision Function, border Gateway Control Function, authentication and Key Agreement, terminal Capabilities hindering service interoperability, may be

37 See Section C.2
38 See Sections B.3.4 and C.3.1

considered a bottleneck, as well as restricting access to the SIP servers (providing access to numbering information)

Apart from the general requirement of open interfaces, interoperability of similar functions in different NGNs may be another issue for regulatory attention. Differing standards, incompatible data formats, and proprietary implementations may cause the occurrence of new bottlenecks. However, it is not a regulatory target *per se* to enforce interoperability on all NGN layers and for all services and applications, but to intervene when practical competition problems occur. Such competition problems can be expected to be more critical when general access to users on network layer ("Internet-style" access) is restricted for other operators (networks).

The ERG clarifies that, in line with Recital 30 FD, standardisation should remain primarily a market-driven process. Nevertheless NRAs need to monitor carefully whether lack of interoperability causes competitive problems. In order to ensure interoperability NRAs should ensure that the reference offer can contain a provision regarding a "minimum set of standards" applicable or refers to a set of specific standards.[39]

Transport interconnection[40]

The possibility to exploit SMP results from the interplay between the three factors: a) physical monopoly for termination; b) charging principles; and c) control of the E-164 number.

Generally, the exploitation of the physical termination monopoly is closely linked with Calling Party's Network Pays (CPNP), the wholesale billing regime mostly used in the PSTN. With this principle, the access operator terminating the call receives a payment from the interconnected network out of a monopoly position. By applying Bill & Keep, a billing regime without payment flows at the last link at the wholesale level, the termination monopoly could be avoided. It is important to note, that with a migration towards IP-based networks market power for termination is not "automatically" avoided. This holds true, as long as CPNP is applied. Moreover, even if customers have different access options available to them, the calling party still encounters a physical terminating bottleneck.

With regard to the probability of abuse of market power of the physical termination bottleneck the reversal of the direction of payment flows makes the big difference. The crucial point is that the end-user pays for these flows, who may change supplier in the case of abuse and therefore his supplier has no incentive to excessively raise prices.

Therefore, where Bill & Keep applies it is unlikely that SMP will be the outcome of a market analysis of the termination market.

---

39 In Italy in a first consultation document about IP Interconnection AGCOM proposed to agree on a set of standards (which include NGN ETSI and ITU standards) to which all parties should be compliant when interconnecting. This does not mean that two Operators can interconnect to each other using a different standard. A final decision has not been taken.

40 See Section C.3.2

With regard to upstream connectivity - whether the Internet access provider participates in an NAP exchange, peers or pays transit – any payments cover upstream and downstream traffic and they will be ultimately borne by his own end-customer creating the right incentives if there is sufficient competitive pressures at the access level.

The wholesale market for connectivity in IP based networks with its peering and transit agreements has so far been considered a market that entails oligopolistic market power but where the 3-criteria test is not fulfilled. An important point that there is generally a choice between different transit partners.

Service interconnection[41]

In IP-based networks, control over the IP-address respectively the E-164 number continues to provide scope for abuse of market power. At present the ERG does not see that other addressing mechanisms such as ENUM will alleviate the problem resulting from control over the E-164 number or any other private addressing mechanism.

It is claimed by many competitive operators that control over signalling information and intelligent features (e.g. presence information) could reinforce market power problems in all-IP networks allow leveraging towards adjacent sectors. If a provider of the transport service does not receive information on the IP-address of the end-user (or of a server) he will not be able to establish the media stream and to provide conveyance.

In practice, a number of third party VoIP providers have decided to share their signalling information (e.g. Freenet and Sipgate) and there is no on net/off net price differential for calling parties. We do not know of excessive payments for access to the signalling information. The cost of the network elements involved are rather limited (e.g. Sipserver).

Regulatory measures[42]

The crucial question for NRAs will be whether interconnection evolves in future networks in a PSTN-manner or IP network manner in its varieties known today and how the implementation is managed.

Providers of ECNS can be obliged to symmetric obligation to negotiate interconnection according to Art. 5 AD. This applied to both PSTN and IP-networks. The provisions have been rarely used. As QoS mechanisms are not yet widely deployed at interconnection points and due to the interdependence between each network involved in the session, one operator could be unwilling to invest in QoS mechanisms if the interconnected networks have no intention to do so. Thus, NRAs can also use symmetric regulation tools in order to enhance QoS development between different networks.

---

41 See Section C.3.3
42 See Section C.4

Regulators could require operators to provide public information about QoS, based on articles 20 and 22 UD. In order to verify that there is no discrimination of QoS between operators interconnecting, NRA may also add other relevant QoS measurements.

Moreover, NRA should have the possibility to recommend or even to set minimum levels of quality of service. This may address possible incentives of operators to degrade their best effort class. ERG welcomes the proposed new para 3 in Art. 22 UD allowing the Commission to adopt technical implementing measures concerning minimum quality of service requirements to be set by the NRA on undertakings providing public communications networks. However, the ERG considers that the power to set minimum quality of service requirements should be entrusted directly to NRAs.

SMP-remedies (based on Art. 15, 16 FD, Art. 9-13 AD) may be applied only in those cases where SMP has been found on a relevant market susceptible to ex-ante regulation. CPNP leads to SMP in termination markets usually implying *ex-ante* price regulation remedy. Once Bill & Keep has been implemented it is unlikely that SMP will be the outcome of a market analysis. NRAs could consider imposing termination rates of zero for the termination segment up to the first router or switch and associated service control functions after the access/concentration network. The possibility to implement Bill & Keep under the current regulatory framework could be explored further by ERG.

Regulatory implications regarding QoS: In order to prevent any anticompetitive behaviour from SMP operators that might intentionally degrade quality of the interconnection with some specific networks, NRAs can use existing tools to impose non-discrimination obligations on SMP carriers in case markets have been defined accordingly.

Regulatory implications regarding interfaces: The implementation of NGNs is expected to lead to a clear distinction between service-specific functions and transport functions common to all services which may support a more rapid and less costly development of new services. The same distinction could allow an SMP operator to provide interfaces for competitors to control services transported over its own network. NRA’s may therefore need to consider interventions which make such interfaces available. The *ex ante* framework allows NRAs to address differences in the product quality by applying non-discrimination obligations to providers found to have SMP. However, where these obligations may not sufficient to deliver a level playing field, NRA’s could consider more stringent requirements, including equivalence.

Costing / pricing[43]

It is generally accepted that NGNs core will lead to a lower overall cost level due to increased economies of scale and scope and that the cost structure will change with a higher proportion of common costs compared to legacy networks. The use of common platforms to deliver multiple services across one network allows exploiting economies of scope thus reducing the costs each service has to bear.

---

43 See Section C5

The Opex and Capex of a NGN are forecast to be significantly lower in the long term than current legacy technologies as NGN core networks are generally seen as cost saving, because they result in a more efficient network design and usage. There are three factors – simpler network structure with fewer levels and fewer nodes at each level plus more efficient equipment (packet switching technology) – that reduce total per minute cost of NGN core networks.

Common and fixed costs of NGNs will represent a high percentage of total costs with a relatively low percentage of costs incremental to individual products or services.

The cost/volume relationship of a NGN seems to be shallower at current volumes than legacy networks suggesting that increases in volumes will have a relatively low incremental cost impact.

NRAs will need to consider adapt modelling and costing approaches in SMP markets. A key feature of a robust NGN model is likely to be the way in which it deals with the capabilities of the technology to deliver multiple services across a network with a high part of common costs. This suggests that NRAs will need to understand the cost orientation and cost recovery (pricing) implications of both SMP and non-SMP services running across the NGN platform.

In general, the cost of efficient service provision should be used as the cost standard for approval of interconnection rates. The pricing should be valid irrespective of whether interconnection is realized via circuit-switched or packet-switched networks, since strict application of the cost standard of long-run incremental costs requires the efficient technology used by the market players to be taken as a basis. Consideration must also be given to the fact that the concept of the cost of efficient service provision does not differentiate the price according to technology used or account for the existence of different prices for the same service. Basing prices on efficient technology also provides incentives for speeding up the migration to this technology.

Based on the hypothesis that the economic rationale for NGNs is partly based on the expectation that the costs of delivering voice services in the long run will be no higher (and probably significantly lower) than using legacy PSTN technologies, then it is reasonable for NRAs, in modelling/evaluating NGN costs and/or associated pricing decisions, to assume that the cost of voice services will be no higher than currently calculated.

Charging mechanisms[44]

Today's PSTN interconnection problems ("termination monopoly") results from the interplay of three factors: a) physical bottleneck for termination; b) control of the E-164 number; c) charging mechanisms.

---

44 See Section C.6

PSTN and mobile networks are governed by the charging mechanism of CPNP, where termination services are being paid for at the wholesale level following the direction of call flows. In IP-based networks, if wholesale payments for transit apply, they flow in the upstream direction. There are no payments for the terminating segment of the broadband access provider. This corresponds to the definition of Bill & Keep in this paper.

As networks migrate towards NGN infrastructure it is unclear *a priori* whether these future networks will be governed by the mechanisms currently used in IP-networks or whether the mechanisms currently applied in the PSTN will be carried over to NGNs.

Coupled with a direction of payment flows the charging mechanism may have implications on the definition of relevant markets and the determination of SMP.

CPNP and Bill & Keep differ with regard to the possibility to exploit the physical bottleneck for termination as has been explained in sections B.2.3 and C.3.2. :

- Under CPNP this bottleneck can be exploited because it entitles the terminating operator to receive a payment out of its physical termination bottleneck.
- With Bill & Keep, this is not possible as the access operator is *not* entitled to such a payment at the wholesale level out of his monopoly position. This major advantage of Bill & Keep affects some of the other issues addressed below.

Bill & Keep is a wholesale billing regime under which each network bears the costs of terminating traffic coming from other carriers. Therefore under Bill & Keep the terminating access network operator does not receive payments at the wholesale level for the termination provided. Instead, it recovers its net costs incurred for termination –- and any payments for upstream connectivity – in other ways, e.g. by billing them to its end customers.

Bill & Keep might be applied for the terminating segment up to the first router or switch and associated service control functions after the access/concentration network. Transit services are not included in the Bill & Keep model as discussed here and operators may charge for that service.[45]

Since termination costs are by definition not accounted for at the wholesale level symmetry of traffic flows cannot be considered a requirement for the applicability of Bill & Keep.

Bill & Keep for the last segment of termination of the broadband access provider requires no regulatory intervention as long as two conditions are fulfilled:

- The transit market on IP-backbones is sufficiently competitive to exert competitive pressures on IP-backbone providers. With an oligopoly of Tier 1 providers allowing choice of transit provider this condition has so far been considered to be fullfilled.

---

45 Applying Bill & Keep for the terminating segment up to the first router or switch and associated service control functions after the access/concentration network can be considered a minimum scenario. It may also be conceivable that the area of application of Bill & Keep might be extended by applying it already from PoIs higher up in the network hierarchy. See C.6.10

- The broadband access market is sufficiently competitive so that access providers are under competitive pressures to be prevented from establishing abusive mark-ups on retail prices.

Application of the CPNP regimes ultimately perpetuates the need for regulation of the termination rates .

The termination bottleneck is an essential ingredient of wholesale regulation. A shift of charging regime could significantly lower the burden of regulation.

Avoiding the bottleneck problem implies, that it is no longer necessary to determine the economically “correct” termination rates. Under Bill & Keep, lengthy and cumbersome regulatory and legal disputes (both, between market players and NRAs but also among market players) on the appropriate level of termination rates may be avoided.

On the other hand, Bill & Keep constitutes an approach which is more closely adjusted to market mechanisms, if end-users can choose the network carrier among various operators. In order for the advantages of Bill & Keep to become effective, competition in the broadband access/Internet access market is a precondition.

Given that Bill & Keep applies for the terminating segment up to the first router or switch and associated service control functions after the access/concentration network, hot potato routing applies on those parts of the network, that are excluded from the application of Bill & Keep, but where transit and peering agreements apply. Transit networks have been excluded from the applicability of Bill & Keep.

In order to qualify for the participation in the Bill & Keep regime a minimum number and location of interconnection points for a specific network operator has to be determined.

Wholesale and retail charging mechanisms are related because interconnection prices affect the structure as well as the level of the interconnecting operator’s costs, impacting on the cost recovery and the retail prices of the services provided to the end-users.

The use of a particular wholesale mechanism does not, however, preclude application of different retail pricing regimes. Both, CPNP and Bill & Keep provide flexibility at the retail level to offer retail schemes based e.g. on minutes, bits, or as buckets of minutes or bits plans as well as flat rates.

Empirical evidence shows that in countries with low mobile retail prices usage is higher than in countries with high retail prices. Retail prices are highest in countries applying CPNP as charging mechanism, whereas Bill & Keep countries exhibit much lower retail prices.

Concluding, it can be said that Bill & Keep has a number of attractive properties. Therefore it merits further study e.g. how to efficiently determine the minimum number of PoI for eligibility to participate in the Bill & Keep regime. This number PoI in turn determining the border of the Bill & Keep domain.

Summary of charging mechanisms and work plan

As networks migrate towards NGN infrastructure it is unclear *a priori* whether these future networks will be governed by the mechanisms currently used in IP-networks or whether the mechanisms currently applied in the PSTN will be partly carried over to NGNs.

Summarizing the preceding comparison of charging mechanisms with regard to a number of criteria, it can stated that Bill & Keep has a number of attractive properties. Assessing all the pros and cons of Bill & Keep, the ERG concludes, that Bill & Keep is a promising interconnection regime. It is supported mainly by theoretical reasoning and a large body of literature and economic modelling. Furthermore empirical evidence seems to suggest higher usage levels and lower usage prices are achieved with less regulatory intervention in Bill & Keep countries than currently applied in the EU.

Some NRAs therefore aim at a shift towards Bill & Keep because it reduces the regulatory burden and relies more on market forces, as is the case already now for today’s unregulated IP connectivity markets. They focus on further studying how to best achieve this goal including finding answers that arise in a transition phase to a new system.

Others, while recognising the merits of Bill & Keep in principle, rather emphasize the risks implied by a change from the well-established regulatory regime of mainstream PSTN and mobile services fearing disruptive change to the industry and therefore see a need for further study.

ERG identifies the following issues meriting further study which were generally supported by the Consultation responses:

Implications for different business models

Depending on how Bill & Keep is introduced the implications of a widespread introduction of Bill & Keep including the change of the cost recovery mechanism may imply that a transition from the current regime is a drastic and disruptive change for PSTN voice operators who have been subject to regulation under the framework. A rapid transition may not allow operators enough time to adjust there business models and retail price structure.

Operators may be affected differently by the introduction of Bill & Keep depending on, among other things, whether they are likely to experience a net cost (more incoming than outgoing traffic) or a net gain (more outgoing than incoming traffic). It has to be evaluated whether this may systematically shift the competitive balance between different operators, particularly incumbents and alternative network operators.

Furthermore implications for specific business models such as CPS have to be studied further, e.g. in case incumbents try to recover potential net-costs in part from CPS operators through higher origination charges.

The ERG acknowledges that changing the wholesale regime impacts on business models and may affect the competitive situation of the market. This is why an appropriate management of the migration towards all-IP infrastructures is important to avoid disruptive changes. Therefore, the ERG identified the implications for different business models, competition and subscribers as one of the issues meriting further study.

Practical implementation issues

- **Migration**

  Currently, different regimes for different types of networks (PSTN or IP) prevail – irrespective of service. As the separate network infrastructures are expected to converge to an all IP network such differences may not be sustainable in the long run.

  Considering the migration to all IP-networks it seems plausible to continue applying the charging mechanism of the networks that are not phased out.

  A meaningful discussion of migration problems, implies having come to a conclusion with regard to the charging mechanism applicable in the long term:

  - In case it is intended to carry over CPNP to NGN voice services this would imply different regimes for different services as a change of charging mechanism cannot necessarily be expected for the unregulated part of IP-networks applying Bill & Keep, Peering and Transit.

    This approach requires that it is possible to clearly distinguish between different services and that usage of services can be measured. If arbitrage problem are to be avoided, it would be necessary to mark different services or to transport them separately. Unless these preconditions are met there is a high risk of adverse selection, moral hazard and arbitrage problems.

    Instead of differentiating regimes according to services one might also envisage differentiation of different QoS classes (best effort vs. QoS level specified).46 Applying such an approach could be done by assigning different services to different QoS classes.

  - In case Bill & Keep was envisaged as long term goal it may be reasonable to further investigate regulatory options to soften the transition to the new regime. Strict application of cost orientation in a CPNP environment in the short term for mobile and/or PSTN networks can be seen as an important step in the migration towards Bill & Keep.

    The length of the migration period can be shorter

    - the lower the absolute level of interconnection rates,

---

46 See also ECC Report 75 (2005).

- the smaller the relative difference between interconnection rates of different networks
- the higher the proportion of flat rates at the retail level is.

- **Defining the border for application of the Bill & Keep regime**

  Applying Bill & Keep for the terminating segment up to the first router or switch and associated service control functions after the access/concentration network can be considered a minimum scenario in terms of the scope of network included within the Bill & Keep domain. However in order to qualify for participation in the Bill & Keep regime a specific network operator is required to access a maximum number of interconnection points in this case.

  It may also be conceivable that the area of application of Bill & Keep might be extended by applying it already from PoIs higher up in the network hierarchy requiring fewer interconnection points per network operator.

  The application of Bill & Keep may essentially require a determination of the topology of points of interconnection as has been the case for the PSTN when determining the number of PoI. The question of how to efficiently determine this minimum number of PoI, i.e. the border of a potential Bill & Keep domain has to be further investigated.

- **How to treat traffic from outside the Bill & Keep area and prevent extensive arbitrage (tromboning, call-back etc.)**

  - Between different countries
  - Between different networks (e.g. fixed/mobile)

  An important issue is how to handle traffic coming from outside the Bill & Keep domain (hereafter: incoming traffic). This traffic could result in problems. For example if the operators inside the Bill & Keep domain want to set a termination rate for incoming traffic, this could be forestalled by competition for receiving this incoming traffic. Every operator inside the Bill & Keep domain would have an incentive to receive incoming traffic and collect a fee for this and then route this traffic towards the final destination and dropping it there for free.

  This problem could be prevented if, subject to the number portability arrangements in place, receiving networks could effectively bill the incoming traffic based on where traffic originated, for example by using the network number of the source network.

  The dimension of these problems increases

  - the larger the traffic volume from outside relative to the traffic exchanged between networks inside the Bill & Keep domain.

- the higher the termination rates outside the Bill & Keep area

Other long term forms of regulation

Are there forms of voluntarily achieving Bill & Keep through a series of other measures and requirements that could be based on symmetric regulation according to Art. 5 AD (e.g. reciprocity) used in combination with termination rates strictly regulated at cost based levels?

Introduction

A.1 Next Generation Networks

The introduction of Next Generation Networks (NGN[47]), leading to a multi-service network for audio (including voice), video (including TV) and data, as well as new plans and investment in next generation access (NGA) sets the communications sector on the verge of a new era. These developments give rise to innovation opportunities at both the service and infrastructure level and may subsequently impact significantly on market structure. Additionally, due to the increased economies of scope of a multi-service network, cost savings are to be expected.

These developments can take different forms. Traditional PSTN operators including incumbents, plan to migrate towards NGNs, relying on the ITU-T[48] and ETSI[49] as relevant standardization bodies. The migration towards NGNs does not only relate to fixed network operators but also to mobile operators.[50] On the other hand, independent ISPs and ITPs (Internet transport providers) continue to develop their IP networks towards multi-service networks, relying more on the IETF as standardization body.

The term NGN was developed and defined by ITU-T[51]. It is, however, also used as a more general slogan for the use of IP technology when converting the telecommunications networks from traditional circuit-switched to packet-switched technology. An IP network that uses some deliberately chosen elements that are specified in NGN standards for improving its transmission performance may, therefore, also be referred to as NGN. Today the term NGN covers a broad performance spectrum from simple implementation of TCP/IP with low level best effort performance to intensive implementation of traffic management methodologies providing high level and stable transmission performance.

Since the liberalization of the telecommunications markets, network interconnection has been one of the basic requirements for enabling competition, because this is the only way a provider can make it possible for its own subscribers to communicate with subscribers of another network.

NGNs do not change the fundamental importance of interconnection to sustainable competition in both network infrastructure and of electronic communication services. However, taking into account the migration towards all-IP networks compared to arrangements in the PSTN,

---

47 A list of definitions and acronyms is presented in the Glossary.
48 See ITU-T information on NGNs in http://www.itu.int/osg/spu/ngn/.
49 See ETSI “TISPAN PUBLISHED NGN SPECIFICATIONS“ in http://portal.etsi.org/docbox/TISPAN/Open/NGN_Published/PUBLISHED_NGN_ SPECIFICATIONS.doc.
50 Thus, the scope of this document covers fixed and mobile networks. When referring to “legacy networks” here, this generally relates to both types of networks.
51 NGN is defined by ITU-T (Rec. Y.2001) as follows: „A packet-based network able to provide Telecommunication Services to users and able to make use of multiple broadband, QoS-enabled transport technologies and in which service-related functions are independent of the underlying transport-related technologies...”. In addition, certain criteria („fundamental aspects“) are specified that must be met by an NGN.

changes, may be expected at the wholesale level, as the markets for network conveyance converge and service intelligence is decoupled from the (transport) network. With the gradual replacement of traditional interconnection by IP interconnection the question of possible future interconnection models becomes relevant.

A.2 Objectives

This "ERG Common Statement on Regulatory Principles of IP-Interconnection/ NGN Core" develops some general regulatory principles focusing on the core network in order to ensure a consistent application of the European regulatory framework. The provisions of the Regulatory Framework – such as the market analysis procedure - require that regulatory actions taken by NRAs are clearly evidence based. In case SMP has been found in a market susceptible to ex-ante regulation, regulatory obligations imposed must be appropriate and proportionate in relation to the nature of the problem identified.

The aim of this document is to present the current situation with regard to IP interconnection in Europe and to outline how the development towards NGNs may affect market regulation[52] in this matter. This document will analyse the effects this evolution might have on interconnection regimes and develops some general principles with regard to regulatory treatment of IP interconnection and interoperability issues reflecting the development towards multi-service NGNs. The ERG intends to increase regulatory certainty for all market players at this time when many stakeholders are considering or actively moving towards NGN implementation.

A.3 Scope and structure of the document

This document is based on the ERG report on IP interconnection (see ERG (07) 09) published in March 2007, tackling IP interconnection and the implications of a move to this, as one of the main challenges emerging out of the developments towards multi-service NGNs in the core network and also takes into account more recent developments. In addition, some work of the Regulatory Accounting PT has been annexed to this document.

The document was published for consultation on the ERG website (consultation period June 4 2008 – July 11 2008). 21 responses were received (one of these confidential). The document has been revised in the light of these responses.

A Consultation Report, summarizing the main arguments in responses received to each question will be published as Part 1 in a separate Supplementary Document. Annexes 2-4 of the Consultation Document (country studies, technical background information, implications on regulatory accounting) will constitute parts 2-4 of the Supplementary Document. The points which received particular emphasis in the responses, including separation of service

52 Issues regarding technical regulation e.g. network integrity or security are not addressed here.

and transport and charging mechanisms, are dealt with briefly in this document and will be considered further in the Supplementary Document.

Since the publication of the ERG IP-Interconnection report, a number of NRAs have continued to work on NGN interconnection, monitoring progress in industry forums, engaging with market players and developing policies and procedural provisions.[53] One NRA has decided to require SMP operators in (old) Markets 8-10 to provide IP interconnection. This obligation has not yet been fully implemented.[54] Another NRA in a draft decision for the market on fixed voice termination defined in a technologically neutral manner determined that SMP operators have to provide interconnection on a regional level at no more than 20 PoI.[55] No other NRA has yet taken a definitive decision on a native IP interconnection product.

In view of the multi-service nature of NGNs and the decoupling of service[56] and transport, the paper will look at IP interconnection in general, and is not confined to voice interconnection.

Issues relating specifically to voice interconnection in future networks will nevertheless be treated in some detail as traditional network operators are mainly preoccupied with migrating their voice service to NGN/IP networks. Furthermore VoIP is the service best known to telecommunications network operators, whereas other services like VoD, IP-TV or presence services are still in early stages of evolution, and are less well understood at the service level. However, VoIP traffic already constitutes a relatively small fraction of overall IP traffic, in the future it is to be expected this fraction will continue to be small in comparison to other services.[57]

Given that the migration towards NGNs is in different stages in different countries, migration issues are not dealt with in detail in this paper[58] and are identified as a topic meriting further study.

Furthermore, the paper will explicitly not look at access to electronic communication services. Access has been addressed by the CP NGN Access. The distinction between access and interconnection can best be explained[59] as follows. Access enables an operator to utilize the facilities of another operator in the furtherance of its own business and in service of its own customers, while interconnection enables an operator to establish and maintain communications with customers of another operator.

---

53 See e.g. NGNuk (2007), AKNN (2007), Bundesnetzagentur (2008).

54 For details, see Italian Country Case Study in Annex 2.

55 If a competitor grants access at less locations, the same interconnection tariffs applies as if KPN had offered access at the 20 regional PoI. See http://www.opta.nl/download/201452ontwerpbesluit+marktanalyse+vaste+gespreksafgifte+%28telefonie%29.pdf.

56 In this document the term service will be used in a wide sense signifying services provided to the end-user. Therefore use of the term service is not confined to the service layer described in some NGN documents but includes parts of the control layer (relevant to services), the service and the application layer.

57 Cisco estimates global consumer Internet traffic until 2011, cf. WIk-Consult (2008), p. 42.

58 This was also pointed out in several responses, see Consultation Report.

59 See WIK report The future of IP Interconnection 29-01-2008 paragraph 1.2.1

After providing an Introduction in Part A, which inter alia outlines the drivers for change towards all-IP networks, Part B briefly sets out some background and facts necessary to analyse the regulatory implications of developments towards NGN:

- it provides a summary of country studies;
- it describes and compares existing interconnection arrangements in the PSTN and the IP worlds;
- it describes different charging mechanisms and derives their impact on the possibility to abuse market dominance;
- it sets out main features of network structure and topology, including number and hierarchy of network nodes and hence possible interconnection points;
- it explains how quality of service may impact on interconnection.

