



Regulatory implications of the introduction of next generation networks and other new developments in electronic communications

Final v.1.0

16th May 2003



DEVOTEAM
SITICOM

The opinions expressed in this study are those of the authors and do not necessarily reflect the views of the European Commission.
© ECSC – EC – EAEC, Brussels – Luxembourg 2003

Document Modification Log

Modifications

Release	Date	Modifications	Owner
Final v. 1.0	16 th May 03	Final Release	SITICOM & Cullen Int.

Authors

Name	Position
Ms. Aurelie Dame	Consultant, Devoteam Siticom
Mr. Jan H. Guettler	Director, Cullen International
Dr. Ken Leeson	Managing Director, Cullen International
Mr. Morten Schultz	Project Director, Devoteam Siticom
Mr. Torben B Jensen	Senior Consultant, Devoteam Siticom

For any request regarding this document, please contact:

Name	Telephone	E-mail
Mr. Morten Schultz	Project Director Devoteam Siticom	Mschultz@devoteam-frontrunner.com
Mr. Jan H. Guettler	Director Cullen International	Jan.Guettler@cullen-international.com

Content

1. Introduction	1
1.1. A Marketplace in Transition	1
1.2. Identifying control points	2
1.3. Self correcting market forces and integration	3
1.4. Jurisdiction and trade agreements	3
1.5. Reader's guide	4
2. The NGN Architecture	5
2.1. NGN principles	5
2.2. Architectural concepts	6
3. Technology and Systems	8
3.1. Migration to NGNs	8
3.2. Developments in access technologies	17
3.3. IP network interconnect	25
3.4. Service provisioning	30
3.5. Interworking of addressing systems	41
3.6. Interworking of signalling systems	53
3.7. Roaming and mobility	62

4. Market developments	74
4.1. Phasing out circuit switched equipment	74
4.2. Standardisation and interoperability	76
4.3. Powerful NGN roles	82
4.4. Interconnecting NGN infrastructures	87
4.5. The new services environment	90
4.6. Customer control through billing	95
4.7. Market views on regulatory approach	98
5. Regulatory implications	100
5.1. Introduction	100
5.2. Control points	101
5.3. Regulatory intervention on control points	105
5.4. Possible regulatory implications of the technical differences between CS and IP	106
5.5. Information society services	109
5.6. Interconnection	110
6. Concluding recommendations	115
6.1. Control Points	115
6.2. Interconnection	116
6.3. Other implications	117

Appendix A. List of acronyms

1. Introduction

This report presents the main findings and recommendations for the study entitled “Regulatory implications of the introduction of next generation networks and other new developments in electronic communications”.

The report provides a plain-language market and technology tutorial to help readers understand where market developments will lead and what is technically feasible in the NGN marketplace. It also raises a number of regulatory issues that warrant public debate and the attention of policy makers and regulators in the European Commission and EU Member States who are working to shape a pan-European information society in coming years.

1.1. A Marketplace in Transition

The term ‘next generation networks’ (NGN) refers here to the future competitive marketplace for communications and information services, a marketplace preceded in time by the generations-long monopoly era, ending circa 1990, and the current era of transition from monopoly to competition.¹

1.1.1. Monopoly era

During the monopoly era the regulator and operator were one and the same. Voice telephony was the only telecommunications service offered. The public switched telephone network exhibited natural monopoly characteristics, where unit costs of operations fall continuously as scale is increased.

1.1.2. Transition to competition

Over time, technology and innovation undermined the rationale for monopoly provision of telephone service by pushing down unit costs, whetting the appetites of consumers for new services, and enabling smaller competitors to enter some segments of the market.

When the transition to competition began, the main “control points” over marketplace activity requiring regulatory attention with regard to competition were vestiges of the monopoly area, primarily residual market power and network “bottlenecks” controlled by the incumbent, such as local access.

To open markets for competition and to promote investment, regulators applied ex ante regulations to establish a level playing field. Ex ante tools include regulations to ensure cost-based, non-discriminatory access to the incumbent’s bottleneck facilities by competitors, and the blocking of alliances seen as anti-competitive.