Part C deals with the regulatory implications and challenges resulting from the expected developments towards NGNs:

- it starts setting out the relevant sections of the current framework and looks at the implications of the changes foreseen in the Review;
- it analyses how the separation of transport and service, the possible introduction of network performance classes and charging mechanisms may impact on the definition of relevant markets;
- it identifies possible bottlenecks and SMP positions at the transport as well as the service level;
- it identifies regulatory measures regarding symmetric regulation, based on USO directive and SMP-remedies
- it identifies probable changes of costs and appropriate pricing mechanisms in case an SMP-position is determined and *ex-ante* regulation imposed;
- it analyses different charging mechanisms (CPNP/ Bill & Keep) as core element of an interconnection regime for termination service and presents a work plan for further study.

Furthermore, a Glossary is provided in Annex 1.

A.4 The starting point: Interconnection in legacy networks and in existing IP networks

Interconnection is the physical and logical linking of public communications networks used by the same or a different undertaking in order to allow the users of one undertaking to communicate with users of the same or another undertaking.[60] The terms interconnection or interconnection services used subsequently include any additional agreements related to interconnection such as supply, transit or termination, in addition to the actual traffic handover at the point of interconnection, because these aspects are dealt with together with network interconnection in terms of market regulation.

Currently interconnection in the PSTN and IP networks are characterized by a number of important differences. The use of different technologies led to different interconnection products. Different charging mechanisms and different regulatory regimes have developed.

PSTNs were designed to be capable of establishing connections (channels) with fixed transmission characteristics. Thus, they provide a fixed transmission performance since they are circuit-switched and use pre-defined PCM channels. When first established, PSTN networks had a clear focus on voice services (additional services such as fax and dial-up Internet access were introduced later). Interconnection between “traditional” circuit-switched networks reflects this focus on telephone services as transport and service are closely linked with each other.

An IP network is an all-purpose network based on packet switched technology using the Internet protocol. It provides a platform for the delivery of multimedia services. In principle, any service can be realised with a specified quality level, if the performance objectives of the service can be met by the network. Since packet switched networks are designed to “only” provide transport resources and to support any service, a separation between transport and service layer can be made. When discussing aspects of interconnection this has to be addressed.

The most common wholesale billing regime in the PSTN is Calling Party’s Network Pays (CPNP), usually covering transport and service. Under this system, the network of the party who placed the call (the originating network) makes a payment to the network of the party that received the call (terminating network). Thus, at the wholesale level the whole call is paid by the caller’s network. The terminating network provider has an incentive to charge high termination rates (sometimes this strategy is called raising rival’s cost), because it is not its own customer who finally has to pay them. The rationale of CPNP is based on the assumption that the costs are caused solely be the calling party’s network.

Interconnection arrangements in IP-based networks exist either in the form of transit, peering or Internet Exchanges (IX). The direction of traffic flows does not play a role for these ar-

---

60 Definition based on Art. 2 b Access Directive. In brief some of the most important terms used throughout the paper such as interconnection, quality of service, network performance, best effort or interoperability are given in the Introduction. For a fuller definition of the terms see Glossary.

rangements. Traffic flows in both directions are added in determining charges. Therefore, there is no need to distinguish between origination and termination for billing purposes. In transit agreements, the Internet/broadband access provider pays for connectivity to the upstream network for upstream and downstream transmission of traffic. In peering agreements, normally there are no payment flows, as long as traffic imbalances do not exceed a certain, specified limit. Those ISP who fulfil the requirements for peering can choose between peering and buying transit services. The market is generally taken to function more or less competitively as long as broadband access providers have a choice of transit providers.

The way transit and peering agreements work implies that the access provider is not entitled to any payment when taking over traffic at his agreed PoI and physically terminating a data flow, e.g. a VoIP call on its network. Such a wholesale regime, where each network bears the costs of terminating traffic coming from other carriers itself, is called Bill & Keep. The carrier will bill these termination cost on its network and any payments for upstream connectivity to its customer. As long as there is sufficient competition for broadband access at the retail level, the access provider has an incentive to keep transit cost low, since too high a cost, if passed on to the end-user, may induce the latter to change supplier.

Interconnection in the PSTN is subject to both symmetric regulation (obligation to interconnect for all market players) and asymmetric regulation (SMP parties mostly subject to *ex-ante* cost-based regulation). The case for regulation was considered strongest for call termination markets. Absent regulation the so called “termination monopoly” of the terminating network operator arises due to the interplay of several factors, namely the physical bottleneck position to terminate the call, coupled with the control of the E.164-number and the CPNP mechanism entitling the terminating network operator to a termination fee from the calling party’s network operator.

In IP networks interconnection developed without any regulatory intervention although the obligation to negotiate for interconnection applies to IP networks as well. These agreements have been largely outside the scope of activity of NRAs and there are no known filings of disputes.

The crucial question for NRAs will be how interconnection evolves in future networks with the different regimes coalescing. Future IP interconnection products in an “all IP” environment – at least for telephone services – could be governed either in a PSTN-like manner or in the varieties of IP interconnection as they are known today.

Among other potential issues in voice (telephony) services, the interplay of the following factors in an NGN world should be carefully analysed:

- Physical bottleneck for termination

  Depending on the degree of competition at the access level, the broadband access provider may have the option of barring traffic to and from particular servers. Thus, this form of market power refers to the ability to deliver any service (not only voice);

- Control of the E-164 number (the subscriber's telephone number)

  It confines market power to the service provider as long as it is not possible for more than one provider to terminate calls to a specific number;

- Role of the charging mechanism (e.g. CPNP)

  The extent to which the market power enabled by the physical bottleneck for termination can be abused depends on the charging mechanism and the direction of payment flows.

A.5 Drivers for change

Future electronic communications networks will be packet-switched, mostly or completely based on the IP. They will be multi-service networks, rather than service-specific networks, for audio (including voice), video (including TV services) and data rather than service-specific networks, allowing a decoupling of service and transport provision.

Developments towards all-IP networks can take different forms. On the one hand, traditional PSTN network operators including incumbents plan to migrate towards NGNs, relying on the ITU[61] and ETSI[62] as relevant standardization bodies and on the other hand, independent ISPs and ITPs continue to develop their IP networks towards multi-service networks, relying more on the IETF as standardization body.[63]

Important features of NGN are:

- separation of transport and service;
- assurance of Quality of service;
- cost savings;
- network topology.

A.5.1 Separation of transport and service

A core feature of NGN architecture is the separation of the main functional levels, i.e. generally a distinction can be made between transport and service. This separation of transport and service constitutes an essential feature of NGN specifications in both the specifications on NGN (ITU-T) and NGI (IETF).[64] A crucial point is the adoption of open and standardised

---

61 See ITU-T information on NGNs in http://www.itu.int/osg/spu/ngn/.
62 See ETSI "TISPAN PUBLISHED NGN SPECIFICATIONS" in http://portal.etsi.org/docbox/TISPAN/Open/NGN_Published/PUBLISHED_NGN_ SPECIFICATIONS.doc.
63 See IETF working groups "Session PEERing for Multimedia INTerconnect –speermint" (http://www.ietf.org/html.charters/speermint-charter.html).
64 See e.g. the description of ITU architecture in ERG (2007), Ch. A1.1 (p. 39). For a comparison of NGN and NGI see Hackbarth/Kulenkampff (2006), Ch. 3.4.

interfaces between each functional level in order to allow third parties to develop and create services independent of the network. More specifically ITU-T Rec Y.2001 (2004) defines a number of necessary criteria for an NGN among them "separation of control functions from bearer capabilities, call/session, and application/service" and "unrestricted access by users to different service providers":

In the present Internet this allowed easier service creation (as it lowered entry barriers) which contributed to the success of the Internet and TCP/IP protocol suite.

The architecture of NGNs can include the principle of separation of transport and service that has so far been characteristic of IP networks. This principle could also support service provision by third parties by providing open interfaces, even though but transport could become service-specific to reflect the different network performance needs of different services (real time etc.).

This separation of transport and services is also expected to be reflected in the respective interconnections services, i.e. service interconnection and transport interconnection.[65]

- Transport interconnection includes the physical and logical linking of networks based on simple IP connectivity irrespective of the levels of interoperability. It is characterised by the absence of the service-related signalling, implying that there is no end-to-end service awareness. Consequently, service-specific quality of service and security requirements are not necessarily assured. At the transport level, only objectives for those parameters[66] are negotiated that affect the transmission performance at the point of interconnection (e.g. availability) and the IP packet transmission performance via interconnected networks.

  This definition does not exclude that some services may provide a defined level of interoperability or the establishment of transport classes to guarantee specific quality parameter at this level.

- Service interconnection in this paper is understood as including solely service-specific aspects.[67,68] It consists of logical linking of network domains, having access and control of resources including the control of signalling (i.e. session based service-related signalling[69]). Depending on the kind of service, different aspects must be considered. For example, in the voice service, the call server interconnection is required for call setup and disconnection.

---

65 See Ch. 1.3.1 (p. 5) stating that with IC taking place at separate levels, this may require defining different IC services.

66 For e.g. bit rate, delay and packet loss ratio.

67 This differs from ETSI/TISPAN's definition of 'service oriented interconnection' also including transport related information. See.below Section 3.2 in Annex 3.

68 This may have an important implication. Given this understanding, service interconnection may not fulfil the definition of Art 2 (a) Access Directive (see Sec. C.1.1) because it does not necessarily include the physical linking of NGN domains (see Sec. C.3.2).

69 Information that allows identification of the end-to-end service that has been requested.

Interconnection between services from different operators requires a minimum set of technical (e.g. defined by a SLA) and commercial conditions to be fulfilled by both operators.

These conditions may include *inter alia*:

(iv) mutual policies for exchange of data (including transcoding the information mapping of quality of service information (if applicable), service control information and network protocols);

(v) agreement of charges;

(vi) agreement on performance and reliability levels.

Other aspects such as security for example, may need to be considered.

If transport interconnection and service interconnection were bundled, as was the case in the PSTN, we may talk about transport+service interconnection (corresponding to the concept of Service oriented Interconnection according to ETSI/TISPAN).

In most NGNs planned by incumbents, services tend to be provided using centralized platforms (Media Gateway, Softswitch). Operators who have market power may not have an incentive to open their networks to competition at the service level and may want to limit use of these capabilities.[70] This impacts on the ability of independent service providers to integrate their services into the NGN platform.

Furthermore, such a configuration of services and the centralization of the control function have implications for the locations at which traffic can be handed over to other networks or received from other networks. This feature will therefore be crucial for the possibility of defining interconnection points.

Therefore, NRAs may have to ensure that interconnection is possible at specific functional levels in a reasonable manner. This may prove to be particularly challenging since a number of “telco” network operators are of the view that a horizontal separation of transport, service (and control) levels is neither appropriate nor in their interest, particularly if they want to guarantee quality of service (see IMS[71], adaptation of IMS to fixed networks, etc.). Most respondents to the Consultation corroborated that their understanding of NGN implies a continuation of vertically integrated provision of transport and services as has been the case in legacy “telco” networks.

These operators argued that a complete separation of service and transport was not reasonable and even counterproductive if the aim is to guarantee QoS, to offer a PSTN substitute and security to the customer. They concluded therefore, that coordination between transport

70 See Katz/Shapiro (1985)
71 See Ericsson’s White Paper, “IMS – IP Multimedia Subsystem” (2004)

and service is required. This conclusion derives from a divergence of their understanding of NGNs from the ITU-T's understanding.

This divergence of their understanding of NGN from the NGN concepts of ITU-T and ETSI-TISPAN becomes explicit in the NGN interconnection concept developed by a German network operator association.[72] In this concept several criteria of the NGN definition of ITU-T are not considered necessary features for NGNs such as "Separation of control functions from bearer capabilities, call/session, and application service" or "unrestricted access by users to different service providers".

If network operators continue applying a vertically integrated approach, implying a simple transfer of the PSTN world to the NGN world, then the same kind of regulation may also need to be transferred respectively. In this case, the potential deregulatory opportunities which could arise from the separation of transport and service may not materialise. If in this case the underlying competition problems remain and/or potential new ones arise, regulation will continue to be required.

The ERG supports the separation as expressed in the NGN concepts of ITU-T and ETSI-TISPAN. Also the OECD supports this notion of separation, stating that "*the separation of networks functional planes should allow for the creation of a horizontal platform for the provision of services, separated from the transport layer. For this separation to be effective, interconnection should be possible at all functional levels. However, there is the risk that operators do not consider horizontal separation appropriate, as it is more difficult to guarantee a certain level of quality of service in interconnected networks, or simply because it is not in their best interest. Most incumbent operators still see NGN as a simple continuation of vertically integrated transport and services, as in the case of legacy networks*.".[73]

The ERG clarifies that this separation does not imply that both levels are completely independent. Even with a logical separation between transport and service, there is co-ordination and interaction between them. What is key for competition, however, is that the separation should allow transport and service to be provided by different parties. Service provision by independent third parties becomes possible, independent of transport technology and type of network access. This approach requires open interfaces for third parties.

The ERG is convinced that such a separation between transport and service would contribute to and promote the development of new and innovative services.

---

72 See AKNN (2007). This "Working group for technical and operational numbering and network interconnection issues" (AKNN) consists of operators and manufacturers.

73 OECD (2007, p. 27)

A.5.2 Assurance of quality of service

Another important distinctive feature of NGN relative to current IP networks is the expected ability to provide a range of service quality levels. However, the implementation of NGN has yet to prove to what extent this is technically and commercially feasible.

This ability could be key in enabling a wide range of services to be carried using a single network infrastructure and hence realizing the economies of scope that could be achieved. Real-time services including conversational and interactive services (for example videoconference calls and on-line gaming) tend to convey ever-richer media content and therefore require ever higher bit-rates. Furthermore, multiple services can be combined into a single user's session. A network operator has to take into account the performance objectives of these services and plan and operate its network accordingly. For this purpose, it may want to implement QoS mechanisms.

The term quality of service (QoS) is a broad concept (see Glossary) as it covers all aspects influencing the user's perception of the quality of the service; related to the network performance (such as reliability or availability) as well as to other factors, such as terminal equipment[74], codecs or customer support. However, the term QoS is sometimes used when referring to part of these aspects and often misunderstood as a synonym for network performance, which is defined and observed as performance of a telecommunication network (or sections of a network) by using objective performance parameters.

In the current practice of interconnection of IP networks, even though there are QoS mechanisms implemented for internal purposes, there is no technical implementation of QoS mechanisms across interconnected networks resulting in a "best effort" transmission performance, i.e. no service prioritisation. This "QoS model" is characteristic of today's IP network but does not necessarily result in a low transmission performance and thus a low quality of service.

PSTN interconnection did not entail any formal QoS guarantee either, but transmission plans were established over many years to provide a comparatively high assurance of voice service quality at the expense of greater complexity.

If a guaranteed QoS over IP-based networks with a PSTN-comparable level of performance is desired, one has to modify and adapt the IP transport model in a way that connections with reliable and fixed transmission characteristics (transport classes) are possible. Moreover, the ability to provide services with guaranteed QoS may not only be applied to a single network but must be maintained over the whole chain of interconnected networks involved in the provisioning of a specific end-to-end service.

It is not clear yet which strategies will be chosen by the different operators. Depending on the strategy (stringent systems versus systems guaranteeing QoS on a statistical basis only),

---

74 For instance by selecting and configuring the terminal equipment the user can substantially affect the resulting end-to-end quality.

specific additional agreements and mechanisms may have to be implemented. Interoperability, i.e. the ability of systems and services to exchange information and mutually use the information that has been exchanged[75], is necessary to make interconnection work in practice.

The implications of different transport classes will need to be analysed in some detail, making NGN interconnection potentially more complex. QoS is therefore potentially a new dimension in the interconnection of NGN, and could be an important focus for NRA because it could enable new forms of discrimination between a larger operator's services and those provided by interconnecting competitors.

A.5.3 Cost savings in NGNs

The development of NGN core networks offers opportunities for increasing efficiency and for innovation. Economies of scope can be expected to result from the convergence of services onto a single network infrastructure which will convey them potentially leading to a reduction of total costs. At the same time, the capacity and processing power of modern network equipment can result in increased economies of scale and in innovation in services and in packaged retail propositions. The greater economies of scale in NGNs are likely to increase the challenges in promoting effective competition in the provision of services solely reliant on core infrastructure.

Therefore, NGN technology is expected to save costs (cheaper equipment, economies of scale and scope). The cost structure of NGN core networks is likely to contain a relatively high proportion of costs common to all services carried by the network, and a relatively low proportion of costs directly driven by the volume of each service (see Sec C.4 and Annex 4).

A.5.4 Network topology

It is yet unclear in which form exactly the implementations of NGNs will occur and they will differ in Member States according to investment cycles, network size and other factors. Operators may rather "pick and choose" among the different architectures/features standardized in ITU and ETSI or other bodies.

This relates for instance to the implementation of transport classes relying on prioritisation versus best effort management with adjusted network dimensioning.

Furthermore, it is not yet clear as of now to what extent features such as the separation of service and transport will remain possible in the architectures finally implemented.

A number of operators intend to implement their NGN using centralized platforms for service provision, affecting the ability of independent service-providers to integrate their services into the NGN platform. Whether independent service providers will be able to do so also depends

---

75 Definition based on ITU-T Rec. Y.101

on the availability of open and standardised interfaces. Furthermore, such a configuration of services and the centralization of the control function have implications for the locations at which traffic can be handed over to other networks or received from other networks.

The separation of transport and services will be crucial for the definition of the interconnection points[76] and (number of) nodes. In any case, an immediate manifestation of the increased capacity of network equipment is likely to be a reduction in the number of nodes in which routing of traffic will occur in NGNs when compared with the number of nodes where traffic is switched today by traditional networks.[77]

However, currently operators consider to continue the implementation of interconnection for voice services in a bundled manner – combining transport and service - during the transitional period towards NGN.

It may have to be taken into account that short-term determinations that are only related to voice services do not reflect the multi-service nature of NGNs and therefore risk hampering or delaying introduction of the NGN principle as an overall concept and not fully exploiting the competitive potential offered by NGN in spreading innovative services rapidly.

A.6 Core elements of interconnection

An interconnection regime contains at least the following core elements:

- number and geographic location of interconnection points as well as functionality and hierarchy of these interconnection points. They are determined by the network architecture (for the network topology and functional layers in NGNs see Sec. B.4.4);
- the network performance needs to be specified (e.g. best effort and/or some additional transport classes specifying the relevant network performance parameters like jitter, delay and packet loss (see Sec. B.4.5);
- furthermore, an interconnection regime is characterized by the charging mechanism used, i.e. who pays for which part of the value chain. This has an impact on the market power that can be exerted by different market parties at different levels of the value chain (see Sec. B.3);
- in case interconnection is regulated according to *ex-ante* cost based regulation, costing and pricing principles have to be discussed (structure of tariffs according to hierarchy, accounting units such as minutes, bandwidth, etc.) (see Sec. C.5).

76 ERG (2007), p. II
77 ERG (2007), Ch. 1.3 (p. 6) and Ch. 3.3 (p. 16)

B Background and technical information

B.1 Main results of country case studies[78]

a) Relevance of IP-interconnection

In most countries, and in general, there are currently no regulations governing IP interconnection, although several parties have completed peering agreements (both direct or using NAP), solely through commercial negotiation. In several countries [CY, FR, GR, IP, AU, PO, PT, SW, RO] IP interconnection is in a relatively early stage of assessment by both regulator and operators. In other countries [NL, DE, NO and IT], considering the increasing use of VoIP and the migration towards all IP networks, IP interconnection increasingly gains relevance. In particular, in [NL], where several parties (the largest cable operators) have set up direct IP-based interconnection through a common exchange, and in the [UK], where a number of network operators have deployed IP and MPLS technology in multi-service core networks. Also, in [NO] it is expected that the incumbent will soon launch an IP interconnection offer. In [RO], national and main city IP-MPLS backbones have already been deployed.

In [IT], there is already an obligation (which is not yet fully implemented) to the SMP Operator to provide an IP interconnection offer (temporarily adopting the same PSTN economical conditions, to the extent that an agreement is reached on the technical specifications to allow direct IP-IC). In addition, there are symmetrical obligations for (all) the operators to adopt the most efficient way to interconnect networks and to give access to their technical interface/protocols and to all the technologies necessary to allow interoperability of VoIP services. Moreover, standard protocols should be adopted whenever possible.

b) Complaints from competitors/disputes

In most countries, no formal disputes or complaints have been presented from competitors in relation to this matter. However, small nomadic VoIP providers in [IT] are asking for the definition of technical interfaces for direct IP interconnection. In [PO] there was one dispute concerning prioritization of IP traffic and in [AU] several operators complained about a deterioration of their broadband services caused by deployment of the incumbent's DSLAMs, but not in the direct context of IP interconnection. OPTA, recently, has seen one or two complaints from smaller market parties regarding IP interconnection with the incumbent, but no formal proceedings were initiated.

c) Actions taken or planned by NRA with regard to NGN core and or IP-interconnection

In several countries [AU, CY, GR, PT, DE] the NRA itself has recently (earlier in [DE]) started the discussion process on NGN/IP Interconnection, including regulatory challenges and options (including wholesale interconnection billing models, migration issues, points of inter-

---

78 Part 2 of the Supplementary Document provides a comprehensive version of the country case study updates.

connection, etc.). In [DE, NL, UK and IR] advisory/industry groups were created to analyse the framework conditions for interconnection of IP-based networks, from both technical and commercial, with the goal of developing an agreed vision of interconnected NGNs delivering a wide range of services. Public consultations have been conducted in [DE, and, mainly, in UK and IT] and in [DE] the “Key Elements of IP Interconnection” were published. In [NL, NO and UK], the regulator is monitoring the development of IP strategy of the incumbent.

In [IT], a proceeding for the definition of technical interfaces for IP interconnection is currently being carried out. A first working paper on IP interconnection has been put to consultation by AGCOM with several responses being received from stakeholders regarding QoS, NAP, ENUM, pricing models, protocol requirements, etc.

d) Number of network nodes and e) Number of interconnection points

In several countries [AU, IR, PO, PT, SW, RO], currently, there are not: (i) specific plans to change the existing number of network nodes or number of PoI (e.g., in [SW] the number of TDM PoI is always 36, 2 per region, and in [IR] the number of nodes that are involved in IP-IC is less than 6); or (ii) information about the way the networks (architecture and topology) will develop; or (iii) migration scenarios.

In other countries [DE, IT, NL, NO and UK] it is expected that the number of PoI in packet-based networks will decline compared to switch-based networks. In [DE], 100 IP core network nodes are considered the upper limit for the number of IP PoI, although some providers consider this figure “highly inefficient” because IP networks operators used to exchange traffic at 1 to 3 points. In [IT] there are 12 PoI for future IP interconnection with the NGN of the SMP operator. In addition, most NAP (commercially agreed) IP peering occur in 2 points (Rome and Milan). In the UK BT expects to interconnect its NGN voice services with other operators at 29 physical locations.

f) Definition of local interconnection

In most countries, there is, currently, no definition of local IP interconnection (or knowledge of local IP interconnection arrangements in place). Also, in [AU] the definition of local interconnection will remain as it is for the time being. However, in [DE] it is assumed that the existing tariff structure (local, transit, double-transit) is likely to become obsolete in future networks and in [SW] local interconnection can apply in the context of LLU (local exchanges level/concentrators level/street cabinet level), but not for optical fibre. In [NO] local IP interconnection is understood to be the exchange of traffic at the lowest level in the network hierarchy. The PoI or accesses are likely to vary for different types of services (e.g. VoIP interconnection might be on the signalling level).

g) Migration scenario

In several countries, information about (official) plans for the migration process into NGN (e.g., Swisscom plans eventually to introduce IMS in its network, but it is waiting) is currently not available. In [AU], an NGN industry working group will be established in 2008 to further discuss aspects of migrating incumbent’s PSTN to a NGN. In this process, for [DE], the transparency on the further development of networks, i.e., on the incumbent's migration plans, is a crucial requirement to assure correct and efficient investment decisions. In [SP] Telefónica has communicated his interest in adapting the Reference Interconnection Offer (RIO) because of the foreseen commercial availability of services based on NGN architecture. In consequence, CMT is currently in a process with operators to analyze the implications and study the convenience of establishing a new NGN interconnection regime.

On the other hand, in [NO, RO], although no migration scenarios have been drafted, the IP-based NGN is being introduced in parallel to the existing PSTN. In [UK], where the migration process is more advanced, the migration to BT's 21CN is still subject to industry discussion, although it is expected that broadband services would be deployed first in dense exchanges in 2008 and fully emulate PSTN services on 21CN, and all customers migrated, over the period 2008-2012.

B.2 Existing interconnection arrangements

B.2.1 Interconnection in the PSTN and mobile networks

Following the current European Regulatory Framework, interconnection in the PSTN and mobile networks has been subject to regulation in all Member States. To ensure any-to-any connectivity, a regulatory requirement to negotiate interconnection is foreseen (Art. 4.1 – current and proposed - Access Directive).

Where an operator is designated as having SMP on a specific market as a result of a market analysis NRAs may impose an obligation "to interconnect networks and network facilities" (Art. 8.2 in combination with Art. 12.1 i Access Directive)

The current Recommendation on Relevant Product and Service Markets considers "Call origination on the public telephone network provided at a fixed location" (Market 2) and "Call termination on individual public telephone networks provided at a fixed location" (Market 3) markets susceptible to ex ante regulation.[79] The Explanatory Note to this Recommendation classifies call termination as the "least replicable element in the series of inputs required to provide retail call services". As in the initial Recommendation, the current Recommendation demarcates the relevant market as wide as each network operator. The Explanatory Note states that "alternatives for demand or supply substitution do not appear currently to provide sufficient discipline on call termination at fixed location" and views the 3 criteria test to be fulfilled.