¹ This definition is used because the successes and failures of regulatory decision making are ultimately felt in the marketplace, where regulators must keep a sharp eye. This definition may be unfamiliar to some readers. For example, in standards organizations NGN refers generally to the physical and logical networks of the future. Even in this setting, however, there is no commonly agreed definition. This became clear during the company interviews conducted for this study.

1.1.3. Next Generation Networks

Decisions by regulators for the NGN marketplace will also have to address anti-competitive “control points,” inherited or devised, that inhibit the normal functioning of a competitive marketplace. The challenge will be to distinguish these *anti-competitive* control points from *competitive* control points. Competitive control points are those developed or discovered in search of a commercial edge that actually promote the normal functioning of a competitive marketplace. These should be left untouched, to better serve consumers and to promote innovation and investment.

In their decision making, regulators will therefore have to:

- identify and assess marketplace control points for their potential impact on competition;
- strike the right balance between using ex-ante regulations and applying ex-post remedies;
- assess the effectiveness of self-correcting forces in a competitive marketplace;
- achieve optimum levels of concentration and vertical integration;
- consider the limits to jurisdiction over services whose components reside in different countries; and
- understand the role of trade agreements in shaping the regulatory landscape.

1.2. Identifying control points

Control points inherited from the monopoly era reside primarily (though not exclusively) in the access and core networks, are relatively easy to identify, and have fairly predictable consequences for competitive activity if left unchecked. By contrast, control points in NGN may reside in any layer of the network hierarchy, from basic access to services and content. They will probably be difficult to identify, and they will have less predictable consequences for competitive activity if left unchecked, or indeed, if regulated.

Not all control points are harmful to competition. On the contrary, most are essential features of a healthy, competitive marketplace, in which companies search continuously for commercial advantages over their competitors. ‘Customer capture’ is the goal, and various methods of attracting and retaining business will be tried.

These may be in the form of intellectual property rights, proprietary solutions, alliances, exclusive supply contracts, strategic investments, simple product differentiation, and ‘walled gardens’ offering pre-determined service functions and content to attract and retain customers. Most of these kinds of activities will provide temporary advantages at best. Some, but few, may lead to a concentration of market power requiring regulatory or anti-trust intervention. Whatever steps are taken, governments should not unwittingly create advantages for some players at the expense of others.

Meeting this challenge will entail distinguishing between potential control points that promote normal competitive activity, and those that may harm competitive activity. This will have to be done in a complicated setting – an NGN marketplace that integrates

commercial activities from diverse market sectors characterised by different levels of competitive activity and subject to different regulatory regimes.

In the sections following this introduction, potential control points in NGN are identified and discussed. The extent to which they will require regulatory attention, or deserve specific ex ante regulatory action or ex post remedies will be subject to debate and cannot be predicted at this point. Nevertheless, specific examples of the kinds of market imperfections likely to attract and deserve the attention of regulators are provided.

1.3. Self correcting market forces and integration

As the transition to competition proceeds, the marketplace will be driven more by competition than by inherited advantage. Thus, the effectiveness of self-correcting forces found in a competitive marketplace, especially as these forces grow in strength over time, will have to be weighed by policy makers seeking the proper level of intervention in the NGN marketplace.

Policy makers will also have to decide how much industry concentration and vertical integration is beneficial for competition and for serving the needs of customers.

1.4. Jurisdiction and trade agreements

Services offered to customers in NGN are, in effect, geographically distributed processing services, with software, content and communications components pulled in from different jurisdictions and accessible from any location where roaming services are offered.

These attributes have the potential of enabling service provider's considerable latitude in selecting the jurisdictions in which they choose to operate and to locate the facilities generating their service offerings.

Will these considerations weaken the ability of regulators in any particular jurisdiction to apply and enforce its chosen set of rules? Will service providers and operators shop around for business-friendly regulatory jurisdictions, and then serve customers in any other jurisdiction they choose? If this is a realistic possibility, how will it affect economic activity in jurisdictions with rules considered less business-friendly? When developing regulatory frameworks, must policy makers factor in a need to 'compete' with other jurisdictions to attract business?