Legacy telephone networks were designed to be capable of establishing connections (channels) with fixed transmission characteristics. Thus, they provide a fixed transmission performance since they are circuit-switched and use pre-defined paths and channels. When

---

79 Different from the current Recommendation the former Recommendation additionally classified "Transit services in the fixed public telephone network" (Market 10) as susceptible to ex ante regulation.

established, PSTN networks had a clear focus on voice services. Nevertheless, it is to be considered that they were also used extensively for narrowband dial-up Internet access later.[80] Interconnection between “traditional” circuit-switched networks reflect this focus on voice services, as transport and the voice service are closely linked with each other.

The most common wholesale billing regime in the PSTN is Calling Party’s Network Pays (CPNP), under which the network of the party who places the call (the originating network)[81] makes a payment to the network of the party that receives the call (terminating network).[82] Thus, at the wholesale level the whole call is paid by the caller’s network. The rationale of CPNP is based on the assumption that the costs are caused solely be the calling party’s network.[83] Under CPNP usage is mostly billed on the basis of Element Based Charging, but there are also countries applying Capacity Based Charging (See Section C. 5).

Absent regulation the CPNP mechanism provides inherent inventives for the physical bottleneck for termination to be exploited because it entitles the terminating operator to receive a payment out of this position.[84] (see Chapter B.2.3). Termination charges are therefore often set by regulatory intervention to bring them in line with costs. Ultimately, the application of CPNP implies a need to maintain regulation on termination rates indefinitely.[85]

The following figure 1 shows in a schematic manner call and payment flows in PSTN networks. It is assumed that end-user A (being a customer of operator 1) makes a phone call to end-user B (customer of competitor 2). The networks of operators 1 and 2 can either be directly interconnected, or be interconnected indirectly via the network of operator 3. The latter situation is more common in practice, as a direct interconnection will often not be possible and competitors interconnect indirectly using the SMP operator’s network to provide transit services. In this case, a call is conveyed via the network of the SMP operator and then terminated on the network of operator 2 (dashed line). Wholesale payments follow exactly the same direction as the call. Operator 1 pays the SMP operator for transit, and operator 2 is paid for the termination provided.[86] The same reasoning applies in case of a direct interconnection between the networks of operator 1 and 2.

---

80 Before the advent of broadband Internet access dial-up Internet accounted for up to 30% of all narrowband traffic in some countries. Fax is another services that is provided over the PSTN.

81 A provider of call-by-call or preselection services also has to buy origination from the access operator of the caller.

82 The principle of CPNP is also applied for regulating “Voice call termination on individual mobile networks” (Market 7).

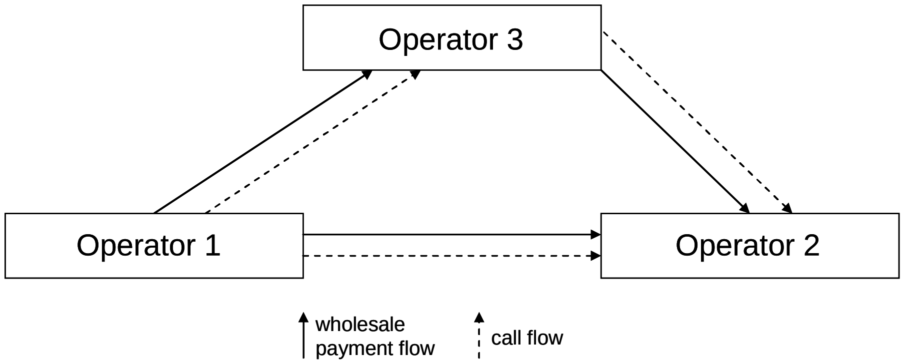
83 This is similar to the retail principle of Calling Party Pays (CPP) where the caller that places the call pays a usage-based price for the call and where the receiving party pays nothing. CPP starts from the assumption that the calling party derives all the utility benefits from the call.

84 See also Marcus (2006a), page 9f, DeGraba (2000), page 7. With reference to the termination problem in mobile radio, see Valetti/Houpis (2005).

85 WIK-Consult (2008), p. 5.

86 In practice, different variants are possible. Operator 2 can be paid either by Operator 1, thus, the latter buys transit and termination. Or, in case of a call-by-call or preselection provider, this provider pays for transit to Operator 2 and buys origination from Operator 1. These differences are neglected here. The figure only displays the direction of payments.

Figure 1: Payment and call flows in PSTN networks



In some cases, Bill & Keep is or has also been used in the PSTN and mobile networks.

- In 2006, the Commerce Commission in *New Zealand* published a Final Determination[87] ordering a Bill & Keep model for local interconnection between Vodafone and Telecom's fixed PSTN.[88]

- The U.S. interconnection regime for local interconnection is based on some simple rules: first, all operators have interconnection obligations; second, wireless operators and non-dominant fixed operators[89] are free to set any rate they wish, including zero, as long as the rates apply reciprocally (the termination rate applied between two CLECs must be the same in both directions). Fixed incumbents[90] are subject to cost-based reciprocal termination fees. Thus, whenever an ILEC is involved this cost-cap becomes effective as a result of the reciprocity requirement. It has to be noted that in the U.S. the same (geographic) number range is used for fixed and mobile telephony.

  These rules had the following effects: (i) termination rates to and from ILECs are very low and (ii) for traffic exchanged among CLECs and mobile operators Bill & Keep mostly applies. The application of partial Bill & Keep is the result of a market outcome and it is not required by the FCC.[91,92]

  Considering that carriers other than ILECs can mutually agree to set the rates at any level it is rational for them to set the rate to zero if traffic tends to be roughly balanced, because

---

87 Commerce Commission (2006).
88 The determination relates to local calls to and from Vodafone's local numbers but not calls to and from Vodafone' mobile numbers.
89 Competitive Local Exchange Carriers (CLECs)
90 Incumbent Local Exchange Carriers (ILECs)
91 "*The U.S reciprocal compensation system is characterized, not by a regulatory obligation to charge nothing for call termination, but rather by the absence of an affirmative obligation to pay something other than zero, framed with a requirement to maintain reciprocity.*", WIK-Consult (2008), p. 61.
92 Marcus (2006, p. 31) explains why a migration towards a uniform termination system based on the notion of Bill & Keep did not emerge in the U.S. despite the attempts of the FCC for several years: "In the highly politicized regulatory environment of the United States, they have been unable to make headway against the determined opposition of those carriers whose financial interests would be impacted by such a migration. The large fixed incumbent operators (called RBOCs) have for the most part been reasonably supportive of a migration to Bill and Keep; small rural fixed operators, whose termination charges tend to be much higher, have been the main opponents.

in that case the net flow of payments will be negligible whether rates are high or low.[93] The reciprocity requirement effectively prevented mobile operators from applying asymmetric payments to fixed network operators.[94]

The U.S. example demonstrates that Bill & Keep can be applied in the mobile sector. Moreover, an international comparison shows that there is a clear inverse correlation between service-based revenue per mobile-user and minutes of use. Countries applying Bill & Keep in the mobile sector (e.g. U.S., Singapore, Hong-Kong) exhibit a significantly higher mobile usage (and subsequently a higher ARPU) than countries with high mobile termination rates.[95]

- The interconnection regime in *Singapore* shows some similarities with the U.S. as mobile operators and non-dominant fixed operators usually apply Bill & Keep and dominant fixed operators interconnect with other networks based on regulated CPNP rates.
- In Europe, French mobile operators decided to apply Bill & Keep for mobile-to-mobile connections until 2004 but it did not apply for fixed-to-mobile connections.[96] In contrast with the situation in the U.S., this selective application opened opportunities for arbitrage, which finally led to the exit of the Bill & Keep system. Fixed operators reacted to high fixed-to-mobile termination rates by implementing gateways which “converted” fixed-to-mobile calls into mobile-to-mobile calls. The big difference between fixed-to-mobile termination charges and zero mobile-to-mobile charges rendered this arbitrage behaviour profitable (even considering the additional costs for implementing the gateways). As a reaction to this arbitrage opportunity, mobile firms withdrew the Bill & Keep arrangement applying for mobile-to-mobile calls.

The reason why similar arbitrage problems did not arise in the U.S. is related to the interconnection rules, namely the reciprocity requirement for all and cost orientation for ILECs. Since significant differences in termination rates did not develop, arbitrage did not became profitable in the U.S. Hence, the French example does not refute the principal applicability of Bill & Keep in the mobile sector. Also, it does not provide evidence that a partial implementation of Bill & Keep necessarily leads to arbitrage problems. Instead, the starting point for avoiding any such problems lies in the general level of termination rates and the differences between them, or, in other words that reciprocity requirements might be necessary as well as some cost-based obligation.

---

93 Laffont/Tirole (2001) have shown that although net payments do not change under these conditions, termination fees do matter. “It is correct that a change in the access charge need not affect the (absence of) net payments between the operators, but the access charge affects each network’s perceived marginal cost and therefore retail prices. It is, therefore, not neutral even if traffic is balanced.”

94 This result is also backed by economic theory. When analysing mobile-to-mobile and fixed-to-mobile arrangements Armstrong/Wright (2007) show that without such a reciprocity requirement operators would set the rates at monopoly levels.

95 WIK-Consult (2008), p. 64-67.

96 See e.g. Loutrel (2006); Cambini/Valletti (2005b), footnote 23.

B.2.2 Interconnection in existing IP-based networks

Interconnection arrangements in IP-based networks – either in the form of peering[97] or transit[98] – are unregulated.

The direction of traffic flows (upstream or downstream) does not play a role in either peering or transit. Traffic flows in both directions are added. Furthermore, there is usually no possibility to determine at the interconnection point the network of origination or termination of a session, and it is therefore not possible to make use of the concepts of origination and termination for billing purposes. Normally, there are no payment flows in peering agreements, as long as traffic imbalances do not exceed a certain specified limit. The precise requirements for the applicability of peering are laid out in the peering policies of the interconnected ISPs. Transit agreements involve payments, covering both, outgoing and incoming traffic (see figure 2 below).

IP networks are related to one another in a hierarchy of different tiers according to whether they only buy transit, buy transit and peer with some IP networks or peer only (see Glossary). Usually, ISPs operating on the same tier apply peering (e.g. between two tier 1 ISPs).[99] At the highest tier, all operators peer with every other operator. The other tiers are not fully meshed. If there is a peering agreement between the Autonomous System 1 (AS 1) and AS 2, then traffic is not only conveyed between these AS involved but also conveyed to AS 3 or AS 4 which have transit agreements with AS 1 and AS 2 respectively.[100] A tier 2 ISP has to buy transit from a tier 1 ISP, paying for both upstream and downstream traffic.[101] As there is no peering agreement between AS 3 and AS 4, they have to buy transit from AS 1 and AS 2 respectively if they want to route traffic between each other.

Very generically, the relation between peering and transit can be schematically illustrated in figure 2:

---

97 See Glosarry. Report of the NRIC V Interoperability Focus Group, „Service Provider Interconnection for Internet Protocol Best Effort Service“, page 7, available at http://www.nric.org/fg/fg4/ISP

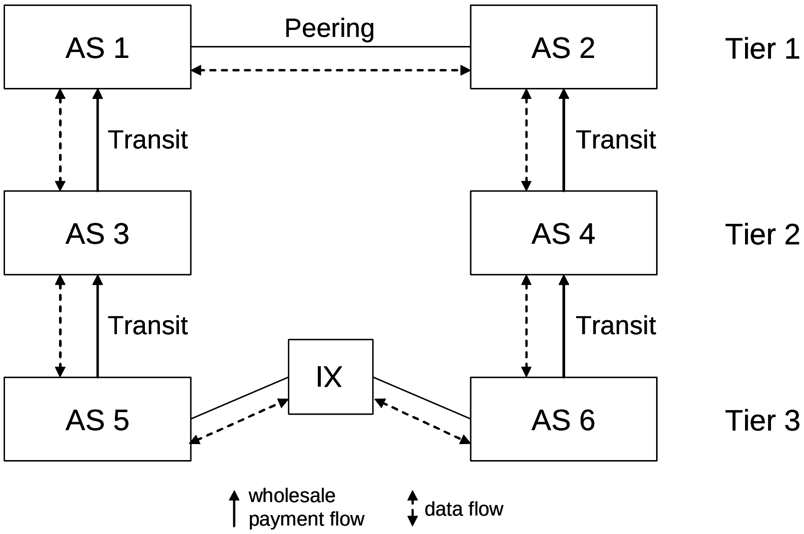
98 Ibid.

99 All ISP operating on the Tier 1 level have peering agreement with each other.

100 Despite this, “a backbone would not, however, act as an intermediary and accept the traffic of one peering partner and transit this traffic to another peering partner.”, FCC (2000), p. 5.

101 Similarly, a Tier 3 ISP would have to buy transit from other Tier 2 ISPs, assuming that (in general) there is a traffic imbalance with more traffic flowing downstream. But this may not necessarily the case if the Tier 3 ISP hosts a lot of content.

Figure 2: Payment and data flows in IP-based networks



The dashed lines in figure 2 show the data flows, arrowheads in both directions indicating data flow up- and downstream. The arrowhead on the solid lines show the corresponding directions of wholesale payment flows.[102]

The major difference between mechanisms illustrated in figures 1 and 2 is that in IP-based networks payments always flow in the upstream direction up to the highest tier, relating to data flows both upstream and downstream, while in the PSTN payment flows always follows the same direction as the call.

If, for example, data is conveyed from AS 5 to AS 6 while data flow downstream from AS2 to AS 6 payment flows from AS 6 to AS 4 and AS 4 to AS 2 flow upstream in the opposite direction (with the data flow being AS 5 → AS 3 → AS 1 → AS 2 → AS 4 → AS 6)[103].

It seems obvious that a fully-meshed infrastructure with peering arrangements between all AS would not be economically efficient.[104] The decision between peering and transit is a matter of network planning and cost optimisation,[105] as transit causes costs for conveying traffic but saves CAPEX investments in own network infrastructure (and hence saves operating costs). With peering, the logic is the reverse: those ISPs who fulfil the requirements for peering can choose between peering and buying transit, whereas those who do not, have to buy transit. In most instances, operators will employ both transit *and* peering arrangements.

Internet Exchanges (IX), or as they are sometimes called Network Access Points (NAP), constitute another institutional setting for the exchange of traffic, where ISPs can voluntarily

102 In the cases of peering and usage of Internet exchanges (IX) there are no payment flows, thus, these lines have no arrows.
103 This is the data flow unless an Internet Exchange – see further below – is used.
104 Today, there are more than 46,000 AS (WIK-Consult (2008), p. 47.
105 Of course, in practice this decision depends on whether an ISP qualifies for peering by fulfilling the requirements of the peering policies.

participate and where they agree to interconnect at a multilateral peering point.[106] Such IX enable the ISPs to interconnect their networks and to exchange traffic directly between them without having to deliver traffic via an upstream provider, hence, reducing costs as there are, usually, no payments for the exchange of traffic. [107] Internet Exchanges are a multilateral form of peering arrangements. Moreover, Internet Exchanges may also improve network resilience. The Internet "world" has historically adopted this interconnection model where many ISPs meet to exchange their traffic with other providers, each bearing the cost of transporting the IP traffic to the IX/NAP.[108]

The Bill & Keep mechanism is widely applied for Internet traffic worldwide. It is applied in the sense that, at the retail level, the end-user's Internet access rates include payment for connectivity and the option to receive and transmit data. Mostly, flat fees are applied, but there are also charging schemes based on data volume.

The Internet access provider handles these volumes on the basis of his peering and/or transit agreements (which may involve payments), but it does not charge, on a session basis any particular service provider for having sent data downstream to its Internet access customer.

The OECD concludes that "*the shift by last mile operators away from time based charging, in many cases for telephony as well, would argue in favour of using Internet charging practices for interconnect.*"[109]

B.2.3 Differences between interconnection in the PSTN and in IP-based networks

There are some important differences between interconnection in PSTN networks and interconnection in IP-based networks:

PSTN networks

As already indicated, PSTN networks have a strong focus on voice services.[110] In PSTN networks, transport and service are "bundled"; i.e. transport interconnection and service interconnection (usually) cannot be realized separately.

The E.164 number is generally assigned to the operator, who provides the end-customer with the access line.[111] If end-user A calls end-user B, A's network operator has no control over how the call is to be terminated. Only a single operator (B's network access operator) is able to terminate calls to end-user B.[112] This confers a special form of market power to the

---

106 Examples are the AMS-IX in Amsterdam, the LINX in the London or the DeCIX in Frankfurt.

107 Costs may be covered by annual or monthly fees, depending e.g. on transmission speeds used.

108 See e.g. FCC (2000) for a discussion of peering, transit and IX.

109 OECD (2007), p. 27.

110 It should be noted that they are used e.g. also for conveying narrowband Internet traffic. Before the advent of broadband Internet such narrowband Internet traffic constituted a significant percentage of the whole traffic in the PSTN.

111 Exception, e.g. MVNO.

112 With the E-164 number being closely linked to the access operator, it is not possible for other operators (other than B's access operator) to build an alternative termination infrastructure to reach end-customer B.

terminating operator. The network operator providing the access has market power for terminating individual calls to a single number.

The market power, called termination monopoly,[113] derives from two factors:

1. It is the terminating network operator alone who possesses the necessary routing information through control of the E-164 number;

2. Given that end-user B's network operator has a physical bottleneck for terminating A's call, under the CPNP regime B's network operator is entitled to a payment for termination from end-user A's network operator - necessarily out of a monopoly position. This bottleneck position will in all likelihood be abused unless regulation applies.

The termination monopoly operates even in markets where competition for call origination is effective, and is by no means limited to large players having significant market power (SMP) on the call origination market.[114] This leads to regulation of even small network operators without SMP in retail markets as each network operator has SMP for termination on its own network.[115]

When it is not possible for a caller to switch to another network carrier to set up the connection to the desired called party, a network operator has market power for terminating individual calls, independent of its market position in the retail market. If unregulated, this could lead to termination rates exceeding marginal costs subsequently leading to inefficient levels of network usage.

IP-based networks

In contrast to PSTN networks, todays IP-based networks do not have a focus on any particular service. The separation of transport and service is a characteristic feature of these networks. Transport interconnection and service interconnection can be provided separately. Consequently, operators providing transport may differ from those providing services.[116]

For a phone call provided in that manner between caller A and called party B by a third party relying on best-effort Internet connectivity for transport, it is no longer B's network operator, but B's voice service provider, who has control of the necessary routing information (e.g. in

---

113 For origination, market power depends on replicability of the access network rather than the exclusive control of the telephone number. The calling party A may have various options for calling B. He may e.g. choose his fixed network telephone or his mobile phone.

114 Rather, operators with a smaller market share will be motivated to set termination charges to even higher levels than will large operators. See Dewenter/Haucap (2005), Laffont/Tirole (2001).

115 Laffont/Tirole (2000), "It is worth recording here the common fallacy that small players do not have market power and should therefore face no constraint on their termination charges. This fallacy results from a misunderstanding of the definition of a market. A network carrier may have a small market share in terms of subscribers; yet it is still a monopolist on the calls received by its subscribers. Indeed, under the assumption that retail prices do not discriminate according to where the calls terminate, the network has more market power, the smaller its market share: whereas a big operator must account for the impact of its wholesale price on its call inflow through the sensitivity of rivals' final prices to its wholesale price, a small network faces a very inelastic demand for termination and thus can impose higher mark-ups …", page 186.

116 See Ch. B.4.2 (NGN architecture and topology).

the third party's SIP server) through the control of the E-164 number. However, it does not complete the physical transport of the call, which is completed by B's ISP. It is assumed here, that B's ISP is also providing the broadband access itself as broadband access and Internet access are offered as a bundled product in most Member States.

Therefore, the market power for terminating the call may be split between the party providing the service and the party providing the transport:

1 B's service provider, has the control of the necessary routing information behind the E-164 number i.e. the IP-address.[117] There is here a potential source of monopoly market power, which could be abused.

2 Under the prevailing billing mechanisms (peering / transit) in IP-based networks, even though B's Internet access provider continues to be the only one in a position to physically terminate the call, it is not entitled to a payment at the wholesale level out of this monopoly position from either B's service provider or A's network/service provider. At the wholesale level B's Internet access provider either peers or pays for transit provided by upstream backbone providers, both for upstream and downstream traffic.[118] At the retail level, it is the end-user B himself who pays for the physical termination of A's call through his monthly Internet access fee (mostly a flat rate) covering B's operator's peering and transit costs (Bill & Keep). The extent of B's broadband access provider's ability to abuse any market power out of its physical monopoly position for termination depends on the level of competition in the broadband access and Internet access markets.

In IP-based networks, control over the IP-address and the E-164 number continues to provide scope for abuse of market power. It is claimed by many competitive operators that control over signalling information and intelligent features (e.g. presence information) could reinforce market power problems in all-IP networks and allow leveraging towards adjacent sectors.[119] If a provider of the VoIP service does not receive information on the IP-address of the end-user (or of a server), it will not be able to establish the VoIP session and to allow the VoIP call.

In practice, a number of third party VoIP providers have decided to share their signalling information by interconnecting their SIP servers (e.g. Freenet and Sipgate) and there is no on-net/off-net price differential for calling parties. We do not know of excessive payments for access to the signalling information. The costs of the network elements involved are limited (e.g. Sipserver).

With regard to the probability of abuse of market power over the physical termination the reversal of the direction of payment flows makes the big difference. The crucial point is that the end-user pays for these flows. Whether the broadband access Internet access provider participates in a NAP/IX, peers or pays transit is not of direct relevance.

---

117 As long as the routing information related to the E-164 number (i.e. the relevant IP-address) is not publicly available, it is not possible for several operators to terminate calls to a single number.

118 See figure 2 in section B.2.2 showing the wholesale payment flows in IP-based networks.

119 See ERG (2007), Ch. 3.5

The wholesale market for connectivity in IP-based networks with its peering and transit agreements has so far been considered a market that could entail oligopoly market power but where the 3-criteria test is not fulfilled. An important point is that there is generally a choice between different transit partners.

Generally, abuse of the termination bottleneck is closely linked with Calling Party's Network Pays (CPNP), the wholesale billing regime mostly used in the PSTN.[120] The access operator terminating the call receives a payment from the interconnected network out of a monopoly position and may abuse it charging an excessive rate. That is why termination rates are usually regulated ex-ante.

By applying Bill & Keep for the terminating segment up to the first router or switch and associated service control functions after the access/concentration network and peering and/or transit for upstream connectivity, which imply a reversal of payment flows at the wholesale level compared to CPNP, the termination monopoly for transport could be avoided.[121] Different from CPNP, the potential to exploit control over the termination bottleneck is not available with Bill & Keep if there is sufficient competition at the retail level and therefore under this condition regulatory intervention can be avoided. It is important to note, that with a migration towards IP-based networks abuse of the termination bottleneck is not "automatically" avoided. That market power remains as long as CPNP is applied.

Table 1: Major differences between PSTN networks and IP-based networks

	PSTN networks	IP-based networks
Services	Focus on voice	Any service – not only voice
Source of market power	Control over access line and E-164 number necessarily linked	Control over the E-164 number and IP address, can possibly be separated from control over access line
Provision of transport and service	By one operator	Provision by different operators possible
Transport interconnection / service interconnection	Bundled	Separate provision possible
Charging mechanism	CPNP	Bill & Keep for the terminating segment and peering/transit for IP backbones. If transit is paid for, the broadband access provider pays for transport up and downstream

Therefore, the possibility to exploit SMP results from the interplay between the three factors:

a) physical bottleneck for termination;

b) control of the E-164 number;

120 See Ch. B.3
121 Ibid.

c) charging mechanisms.

B.3 Network structure and topology

B.3.1 General principles of IP architecture and topology

An IP network is an multi-service network based on packet switched technology using the Internet Protocol. It provides a platform for the delivery of multimedia services. In principle, any service can be realised with a specified quality level, if the performance objectives of the service can be met by the network.

As opposed to legacy telephone networks, the underlying philosophy of IP networks is that the network provides the transport resources and that the service is generated in the terminals and/or edge servers. Thus in principle an IP network has no “service awareness”. Therefore, in a pure IP network there is no service-specific data transmission.

The transmission of data is organized in a way that the network tries to transport all data as long as resources are available; and no data transmission request is rejected. Therefore, the transmission performance is not stable but dependent on traffic load. This performance characteristic is referred to as “best effort” transmission. The consequence is that there are no bearer services with specified transmission characteristics as in circuit-switched networks. The transmission performance depends on the IP network design and management.

To follow the argumentation and conclusions of this document no detailed technical knowledge is needed. However, additional technical background information on the design and management of IP networks is given in Annex 3.

Since IP networks (and packet switched networks in general) are designed to “only” provide transport resources and to support any service, a separation between transport and service can be made. Therefore, when discussing aspects of interconnection this has to be addressed.

B.3.2 NGN

In this context, an NGN is to be seen as a specific form of implementing and managing an (all) IP network. The aim of NGNs is to use the advantages of IP technology while having the same level of control over services and user profile management that are achieved in circuit-switched networks.

The general concept of an NGN is to provide a multi-service QoS-enabled, service-aware, secure, global network based on packet mode technology[122], able to support voice, data and video/TV. Details on the specification and definition of NGNs in standardisation are given in Annex 3.2.

The underlying mechanisms to control and manage the network at the transport layer are the same as for IP-networks (see Annex 3.1). However, compared to a “simple” IP network, a NGN has the following important additional features:

- the access to the NGN is controlled, i.e., there is an admission control, user profile management and dedicated bandwidth allocation for different services;
- the transmission of data is service-specific and managed through bandwidth allocation by specific “NGN”-protocols and policies;
- there are standardized interfaces at the transport and service layer that allow third parties to connect to NGNs, use its resources and offer their own services;
- through the implementation of stringent policies and signalling mechanisms, end-to-end services are controlled and the necessary network resources are allocated and maintained during the use of service.

With respect to interconnection, the consequence is that interconnection agreements for NGNs can be more complex and NGN-specific aspects have to be considered.

B.3.3 NGN network topology and the implications for interconnection

Current IP-based networks (e.g. public Internet and non-NGN IP networks) grew in a more relaxed regulatory environment. On the other hand, circuit-switched networks developed in the context of a fewer number of operators and a higher degree of centralised control[123]. “Real world” NGN (i.e. “telco” operators) seem to develop in a “walled garden” style - service control and routing/switching is shared by different network elements but, typically, there is strict control of transport resources by service logic within individual networks.

With regard to transport interconnection, the transition towards NGNs has further structural implications.[124] Even if there is not yet a definitive understanding of the future network structure and regime, it may for example entail a rearrangement of core network nodes, changes

---

122 A logical evolution from separate network infrastructures into a unified network for electronic communications.

123 They establish a single physical path for the duration of a call or session, via a signalling network that provides end-to-end traffic management and billing information.

124 Furthermore, it is commonly assumed that Next Generation Networks will be operated at significantly lower costs than other fixed networks by passing to a single infrastructure based on IP for transporting any kind of flow, voice or data, and for any access technology (DSL, FTTH, WiFi, etc.). NGNs can provide operators ample flexibility in their cost base to reduce OPEX and CAPEX (see: Credit Suisse First Boston, IP: The Holy Grail for Telcos, March 2005). This issue of NGN implications for network costs will be further dealt with in the NGN document taking into account the results of the Regulatory Accounting Working Group.

in the number of network hierarchy levels and consequently a geographic rearrangement of points of interconnection (PoI).

In this context the number of network nodes and points of interconnections in IP networks and NGN and the definition of local interconnection in IP networks and NGN are particularly relevant .

B.3.3.1 Number of network nodes and points of interconnection (PoI)

The number of network nodes/PoI at each hierarchy level for NGN is not yet decided upon (or relevant information is not available) in most countries. The empirical basis is not broad enough to derive substantial conclusions, as the sample is too small to derive stable relationships between the number of nodes at different hierarchy level.[125] This situation highlights the need for more transparency in SMP operators' NGN plans.

- As the Consultation has confirmed one can assume that the number of PoI and nodes in which routing of traffic will occur will be less than in current PSTN networks.[126] Whether and to what extent this will be the case may depend on the specifics of each network and may therefore differ among countries. Although it is generally agreed by respondents that the number of PoI will decrease, it is not clear whether there is any one best number for any particular type of NGN implementation.

  As traffic costs become less dependent on distance, bigger interconnection pipes are more efficient than smaller ones and a greater centralisation of interconnection points may be appropriate.[127] VoIP is deemed more conducive to this centralisation of traffic and also reducing the number of interconnection points is justified due to the reduced number of (bigger) switches.

  Moreover, it is argued that if the network provides sufficient bandwidth such that no traffic overload occurs then no specific measures for traffic engineering due to interconnection is needed, then it makes sense to have only one point of interconnection (or a small number to ensure resilience of the interconnection). This can be observed for best effort services in current IP networks.

  On the other hand, issues of network resilience (the need for a minimum of network nodes and resources in case of “force majeure”) and streaming services leading to interconnection points located closer to the end-user due to higher (IP) traffic demands, may attenuate this trend or even increase the number PoI.

---

125 See in particular country case studies in Annex 3 (I, NL, NO, UK).

126 Such a reduction could lead to stranded investments; these are not restricted to either incumbents or competitors.

127 At a technical level, the „reasonable“ number of PoI will depend on parameters like transport costs, traffic volume, distribution and symmetry and (much) less on distance. These factors influencing the number of PoI were also mentioned in several responses.

Issues concerned with streaming services (e.g. multicast) are more likely to relate to bit-stream services and not interconnection as they are provided in a bundled manner by access providers. The number of nodes required depend in particular on the location of the servers in the network hierarchy.[128]

- Transport and service interconnection might occur at different nodes and hierarchy levels. Considering the distinction between transport and service, transport interconnection could take place at a greater number of locations than service interconnection.

B.3.3.2 Definition of lowest level of IP interconnection

Generally, the lowest level of the core network constitutes the lowest level for routing.

In the PSTN this corresponds to the local switch. The local PoI in the PSTN interconnection structure is the local (circuit) switch where a specific set of lines – directly or indirectly (via remote switch or concentrator) – is connected to this switch. This (local) interconnection point is the unique point of access to these lines and does not provide interconnection to any set of lines (e.g. in another switch).[129]

In an all-IP network the broadband remote access server (BRAS[130]) defines the border between access concentration and routing**.** Therefore the core network is restricted to the network being fed by different access and concentration networks (mobile, cable). Its edge is the routing node closest to the end-user, that is the router or switch through which all communications of a end-user must be routed and have the capacity to switch communications between users connected to this node. The MSAN, base station or cable head-end do not have such switching capabilities. [131]

The currently classically used hierarchy concept of three physical levels, “local, regional, national”, may not be applicable in an IP network. The amount of physical levels in a NGN network may change from formerly 3 in a PSTN network to only two physical levels. However that does not mean that double tandem or single tandem termination might necessarily disappear or merge. The highest level in a NGN network could logically serve both double tandem and single tandem termination when the highest level is composed of more than 1 loca-

---

128 That may even lead to a higher number of nodes for bitstream access. This issue may be addressed separately from this document.

129 For the purposes of cost accounting a local PoI does not necessarily constitute the border of the core network. Backhaul networks are for the purposes of cost accounting taken to belong to the core network, because costs in backhaul network can be considered traffic dependent as can be costs in core networks. The end of the access network is the dedicated interface to serve a single end-user. The costing boundary between access and core network services needs revision taking account of NGN technologies. See Annex 4, p. 121

130 TV services with servers being located in the concentration network may be a bitstream issue rather than an interconnection issue. Bitstream requires the competitor to access the wholesale service at layer 2 or layer 3 of the communication protocol stack.

131 In the UK, access and interconnection at the level of the local access network is currently at the MDF sites / MSAN nodes. However, in other countries where the relation between MSANs and Metro core locations is different this would lead to non-efficient costs ultimately to be covered by the users. For example [NL] 1365 local interconnection points in the PSTN network versus 24.000 to 25.000 MSAN locations and 137 Metro-core locations in the NGN network.

tion and these locations each serve a specific region of the country under which a certain amount of metro nodes are implemented. The lowest interconnection level possible will be the metro node level in most countries, but this will depend on the design of the network in question.

Also, one should consider the "hot-potato" routing within the IP networks[132], i.e., the practice of passing traffic off to another network as quickly as possible (towards the path with the lowest delay). So, normally the originating network/router will not hold onto the packets until it is as near to the destination as possible, as is done in the circuit switched local interconnection model. This will then drive the interconnection towards the higher levels of the core networks.

Further arguments to confirm this results are the following

- services will be increasingly nomadic and customers will no longer be attached to the network on the same location or services might even not be driven by location (e.g. mobile). However this does not mean that POI's cannot serve a certain region (for fixed orientated networks) or a group of customers (mobile orientated networks). This increases the interconnection requirements and the need of interconnection agreements that do not exist today;

- the influence of distance on costs is less relevant in NGN than in PSTN. The transmission at lower costs and the possibility of choosing the location freely may make local interconnection unnecessary;

- the design of NGN could drive away from the idea of a unified network as understood in the context of PSTN (e.g. 64 kbps circuit switching, SS7 signalling/interconnection), increasing national differences and growing complexity;

- other technical reasons mentioned include the tendency to increase the capacity of routers in network architecture, the separation of signalling and data transport and integrated networks carrying more non-local traffic than PSTN.

- Whereas in TDM networks the functionality of switches is to support voice services only, in IP networks, it is envisaged that PoIs for transport interconnection will have to support many services. Service interconnection may be restricted to higher levels of interconnection (e.g. VoIP).

These points were supported in many responses. Furthermore, several respondents considered the SBCs as the lowest level.

ERG considers that the maximum efficient number of PoI offered in NGNs should be used for applying the lowest interconnection rate accordingly, even in case where not all of these

---

132 „Hot-potato routing“ seems to be the normal behaviour of most IP peering agreements.

points are physically offered for interconnection. This view is supported by a number of respondents and is also in line with OPTA's draft decision on Market 3.[133]

B.3.4 Interconnection and interoperability

Service and transport interconnection between different core networks is necessary for end-to-end global services.[134,135] In summary, and generically, there will be two types of NGN interconnections[136]: transport interconnection and service interconnection.

Since NGNs will not be constituted by a single element network or service, but by a range of diverse components requiring seamless interconnection and full service interoperability.

Service characteristics and required interoperability standards, that include support of voice services e.g. SIP-I (Session Initiation Protocol-ISUP) and requirements of new multimedia services are dependent on QoS, device capability, security, number portability/translation (including ENUM) and authentication across service provider network boundaries[137].

In order to allow full interoperability of IP based services offered to the customers, interconnection need to be assured at the transport level[138] as well as at the service level. Each service may use network elements from lower levels in different combinations, some of these elements are common to two or more services and others are unique to a single service.

To enable these heterogeneous elements/networks to interoperate, the elements/interfaces/protocols involved in the support and delivery of services need to be standardised.

---

133 OPTA applied a similar reasoning in a draft decision (market analysis fixed voice termination). OPTA determined that all parties with SMP should grant terminating access on a regional level. The efficient number of POIs in the Netherlands has a maximum of 20. KPN provides 20 points of interconnection at a regional level for the termination of fixed services. If a competitor grants access at less locations, the same interconnection tariffs applies as if KPN had offered access at the 20 regional PoI. In addition, KPN exclusively is obliged to grant terminating access on the local level (OPTA, 2008, http://www.opta.nl/download/201452+ontwerpbesluit+marktanalyse+vaste+gespreksafgifte+%28telefonie%29.pdf.

134 The global NGN architecture consists of interconnected core networks belonging to different carriers, with endpoints connected through attached access networks, and gateways (border gateways control access into and out of each core network, monitoring and regulating the data flows on each interface) to non-NGN networks.

135 A similar situation as in legacy circuit switched networks, with SS7 (MTP and ISUP), ISDN, etc.

136 An NGN interconnection mode can be direct or indirect. Direct interconnection refers to the interconnection between two network domains without any intermediate network domain. Indirect interconnection at one layer refers to the interconnection between two network domains with one or more intermediate network domain(s) acting as transit networks. The intermediate network domain(s) provide(s) transit functionality to the two other network domains. Different interconnection modes may be used for carrying service layer signalling and media traffic.

137 On this basis, ETSI identifies the minimal requirements and functionalities that operators and providers have to assure in their NGN for service "aware" interconnection. Main requirements include the following:

- definition of both application level and gateway functionalities; the only elements acting as a interconnection boundary between operators' network domains;
- operators shall control the flow of service request through the signalling protocol which is used to control the set-up of the communication;

ETSI has adopted SIP signalling protocol (ETSI ES 283003/TS 124229 standards) as the unique protocol for NGN interoperability (the specification of interworking function between SIP-based networks and traditional circuit-based telephony network is defined in ETSI ES 283027/TS 129163 standards).

138 See ECC 75 Report (2005). See also IETF working groups "Session PEERing for Multimedia INTerconnect –speermint" (http://www.ietf.org/html.charters/speermint-charter.html).

Hence, operators should be encouraged to give access to the technical interfaces, protocols and all other technologies necessary for the interoperability of IP based services, and to use standard interfaces and protocols.

This may prove to be particularly challenging since a number of “telco” network operators consider that a horizontal separation of transport and service (and control) levels is neither appropriate nor in their interest, particularly if they want to guarantee quality of service (e.g. by using IMS). Their understanding of NGN seems to imply a continuation of vertically integrated provision of transport and services as has been the case in legacy “telco” networks.[139] Interoperability is of particular relevance for those operators who only provide services without operating networks as it enables them to integrate their services into other operators’ networks.

B.3.5 Quality of service issues related to interconnection

The term quality of service (QoS) is a broad concept (see definitions section) as it does cover all aspects influencing the user’s perception of the quality of the service related to the network performance (such as reliability or availability) as well as to other factors (such as customer support for instance). However, the term QoS is sometimes used when referring to part of these aspects and often loosely used as a synonym for network performance.

B.3.5.1 Why QoS needs to be addressed

Legacy telephone networks use circuit-switched technology. Fixed resources (e.g. fixed bit rate, fixed routing) are reserved for each call, such that transmission performance is stable and thus quality of the service is guaranteed. Circuit-switched technology (PCM channels, 64 Kbps) is well suited for telephony services but cannot easily provide other types of electronic communication services.[140] Since the circuits and network components were well standardized and each operator uses the same technology, the interconnection arrangements for legacy telephone networks were rather simple. One had only to agree on the kind of bearer services to be interconnected, signalling, capacity and availability requirements. QoS matters were automatically taken into account by assigning dedicated channels to a single connection.

IP networks use packet-switched technology that by default does not provide fixed transmission channels for voice communications. They provide transport capabilities irrespective of the services that use the network while the intelligence and complexity that is necessary for the provision of services relies on the end-devices. This approach of data transmission via an

139 See comments of France Telecom, Deutsche Telekom, Vodafone and the GSM-E comments to the ECC 75 report.

140 They were used for other services though (e.g. fax, dial-up Internet access) because theses services had similar network performance objectives.

IP connection, which does not provide any guarantee of level of performance, of priority, or of data being delivered at all, is referred to as “best-effort” transmission.

In both circuit-switched networks and “best effort” IP networks, transmission performance varies with the traffic load. The difference between circuit-switched networks and IP networks is the different behaviour in dealing with high traffic load. While circuit-switched networks denying further connection when all channels are busy (“busy tone”), an IP network still transports IP packets up to the link capacity. Although some IP packets are still transported in heavily loaded IP networks, the performance of individual connections is gradually degraded. In an IP environment, networks do not have standardized bearer services at interconnection points because network performance is provided thanks to “best effort” transmission. This has allowed the Internet to remain as simple as possible.[141]

If a network is designed and operated properly, network performance might remain stable throughout the user’s session. Given that the level of the transmission performance is below the sensitivity threshold of a service, the user will not perceive any quality impairments. In fact, “best effort” does not automatically mean that the quality of a service is low. Unidirectional services, such as IPTV or web browsing may only require high bitrates and can generally tolerate some variations of the bit rate thanks to the use of buffer techniques in the terminal device.

On the other hand, telephony, videoconference and other bidirectional services (like online gaming) may require performance objectives guaranteed by the network that cannot be met throughout the whole user’s session by “best effort” transmission (for example guarantee a low latency in the delivery of IP packets, i.e. a highly “responsive” connection).

If a guaranteed QoS over IP-based networks with a level of performance comparable to the PSTN is desired, one has to modify and adapt the IP transport technology in a way that connections with reliable and fixed transmission characteristics (transport classes) are possible. Moreover, the ability to provide services with guaranteed QoS may not only be applied to a single network but must be maintained over the whole chain of interconnected networks involved in the provisioning of a specific end-to-end service. Different strategies for doing so are possible (see Annex 3.1 for technical background information).

Improving the overall performance of current IP networks and still sticking to the best effort policy may not cope with all situations that might occur in an all-purpose IP network. Quality of service is therefore potentially a new dimension in the interconnection of IP/NGN, and could be an important focus for NRAs because it could enable new forms of discrimination between a larger operator’s services and those provided by interconnecting competitors (see Section C.3).

---

141 In the sense of easily enabling the implementation of new services.

B.3.5.2 End-to-end QoS perspective and its support in IP networks

One should keep in mind that QoS is an end-to-end issue. A specification of QoS refers to the quality that is perceivable between these end points. To stress the fact that a quality statement is really referring to the quality of a service and not to sub-parts of it, the term end-to-end quality is often used.

Depending on the service under consideration, the end points may vary and be of a different nature (e.g. for voice it can be mouth to ear, for data transmission it can be UNI to UNI).

The perceivable end-to-end quality of a service is the result of many different factors that also have interrelations, such as:

- Parameters of the terminal equipment:
  - in case of a voice service, technical characteristics of the codec play an important role in the user's perception such as its ability to compress speech without causing audible speech quality degradation and its ability to cope with disruptions in the network connection (compensation for packet loss etc.);
  - terminal settings: settings of the protocol stack (retry delays, buffer sizes, etc.)
- Network performance parameters:
  - IP traffic is influenced by delay, delay variation, packet loss142. Dependant on the kind of service stringent performance objectives have to be met;
  - if multiple interconnected networks are involved additional aspects must be considered, e.g. allocation of performance budgets, implementation of QoS mechanisms, agreement with interconnected networks upon consistent QoS policies.

End-to-end QoS cannot be divided among different sections of networks or even between separate networks of different providers. That is because end-to-end quality is always the result of the interaction of the different influences (quality impairments) that occur within the chain of the whole end-to-end connection. This should be kept in mind when dealing with QoS matters. However, performance parameters objectives can be given and divided among the network section and terminal elements. For example, packet delay limits can be allocated to each section and element (also for the purpose of interconnection agreements)[143].

B.3.5.3 QoS across networks

When trying to support QoS via multiple interconnected networks, one has to ensure that certain transmission performance objectives can be met by single and interconnected net-

---

142 See ITU-T Rec. G1010

143 A framework for this task is given in ITU-T Rec. Y.1542.

works. For doing so a variety of transport classes have to be specified. These or comparable classes must be supported by all networks and of course a policy is needed on how to handle different classes via interconnected networks. In this way, end-to-end QoS (over interconnected networks) with guaranteed quality levels can be maintained at all instances of time.

Strategies to achieve end-to-end QoS

The challenge of providing services with guaranteed QoS over IP-based networks is to modify and adapt the IP transport technology in a way that connections with reliable and fixed transmission characteristics (transport classes) are possible.

ITU and IETF have both produced Standard[144] metrics that can be used for measuring the performance of an IP flow and to set performance objectives. One possible way of meeting these performance objectives is for example to prioritize “QoS sensitive” packets for transmission in case some packets must be dropped due to congestion. Other possibilities are to use stringent admission control policies to avoid congestion or to use more sophisticated traffic management strategies as explained in Annex 3.1. By use of such mechanisms, the end-user’s quality perception can be improved, at least from a statistical point of view.

QoS mechanisms for network performance are numerous and can be sorted in two classes:

- stringent QoS mechanisms giving strict guarantees on the availability and the quality of the service: RSVP is one of them but never saw widespread use because of its complexity and lack of flexibility.

- statistic QoS mechanisms giving statistically good results with no absolute guarantee: DiffServ and MPLS among them are widely used within carrier networks and within Virtual Private Networks (VPN). Moreover, new statistical mechanisms focus on the management of the IP flow rather than packet labelling (e.g. research studies conducted by Eurescom).

NGN as specified in standards

NGNs as specified in ITU-T, ETSI, 3GPP have a strong and clear focus on network performance models including use of techniques such as prioritisation, resource reservation and admission control techniques (cf. RACS, NASS and PDF functions) to provide guaranteed quality for a multitude of services. They are tending to be more stringent with respect to QoS and service delivery.

Hence, NGN interconnection taking into account all requirements of the standards will automatically require the agreement and support of such transport classes/Classes of Services and mechanisms between the involved operators to ensure the transmission of data both within and across networks in a uniform and predictable manner, allowing effective and effi-

144 Y.1541 defining sets of performance objectives for technical parameters applying to different transport classes, RFC 2679 defining metric for one-way delay of packets across Internet paths and RFC 2680 defining a metric for one-way packet loss across Internet paths.

cient any-to-any interconnection. The prerequisite is strategies that the different QoS protocols and mechanisms implemented across networks can be mapped to each other and that interconnection partner have full access to interfaces of resource management sources.

Deploying NGN and dealing with other IP networks

Operators intending to deploy NGN architectures follow a step-by-step implementation of NGN based on releases of the standardisation bodies[145]. That means that deploying full NGN architectures will take a long time and that heterogeneous networks will in the meantime have to coexist and deal with the QoS issue.

In the current practice of interconnection of IP networks, the technical implementation of QoS mechanisms is not carried out. The existing and upcoming NGN standards bring a pool of ideas for traffic management. It is likely that at least in the start of the migration phase, early NGN networks will cover a broad performance spectrum from simple implementation of TCP/IP with best effort to intensive implementation of traffic management methodologies providing high level and stable transmission performance.

With such an approach, network design might be inspired from NGN standards, although it is not a fully implemented NGN. That is because the network operators will have to take a balanced approach which traffic management methodologies should or should not be used taking into account cost and efficiency constraints.

It can be profitable for an operator to take this approach since end-to-end performance levels sufficient for many services can be obtained. However, for some services there might very high performance requirements (e.g. interactive real-time communication services with high data throughput like high quality video conferencing), where it might be necessary to meet stringent NGN standards.

The technical interconnection of NGN is potentially more complex than that of current public IP networks because NGNs are designed to have full control over the services whereas IP networks are following a more simple and flexible concept. Thus, in NGNs, the range of transport classes and access to control functions supported, and the broad range of technical choices that exist should support any service feature. Quality of service is therefore potentially a new dimension in the interconnection of NGN.

145 At present: ITU-T release 1, ETSI TISPAN release 2 and 3GPP release 8 (common IMS).

C Regulatory challenges and implications

C.1 Existing and proposed Framework

C.1.1 Existing Framework

The existing Regulatory Framework, which came into force in July 2002 is currently under review. It will continue to be the reference legislation until new legislation comes into force and the old one is repealed (or amended). The existing Access Directive in particular contains provisions defining the legal basis for access[146] and interconnection[147] agreements, wherein:

- Article 3.1 requires Member States to ensure that there are no restrictions preventing undertakings to negotiate between themselves agreements on technical or commercial arrangements for access/interconnection, in accordance with Community law; undertakings do not need to be authorised to operate in the Member State where access or interconnection is requested.[148]
- Article 4.1 requires operators to negotiate interconnection with each other for the purpose of providing publicly available electronic communications services, in order to ensure provision and interoperability of services throughout the Community.[149]
- Article 5.1 and 5.2 empower the NRAs to set access and interconnection obligations, by explicitly mentioning promotion of efficiency, sustainable competition, and benefit to end-users together with operational and technical conditions.[150]

---

146 Access is defined in AD Art. 2 (a): “access means the making available of facilities and/or services, to another undertaking, under defined conditions, on either an exclusive or non-exclusive basis, for the purpose of providing electronic communications services. It covers inter alia: access to network elements and associated facilities, which may involve the connection of equipment, by fixed or non-fixed means (in particular this includes access to the local loop and to facilities and services necessary to provide services over the local loop), access to physical infrastructure including buildings, ducts and masts; access to relevant software systems including operational support systems, access to number translation or systems offering equivalent functionality, access to fixed and mobile networks, in particular for roaming, access to conditional access systems for digital television services; access to virtual network services;”

147 Interconnection is defined in AD Art. 2 (b): “.interconnection. means the physical and logical linking of public communications networks used by the same or a different undertaking in order to allow the users of one undertaking to communicate with users of the same or another undertaking, or to access services provided by another undertaking. Services may be provided by the parties involved or other parties who have access to the network. Interconnection is a specific type of access implemented between public network operators;”

148 The norm addresses possible limitations in the national laws and lays down the principle of, in first instance, leaving to parties, even outside the national border, autonomy in setting up agreements, as long as they are in conformity with Community law.

149 The rule mandates interconnection – also for ensuring provision of services and interoperability. This rule encompasses regulations which can be imposed using articles 5-8 of the Access Directive, and thus is the explicit legal base for enforcing obligations, also with regard to IP interconnection.

150 Economic efficiency and competition are on a pair with detailed operational and technical conditions in maximizing end-user welfare – this allows, for instance – NRAs to impose measures which may address the reluctance of operators in allowing for more expedient and innovative forms of interconnection, such as “native” IP-IP instead of a bulky PSTN-IP-PSTN transaction.

- Article 5.4 empowers NRAs to intervene with regard to access and interconnection at their own initiative where justified or, in the absence of agreement between undertakings, at the request of either of the parties involved,

- According to recital 6 (Access Directive) NRAs should have the powers to secure, where commercial negotiations fail, adequate access and interconnection and interoperability of services in the interest of end-users. This indicates that in the first place it is up to operators to reach agreements on interconnection (including the appropriate billing regime).

Besides these provisions that may be imposed on all operators, more specific obligations in terms of definition of interconnection services and processes and possibly related services[151] may need to be imposed on SMP operators, following a thorough market analysis (according to Articles 9-13 of the Access Directive).

In a broader context the Framework Directive must also be considered. Political objectives and regulatory principles are laid out in Article 8, thereby providing NRAs with a framework when developing principles for the regulatory treatment of IP interconnection in the context of the migration process from PSTN to IP-based networks and also for evaluation of different conceivable regulatory options.

According to Article 8 (2) of the Framework Directive, NRAs shall promote competition by *inter alia*:

- ensuring that there is no distortion or restriction in the electronic communications sector (Art. 8 (2) (b));

- encouraging efficient investment in infrastructure and promoting innovation (Art. 8 (2) (c)).

Further, according to Article 8 (3) of the Framework Directive, NRAs shall contribute to the development of the internal market by *inter alia*:

- removing remaining obstacles to the provision of electronic communications networks, associated facilities and services and electronic communications services at European level (Art. 8 (3) (a));

- encouraging the establishment and development of trans-European networks and the interoperability of pan-European services, and end-to-end connectivity (Art. 8 (3) (b));

- ensuring that, in similar circumstances, there is no discrimination in the treatment of undertakings providing electronic communications networks and services (Art. 8 (3) (c)).

---

151 These related services include all services requested to build a full operational interconnection, e.g. access to interconnect location, bandwidth capabilities, collocation.

C.1.2 Proposed Framework

The existing regulatory framework of July 2002 is currently undergoing a review process. This process was started within the i2010 initiative as part of a renewed Lisbon strategy in June 2005 and is seen as one of the main challenges for establishing a European Information Area. On Nov. 13, 2007, the Commission published several documents which, among other things, contain modification proposals for the regulatory package. As not all of them are relevant in the context of this paper, the focus lies on those modifications which are considered relevant for the topics dealt with here, mainly modifications of the Framework Directive ("FD"), the Access Directive ("AD") and the Universal Service Directive ("UD").