In fact, it might be argued that national or EU-wide ex ante regulations will be largely futile unless they are developed in the context of agreed and enforceable trading rules. Over the past several years, agreements among countries on the treatment of basic telecommunications and value added services have been negotiated through the World Trade Organisation (WTO). Additional negotiations on these and related services are now underway.

Many WTO Member States have already made commitments to allow cross-border provision of basic telecoms, value added and related services, and the question arises as to whether such commitments will carry over and apply to similar services delivered in an NGN environment, or whether a need will be seen for new or additional commitments.

Furthermore, the growth of e-commerce in recent years, destined to play a central role in the NGN marketplace, raises additional issues for trade negotiators and their colleagues in charge of e-commerce policies for the EU and Member States. If additional

commitments are made in this area, basic WTO principles would require that the underlying EU or national level regulations for e-commerce be transparent, non-discriminatory, and contain the least-trade restrictive measures available. Indeed, trade negotiations may offer an effective venue for addressing some of the issues raised by the geographically distributed components of NGN.

1.5. Reader's guide

Prior to discussing a number of regulatory issues, this report has been structured to provide the reader with an insight into the technologies associated with NGN as well as an understanding of the market views on NGN related developments.

The report starts by outlining the technology framework around which the following observations and analysis have been developed. Chapter 2 describes the overall NGN architecture and the main functional principles guiding NGN developments whereas chapter 3 provides a technology tutorial using diagrams and plain-language explanations to show component parts and how they are likely to interact to provide advanced communications and information services. Chapter 3 has been designed to enable the motivated reader to achieve some fluency in the terminology and technical concepts under discussion now and in the near future.

Readers with a non-technical interest in the report's findings and conclusions might want to proceed straight from chapter 2 to chapter 4 which provides a summary of the views expressed by representatives of 30 companies interviewed for this study.

Chapter 5 provides an assessment of the regulatory implications of the differences between conventional circuit-switched technology and Internet protocol (IP) technology. This section specifically discusses the general nature and implications of control points and also provides a number of examples of functions that could provide a basis for such control points.

Finally, chapter 6 presents a number of concluding recommendations.

2. The NGN Architecture

The purpose of this section is to provide a general technical and functional interpretation of NGNs as a base for discussions in the following sections.

It is generally accepted that Next Generation Networks do not represent specific technologies but rather a vision and a market concept enabled by different technologies. This chapter, as well as a chapter 3, explores the technology framework and a number of specific technologies which are often associated with the NGN market. The main purpose is thus to set the scene for the market observations and the regulatory discussions that follow in chapters 4 and 5.

2.1. NGN principles

From a high-level perspective, Next Generation Networks rely on three main principles.

First of all, NGNs are implemented in such a way that the functions performed by the network are separated into functional planes. The functional planes include access, transport, control & intelligence, and service. Layers are independent in the sense that they can be modified or upgraded regardless of other functional layers. This layered architecture provides a flexible and scalable network, reducing time to market for the implementation of new services.

Moreover, the functional planes are separated by open interfaces in order to facilitate the interconnection to other operators' networks but also the integration of third-parties' services and applications. Provided that commercial agreement is reached between the different parties, such a principle can widen the operator's coverage and service scope and can also provide end-users with an access to a greater number of services.

Last but not least, NGNs are multi-service networks, meaning that an NGN can be used to provide multiple services, as opposed to legacy networks that are only used for specific services. This multi-service network enables operators to implement converged and new services. From the users' perspective, the convergence of services will enable the emergence of the seamless service concepts, where users can access their "home" services from any type of access network.

2.2. Architectural concepts

The NGN architecture can be illustrated as shown in the figure below.