Framework Directive

- An additional sentence has been inserted in Art. 5 para 1 FD with regard to information to be provided by undertakings on network development. According to the new provision, undertakings providing electronic communication networks and services (ECNS) can be required to submit information concerning future network or service developments that could have an impact on the wholesale services made available to competitors.

  It is evident that the modification aims at increasing the transparency of incumbent operators' plans concerning the development of NGNs for their competitors and wholesale partners; the Commission's comments (p. 10) explicitly mentions NGN architecture as an example for network developments. Apparently, the current more general wording in Art. 5 FD which obliges Member States to ensure that ECNS providers deliver all information, including financial information necessary for NRAs to ensure conformity with the provisions of, or decisions made in accordance with, the FD and other specific directives have not proven to be sufficient to justify information requests by NRAs directed to incumbent operators.

- A new section on security and integrity of networks and services has been introduced as Art. 13a and 13b FD. Pursuant to the new stipulations,

  - Member States have to ensure that undertakings providing public communications networks or publicly available electronic communications services

    - have to take appropriate technical and organisational measures to maintain the security of their networks and services; and

    - that those measures must comply with a state-of-the-art security level corresponding to existing risks and must try to avoid or decrease the consequences of security breaches for users or interconnected networks; and

    - that those undertakings take appropriate measures to maintain network integrity in order to ensure a continuous availability of services provided over those networks.

- The NRA must be informed of security or integrity breaches with significant impact on network operation or service provision and will report to the Commission every 3 months on information received and measures taken.

- The Commission may adopt technical enforcement measures including form and procedure of the aforementioned reporting obligations.

- In addition, Member States ensure that NRAs may issue binding instructions to undertakings providing public communications networks or publicly available electronic communications services to transmit information necessary to judge the security of their networks and services including documentation of their security measures and to charge an independent institution with a security check the results of which are to be submitted to the NRA.

Those new provisions are relevant for this paper insofar, as the provisions on network integrity for fixed telephony service providers in Art. 23 UD are now extended to all kinds of electronic communication networks and services, including mobile and IP networks (see Commission comments, p. 10).

- Concerning definitions, the term “associated facilities” in Art. 2 e) FD has been widened and now also covers explicitly systems for number or address translation as well as ducts, masts, street cabinets and buildings.

Access Directive

- Art. 5 para 2 AD which allows NRAs when imposing access obligations pursuant to Art. 12 AD on an operator to lay down technical or operational conditions by the operator to be met by the access provider, as well as the beneficiary, where necessary to ensure normal network operation, was shifted to Art. 12 para 3 AD.

- Art. 5 para 4 AD, which stipulated Member States have to empower NRAs to intervene at their own initiative where justified or, in the absence of an agreement between the parties, at the request of either party involved to secure the policy objectives of Art. 8 FD, has been deleted.

  According to the comments given by the Commission (p. 11), *“those provisions”* (which apparently refers to Art. 5 para 4) were deleted due to an overlap with various other provisions. As the reference to the policy objectives of Art. 8 FD covered, among other things, the contribution of NRAs to the internal market development by encouraging, *inter alia*, the interoperability of pan-European services, and end-to-end connectivity, the deletion of Art. 5 para 4 AD is likely to have a disadvantageous impact on NRA’s efforts to ensure interoperability of services in an NGN environment. This, will become more important as on the one hand a greater variety of standards (IETF, ITU, proprietary) exists, while on the other hand, NRAs do not necessarily want to extend SMP regulation to functioning IP peering/transit arrangements. Therefore, the ERG holds the view that the power of NRAs to act on their own initiative to ensure end-to-end connectivity/interoperability should be maintained in Art. 5 para 4 AD.

- The list of potential access obligations in Art. 12 para 1 AD has now been complemented by including an obligation to grant access to associated services with regard to identity, location and presence of the user. Pursuant to the Commission's comments (p. 12), the new provision shall exclude discrimination concerning interconnected IP networks.

Universal Service Directive

- Concerning definitions, the former term "public telephone network" in Art. 2 (b) UD was considered not to be needed any more (Commission' s comments, p. 11) and therefore deleted.

- The definition of publicly available telephone services in Art. 2 (c) UD (*"a service available to the public for originating and receiving national and international calls directly or indirectly via CbC or CPS or resale and access to emergency services through a number or numbers in a national or international telephone numbering plan, and in addition may, where relevant, include one or more of the following services: the provision of operator assistance, directory enquiry services, directories, provision of public pay phones, provision of service under special terms, provision of special facilities for customers with disabilities or with special social needs and/or the provision of non-geographic services"*) must be read together with a modification to Art. 26 UD as Art. 26 para 2 UD now limits the obligation to provide access to emergency services to operators who offer a telephone service for outgoing calls through a number or numbers in a national or international telephone numbering plan.

- In Art. 22 UD a new para 3 is inserted allowing the Commission to adopt technical implementing measures concerning minimum quality of service requirements to be set by the NRA on undertakings providing public communications networks.

  This is a useful addition. However, the ERG considers that the power to set minimum quality of service requirements should be entrusted directly to NRAs. As minimum quality of service requires measures on both the end-user and the network level, it should be clarified that NRAs can require minimum quality of service on the network level as well.

  In case it is not possible to do so in the UD, a second best option would be to empower NRAs in Art. 5 AD to set on their own initiative minimum quality of service requirements on operators of public communications networks.

Data Protection Directive

- The definition of "call" (Art. 2 (e)) does not make reference to the criterion of real time as in the current Directive.

C.2 Relevant markets

The new Recommendation on relevant markets issued by the European Commission in November 2007 led to the following markets susceptible to ex-ante regulation:

1. Access to the public telephone network at a fixed location for residential and non-residential customers.
2. Call origination on the public telephone network provided at a fixed location.
3. Call termination on individual public telephone networks provided at a fixed location[152].
4. Wholesale (physical) network infrastructure access (including shared or fully unbundled access) at a fixed location.
5. Wholesale broadband access.
6. Wholesale terminating segments of leased lines, irrespective of the technology used to provide leased or dedicated capacity.
7. Voice call termination on individual mobile networks.

In line with the document's focus on interconnection in NGNs environment, it deals mainly with markets 2, 3 and 7 from the above list have to be dealt with.

These may be contrasted with access-products like bitstream, where one operator uses the facilities of another operator instead of building its own infrastructure reflecting an operator's "make or buy" decision. These markets are not dealt with in this paper.

A number of NRAs are in the process of reviewing Markets 2, 3 and 7 with regard to IP:

In Italy, for example, AGCOM has introduced in 2006 an IP interconnection obligation to SMP Operator in Markets 8,9 and 10 according to the principle of technological neutrality. In this context, Telecom Italia was obliged to temporarily apply the same PSTN economical conditions and provide access to the IP proprietary technical interfaces/protocols which itself use for its own services. At the same time, a specific proceeding was started to reach the definition at national level, and on the basis of ETSI, ITU standardization work, of standardized technical specifications for IP interconnection. This last proceeding is not yet concluded. In addition, using Article 5.1 and 5.2 AGCOM has set a symmetrical obligation to adopt the most efficient way to interconnect networks to allow the interoperability of VoIP services (this means that whenever IP interconnection is more efficient than CSS7 interconnection, IP interconnection should be adopted). In addition, AGCOM has introduced the obligation for operators to give access to their technical interface/protocols and to all the technologies necessary to allow interoperability of VoIP services. Standard protocols should be adopted whenever possible. A new market analysis for Markets 8, 9 and 10 should be started in may 2008 and will include IP interconnection.

These market analyses will be carried out as appropriate to national circumstances, taking utmost account of the Recommendation. This document cannot anticipate any results of fu-

152 Different from the current Recommendation the former Recommendation additionally classified "Transit services in the fixed public telephone network" (Market 10) as susceptible to ex ante regulation.

ture market definitions. The aim of this chapter is to highlight further developments of these markets without assessing if future markets would be susceptible to *ex ante* regulation.

The markets that are regulated today relate to the public telephone network. Telephone services provided on the public telephone network are not confined to voice service, for which the public telephone network was originally designed but include fax and Internet access.

It should be noted that the definition of the term “Public telephone network” is deleted in the Review proposals.

In a first step NRAs will have to examine whether these markets, which had been traditionally reduced to interconnection for narrowband telephone services contain interconnection for broadband telephone services.

These markets would still be service-specific. Considering the background information given in this report the division between transport and service may lead to different market definitions.

One possible way would be to define markets for origination and termination regardless of the services that are originated and terminated. Markets for “transport interconnection” could emerge out of this approach. Pure transport interconnection would be similar to current interconnection arrangements on IP networks (transit and peering) at the transport level but do not distinguish between origination and termination. Therefore, only a single market for connectivity rather than separate markets for origination and termination would be defined. Accordingly, the Commission speaks about a market for “Wholesale Internet connectivity” in its Explanatory Note to the new Recommendation. Originating and terminating Internet access providers do not receive any payment for terminating and originating traffic, which limits the possibility of exerting market power by asking for excessive origination/termination rates. Therefore, the connectivity market has in the past not been part of the recommendation as the “3 criteria” test was not deemed to have been fulfilled. This shows that different billing mechanisms which present different opportunities to abuse market power may lead to the delineation of different markets and different outcomes in terms of SMP.

Thinking beyond today the implementation of different transport classes could lead to further connectivity sub-markets. Currently, network operators have no interconnection offers for transport classes in place which fulfil network performance parameters but this may change in the future as stated above. If such offers were established one would have to check whether higher network performance is to be considered a substitute for best effort or not.

Another possible way would be to define multiple markets for origination and termination with reference to specific services. This could be the case if different market conditions are determinable.

Additional interconnection markets could arise at the service level. Access to SIP servers (e.g. routing information behind E-164 number) or services concerning AAA (authentication, authorization, and accounting) are often named in this context.

This shows that interconnection markets will evolve over time and leads to the question if the wording used in Markets 2 and 3 of the Recommendation is future-proof.

- What are the implications of the deletion of the definition of the term “public telephone network” from the Review proposals?
- Is the term “call” sufficient for future markets that do not feature telephone services?

It is likely that the term public telephone network will converge to the broader definition of electronic communication networks. This would confirm the multi-service network philosophy of NGN.

The term “call” is only defined in the Directive on privacy and electronic communications.[153] It may need to be described in a more general way such as call: generic generic term to describe the establishment, utilization, and release of a connection (bearer path) or data flow, where the flow is the bearer traffic associated with a given connection or connectionless stream having the same originating node, destination node, class of service, and session identification and connection is the bearer path, label switched path, virtual circuit, and/or virtual path established by call routing and connection routing. Such a definition seems to be more appropriate for further developments of interconnection markets.

C.3 Bottlenecks and SMP positions

As illustrated in Section A and B, NGNs are conceptually characterized by a horizontally layered structure differentiating between transport and services with both spheres comprising a multitude of functions and reference points (i.e. interfaces) defined by corresponding standards.[154] While the Internet model is built around a “dumb” network layer concentrating the intelligence at the edge of the network (e.g. in clients and servers), the NGN model incorporates intelligent NGN-specific functions in addition to the edge device intelligence. NGN building blocks are defined in a generic way and practical implementation is left to vendors and operators which may have the possible effect that today’s service-related regulation will no longer be the adequate answer to competition problems. The NGN concept allows for multi-vendor and (possibly) multi-operator environments making functional interoperability on the various levels a key issue.

C.3.1 Interoperability issues

When examining interoperability in an NGN environment a distinction must first be made differentiated between vendor interoperability and operator interoperability.

---

153 See also Sec. C.1.2
154 E.g. as specified by ITU, ETSI, or 3GPP

Operator interoperability is about open, standardised interfaces allowing one operator to access another operator's NGN and/or its related functions which may result in additional standardisation requirements. This issue is relevant between interconnected NGNs or between NGNs and interconnected legacy networks and may become critical in terms of competition at the various levels of the value chain where old bottlenecks may be fostered and new ones may be established. In the following, operator interoperability issues on the various NGN layers are examined by identifying potential bottlenecks that may become relevant in an NGN scenario from a regulatory point of view. In addition, issues that may be relevant in relation to operators having significant market power in a specific market are discussed in the following.

Vendor interoperability is characterized by open, standardised interfaces allowing operators to combine equipment from different vendors within their own NGNs. Moreover, it can be considered a precondition for operator interoperability. Vendor interoperability is an issue mainly relevant to single operators deploying their own NGNs and mainly to be addressed in standardisation bodies, where NRAs do not intervene directly. However, the absence of vendor interoperability has the potential of introducing new bottlenecks in interconnection, when a competing operator's equipment is not compatible and interoperable with the equipment used by the incumbent operator. Therefore, vendor interoperability may become a regulatory issue as NRAs have to consider the possibility of market power being leveraged to a variety of markets by deploying NGN systems with vendor interoperability missing. Nevertheless, the ERG does not see a need to act *ex ante* but dispute resolution mechanism may need to be applied.[155]

Several respondents consider interoperability to be a market issue which should be addressed by standardisation bodies whereas others stressed that the Framework's principles of end-to-end connectivity of services and interoperability of networks need to be maintained.

The ERG clarifies that, in line with Recital 30 FD, standardisation should remain primarily a market-driven process. Nevertheless NRAs need to monitor carefully whether lack of interoperability causes competitive problems. In order to ensure interoperability NRAs should ensure that the reference offer can contain a provision regarding a "minimum set of standards" applicable or refers to a set of specific standards.[156].

C.3.2 Transport-related bottlenecks

The major bottleneck of today's legacy telecommunication networks is the working access line connecting the individual customer. This situation is not expected to change in the future, as the customer has to be connected no matter whether it is a traditional or a next generation

155 See Section C.3.4 below

156 In Italy in a first consultation document about IP Interconnection AGCOM proposed to agree on a set of standards (which include NGN ETSI and ITU standards) to which all parties should be compliant when interconnecting. This does not mean that two Operators can interconnect to each other using a different standard. A final decision has not been taken.

network. However, with access technologies evolving regulatory intervention (remedies) may have to be adjusted accordingly.[157] Furthermore, these developments may influence the situation with regard to wholesale access markets, where the economic importance of service competition could increase relative to infrastructure competition.[158]

NGNs follow the concept of a universal core network utilising different fixed and wireless technologies in the access. Controlling these access networks is a necessary prerequisite for enabling operators to provide their customers with QoS-enabled products. As the service layer (see next section) exercises control on the transport layer, e.g. by reserving bandwidth for QoS-enabled services, bottlenecks may arise when service-related functionalities are not allowed (for technical or policy reasons) to interact with transport-related functionalities. Therefore, an integrated operator controlling both service and transport may have significant competitive advantages compared to an operator relying on wholesale offerings.

Impact of the charging mechanism on the relevance of transport bottleneck

It was shown in Section B.2.3 that the termination bottleneck and the possibility to exploit SMP results from the interplay of three factors: a) physical bottleneck for termination; b) control of the E-164 number; c) charging mechanisms.

The potential for abuse of the physical bottleneck for termination is closely linked to the charging mechanism.[159] With CPNP, this bottleneck can be exploited because it entitles the terminating operator to receive a payment out of its position of control over this bottleneck. Furthermore, a high termination fee does not hurt the terminating operator’s market position in competing for customers since this fee is not levied on its *own* customers but ultimately levied on the customers of its interconnecting party, i.e. the calling party. Termination fees may work as a collusion device allowing access providers to keep retail prices high[160]. Therefore application of CPNP generally leads to the determination of SMP in the relevant termination market with subsequent remedies being necessary to apply.

Thus, contrary to what is claimed in some responses, NGNs are in themselves unlikely to change market power in the market for termination of voice services as long as CPNP is the charging mechanism used. Application of the CPNP regimes therefore ultimately perpetuates the need for regulation to obtain cost-oriented termination rates. Moreover, even if customers have different access options available to them, the calling party still encounters a physical terminating bottleneck.

---

157 E.g. the terms and conditions for unbundling may have to be adapted in order to address needs in the advent of new developments of hybrid fibre-copper access networks in combination with VDSL roll-out

158 The ERG Common Position on Regulatory Aspects of NGA points out that the problem of bottlenecks will – at least – remain in Next Generation Access Networks. *“Given that next generation access networks may be more likely to reinforce rather than fundamentally change the economics of local access networks, NGA may be likely to, at least, provide the same competition challenges to regulators as current generation wireline access networks.”*, ERG (2007a), p. VII.

159 See also WIK –Consult (2008, p. XI. “*As long as Europe operates under CPP/CPNP retail and wholesale arrangements, the migration to IP-based NGNs will not ameliorate the termination monopoly (the tendency to charge high prices for termination) until and unless mechanisms emerge to enable more than one telephone service provider to terminate calls o a single telephone number*.”

160 See Armstrong (1998)

Furthermore, contrary to what is claimed in some responses, at present the ERG does not see that other addressing mechanisms such as ENUM will alleviate the problem resulting from control over the E-164 number or any other private addressing mechanism.[161]

As can be seen from peering and transit agreements and the application of Bill & Keep provisions for the termination of traffic on the last segment of the access provider's network it is not capable of abusing the same market power growing out of the physical bottleneck if money flows are reversed and competition is effective at the retail level. If the broadband access provider is facing competition for its end-users, its incentive is to pay low transit fees.[162].

This analysis is in conformity with the Commission's position on the Internet connectivity market not included in the list of markets susceptible to ex-ante regulation as set out in the Explanatory Note emphasizing the relevance of direction of payment flows for the absence of market power. It is noted that it is the end-user who pays implicitly for sending and receiving packets and there was no wholesale payment for incoming traffic, where the charge is passed to the traffic sender via the network. Moreover, it did not consider it necessary to include the Internet connectivity market in the list of markets to be subject to ex-ante regulation[163].

Bill & Keep avoids the need for determining termination rates and thereby significantly reduces the need for regulatory intervention as long as two conditions are fulfilled:

- The transit market on IP-backbones is sufficiently competitive to exert competitive pressures on IP-backbone providers.[164] With an oligopoly of tier 1 providers allowing choice of transit provider this condition has so far been considered to be fulfilled.

- The broadband access market is sufficiently competitive so that access providers are under competitive pressures to be prevented from establishing abusive mark-ups on retail prices.

Therefore where Bill & Keep applies it is unlikely that SMP will be the outcome of a market analysis for the termination market.

---

161 Public ENUM does not seem to be embraced by network operators.

162 See Explanatory Memorandum (2007), p. 36. „Since any terminating charge is incorporated into the overall amount that is charged by the ISP (and faced by the end-user), and end-users can switch between competing ISPs, ISPs have an incentive to minimise the termination charges that they pay."

163 See Commission (2007), p. 37 „There are a number of differences between the typical arrangements for terminating calls on the public telephone network and delivering packets to destination addresses on the public Internet. In the latter case, end-users are implicitly paying to both send and receive packets. It is not automatically or typically the case that incoming traffic is charged for and that this charge is passed to the traffic sender via the sender's network. As indicated above, traffic connectivity can be arranged in a number of ways. Entry barriers to this market are low and although there is evidence of economies of scale and that the ability to strike mutual traffic exchange (peering) agreements is helped by scale, this alone cannot be construed as inhibiting competition. Therefore, unlike the case of call termination in section 4.2.1, there is no a priori presumption that ex ante market analysis is required. Therefore, no market for wholesale Internet connectivity (or delivery of incoming packets) is identified for the purposes of the Recommendation."

164 NRAs may still need to monitor whether this condition is fulfilled. This was also mentioned in some responses.

QoS for NGN services

On the NGN core network, traffic is exchanged on IP layer as is the case in many of today's legacy data networks. While this traffic exchange based on IP packets today is mainly performed on a "best effort" basis, NGNs are expected to offer (and guarantee) service and quality levels for value added and real-time services as e.g. voice, video or gaming. Therefore, interconnection in NGNs is enhanced allowing operators to differentiate services by defining network performance parameters e.g. minimum delay, jitter, packet loss or bandwidth needed for the provision of a certain service.

Such QoS control might imply the introduction of a premium transport class that requires access to associated transport-related functionalities, as without proper access to these functionalities interconnection will be possible on a best effort basis only. It is evident that the introduction of interconnection transport classes providing better quality than best effort is a potential bottleneck. Possible strategies to prevent such a situation would be

- imposing non-discrimination in the case of an SMP operator in the relevant market, or
- imposing minimum quality (independent of SMP) in the case it is proven that there exists a strategic incentive to deteriorate best-effort transport – applicability of Art 5 AD, Art 22 refers to transport layer.

In the discussion regarding quality issues in an NGN context, it is often questioned whether QoS will be demanded and paid for at all by the end customer because they do not perceive any differences. This question has to be looked at with a broader view as it seems likely that demand may rather come from providers than from end customers. Service providers know best about the quality demand for their services, e.g. what delay and jitter is acceptable for video telephony or what bandwidth is needed for video streaming.

In traditional telecommunications networks, the issue of providing quality across network borders and within third-party networks is usually addressed by setting up mutual service level agreements (SLA). In NGNs, that problem remains, as current technology is still not very well equipped to support end-to-end quality across IP network borders.[165] This may result in NGN operators having to agree on mutual SLAs again to guarantee network boundary quality.[166] NRAs will have to monitor the situation and prepare for regulatory intervention when network border quality becomes a bottleneck for alternative providers. This may be the case regarding interconnection between NGNs, and between NGNs and legacy networks alike.

Further potential bottlenecks are the various associated facilities used in IP interconnection, including firewalls, proxy servers, and network address translators.

---

165 Cp. MPLS interconnection issues currently being subject to work in progress.
166 It has to be mentioned in this context that agreeing on SLAs for IP-based traffic is significantly more complex than in legacy telco networks because more different elements need to be considered.

C.3.3 Service-related bottlenecks

The service layer comprises several functions with regard to service and control, customer profile management, and applications (e.g. for realizing APIs). These functions are invoked by applications and will utilize transport/service-related functions for provisioning of services to the end-users.

One critical factor is the question whether service-related functions are open or closed to competing operators. The extent to which information, functions and interfaces will be open to competitors will be important for the regulatory assessment regarding NGN impact on the services to be offered to the consumers. However, this question should not be answered based on generic NGN principles or standards; instead, this should be evaluated by NRAs as soon as practical NGN implementations are known in detail. To illustrate the complexity, the following list shows some example functions situated on the service layer:

- Home Subscriber Server;
- User Profile / User Identity;
- Location Information;
- Call Session Control Function;
- Charging Collection Function / Online Charging System;
- Policy Decision Function;
- Border Gateway Control Function;
- Authentication and Key Agreement;
- Terminal Capabilities.

Barring the access to this kind of functions as well as restricting access to the SIP servers (providing access to numbering information), may be considered a bottleneck, if hindering service interoperability (end-to-end connectivity) and interconnection.

Apart from the general requirement of open interfaces, interoperability of similar functions in different NGNs may be another issue for regulatory attention. Differing standards, incompatible data formats, and proprietary implementations may cause the occurrence of new bottlenecks. NRAs will have to define, where and to what extent interoperability on the NGN service layer is necessary in order to prevent competition problems. It is not a regulatory target per se to enforce interoperability on all NGN layers and for all services and applications, but to intervene when practical competition problems occur. Such competition problems can be expected to be more critical when general access to users on network layer ("Internet-style" access) is restricted for other operators (networks) in one or the other aspect.

Regarding applications (including APIs[167]) and services offered within an NGN, it would be desirable that the service layer is designed to enable a more efficient and modular creation of new services and to allow for third-party service development and offering. Moreover, regarding the universal compatibility and re-usability of applications created for one NGN to be used on another. NGNs from different operators and vendors should provide standardised interfaces and functions to enable service providers to offer their services in more than one NGN and to benefit from economies of scale effects.

The high penetration figures of popular Internet applications to a significant extent rely on the universal compatibility of applications in the Internet TCP/IP model as global access to users via the network layer is generally open.

Voice services in a NGN environment need transport and service interconnection as the former provides connectivity, while the latter provides the necessary service-related control and signalling functionalities.[168] Regarding termination of calls based on VoIP this means in principle, that voice data packets – regarding transport– might be sent to each and every public (IP) address on the (Inter)net. Though this ubiquitous connectivity in principle has the potential of breaking the termination monopoly for voice calls, the control functions on the service layer generally will prevent such procedure. VoIP calls are set up using higher-level protocols, e.g. SIP, that provide a translation from an individual customer's E.164 number or Internet-style "user name" to an actual IP address that is needed for a call to be terminated. As this IP address is only known by the customer's VoIP provider[169], the termination monopoly is set to remain also in the NGN world as the called party's VoIP provider is still needed to terminate a call, even though only with regard to signalling matters.

Furthermore, NGN voice telephony in most cases will still involve the usage of E.164 numbers that are often demanded as identifier (CLI or Calling Line Identification) when interconnecting with legacy voice telephony networks and, most importantly, when calling emergency services, 112. This further fosters the importance of (VoIP) telephony providers having E.164 numbers allocated.

"Telco-style" approaches to the integration of new services into an IP-based platform are significantly different from those that are typical for the Internet today. For the Internet, services are generally implemented at the edge of the network, often by the addition of new servers, or of new software into end systems. VoIP is an obvious example, where a service provider might achieve market entry by simply deploying SIP servers and (for pri-

---

167 Applications in the NGNs utilise application programming interfaces (API), originally known from software development allowing programs to access other software systems. Today, APIs are also provided by several services on the Internet. Cp. http://www.flickr.com/services/api/, http://www.youtube.com/dev, 19.12.2006.

168 The distinction between transport and service interconnection has an important implication. Understanding service interconnection as including solely service-specific aspects, service interconnection may not fulfil the definition of Art. 2 (a) AD because it does not necessarily include the physical linking of NGN domains.

169 Typically, the end customer is provided with an IP address from the access provider's IP address pool. This IP address may change from one log-in to another (dynamic IP address) or may stay the same from one log-in to another (static IP address). However, the IP address will change in any case when the customer logs in to a different access provider's network utilising a different IP address space.

vacy/security reasons) Session Border Controllers and connecting them to the Internet, without making any special arrangements with network operators.

This is viable only to the extent that the underlying networks already provide the kind of connectivity that is needed and this model of service deployment provides no inherent solution if needed transport services are not inherently present in the network – for example, support for different classes of service.

By contrast, the NGN concepts that many incumbent operators are implementing (or propose to implement), such as BT's 21CN, seek to integrate currently separate networks into a single network and service platform with a centralized control platform containing corresponding functions for the different types of services, as well as a service/application creation environment. This approach can support access by a range of devices and corresponding access networks by means of corresponding access gateways; at the same time, it can also support the integration of mobile access by means of the IP Multimedia System (IMS).

The two concepts (Internet versus NGN) imply different interconnection strategies, both at the transport and at the service layer. For example, a pure service provider with little or no infrastructure can easily offer services at a national or even international level using the public Internet, as is the case with Skype or SIPgate. By contrast, the integration of new services into an NGN (based, for instance, on IMS) will require strong coordination with the NGN operator, and thus depends on the latter's willingness to open up the NGN to independent service providers. NRAs should observe the evolution and effective application of the concept(s) to prevent the build up of "walled garden(s)".

While incumbents do not expect new bottlenecks to emerge or even for existing ones to disappear in an NGN context, other responses saw potentials for new bottlenecks such as authentication or presence services.

From the ERG's point of view it is important to monitor carefully whether new bottlenecks leading to SMP positions will emerge or whether certain bottlenecks will disappear. Observing the ongoing migration towards IP-based networks and the related technological changes is necessary in order to be prepared in case new bottlenecks should arise.

C.4 Regulatory measures

C.4.1 Symmetric regulation (Art. 5 Access Directive)

Generally, providers of electronic communications service can be obliged to symmetric obligation to negotiate interconnection according to Art. 5 AD. This applied to both PSTN and IP-networks. The provisions have been rarely used.

A notable exception has been AGCOM, which has recently set a symmetrical obligation referencing to Article 5.1 and 5.2 to adopt the most efficient way to interconnect networks to allow the interoperability of VoIP services (this means that whenever IP interconnection is more efficient than CSS7 interconnection, IP interconnection should be adopted). In addition, AGCOM has introduced the obligation for operators to give access to their technical interface/protocols and to all the technologies necessary to allow interoperability of VoIP services. Standard protocols should be adopted whenever possible.

As QoS mechanisms are not yet widely deployed at interconnection points and due to the interdependence between each network involved in the session, one operator could be unwilling to invest in QoS mechanisms if the interconnected networks have no intention to do so. Thus, NRAs can also use symmetric regulation tools in order to enhance QoS development between different networks, as operators will have real benefit from QoS implementation when QoS mechanisms are widely supported and in place at interconnection points.

C.4.2 Measures based on USO directive

- Regulators could require operators to provide public information about QoS, based on articles 20 and 22 of the Universal Service Directive. Obviously, such evaluation would require agreed and comparable indicators to be used by all concerned operators. In order to verify that there is no discrimination of QoS between operators interconnecting, NRA may also add other relevant QoS measurements.

- Moreover, NRA should have the possibility to recommend or even to set minimum levels of quality of service. This relates to the fact that there might only be a willingness to pay for a premium transport class in case the best effort class quality is “bad enough”. Therefore operators might have an incentive to degrade their best effort class (see new provision in Art. 22 para 3 UD).

  Whereas most respondents to the consultation claimed that any minimum QoS should be proposed by the industry and established standardization bodies, others saw a role for NRAs in setting minimum QoS. Diverging views were expressed regarding the question whether the setting of quality parameter should take place on the level of NRAs or rather at the level of the Commission.

  ERG therefore welcomes the proposed provision in Art. 22 para 3 UD, but considers that the power to set minimum quality of service requirements should be entrusted directly to NRAs.[170]

---

[170] The EP adopted an amendment to that effect in its vote on September 24.

C.4.3 SMP-remedies (based on Art. 15, 16 Framework Directive, Art. 9-13 Access Directive)

The following imposition of remedies applies only in those cases where SMP has been found on a relevant market susceptible to *ex-ante* regulation, i.e. the market has passed the 3-criteria test.

Regulatory implications regarding termination bottleneck

- CPNP leads to SMP in termination markets usually implying *ex-ante* price regulation remedy

  In cases where NRAs want to shift to a Bill & Keep regime they could consider imposing termination rates of zero for the terminating segment up to the first router or switch and associated service control functions after the access/concentration network.

  The possibility to impose Bill & Keep under the current regulatory framework could be explored further by ERG as well as other means to move towards Bill & Keep.

Regulatory implications regarding QoS

- While core networks are migrating to NGN architectures, new forms of discrimination based on QoS discrepancies could emerge between a larger operator’s services and those provided by interconnecting operators. NRAs should prevent any anticompetitive behaviour from SMP operators that might degrade quality of the interconnection with some specific networks to benefit their own quality service. In order to avoid this possible deviation, NRAs can use existing tools to impose non-discrimination obligations on SMP carriers in case markets have been defined accordingly.

Regulatory implications regarding interfaces

- The capacity and converged nature of NGNs offer potential for development of new services and service propositions. At the same time, the implementation of NGNs is expected to lead to a clear distinction between service-specific functions and transport functions common to all services. Such distinction could support more rapid and less costly development of new services than legacy networks currently allow. The same distinction could allow an SMP operator to provide interfaces for competitors to control services transported over its own network, which could promote effective competition in two ways: a) reduce barriers to the entry of new players, who could enter the market with limited investment in infrastructure, and b) allow competitors with more substantial investments in infrastructure to increase scale by using the SMP operator’s network to extend the market addressable by their services.

  NRA’s may therefore need to consider interventions to make such interfaces available. The interfaces are likely to be technically sophisticated and will be subject to change as new services are developed. Opportunities could therefore arise for forms of discrimination other than price. For example, the owner of the transport network could control the relevant functions using a separate or enhanced version of the interfaces from the one it offers to its competitors, and therefore frustrate equality of access to the network.

The *ex ante* framework allows NRAs to address differences in the product quality by applying non-discrimination obligations to providers found to have SMP. However, where these obligations may not be sufficient to deliver a level playing field, NRA's could consider more stringent requirements, including equivalence.

Equivalence is a remedy in respect of a wholesale service, which requires the SMP operator to provide the same service both to its external wholesale customers and to its internal downstream operations. In its stricter form, known as equivalence of inputs, the SMP operator is required to provide the same service to both external and internal users on the same timescales, terms and conditions (including price and service levels) by means of the same systems and processes, and includes the provision to both external and internal users of the same commercial information about such products, services, systems and processes. Under a less strict form, equivalence of outcomes, the wholesale input supplied to the SMP operator's own downstream divisions is equivalent to the comparable product or service supplied to other users, but is not necessarily supplied in an identical manner.

C.5 Costing and pricing

The cost structures of NGN when compared to legacy networks will probably differ in a number of ways.

It is generally accepted that NGNs core will lead to a lower overall cost level due to increased economies of scale and scope and that the cost structure will change with a higher proportion of common costs compared to legacy networks. The use of common platforms to deliver multiple services across one network exploits economies of scope allowing the opportunity to recover an overall lower cost base across a range of services.

- The Opex and Capex of a NGN are forecast to be significantly lower in the long term than current legacy technologies as NGN core networks are generally seen as providing a more efficient network design and usage. The main reasons for this can be summarised as:

  - less physical layers (fewer network hierarchy levels);

  - fewer network components and interfaces (rationalisation of network components), in particular fewer nodes per layer (depending on the technology adopted);

  - higher capacity of NGN equipment, because of packet switching technology, resulting in lower per unit cost (€/bit).

These three factors – simpler network structure with fewer levels and fewer nodes at each level plus more efficient equipment (packet switching technology) – lead to a reduction in the total per unit cost of NGN core networks.

Common and fixed costs of NGNs will represent a high percentage of total costs with a relatively lower percentage of costs incremental to individual products or services compared to legacy networks.

The cost/volume relationship of a NGN seems to be shallower at current volumes than legacy networks suggesting that increases in volumes will have a relatively low incremental cost impact.

Taking these factors into account, NRAs will need to consider how modelling and costing approaches are adapted in supporting costing and pricing decisions in SMP markets

The technological neutral guidance on regulatory accounting principles published by ERG[171] will continue to apply for the calculation of NGN costs. While the accepted costing principles of cost causality and recovery of efficiently incurred costs remain valid, their application might need to be adapted to take proper account of new cost drivers. Also, allocation keys for the relatively higher share of common costs need to be adjusted where necessary to ensure competitive neutrality, e.g. of bundled offers. The use of current cost accounting (CCA) together with LRIC or efficiently incurred FAC/FDC is a preferable framework for estimating efficient costs. CCA when used to inform pricing decisions sends economically sound signals to the market as it calculates the costs relevant for decision making in a competitive market. NGNs will become or are already the accepted modern equivalent asset (MEA) for core networks.

A key feature of a robust NGN model is likely to be the way in which it deals with the capabilities of the technology to deliver multiple services across a network with a high proportion of common costs. This suggests that NRAs will need to understand the cost orientation and cost recovery (pricing) implications of both SMP and non-SMP services running across the NGN platform.

The use of traditional costing methodologies, such as LRIC, will need to be adapted to recognise the different cost characteristics of NGNs. As all services (voice, data, video) are converted into a similar format (i.e. packets) in NGNs traditional parameters (cost drivers) such as minutes and distance are no longer the most relevant. Hence, the calculation must reflect per bit cost or capacity required by the service measured e.g. with contended bandwidth (defined as the absolute bandwidth required for each service taking into account, delay, priority, QoS) or number of packets/sec or number of simultaneous calls. Contended bandwidth measures the dimensioning needs for the network and can therefore be identified as the main (but not only) cost driver. The change of parameters as well as the change of network

---

171 ERG (05) 29 – ERG CP „Guidelines for implementing the Commission Recommendation C(2005) 3480 on Accounting Separation & Cost Accounting Systems under the regulatory framework for electronic communications", see ERG website www.erg.eu.int; and IRG (05) 40 IRG PIBs on CCA, see IRG website: www.irg.eu.

structure is likely to result in a change of the current narrow-band IC pricing structure, i.e. the minute-based 3-stage (local, single, and double transit) structure will not be appropriate for a packet based network (see below pricing regime).

There is a possibility that an operator may be left with stranded legacy assets as NGNs are introduced. However, following the technology neutral costing principles, these costs are not relevant for regulatory accounting and are not accepted. Any “double counting” should be avoided and legacy assets should not be costed. Also, the sub-optimal use of capacity in the migration period (due to running in parallel the legacy and the next generation network) may not lead to a cost increase as this would be inefficient while only the cost of an efficient operator should be taken into account. In general the cost of efficient service provision should be used as the cost standard for approval of interconnection rates. The pricing should be valid irrespective of whether interconnection is realized via circuit-switched or packet-switched networks, since strict application of the cost standard of long-run incremental costs requires the efficient technology used by the market players to be taken as a basis. Consideration must also be given to the fact that the concept of the cost of efficient service provision does not differentiate the price according to technology used or account for the existence of different prices for the same service. Basing prices on efficient technology also provides incentives for speeding up the migration to this technology.

Based on the hypothesis that the economic rationale for NGN's is partly based on the expectation that the costs of delivering voice services in the long run will be no higher (and probably significantly lower) than using legacy PSTN technologies then it is reasonable for NRAs, in modelling and evaluating NGN costs and/or associated pricing decisions, to assume that the cost of voice services will be no higher than currently calculated.

Pricing regime under CPNP

The following section applies only in a CPNP world where SMP has been found in the relevant termination markets and a price control measure according to Art. 13 AD was imposed. Once in a Bill & Keep regime, prices do not need to be set by regulators any longer.

It is important that the pricing methods send the right economic signals in the relevant markets and reflects properly the underlying cost drivers. For example, prices should not be set at a level that distorts investment decisions or allows exploitation / leverage of market power.

This section discusses established wholesale pricing methods and the ways in which they may apply in a NGN environment. It would appear that some methods would suit the changing economics of an NGN better than others. For example as the share of common costs is likely to be higher than in legacy network costs, this may lend itself more easily to capacity based charging at the wholesale level and flat rate pricing on the retail level than in an EBC regime. It could be that a specific charging mechanism fits “naturally ” for certain services.[172] However, recent developments across most European retail markets show that flat rates are also sustainable under EBC wholesale regimes.

---

172 Cf. section C.6.

EBC vs. CBC as possible regimes for a NGN pricing structure under CPNP

Wholesale charges in a CPNP model can either take the form of Element Based Charging (EBC) or Capacity Based Charging (CBC). Both systems constitute cost-based systems in the sense that NRAs refer to specific efficient cost standards when they determine wholesale rates thereby assuring efficient incentives for investment, i.e. dimensioning the capacity of the network required to handle the traffic at peak time. Usually, the efficient costs consist of LRAIC plus a mark-up for common costs including an appropriate rate of return on capital employed.

EBC[173]

EBC is the pricing system predominantly used for the regulation of narrowband interconnection rates in Europe.[174] Under EBC, the interconnection rates depend on the number of network elements that are required for the completion of a call. Before the application of EBC interconnection rates were largely distance-related. By changing to EBC, it was intended to better reflect the underlying cost drivers and structure of networks, whose costs are becoming progressively less sensitive to distance.

The price structure reflects the network structure in terms of hierarchy and topology, and in most countries tariffs are differentiated in local, single transit and double transit. The category applied depends on the number of switches passed. Each call – either local, single or double transit – is characterized by a certain profile of network usage implying the specific network elements passed by the call. As in a circuit switched network, each connection passes a predefined path (routes), the network elements are known in advance (by way of the routing matrix) and the cost of a call is the sum of the costs of the elements passed. To determine the total network costs the capacity needs to be dimensioned to carry the traffic at the busy hour (peak load pricing also known as time of day weighting). For this, analytical bottom-up cost models are applied in many countries.

Under EBC, usage of interconnection services is usually charged on a per-minute basis as minutes can be measured relatively easily.

EBC may be less appropriate in the context of IP-based networks where a relatively higher portion of common costs makes the allocation of costs to network elements less easy.

If EBC is carried over to NGN, the price structure would need to be adjusted according to the NGN architecture, i.e. most likely there would only be a 2-level price structure and the number of network nodes (= PoI) would be less. Moreover, the usage would likely not be measured in minutes any more, but in (contended) bandwidth used (bits per second).

---

173 See Vogelsang (2006, Ch. 3.2.1/3.2.2) for a comprehensive discussion of EBC and CBC as well as their specific properties.

174 More extensive descriptions of the implementation of EBC can be found in the numerous IC rates decisions of NRAs, e.g. BNetzA (2006) or Ofcom (2005) www.ofcom.org.uk/consult/condocs /regfinch/main/regfinch.pdf

CBC

In a CBC regime the interconnection charge does not depend on the volume of traffic exchanged between the operators; it only depends on the traffic bandwidth (i.e. number of channels) that both parties have agreed on the contract.

This modality implies that both interconnected operators allocate the necessary network resources to guarantee the interconnection demand, with a charging structure based on the number of interconnection links agreed between both parties.

The central feature distinguishing CBC from EBC is that, under the CBC, system bandwidth (channels or bit/s) is being bought in advance by competitors. This leads to a change in risk sharing between incumbent and competitor as CBC requires competitors to plan the dimensioning of capacities more carefully. This may pose a problem in particular for smaller operators as they have a smaller customer base. Such a risk distribution between incumbent and competitors might lead to a higher degree of market concentration.

In a CBC regime, the price would therefore depend on the capacity required at peak time as this determines the share of costs to be borne by the competitor according to the bandwidth booked. The price would be expressed in €/Kbit/s.

CBC regime allows operators to request a specific capacity for interconnection and pay a flat rate charge that reflects the fixed cost nature of the interconnection capacity. As interconnection capacity is dimensioned to peak-hour traffic, CBC rates reflect true economic costs and do not require spreading such fixed costs over projected traffic minutes to arrive at a per-minute charge.

In this way, CBC allows operators to benefit from the economies of scale of incumbent operators and decreases unitary prices compared to time-based interconnection. Further, it provides additional tools and incentives to operators to increase the flexibility in retail tariffs.

Compared to time-based interconnection, alternative operators benefit, and achieve higher levels of usage, from lower effective termination prices when they contract capacity based interconnection. Beyond a certain threshold of minutes, the per minute cost decreases substantially. This implies that (except for the smaller or niche operators), it is a very convenient system for interconnection.[175, 176]

C.6 Charging mechanisms

In this section the pros and cons of different charging mechanisms (Bill & Keep and Calling Party's Network Pays) as core elements of an interconnection regime for termination service are systematically compared with regard to a number of criteria. The charging mechanism

175 Examples of existing CBC regimes in Spain, Italy and Poland can be found in Country Studies in Annex 3.

176 OECD (2007, p27) The OECD further refers to the difficulties of determining capacity charges such as changing peak usage profiles depending on the services being used.

that usually applies in PSTN and mobile networks is CPNP, where termination services are paid for at the wholesale level following the direction of call flows. In IP-based networks if wholesale payments for transit apply, they flow in the upstream direction. There are no payments for the terminating segment of the broadband access provider. The different charging mechanisms predominantly used in practice and their implications have been described in section C2.

As networks migrate towards NGN infrastructure it is unclear *a priori* whether these future networks will be governed by the mechanisms currently used in IP-networks or whether the mechanisms currently applied in the PSTN will be carried over to NGNs.

The charging mechanism, deals with "who pays" as opposed to "what is being paid for" which was the topic of the previous section. Coupled with a direction of payment flows it may have implications on the definition of relevant markets and the determination of SMP as has been discussed in sections C2 and C3.

A number of NRAs have taken up these issues initiating discussion on the appropriate charging mechanisms for NGNs.[177] In Austria and Slovenia, for example, a dialogue has been initiated covering the future billing method to be used when settling interconnection costs between NGNs and the question whether Bill and Keep might be a preferable billing method within an NGN environment. In Germany, the Federal Network Agency published its "Key elements of IP Interconnection" which reflect the current state of the debate acknowledging that many issues are still open due to lacking practical experiences with interconnection of IP-based networks. And in the UK, NGNuk commissioned a report[178] "NGN Interconnection: Charging Mechanisms and Economic Efficiency". The EU-Commission has commissioned a study on "The Future of IP Interconnection: Technical, Economic, and Public Policy Aspects" which was published in January 2008.[179] This study, *inter alia*, elaborated on the relationship between the wholesale billing regime and the existence of the termination problem.

Some NRAs consider Bill & Keep a promising concept that should be aimed at in the medium to longer term. They focus on further studying how to best achieve this goal including finding answers that arise in a transition phase to a new system. Some NRA's, while recognising the merits of Bill & Keep, rather emphasize the risks implied by a change from the well-established regulatory regime of mainstream PSTN and mobile services and therefore see the need for further study.

Next to the objectives and principles mentioned in the Regulatory Framework (see Section C.1) the following economic criteria should be adhered to when NRAs evaluate different charging mechanisms:

- Sustainable competition should be intensified.
- Efficient investment should be encouraged.

---

177 For details see Annex 2.
178 NGNuk (2007)
179 WIK-Consult (2008)

- Incentives for efficient network use should be given.
- Transaction costs of market players as well as for NRAs implied by a particular interconnection regime should be minimized.[180]
- Interconnection regimes should avoid potentials for regulatory induced arbitrage.[181] Exploitation of such potentials might reduce market efficiency.
- Network externalities should be internalised.[182]

The two systems CPNP and Bill & Keep[183] exhibit very different properties with regard to these points, which will be discussed after giving a definition for CPNP and Bill & Keep.

C.6.1 Definition of charging mechanism discussed

CPNP

The regime most commonly employed so far in traditional circuit-switched networks is CPNP. Under CPNP, the network operator of the calling party pays for the whole call and the receiving party's operator pays nothing for incoming traffic. Instead, the called party's operator receives a payment for the termination service provided to the calling party's network.

Bill & Keep

Bill & Keep is a wholesale billing regime under which each network bears the costs of terminating traffic coming from other carriers. Therefore under Bill & Keep the terminating access network operator does not receive payments at the wholesale level for the termination provided. Instead, it recovers its net costs incurred for termination –- and any payments for upstream connectivity – in other ways, e.g. by billing them to its end customers.

Bill & Keep might be applied for the terminating segment up to the first router or switch and associated service control functions after the access/concentration network. Transit services

---

180 Transaction costs can take different forms such as: costs of negotiations, measurement costs (traffic), costs incurred through delaying the achievement of interconnection agreements, costs of interconnection disputes (market players, NRS, courts), of determining interconnection rates (e.g. cost modelling), of adapting to a new interconnection regime / possible changes with regard to retail pricing systems. additionally to transaction costs in a narrow sense (like costs of negotiation) rent seeking behaviour also causes transaction costs. This list is not meant to be exhaustive. See also Vogelsang (2006), ch. 2.1.4. Marcus (2006) elaborates on this issue focussing on intercarrier compensation charges and intercarrier compensation accounting as well as dispute resolution, ch. 3.3.1 & 3.3.2.

181 It should be noted that in competitive markets arbitrage is efficiency-enhancing. The term is regulatory induced arbitrage refers to rent-seeking behaviour that seeks to take advantage of cost or revenue disparities that are due solely to regulation (DeGraba, 2000, p.1).

182 Telephone networks are a typical example of positive network externalities. Network externalities may arise when the decision of an individual to join a network does not take into account the (positive) effect this has on other users of the network.

183 With regard to the properties of Bill & Keep see e.g. ERG (2007), Ch. 4.2.3.; Marcus (2006a); or – very extensively – Vogelsang (2006), Ch. 3. For an analysis of CPNP properties, see also Vogelsang (2006), Ch. 3.

are not included in the Bill & Keep model as discussed here and operators may charge for that service.[184]

Since termination costs for the last segment are by definition not accounted for at the wholesale level symmetry (e.g. of traffic flows) cannot be considered a requirement for the applicability of Bill & Keep.[185]

C.6.2 On the relationship of retail and wholesale charging mechanisms

Wholesale and retail charging mechanisms are related because interconnection prices affect the structure as well as the level of the interconnecting operator's costs, impacting on the cost recovery and the retail prices of the services provided to the end-users[186].

Retail billing mechanisms can be Calling Party Pays (CPP) structurally corresponding to CPNP[187] in which the calling party bears all the cost of a call or Receiving Party Pays (RPP), in which the receiving party pays part of the call rather corresponding to Bill & Keep[188]. Nevertheless these structural similarities do not preclude flexibility in combining different wholesale and retail regimes.

Generally, applying CPNP as billing principle using EBC on a minute basis tends to set a floor for the retail price of a call because termination fees are perceived as costs for the network operator providing the voice service even if aggregated payments for termination net out due to traffic symmetry. Such a structure of wholesale pricing has therefore frequently lead to a differentiation of on-net and off-net prices (in mobile networks and for VoIP based on PSTN termination charges)[189]. Furthermore, high CPNP wholesale payments usually lead to an even higher CPP retail price and may preclude retail flat rates or buckets of minutes plans[190].

---

184 Applying Bill & Keep for the terminating segment up to the first router or switch and associated service control functions after the access/concentration network can be considered a minimum scenario. It may also be conceivable that the area of application of Bill & Keep might be extended by applying it already from PoIs higher up in the network hierarchy. See C.6.10. The suggestion to limit application of Bill & Keep to the terminating segment was first made by DeGraba (2000) in his COBAK proposal. Contrasting with the definition of Bill & Keep used in this paper some authors consider Bill & Keep to be applied all over the network implying that transit remains unpriced (e.g. Kennet/Ralph 2007).

185 Peering agreements that usually do require some degree of traffic symmetry should not be confounded with the definition of Bill & Keep as used in the document. The requirements for the applicability of peering are laid out in network operators' peering policies. In a number of papers Bill & Keep and Peering seem to be used synonymously (e.g. Kennet/Ralph 2007), explaining why it is argued that application of Bill & Keep requires some symmetry.

186 Vogelsang (2006), Ch. 3.3.1 and 7.3.3.

187 With CPP, the calling party pays for the whole call, similarly, the network operators pays for the whole call at the wholesale level. Neither the receiving end-user nor the receiving operator pay anything. The CPNP termination payment paid by the calling party's network can be considered a compensation for the fact that the receiving end-user does not make a payment to his network operators under CPP.

188 With RPP, the called party pays his operator for receiving calls whereas under the wholesale mechanism of Bill & Keep the terminating operator charges his end-user.

189 WIK-Consult (2008), p. 63., see also Laffont et al (2003) on the off-net pricing principle.

190 DeGraba (2000, p. 27/28), described the relation between the wholesale billing principle and the arrangements at the retail level: *"One source of inefficiency is that existing termination charges create an "artificial" per-minute cost structure for carriers that will tend to result in inefficient per minute retail prices. In unregu-*

Offering flat rates or bucket plans at the retail level implies a risk of loss-making if wholesale charges are usage-based because it is more difficult for the network operator to forecast the level of usage. Therefore, retail schemes on a flat or capacity basis are structurally more compatible with CPNP based on CBC than CPNP based on EBC.

Some responses consider CPNP to be superior because it allowed network operators to recoup their costs whereas this was not possible under Bill & Keep due to strong retail competition. This argument does not appear to be convincing because competition drives down retail prices only to the competitive level allowing recovery of efficiently incurred cost including a reasonable rate of return (see Section C.6.7). The costs that need to be recovered are the net costs and these will depend on, among other things, the traffic balance. These net costs could be recovered at the retail level.