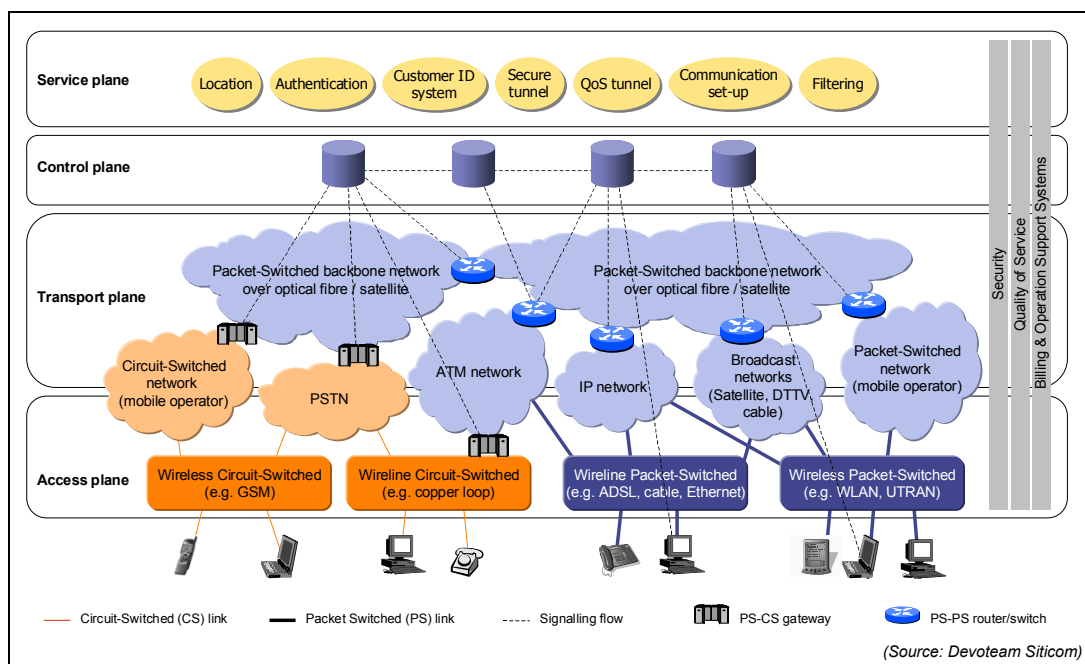


Figure 1 Typical representation of the NGN architecture

Other technologies not mentioned in the above figure will also play a role in NGNs whereas some of the technologies mentioned in the figure might disappear or not have the predicted impact.

Operations Support Systems (OSS) and billing systems, Quality of Service (QoS) management systems and security systems are used in all of the four planes described below.

The architecture is composed of functional planes that perform tasks at different levels:

- The **access plane** provides the infrastructure, i.e. the access network, between the end-user and the transport network. The access plane may be wireless or wireline, and it can be based on different transmission media e.g. copper wires, cable and fibre optic. Technologies in the access plane can be circuit-switched or packet-switched. The access network is connected to network nodes at the edge of the backbone network.
- The **transport plane** provides transport between network nodes to which the access networks are connected. The transport plane consists of one or several backbone networks based on packet or cell switched network nodes. Links are mainly based on optical fibre links but can also be satellite or terrestrial links. The transport plane is capable of handling different kinds of traffic, e.g. voice conversation, streaming video, interactive data, and batch data. Gateways at the edge of the transport network converts traffic to and from legacy networks, e.g. telephony, Internet and real-time data applications.
- The **control plane** includes both service control and network control elements. As such, the control plane controls all other planes shown in the figure above: access, network, and services. The control plane can for instance be responsible for the

control of communication sessions, e.g. establishing or disconnecting voice calls or multimedia sessions, intelligent service provisioning or resources provisioning depending on the service requested. A core principle of the NGN architecture is to separate the control logic from the underlying switching hardware.

- The **service plane** offers elementary service functions that can be used by service providers to build more complex or comprehensive services. Examples of such elementary functions are shown in Figure 2 below. The service plane also provides interfaces towards service providers who want to use these elementary service functions to access the underlying infrastructure. Such access will depend on commercial agreements between service providers/third parties and network operators. The interfaces may be implemented in different ways, e.g. in the form of APIs for service specific software to be run on servers within the network, or in the form of open standardised interfaces between the network and application servers. Such interfaces will enable the unbundling of services and underlying technologies.