Other respondents referred to the competitive distortions resulting from on-net / off-net price discrimination making Bill & Keep more attractive. Economic literature[191] argues that with on-net retail prices being priced below off-net call, customers of the larger network benefit from lower average retail prices compared to customers of smaller networks. This may occur because for terminating on-net traffic only marginal costs are relevant, whereas an operator, who gets his traffic terminated off-net, incurs average costs as relevant costs. High termination rates can create barriers to entry for smaller networks who cannot replicate the incumbent's pricing strategy profitably.[192]

Flexibility of retail pricing

However, flat rates have become increasingly popular at the retail level, reflecting end-user preferences for this type of tariff scheme[193]. In Germany for example, consumers responded enthusiastically to packages combining telephone lines with broadband Internet access and offering unlimited call minutes and Internet use while regulated CPNP applies at the wholesale level.[194]

Flat rate schemes involve a fixed fee per month independent of actual usage. This flat fee can be thought of as covering the total cost of outgoing calls (corresponding to CPNP) or as covering a share of the costs for incoming calls and a share of the cost for outgoing cost (corresponding to Bill & Keep). Where traffic flows can be considered symmetric, costs add up to the same amount for both cases. The end-users will focus mainly on the price level of

---

lated, competitive markets, such as the markets for CMRS services and Internet access services, retail pricing is moving away from per-minute charges and towards flat charges or two-part tariffs that guarantee a certain number of free minutes. This suggests that few costs are incurred on a per-minute basis, and that flat-rated pricing will lead to more efficient usage of the network. The existing reciprocal compensation scheme, which requires the calling party's network to pay usage sensitive termination charges to the called party's network, imposes an "artificial" per-minute cost structure on carriers which, if retail rates are unregulated, will likely be passed through to customers in the form of per-minute retail rates. Such usage sensitive rates thus would likely reduce usage of the network below efficient levels."

191 Vogelsang (2003, Ch. 3.2.4). See also Harbord/Pagnozzi (2008) who address termination rates in mobile networks.

192 See also Vogelsang stating that overall, on-net / off-net price discrimination "seems to have few desirable and many detrimental effects, largely because it increases the market power of the dominant incumbent".

193 See e.g. Marcus (2006a), p. 8.

194 Annual Report 2006, p. 73, www.bundesnetzagentur.de the flat rate. Therefore Bill & Keep is compatible with flat rate pricing at the retail level. This is also illustrated by looking at the retail price structure of Internet access.

Concluding the use of a particular wholesale mechanism does not preclude application of different retail pricing regimes[195]. Both, CPNP and Bill & Keep provide flexibility at the retail level to offer retail schemes based for example on minutes, bits, or as buckets of minutes or bits plans [196] as well as flat rates.

Customer acceptance of tariff schemes

Some respondents argued that European customers would not accept RPP. They seem to assume implicitly that Bill & Keep leads to RPP at the retail level. Given the trend towards flat rate regimes it should be mentioned that Bill & Keep would not necessarily lead to RPP at the retail level. It might rather be applied in combination with flat rates implying no change for end-users as has been the case for Internet access.

Furthermore, as the example of the mobile sector in the U.S. and several other countries shows, there are examples of RPP systems which are accepted by the end users.[197],[198] Mobile usage is often significantly higher in countries where RPP is applied than in those where CPP is used.[199]

C.6.3 Utility derived from a call

The CPNP charging mechanism assumes that the party that originated a call presumably wanted the call to complete, and that the originating party can therefore be considered the prime beneficiary of the call.[200] Correspondingly, the receiving party has been thought of as a passive party, involuntarily receiving a call from the originator. This rationale ignores the utility the called party derives.

---

195 Vogelsang (2006), Ch. 3.3.1 and 7.3.3. See also Ch. 3.3.1.6, Table 3-2 outlining the compatibility between wholesale and retail charging mechanisms. According to this, Bill & Keep is not only compatible with RPP but – to a lesser extent- also with CPP. CPNP and RPP are viewed as less compatible as this may involve a “double payment” to the receiving network operator on the retail and the wholesale level. Littlechild (2006) points out that the combination of Bill & Keep and CPP does also exist in practice.

196 Vogelsang (2006), Ch. 3.3.1.6, Table 3-2.

197 It is not intended here to provide an answer whether this would also be the case in Europe.

198 Littlechild (2006, p. 242) reasons that the alleged disadvantages of RPP do not withstand investigation. “Some mobile operators in RPP countries are now offering customers the option of calling plans with free in-coming calls.” Thus, Bill & Keep would not only avoid the bottleneck monopoly, avoid distortions of CPP but also allow the choice between CPP and RPP.

199 See also Marcus (2006a; Ch. 2.3)

200 According to a related argument the originating party is considered the sole cost-causer of the call. Knieps/Vogelsang (2007) in their summary of the discussion on cost causation consider that while the caller takes the initiative, both, calling and called party are responsible for continuing the call. In this “new” per-spective both parties can be considered cost causers and should therefore share the interconnection costs (similar DeGraba, 2000; recital 53).

It seems more realistic to assume that in most instances, both, the originating and the receiving parties derive some benefit from the call.[201] This "utility sharing" cannot be reflected under a CPNP wholesale regime, where the calling party bears *all* of the cost for the call. It could be better reflected under Bill & Keep because the costs of the call are shared between the calling party (finally bearing the costs for originating the call) and the called party (finally bearing the costs of termination).

In any event, this "utility sharing" argument may not hold for calls to value added services, because here it is the caller who derives all the benefit from the call.[202] Therefore it has to be discussed how to treat calls to service numbers.

Given these differences between CPNP and Bill & Keep, it is possible that the latter mechanism could be more capable of internalizing positive usage externalities. The positive usage externality relates to the utility a user derives from receiving a call. If this user pays for this utility, the externality – "produced" by the originating party – is internalized.[203] This may lead to a more efficient volume of calls.[204]

C.6.4 Network usage

Wholesale charging mechanism may also differ with regard to the incentives they produce for efficient network usage. If termination rates exceed marginal costs this will lead to a suboptimal level of network usage.[205] As outlined in C.6.2 above termination rates feed through to retail prices (cet.par.) implying an inverse relation between the level of termination rates and network usage.

CPNP in the form of CBC may provide incentives for efficient network usage. Operators buying capacities upfront will try to exploit this capacity in an optimal way.[206] By contrast, under a EBC system efficient network usage is less likely.[207]

---

201 Exactly determining the degree of utility derived by both parties may be not possible in practice (see Gabel 2002). For the purpose of this document it is sufficient to assume that both parties derive some benefit, with the originating party often likely to derive a greater utility.

202 Unwanted calls are another example where the receiving part (mostly) derives no utility from a call. Here, the called party has the option of simply hanging up.

203 Despite the practical problems of determining the utility distribution among calling and called party it is worth mentioning that the recent literature took on the issue of call externalities. Harbord/Pagnozzi (2008, p. 22) claim that the inclusion of call externalities in the analysis has "significant effects on competition, the equilibrium structure of retail prices, and optimal regulatory policy" leading to profound changes of the analysis of welfare-optimal prices and termination rates. And Berger (2005, p. 9) shows that "with two-part tariffs and discriminatory prices, cost-based access pricing can never be optimal from the social viewpoint, if the call externality is taken into account." A different perspective is taken by Gabel (2002) who assumes that the externality caused by the called party to the calling party is larger than in the opposite direction. He therefore considers that CPNP is more likely to internalize positive network externalities than Bill & Keep.

204 See Vogelsang (2006), e.g. Ch. 3.1.1.1. and also Valletti and Cambini (2005) assuming a complementarity between outgoing and incoming calls.

205 Correspondingly, too low termination rates would induce an „excessive" level of network usage.

206 Ibid, Ch. 3.2.2.2.

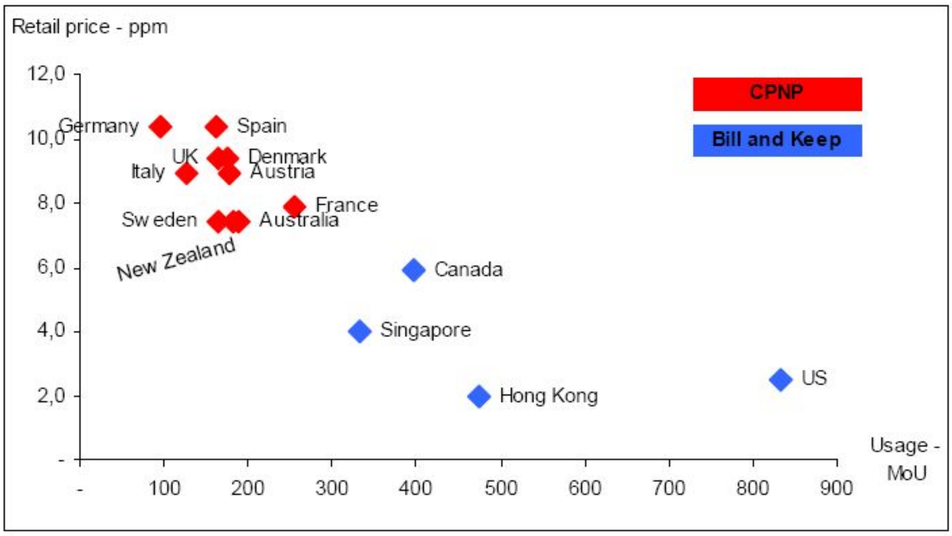
207 This is due to the fact that under EBC interconnection rates are usage based whereas network costs are rather determined by the peak capacity required. See Vogelsang (2006), Ch. 3.2.1.5.

The risk of a non-optimal level of network usage is circumvented under Bill & Keep, as this system does not require efficient termination costs to be determined. Moreover, the flexibility under Bill & Keep to apply different tariff schemes at the retail level may also be conducive to an efficient network usage because operators can offer tariff schemes best suited to customer needs.[208]

It seems plausible that charging mechanism leading to efficient network usage will also provide more efficient investment incentives (see below).

This plausibility is corroborated by empirical findings on the relationship between the level of retail prices for mobile usage and the minutes of usage (MoU):[209]

Figure 3: Usage and average retail prices, Q1 2007: bill-and-keep vs CPNP



Source: Merrill Lynch Global Wireless Matrix 1Q07, 15 June 2007

This figure[210] illustrates that in countries with low mobile retail prices usage is significantly higher than in countries with high retail prices. Retail prices are highest in countries applying CPNP as charging mechanism, whereas Bill & Keep countries exhibit much lower retail prices.[211]

---

208 See Vogelsang (2006), Ch. 3.2.3.4., also Charles River Associates (2002; recitals 42, 43, 50) conclude that Bill & Keep is superior not only with static efficiency (allocative, productive) but also with regard to dynamic efficiency.

209 See also Ofcom (2006, paragraph 6.6) comparing Bill & Keep countries (USA, Canada etc.) with CPNP countries in and outside Europe and has drawn conclusion in line with these finding stating that Bill & Keep "tends to yield significantly higher minutes of use per subscriber" and that "average revenue per minute is lower".

210 See Harbord/Pagnozzi (2008)

211 These empirical relations support theoretical findings. According to these, termination charges set a floor for retail prices as they are perceived as costs (see Laffont/Tirole, 2001). Similarly, also DeGraba's observation that existing termination charges imply "artificial" cost structures for carriers ultimately leading to inefficient per-minute retail prices seems supported by the empirical findings (see. DeGraba, 2000, recital 95).

As pointed out by Merrill Lynch[212] the minutes of use are potentially overstated in Bill & Keep countries by approximately 20 %.[213].

ERG considers that these effects[214] while possibly weakening the quantitative strength of the argument do not change the conclusions implied by this straightforward empirical finding that higher termination rates translate into high retail prices for originating calls and finally into lower usage.

Given these results, Bill & Keep seems to be associated with incentives for efficient network usage.

In the nascent stage of a market where the adoption hump and an increase of penetration rates are the main goals Bill & Keep may be considered less appropriate than CPNP to induce a rapid increase in the penetration rate of services as it might lead to RPP.[215] CPNP accompanied by CPP at the retail level leads to rather low fixed fees and higher usage fees making it cheaper to subscribe to mobile services.

C.6.5 Termination monopoly

CPNP and Bill & Keep differ with regard to the possibility to exploit the physical bottleneck for termination as has been explained in sections B.2.3 and C.3.2. :

- Under CPNP this bottleneck can be exploited because it entitles the terminating operator to receive a payment arising out of its control over the physical termination bottleneck.[216]

---

212 Merrill Lynch (2004, slide 136): http://www.comcom.govt.nz/IndustryRegulation/Telecommunications/Investigations/MobileTerminationRates/ContentFiles/Documents/Global%20Wireless%20Matrix%202Q04%20Sept%2004.pdf. Minutes of Use per month per average user is calculated by dividing total minutes of use on the operator's network by the average subscriber base during the quarter

213 The figure overstates the differences between CPNP and Bill & Keep countries. Littlechild (2006, p. 256), comparing CPP with RPP, points out that the "Minutes of Use" and depending on pricing structure and level revenues respectively may be overstated in RPP countries and understated in CPP countries. This is due to the fact that on-net mobile-to-mobile calls are counted twice in "non-CPP" countries because the same minute of an on-net call is billed to both the caller and the receiver. On the other hand, such on-net mobile-to-mobile calls are only counted once in CPP countries. Here, terminating mobile-to-mobile minutes are not counted because they are not associated with any revenue.

214 Furthermore international comparisons generally suffer from the need to standardise data, that are based on different ways of defining metrics. Another issue refers to the difference between number of users versus number of subscriptions. In the EU, penetrations rates of approx 114% are based on the number of subscriptions rather than number of users, implying that minutes per user are higher than minutes per subscriptions. The high number of subscriptions may also reflect economic inefficiencies due to on-net/off-net discrimination and handset subsidies. Correspondingly, penetration rates based on a number of users rather than subscription would be lower than 100 % (83 % according to Eurobarometer of December 2007 vs. approximately 82 % for the US according to CDC survey data). Overall, these effects do not change the conclusions that higher termination rates translate into high retail prices for originating calls and finally into lower usage. See also WIK-Consult (2008, footnote 115). Harbord/Pagnozzi (2008, p. 32) summarize their findings stating that "(i)t is difficult to avoid the conclusion, that, …., Bill & Keep leads to more intense price competition and hence lower prices for mobile subscribers".

215 Besides this, Vogelsang (2006, Ch. 2.1.6) points out, that having flexible retail prices is a much more efficient means to increase penetration than extensive cross-subsidies (e.g. by using high termination rates to cross subsidise mobile subscription prices).

216 See also Marcus (2006a), page 9f, DeGraba (2000), page 7. With reference to the termination problem in mobile radio, see Valletti/Houpis (2005). When explaining the reasons for the termination problem, reference

- With Bill & Keep, this is not possible as the access operator is *not* entitled to such a payment at the wholesale level. This major advantage of Bill & Keep affects some of the other issues addressed below.

Application of the CPNP regimes therefore ultimately perpetuates the need for regulation of the termination rates.

Bill & Keep for the terminating segment of the broadband access provider requires no regulatory intervention as long as two conditions are fulfilled:

- The market for transit on IP backbones is sufficiently competitive to exert competitive pressures on IP backbone providers. With an oligopoly of tier 1 providers allowing choice of transit provider this condition has so far been considered to be fulfilled.

- The broadband access market is sufficiently competitive so that access providers are under competitive pressures preventing them from establishing abusive mark-ups on retail prices.

The termination bottleneck is currently an integral difficulty associated with wholesale regulation. A shift of charging regime could significantly lower the need for detailed regulation.

C.6.6 Level of regulation required, transaction cost

The necessity to regulate termination rates results from the physical bottleneck for termination. Bill & Keep avoids this need as it prevents effective abuse of this. Therefore, Bill & Keep could pave the way for a reduction of detailed regulation[217] and to a significant reduction of transaction costs.[218]

Avoiding the bottleneck problem implies that it would no longer be necessary to determine the economically "correct" termination rates[219]. Given the complexities of this task, this will always involve some uncertainties for market players with regard to the final decision to be made by the NRA.[220, 221] Their decision always bears a certain risk of regulatory distortion. Such distortion may result from imperfect information on the part of the NRA. This will occur

---

is sometimes made to CPP (i.e. end customer level) and sometimes CPNP (.i.e. wholesale service level) in the literature. This difference can be ignored since the systems at both levels - to a certain extent - are interconnected. Vogelsang (2006, page 153) links both levels and states that "Calling Party Pays" causes a termination problem as long as there are not several independent accesses to the individual call receivers.

217 Considering this, it seems remarkable that Bill & Keep is often opposed by those operator, who usually call for more deregulation.

218 Bill & Keep may require a determination of number and location of points of interconnection. Nevertheless, these determinations are also to be made under a CPNP regime. Furthermore, it is obvious that Bill & Keep does not necessarily avoid any traffic monitoring activities because such monitoring may be necessary e.g. for retail pricing.

219 It will have to be checked whether under these conditions a necessity remains to regulate services such as interconnection links.

220 See WIK-Consult (2008), Ch. 3.2.3.; DeGraba (2000), p. 26/27.

221 The complexities of determing "correct" termination rates were also addressed in the Consultation. According to this argument Bill & Keep could create a level-playing field for all operators.

when the regulator does not set “the right tariffs” (for FTR and MTR), and so does not fully address the competition problems. This is a real risk because “the right tariff” is a difficult concept in the case of two-way access (interconnection). To illustrate this point it is referred to WIK(2008). Part 3.2 of this report explains the problems with the cost-based approach. The report quotes Littlechild, which states regulation of terminating is a “Sisyphean task”, “endless and unrewarding” and “the idea of an optimal price is a chimera.”

Under Bill & Keep, lengthy and cumbersome regulatory and legal disputes (both, between market players and NRAs but also among market players) on the appropriate level of termination rates may be avoided. Such disputes increase uncertainty for the market players involved.[222]

A wholesale charging mechanism such as CPNP where termination rates need to be determined also causes significant regulatory costs. Littlechild estimates the costs for price controls for mobile termination rates at “nearly $50 m”.[223]

C.6.7 Market forces

Under CPNP, there is no market pressure to keep the termination costs low as these costs are not borne by the provider’s own end-users. To the contrary, there might even exist an incentive to raise rival’s costs and keep wholesale prices high as a collusive device.[224]

On the other hand, Bill & Keep constitutes an approach which is more closely adjusted to market mechanisms, if end-users can choose the network carrier among various operators. Since the provider has to bill termination cost to its own end-users, it has no incentive to pay and bill high cost for connectivity, because otherwise it may risk losing them.

In order for the advantages of Bill & Keep to become effective, competition in the broadband access/Internet access market is a precondition. Such competition is a crucial factor, because it affects an access provider’s ability to abuse its market powers.[225]

C.6.8 Investment incentives

Routing in IP-networks follows the “hot-potato” routing principle[226], i.e., the practice of passing traffic off to another network as quickly as possible (towards the path with the lowest delay). So, normally the originating network/router will not hold onto the packets until it is as

---

222 All these arguments (uncertainties for market players, risk of regulatory distortion, lengthy and costly disputes) show that the ERG explicitly considers transaction costs of market.

223 Littlechild (2006), slide 3.

224 See Armstrong (1998), Laffont/Tirole (1998a, 1998b).

225 It is contested by one response that Bill & Keep relies more on market forces which was only the case if Bill & Keep was imposed by the market and not imposed by regulatory means. This argument seems misleading because different charging mechanisms (CPNP / Bill & Keep) need to be compared for a situation where regulation is needed.

226 „Hot-potato routing“ seems to be the normal behaviour of most IP peering agreements.

near to the destination as possible, as is done in circuit switched local interconnection model. This also reflects that in IP-networks in general the location of the IP-address is not known to the calling party. This approach would tend to drive the interconnection towards the higher levels of the core networks. Hot potato routing applies on those parts of the network, that are excluded from the application of Bill & Keep, but where transit and peering agreements apply. Transit networks have been excluded from the applicability of Bill & Keep. The network operator requesting termination can either use its own network to convey its traffic to these points of interconnection, or transit services can be bought from other operators.[227] Operators decide on a commercial basis on the appropriate wholesale charging mechanism. This can be transit or peering.[228] So far an under-investment problem has not occurred in these networks.

Some respondents argue that the "hot-potato" routing principle leads to insufficient or even lack of incentives for investment in infrastructure and QoS because operators are not compensated for their investments. According to this view this situation would ultimately lead to a process of adverse selection. Given that Bill & Keep applies for the terminating segment up to the first router or switch and associated service control functions after the access/concentration network, however, this view ignores that hot potato routing applies on those parts of the network only, that are excluded from the application of Bill & Keep, but where transit and peering agreements apply.

In order to qualify for participation in the Bill & Keep regime a minimum number and location of interconnection points for a specific network operator has to be determined.[229] Applying Bill & Keep for the terminating segment up to the first router or switch and associated service control functions after the access/concentration network can be considered a minimum scenario in terms of the scope of network included within the Bill & Keep domain. However, in order to qualify for participation in the Bill & Keep regime a specific network operator is required to access a maximum number of interconnection points in this case.

It may also be conceivable that the area of application of Bill & Keep might be extended by applying it already from PoIs higher up in the network hierarchy requiring fewer interconnection points per network operator.[230]

The application of Bill & Keep may essentially require a determination of the topology of points of interconnection. The same applied in the PSTN when determining the number of PoI to be eligible for local interconnection. Furthermore as has been outlined throughout the document the move towards NGNs will require change in the number of interconnection

---

227 Assuming that competitors have to increase their network in order to qualify for participating in Bill & Keep , this investment may be not economically efficient (duplicating infrastructures), also leading to a higher concentration in the market. See Vogelsang (2006), Ch. 3.2.3.4.

228 For the commercial applicability of transit rules a sufficient degree of competition between transit providers is required.

229 DeGraba (2000); Vogelsang (2006), see e.g. Ch. 7.3.2.3 or Ch. 8 or Berg et al. (2006), Ch. 2.4.3 or 4.1.

230 There may be a trade-off between requiring a high number of PoI and inefficient network duplication. It holds (cet. par.): the higher the number of points of interconnection (= closer to the customer) required for participating in Bill & Keep the the more infrastructure is required to reach these interconnection points. Vogelsang (2006, p. 63 footnote 47) points out that hot potato routing may even be efficient, if the location of the called party in the other network is not known.

points anyways.[231] The question of how to efficiently determine this minimum number of PoI, i.e. the border of a potential Bill & Keep domain has to be further investigated.

Furthermore, considering that under Bill & Keep the terminating operator cannot charge the originating operator but generally recovers its net-costs from its own end-users, the terminating network has an incentive to have an efficient network structure. If it tried to levy any inefficient network costs on the end-users, it might risk losing these customers.[232]

C.6.9 Further issues

Existing business models

It seems obvious that any change of the charging mechanism may have an impact on attractiveness and competitiveness of business models. This holds true also for providers of call-by-call and pre-selection.

Accordingly, some respondents called for a further analysis of the effects of Bill & Keep on different business model such as call-by-call/pre-selection, value-added and free-phone services or pre-paid cards. Service providers of these services claim that their services cannot be billed under a Bill & Keep regime. The ERG acknowledges that changing the wholesale regime impacts on business models. This is why an appropriate management of the migration towards all-IP infrastructures is so important to avoid disruptive changes. Therefore, the ERG identified the implications for different business models as one of the issues meriting further study (see.C.6.10 below).

Under Bill and Keep, specifically due to the CPS operator not paying for call termination, the costs for the CPS operator will be significantly reduced.[233] As the CPS operator does not have the customer on its network it will not incur a cost associated with calls terminating on its customers. In this case the terminating SMP operator would be disadvantaged under a Bill and Keep model.

It is claimed that the SMP operator may include the cost of termination as a mark-up on the origination charge for CPS traffic or bill his end-users using RPP. However other options may be discussed depending on the relevance of the CPS model in the different Member States. It is possible that this trend towards flat rates could affect the viability of call-by-call and pre-selection more significantly than an eventual transition towards Bill & Keep.[234]

---

231 Whereas some respondents emphasize the transaction costs of determining this number and the artificial network structure resulting thereof (see 6.6 above), one respondent considers this a measure to counter the free rider problem in case of unbalanced traffic between operators. ERG considers that the migration to NGN will require a determination of PoI anyways and will further study the issue.

232 Provided there is sufficient competition on the access level allowing end-uses to switch to another operator.

233 See also Kennet/Ralph (2007) stating that without an alternative mechanism long distance carriers would simply have free access to origination and termination services.

234 To the contrary, considering that the terminating network operators would have to cover their costs from their own end-users, the trend towards retail flat rate scheme may even gain further impetus.

SPIT

A CPP system (with CPNP at the wholesale level) is generally viewed as less susceptible to the problem of SPIT (SPAM over Internet Telephony) as the calling party has to pay for the whole call. On the other hand RPP/Bill & Keep may (cet.par.) lead to a higher number of such unwanted calls because the costs of calling consumers for marketing and sales, would be reduced. While receiving parties can hang-up on such calls, doing so frequently may result in disutility and in fewer calls being answered generally. This was stressed by those respondents who are sceptical of Bill & Keep.

However, the costs of voice traffic for engaging in such activities seems negligible compared with the costs of labour. Furthermore receiving customers can hang up and will do so more when they are charged for the call. In some countries consumer protection measures are in place and so called opt-out or opt-in systems exist.[235]

C.6.10 Summary of charging mechanisms and work plan

As networks migrate towards NGN infrastructure it is unclear *a priori* whether these future networks will be governed by the mechanisms currently used in IP-networks or whether the mechanisms currently applied in the PSTN will be partly carried over to NGNs.

Summarizing the preceding comparison of charging mechanisms with regard to a number of criteria, it can stated that Bill & Keep has a number of attractive properties. Assessing all the pros and cons of Bill & Keep, the ERG concludes, that Bill & Keep is a promising interconnection regime. It is supported mainly by theoretical reasoning and a large body of literature and economic modelling. Furthermore, empirical evidence seems to suggest higher usage levels and lower usage prices are achieved with less regulatory intervention in Bill & Keep countries than currently applied in the EU.

Some NRAs therefore aim at a shift towards Bill & Keep because it reduces the regulatory burden and relies more on market forces, as is the case already now for today's unregulated IP connectivity markets. They focus on further studying how to best achieve this goal including finding answers that arise in a transition phase to a new system.

Others, while recognising the merits of Bill & Keep in principle, rather emphasize the risks implied by a change from the well-established regulatory regime of mainstream PSTN and mobile services fearing disruptive change to the industry and therefore see a need for further study.

---

235 If the aim is to protect customers from paying for receiving unwanted calls one might conceive of a retail price of zero for the first minute. However, this would not avoid the annoyance for customers from such calls. It is acknowledged that SPIT may become more acute than today. Nevertheless, it may be more appropriate to tackle this issue with consumer protection measures as is the case in the Netherlands than with the wholesale charging mechanism.

ERG identifies the following issues meriting further study which were generally supported by the Consultation responses:

Implications for different business models

Depending on *how* Bill & Keep is introduced the implications of a widespread introduction of Bill & Keep including the change of the cost recovery mechanism may imply that a transition from the current regime is a drastic and disruptive change for PSTN voice operators who have been subject to regulation under the framework. A rapid transition may not allow operators enough time to adjust there business models and retail price structure.

Operators may be affected differently by the introduction of Bill & Keep depending on, among other things, whether they are likely to experience a net cost (more incoming than outgoing traffic) or a net gain (more outgoing than incoming traffic). It has to be evaluated whether this may systematically shift the competitive balance between different operators, particularly incumbents and alternative network operators.

Furthermore implications for specific business models such as CPS have to be studied further, e.g. in case incumbents try to recover potential net-costs in part from CPS operators through higher origination charges.

The ERG acknowledges that changing the wholesale regime impacts on business models and may affect the competitive situation of the market. This is why an appropriate management of the migration towards all-IP infrastructures is important to avoid disruptive changes. Therefore, the ERG identified the implications for different business models, competition and subscribers as one of the issues meriting further study.

Practical implementation issues

- **Migration**

Currently, different regimes for different types of networks (PSTN or IP) prevail – irrespective of service. As the separate network infrastructures are expected to converge to an all IP network such differences may not be sustainable in the long run.

Considering the migration to all IP-networks it seems plausible to continue applying the charging mechanism of the networks that are not phased out.

A meaningful discussion of migration problems, implies having come to a conclusion with regard to the charging mechanism applicable in the long term:

- In case it is intended to carry over CPNP to NGN voice services this would imply different regimes for different services as a change of charging mechanism cannot necessarily be expected for the unregulated part of IP-networks applying Bill & Keep, Peering and Transit.

  This approach requires that it is possible to clearly distinguish between different services and that usage of services can be measured. If arbitrage problem are to be avoided, it would be necessary to mark different services or to transport them separately. Unless these preconditions are met there is a high risk of adverse selection, moral hazard and arbitrage problems.

  Instead of differentiating regimes according to services one might also envisage differentiation of different QoS classes (best effort vs. QoS level specified).236 Applying such an approach could be done by assigning different services to different QoS classes.

- In case Bill & Keep was envisaged as long term goal it may be reasonable to further investigate regulatory options to soften the transition to the new regime. Strict application of cost orientation in a CPNP environment in the short term for mobile and/or PSTN networks can be seen as an important step in the migration towards Bill & Keep.

  The length of the migration period can be shorter

  - the lower the absolute level of interconnection rates,
  - the smaller the relative difference between interconnection rates of different networks
  - the higher the proportion of flat rates at the retail level is.

- **Defining the border for application of the Bill & Keep regime**

  Applying Bill & Keep for the terminating segment up to the first router or switch and associated service control functions after the access/concentration network can be considered a minimum scenario in terms of the scope of network included within the Bill & Keep domain. However in order to qualify for participation in the Bill & Keep regime a specific network operator is required to access a maximum number of interconnection points in this case.

---

236 See also ECC Report 75 (2005).

It may also be conceivable that the area of application of Bill & Keep might be extended by applying it already from PoIs higher up in the network hierarchy requiring fewer interconnection points per network operator.

The application of Bill & Keep may essentially require a determination of the topology of points of interconnection as has been the case for the PSTN when determining the number of PoI. The question of how to efficiently determine this minimum number of PoI, i.e. the border of a potential Bill & Keep domain has to be further investigated.

- **How to treat traffic from outside the Bill & Keep area and prevent extensive arbitrage (tromboning, call-back etc.)**

  - Between different countries

  - Between different networks (e.g. fixed/mobile)

  An important issue is how to handle traffic coming from outside the Bill & Keep domain (hereafter: incoming traffic). This traffic could result in problems. For example if the operators inside the Bill & Keep domain want to set a termination rate for incoming traffic, this could be forestalled by competition for receiving this incoming traffic. Every operator inside the Bill & Keep domain would have an incentive to receive incoming traffic and collect a fee for this and then route this traffic towards the final destination and dropping it there for free.

  This problem could be prevented if, subject to the number portability arrangements in place, receiving networks could effectively bill the incoming traffic based on where traffic originated, for example by using the network number of the source network.

  The dimension of these problems increases

  - the larger the traffic volume from outside relative to the traffic exchanged between networks inside the Bill & Keep domain.

  - the higher the termination rates outside the Bill & Keep area

Other long term forms of regulation

Are there forms of voluntarily achieving Bill & Keep through a series of other measures and requirements that could be based on symmetric regulation according to Art. 5 AD (e.g. reciprocity) used in combination with termination rates strictly regulated at cost based levels?

Annex 1: Glossary

Bitstream

Bitstream access is a wholesale product enabling operators with an own backbone the provision of services. A provider of bitstream access provides the requesting party with the broadband access line and transport over the concentrator network up to the point of handover. This implies that bitstream constitutes a concept of one-way access whereas interconnection rather is a concept of providing mutual access. Bitstream access enables an operator to use other operators' infrastructure instead of employing own infrastructure. In this respect, it reflects an operator's make or buy decision.

Best effort

Best effort depicts the fact that a data transmission via an IP connection without any guaranteed level of performance nor priority nor that the data will be eventually delivered. This is related to the nature of IP transmission in contrast to circuit-switched networks, transmission performance IP networks varies with the traffic load. Since the traffic load is generated by the final user, there is a level of uncertainty on what the resulting transmission performance is at a specific moment. Therefore, the network operator cannot give any guarantee in principle. Best effort transport does not necessarily mean that the service that uses this transport class has a low QoS, but less reliability of transmission performance.

Billing / charging mechanism

Bill & Keep is a wholesale billing regime under which each network bears the costs of terminating traffic coming from other carriers. Therefore under Bill & Keep the terminating access network operator does not receive payments at the wholesale level for the termination provided. Instead, it recovers its net costs incurred for termination –- and any payments for upstream connectivity – in other ways, e.g. by billing them to its end customers.

Calling Party's Network Pays is a wholesale billing regime where the network of the party who placed the call (the originating network)[237] makes a payment to the network of the party that received the call (terminating network). Thus, at the wholesale level the whole call is paid by the caller's network.

Interconnection

Interconnection is the physical and logical linking of public communications networks used by the same or a different undertaking in order to allow the users of one undertaking to communicate with users of the same or another undertaking.[238]

---

237 A provider of call-by-call or preselection services also has to buy origination from the access operator of the caller.

238 Definition based on Art. 2 b Access Directive

This implies that realising any-to-any communication presupposes interconnection in a multi-network environment. Furthermore, interconnection constitutes “a specific type of access implemented between public network operators”. Transit and peering, for example, are specific forms of interconnection arrangements.

Since the use of packet switched (i.e. IP) technology to the separation of transport and service layer, interconnection has to be distinguished into transport interconnection and service interconnection (these terms are defined separately in this Glossary).

Interoperability

The ability of systems and services to exchange information and to mutually use the information that has been exchanged.

In the context of interconnection interoperability is needed in order to ensure that services can interact. To provide Interoperability transport interconnection and service interconnection agreements are needed.

Network performance

"The performance of a portion of a telecommunications network that is measured between a pair of network-user or network-network interfaces using objectively defined and observed performance parameters."[239]

“The ability of a network or network portion to provide the functions related to communications between users.

NOTE 1: Network performance applies to the Network Provider’s planning, development, operations and maintenance and is the detailed technical part of QOS, excluding service support performance and human factors.

NOTE 2: Network performance is the main influence on serveability performance.

NOTE 3: Network performance measures are meaningful to network providers and are quantifiable at the part of the network to which they apply. Quality of service measures are only quantifiable at a service access point.

NOTE 4: It is up to the Network Provider to combine the Network Performance parameters in such a way that the economic requirements of the Network Provider, as well as the satisfaction of the User, are both fulfilled.”

NGN

“A packet-based network able to provide telecommunication services and able to make use of multiple broadband, QoS-enabled transport technologies and in which service-related

---

239 ITU-T Rec. E.417

functions are independent from underlying transport-related technologies. It enables unfettered access for users to networks and to competing service providers and/or services of their choice. It supports generalized mobility which will allow consistent and ubiquitous provision of services to users.” [240]

An IP network that uses some deliberately chosen elements that are specified in NGN standards for improving its transmission performance is also referred to as NGN. Today the term NGN covers a broad performance spectrum from simple implementation of TCP/IP with low level best effort performance to intensive implementation of traffic management methodologies providing high level and stable transmission performance.

Non-NGN IP network

An IP network that is designed and managed according to the requirements set out in the NGN standards of ITU-T, ETSI or 3GPP and that has not implemented the respective interfaces in total. A non-NGN IP network may use some deliberately chosen elements that are specified in NGN standards for improving its transmission performance. The term non-NGN IP network covers a broad performance spectrum from simple implementation of TCP/IP with low level best effort performance to intensive implementation of traffic management methodologies providing high level and stable transmission performance.

Peering

Peering “is an 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. Peering is usually a bilateral business and technical arrangement, where two parties agree to accept traffic from one another, and from another’s customers (and thus from their customers’ customers).”[241]

Performance versus QoS:

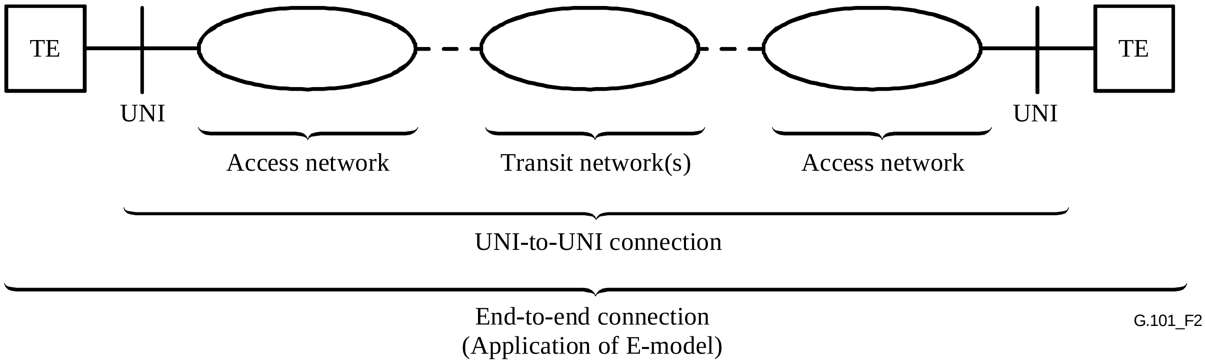
While QoS and performance parameters are different in essence, it is clear that there exist intrinsic relationships between QoS and performance parameters, the former having a direct or indirect, and sometimes even inverse, influence on the latter. Furthermore, some performance measures can have a direct QoS meaning, while some others have to be combined in order to have a QoS signification. When doing so the different focus of performance parameters and the interfaces they are measured at have to be taken into account in order to map them correctly to service related interfaces as seen by a user.

For voice services, for example, QoS (in terms of speech quality) means the quality “from mouth to ear”. For planning purposes, the E-Model (ITU-T Rec. G.107) can be used to estimate the resulting speech quality of an end-to-end connection.

---

240 ITU-T Rec. ITU Y.2004
241 Report of the NRIC V Interoperability Focus Group: “Service Provider Interconnection for Internet Protocol Best Effort Service“, page 7, available at http://www.nric.org/fg/fg4/ISP.

Figure 4: Different aspects of QoS



Network performance is measured from UNI to UNI (user network interface).

Performance parameters are used to measure objectively the performance of specific network and terminal elements that have an influence on the resulting end-to-end quality of a service. Performance is measured and expressed in performance parameters.

In IP networks and IP connections the performance parameters IP packet transfer delay, IP packet delay variation, IP packet loss ratio, IP packet error ratio are important[242]. Based on these parameter the performance of an IP data flow can be assessed. In operating IP networks the IP Packet Error Ratio is very low and can be can neglected when investigating the performance of real IP packet transmission. This leaves three parameters that are commonly referred to as delay, jitter and packet loss.

QoS of service is a broader concept as it covers the whole service.[243] Thus all aspects that have an influence on the quality of a service have to be taken into account.

Quality of Service (QoS)

"The collective effect of service performance which determine the degree of satisfaction of a user of the service.

NOTE 1: The quality of service is characterized by the combined aspects of service support performance, service operability performance, serveability performance, service security performance and other factors specific to each service.

Note 2: The term "quality of service" is not used to express a degree of excellence in a comparative sense nor is it used in a quantitative sense for technical evaluations. In these cases a qualifying adjective (modifier) should be used."[244]

Thus, the end-to-end quality of a service depends on many factors, such as network performance on the transport layer between the network termination points, in the interconnec-

242 The parameters are defined in ITU-T Rec. Y.1540.
243 The QoS of a voice service for example relates to the entire transmission path from mouth to ear.
244 ITU-T Rec. E.800

tion point, i.e. transmission conditions in the access and core network and, in particular, the influence of terminal equipment used as well as codecs. By selecting and configuring the terminal equipment, the user can substantially affect the resulting end-to-end quality.

Remark: When talking about QoS the whole service is covered from end-to-end. A specification of QoS refers to the the quality that is perceivable between theses end points. Depending on the service under consideration the end points may vary and be of different nature (e.g. for voice it can be mouth to ear, for data transmission it can be UNI to UNI).

However, in many publications and even sometimes in standardisation the term QoS is used laxly. There is a tendency to use the term QoS whenever telecommunication aspects are investigated that have an influence on services quality. Therefore care should be taken when consulting such documents whether the information is really addressing QoS or a related aspect like performance.

Service interconnection

Service interconnection in this paper is understood as including solely service-specific aspects.[245,246] It consists of logical linking of network domains, having access and control of its resources including the control of signalling (i.e. session based service-related signalling[247]). Depending on the kind of service, different aspects must be considered. For example, in the voice service, the call server interconnection is required for call setup and disconnection.

Interconnection between services from different operators requires a minimum set of technical (e.g. defined by a SLA) and commercial conditions to be fulfilled by both operators.

These conditions may include inter alia:

1. mutual policies for exchange of data (including transcoding the information mapping of quality of service information (if applicable), service control information and network protocols);
2. agreement of charges;
3. agreement on performance and reliability levels.

   There may be other aspects that need to be considered, like security.

Tier 1 network

A Tier 1 network is an IP network (typically but not necessarily an Internet Service Provider) which connects to the entire Internet solely via Settlement Free Interconnection, commonly

---

245 This differs from ETSI/TISPAN's definition of 'service oriented interconnection' also including transport related information. See below Section 3.2 in Annex 3.

246 This may have an important implication. Given this understanding, service interconnection may not fulfil the definition of Art 2 (a) Access Directive (see Sec. C.1.1) because it does not necessarily include the physical linking of NGN domains (see Sec. C.3.2).

247 Information that allows identification of the end-to-end service that has been requested.

known as peering. Another name for a Tier 1 network is "transit-free", because it does not receive a full transit table from any other network.

Although there is no formal definition of the "Internet Tier hierarchy", the generally accepted definition among networking professionals is:

- Tier 1 - A network that can reach every other network on the Internet without purchasing IP transit.
- Tier 2 - A network that peers with some networks, but still purchases IP transit to reach at least some portion of the Internet.
- Tier 3 - A network that solely purchases transit from other networks to reach the Internet.

Transit

Transit "is an agreement where an ISP agrees to carry traffic on behalf of another ISP or end-user. In most cases transit will include an obligation to carry traffic to third parties. Transit is usually a bilateral business and technical arrangement, where one provider (the transit provider) agrees to carry traffic to third parties on behalf of another provider or and end-user (the customer). In most cases, the transit provider carries traffic to and from its other customers, ant to and from every destination on the Internet, as part of the transit arrangement."[248]

Transport interconnection

Transport interconnection includes the physical and logical linking of networks based on simple IP connectivity irrespective of the levels of interoperability. It is characterised by the absence of the service-related signalling, implying that there is no end-to-end service awareness. Consequently, service specific QoS and security requirements are not necessarily assured. Only transport-specific performance objectives[249] for performance parameters that are affecting the transmission performance at the point of interconnection (e.g. availability) and the IP packet transmission performance via interconnected networks are negotiated.

This definition does not exclude that some services may provide a defined level of interoperability or the establishment of transport classes to guaranty specific quality parameter at this level.

248 Report of the NRIC V Interoperability Focus Group: "Service Provider Interconnection for Internet Protocol Best Effort Service", page 7, available at http://www.nric.org/fg/fg4/ISP.
249 For e.g. bit rate, delay and packet loss ratio.

Literature

3GPP: Specs in http://www.3gpp.org/specs/specs.htm

AKNN (2007): Konzept für die Zusammenschaltung von Next Generation Networks. Version 1.0.0 as of May 8, 2007

Armstrong, Mark (1998): Network Interconnection in Telecommunications, in: Economic Journal, Vol. 108. pp. 545 – 564

Armstrong, Mark (2002): The Theory of Access Pricing and Interconnection, in: Cave, Martin / Majumdar, Sumit / Vogelsang, Ingo (eds.): Handbook of Telecommunications Economics, Elsevier Publishers, pp. 295 – 384

Armstrong, Mark and Julian Wright (2007): Mobile Call Termination, MPRA Paper number 4858, available at at http://mpra.ub.uni-meunchen.de/4858/

Berg et al. (2006): Final Report of the Project Group Framework Conditions for the Interconnection of IP-Based Networks, www.bundesnetzagentur.de, December 15, 2006

Berger, Ulrich (2005), Bill-and-Keep vs. Costs-Based Access Pricing Revisited, Economic Letters, 86(1), 107-112

Bundesnetzagentur (2008): Eckpunkte der Zusammenschaltung IP-basierter Netze, February 11, 2008

Cambini, Carlo and Tommaso M. Valletti (2003): Network competion with price discrimination: Bill and Keep is not so bad after all, Economic Letters, 81(2), 205-213

Cambini, Carlo and Tommaso M. Valletti (2004): Access charges and quality choice in competing networks, in: Information Economics and Policy, Vol. 16, pp. 391 - 409

Cambini, Carlos and Tommaso M. Valletti (2005): Investment and Network Competition, Rand Journal of Economics, Vol 36, pp. 446 – 467

Cambini, Carlos and Tommaso M. Valletti (2005b): Information Exchange and Competition in Communications Networks, CEPR Discussion Paper No. 5031, May 2005

Carter, Michael and Julian Wright (1999): Interconnection in Network Industries, in: Review of Industrial Organisation, Vol. 14., pp. 1 – 25

Carter, Michael and Julian Wright (2003): Asymmetric Network Interconnection, in: Review of Industrial Organization, Vol. 22, pp. 27 – 46

Charles River Associates (2002): Comments on the Draft Access Determination: Interconnection Pricing, September 9, 2002

Commerce Commission (New Zealand) (2006): Final Determination on the application for determination for 'Interconnection with Telecom's fixed PSTN, 28 September 2006

Credit Suisse First Boston (2005): IP, The Holy Grail for Telcos, March 2005.

DataConnection "White paper" on Session Border Control in IMS: http://www.dataconnection.com/network/download/whitepapers/SBCinIMS.pdf.

DeGraba, Patrick (2000): Bill and Keep at the Central Office as the Efficient Interconnection Regime, OPP Working Paper Series, No. 33, FCC, December 2000.

Dewenter, Ralf and Julius Haucap (2005): The Effects of Regulating Mobile Termination Rates for Asymmetric Networks, in: European Journal of Law and Economics, Vol. 20, pp. 185 – 197

ECC Report 75 (2005): A Model for Interconnection in IP-based Networks, October 2005

Eickers, Gerd (2005): Technische Möglichkeiten der Zusammenschaltung bei VoIP, www.bundesnetzagentur.de, Bonn, September 2005

ERG (2007): Report on IP-Interconnection, Project Team on IP-Interconnection and NGN, (ERG (07) 09), March 2007

ERG (2007a): ERG Opinion on Regulatory Principles of NGA, (ERG (07) 16rev2), October 2007

Ericsson (2004): White Paper, IMS – IP Multimedia Subsystem

Ericsson (2004): "IMS – IP Multimedia Subsystem", White Paper, October 2004.

ETSI (2005): "TISPAN PUBLISHED NGN SPECIFICATIONS": http://portal.etsi.org/docbox/TISPAN/Open/NGN_Published/PUBLISHED_NGN_SPECIFICATIONS.doc. (e.g. ETSI TR 102 055: "TISPAN: ENUM scenarios for user and infrastructure ENUM", May 2005).

FCC (2000): The Digital Handshake: Connecting Internet Backbones, OPP Working Paper No. 32, www.fcc.gov/Bureaus/OPP/working_papers/oppwp32.pdf, September 2000

Gabel, David (2002), Interconnection Payments in Telecommunications A Competitive Market Approach,

Hackbarth, Klaus and Gabriele Kulenkampff (2006): Technische Aspekte der Zusammenschaltung in IP-basierten Netzen unter besonderer Berücksichtigung von VoIP, Study for the Federal Network Agency, Germany

Harbord, David and Marco Pagnozzi (2008), On-net / Off-net Price Discrimination and 'Bill-and-Keep' vs. 'Cost-Based' Regulation of Mobile Termination Rates, January 2006, www.market-analysis.co.uk/PDF/Topical/harbord_pagnozzi_b&k_22jan2008.pdf

Horrocks, John (2006): Regulation of NGN: New Approaches to Interconnection, ITU Workshop March 23-24 2006 (http://www.itu.int/osg/spu/ngn/presentations.html)

IETF (Internet Engineering Task Force) specifications (http://www.rfc-editor.org/index.html) and IETF working groups (e.g. "Session PEERing for Multimedia INTerconnect –speermint" - http://www.ietf.org/html.charters/speermint-charter.html).

ITU-T Recommendations: E.164, E.360.1, E417, E.800, G.101, G.711, H.323, Q.3401, Y.1540, Y2001, Y2011, etc. See also: http://www.itu.int/osg/spu/ngn/.

ITU-T (2002): "GLOBAL IMPLEMENTATION OF ENUM: A TUTORIAL PAPER", February 2002.

ITU-T (2005): "NGN FG Proceedings Part II" – NGN GSI (Global Standards Initiative), 2005.

Katz, Michael L. and Carl Shapiro (1985): Network Externalities, Competition, and Compatibility, The American Economic Review, Vol. 75, pp. 424-440

Kennet, D. Mark; Eric K. Ralph (2007), Efficient Interconnection Charges and Capacity-Based Pricing, International Economics and Economic Policy, Vol. 4, Number 2, August 2007

Knieps, Günter; Ingo Vogelsang (2007), Digital Economy and Regulatory Issues – Introduction, in: International Economics and Economic Policy, Vol. 4, Number 2, August 2007

Laffont, Jean-Jacques and Jean Tirole (1998a): Network Competition I: Overview and Nondiscriminatory Pricing, in: Rand Journal of Economics, Vol. 29, pp. 1 – 37

Laffont, Jean-Jacques and Jean Tirole (1998b): Network Competition II; Price Discrimination, in: Rand Journal of Economics, Vol. 29, pp. 38 – 56

Laffont, Jean-Jacques and Jean Tirole (2000): Competition in Telecommunications, MIT-Press

Littlechild, Stephen C. (2006): Mobile Termination charges: Calling Party Pays versus Receiving Party Pays, in: Telecommunications Policy, Vol. 30, pp. 242 – 277

Loutrel, Benoit (2006): Bill and Keep in the French Mobile Industry: A Case Study, WIK Workshop "Bill and Keep: A New Model for Intercarrier Compensation Arrangements?", 4/5 April 2006. http://www.wik.org/content/bill_keep/konf_billkeep_main.htm

Marcus, Scott (2006): Framework for Interconnection of IP-Based Networks – Accounting Systems and Interconnection Regime in the USA and the UK, Study for the Federal Network Agency, Germany

Marcus, Scott (2006a): Interconnection in an NGN Environment, ITU/02, March 2006

Merill Lynch (2004), Global Wireless Matrix 4Q04, Global Securities Research & Economics Group, Merrill Lynch, 29 September 2004 (Corrected), http://www.comcom.govt.nz/IndustryRegulation/Telecommunications/Investigations/MobileTerminationRates/ContentFiles/Documents/Global%20Wireless%20Matrix%202Q04%20Sept%2004.pdf

NGN^{uk} (2007): Richard Cadman, NGN Interconnection: Charging Principles and Economic Efficiency, 12 July 2007

OECD (2007): Convergence and Next Generation Networks: Background Report, Working Party on Communication Infrastructure and Service Policy, DSTI/ICCP/CISP(2007)2/REV1

Ofcom (2006): Mobile Call Termination: Proposal for Consultation

Valletti, Tommaso M. and George Houpis (2005): Mobile termination: what is the "right" charge?, http://www.itu.int/osg/spu/ni/multimobile/legacyregulation/index.html

Vogelsang, Ingo (2003): Price Regulation of Access to Telecommunications Networks, Journal of Economic Literature, Vol. 41, issue 3, pages 830-862

Vogelsang, Ingo (2006): Abrechnungssysteme und Zusammenschaltungsregime aus ökonomischer Sicht, Study for the Federal Network Agency, Germany

WIK-Consult (2008): The Future of IP Interconnection: Technical, Economic, and Public Policy Aspects, Final Report, Study for the European Commission, 29 January 2008

Wright, Julian (2002): Bill and Keep as the Efficient Interconnection Regime?, Review of Network Economics, Vol. 1, Issue 1, March 2002