Transport- and traffic-related elementary functions	Content- and service-related elementary functions	AAA- & billing-related elementary functions
<ul style="list-style-type: none"> • Create tunnels in or across networks • Secure network traffic • Guarantee a certain QoS • Restrict communication with other network and nodes • Redirect traffic 	<ul style="list-style-type: none"> • Set-up and manage a voice or multimedia communication • Provide a Virtual Home Environment • Resolve name and numbers • Determine a user's location • Determine a user's status • Restrict access to certain types of content and services 	<ul style="list-style-type: none"> • Authenticate a user • Authorise a user • Collect information on the user's use of (network) resources • Produce a bill

Figure 2 Examples of elementary service functions

The elementary service functions listed above do not control nor provide the infrastructure resources associated with the elementary service required (e.g. multimedia communication set-up) but only trigger the activation of the associated control elements and therefore of the resources and other elementary services needed.

The description above highlights that potential control points might appear in any of the four functional planes. For instance:

- market power in the access plane could provide access operators with some control on customer ownership;
- market power in the transport plane could provide network operators with some control on interconnect conditions;
- market power in the control plane could provide network operators or service providers with some control on how resources and services are provided (communication set-up, control of bandwidth and QoS, etc)
- market power in the service plane could provide service providers / network operators with some control on how easily a third-party can access and use the underlying infrastructure or “elementary service functions” to provide more elaborate services to their customers.

3. Technology and Systems

This tutorial chapter provides a detailed introduction to a number of key concepts, which are likely to influence future developments in Next Generation Networks (NGN).

As already shown, implementing the NGN architecture involves putting together a jigsaw of technology components working at various functional layers and this chapter explores and explains the specifics of the technology components implemented in the NGN architecture described in chapter 2.

Many of the technology components explained below could, either by themselves or in conjunction with other components, constitute potential control points in the NGN market place. While the specific discussion of such control points is left to chapters 4, 5 and 6, this chapter provides the technical basis for such discussions. Readers with non-technical interest may proceed to chapter 4.

3.1. Migration to NGNs

3.1.1. Introduction

The drivers for migrating to NGN are the same for all the players: if possible, reducing their network infrastructure and maintenance costs, but most of all, enabling faster service deployment for the provisioning of enhanced services and therefore creating new sources of revenue. The high flexibility, low cost and wide support throughout the world of the Internet Protocol nominates it as the best option for building NGNs, even though it has some limitations that need to be overcome, as for instance the lack of guaranteed QoS.

Each network operator will potentially choose a different migration path depending on their actual assets. This path will therefore involve different technologies and happen at different speeds. The figure below shows how NGN technologies relate to more traditional technologies and gives an indication of the migration depending on the operator's background.

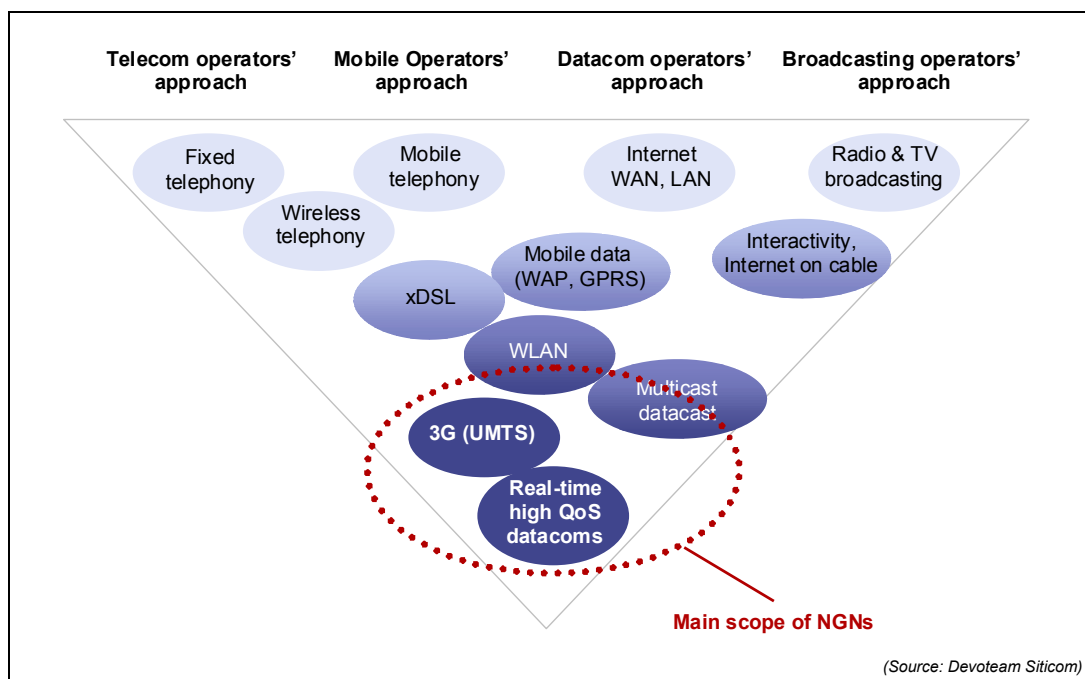


Figure 3 NGN perspectives and technologies

The main categories of operators described above can also be used to define different types of network migration: the migration of circuit-switched networks used for telephony services, the migration of packet-switched networks used for data services, the migration of mobile networks, and the evolution of broadcasting networks. Each of these is described in the following sections.

The section on market developments provides further details and interviewees' views on the technology transition (see section 4.1).

3.1.2. Circuit switched networks migration

3.1.2.1. Current situation

The Public Switched Telecommunications Network (PSTN), based on circuit-switched technology and built by incumbents and competitive carriers, provides telephony services with high QoS, enhanced with services provided by the associated Intelligent Network (IN) (e.g. call back on no answer service). These carriers have also implemented packet-switched data networks and services on top of their circuit-switched voice network, relying on different technologies such as Frame Relay, ATM or IP. Therefore, they currently have a layered architecture implemented, using different technologies for the delivery of different services. Internet traffic originated on the PSTN can for instance be routed to the IP network as early as possible.

For circuit-switched network operators, a multi-service network for converged services and more revenues represents the main driver for the migration to NGNs. However, a migration also incurs various direct and indirect costs including network upgrades, engineer retraining costs, organisation processes changes, etc. Moreover, some circuit-switched networks are still in a very good state and provide top quality telephony services, which is currently difficult to replicate on large-scale IP networks. Some new entrants have even claimed that the cost per line of CS equipment in the access network is less than that of IP, while still providing greater quality, which could be a further reason

for delaying the migration to a packet-switched IP network. In addition, a circuit switched access network would seem to provide better conditions for control over the customer by the network operator.

Notwithstanding these points, as packet-switched technologies mature and enable the integration of telephony services and other multimedia services, we will likely observe a gradual migration from circuit-switched to packet-switched technology but with a long period of co-existence of 10 years or more to allow network operators to phase out circuit switched equipment. This implies that issues such as call set-up, QoS and billing across networks are solved.

Two types of migration scenarios could be foreseen based either on a replacement or on an overlay strategy:

- The replacement strategy consists of replacing traditional PSTN network devices by next generation devices. This can be done either in the core of the network in order to provide further capacity and to enable a better utilisation of the network resources, or at the edges of the network, in order to provide advanced services to the customer.
- The overlay strategy is described in greater details below.

3.1.2.2. The overlay strategy

With an overlay strategy, the NGN network will integrate current circuit switched and packet switched technologies. The modern packet-based overlay network will provide advanced services whilst the circuit switched PSTN network will continue to provide basic telephony services. Both networks are interconnected via gateways as required by specific types of services (e.g. VoIP call originated from an IP phone and terminated on the PSTN, or Internet data traffic originated from the PSTN). When the overlay network eventually becomes capable of providing sufficient QoS, all traffic could be diverted from the circuit switched PSTN to the packet-switched overlay network as shown in the figure below.

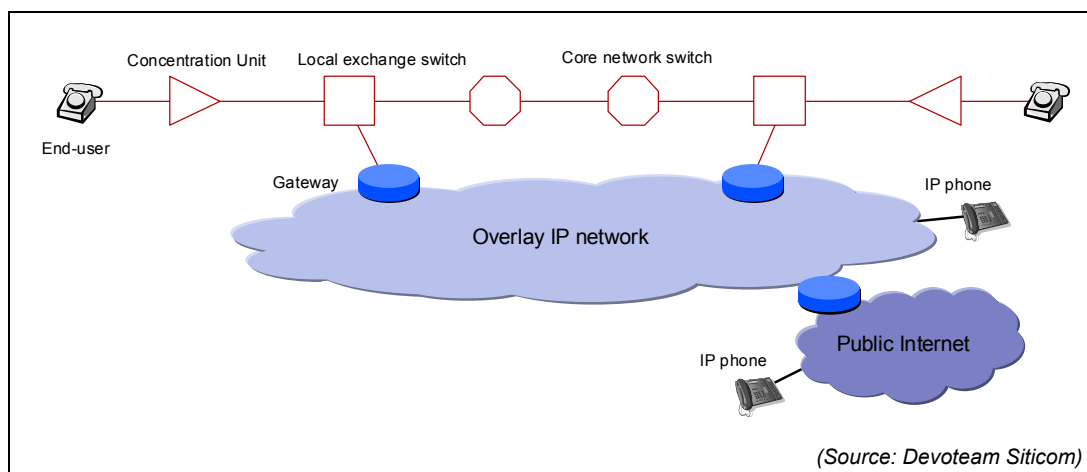


Figure 4 Overlay scenario

The overlay backbone in such a migration would be based on IP from the start, possibly used together with other technologies in order to provide sufficient QoS for some specific applications (e.g. IP over ATM). Session Initiation Protocol (SIP) has shown to be the best signalling protocol for implementing call control on IP networks as it can be integrated easily with other Internet protocols, as opposed to H.323 (see section 3.6,

“Interworking of signalling systems”). However, due to the early availability of H.323 products compared to SIP, some operators have already implemented H.323 solutions and will therefore not migrate to SIP in the short term.

The replacement and overlay strategies are not exclusive and some operators actually transport voice traffic on packet-switched trunks in the core network while also building overlay networks.

3.1.3. Packet switched networks migration

3.1.3.1. Current situation

Packet-switched data networks rely on several technologies and a number of protocol stacks are used depending on the service delivered and the functionality offered by each protocol, e.g. “IP over SDH over DWDM”, or “IP over ATM over SDH over DWDM”, or “IP over Ethernet over SDH over DWDM”. The migration towards NGN for these types of networks means a simplification of the network and more flexibility. Building an NGN network also means that the network needs to support converged services such as voice or real-time applications. This implies that certain features are implemented.

The migration of packet-based networks towards NGN can follow several strategies that could possibly be combined. A migration could also involve a move towards the next version of the Internet Protocol, IPv6. This section only describes the migration to Internet Protocol version 6.

3.1.3.2. Migration to IPv6

IPv6 is the upgraded version of the current Internet Protocol version, IPv4. It has already been fully specified by the IETF but has not been widely implemented yet.

The key drivers towards IPv6 are the address space provided and the mobility features inherently implemented in IPv6. Address fields contained in IPv6 headers are coded on 128 bits instead of 32 bits in IPv4 addresses, which could solve the address scarcity problem that is currently being encountered. Some mechanisms like Network Address Translation (NAT) are currently being used in order to solve the problem of address scarcity, but these add complexity in the implementation of peer-to-peer applications and are pointed out as a mechanism used by network operators to retain control over end-users. Some solution providers have however indicated that certain software developments could overcome the NAT limitations in peer-to-peer applications and therefore limit the need for operators to deploy IPv6 on their networks.

Moreover, as opposed to an IPv4-only enabled host, an IPv6-enabled host can automatically create its own address when connecting to a network, based on a network identifier acquired from a local router and its own specific identifier. This facilitates configuration processes and mobility applications.

IPv6 packets can also have extensive headers, which provides further advantages like the possibility of avoiding tunnelling in mobility applications (see section 3.7.4, “The IP mobility challenge”).

On the other hand, the implementation of IPv6 requires hardware and software upgrades, in the operator’s network as well as in the applications and programs used at the user’s side. It is unlikely that operators will implement IPv6 in the core network as long as the “edge”, that is the end-user equipment, is not IPv6-enabled. Enabling the edge involves

an upgrade of commonly used operating systems and software into IPv6-enabled software, but also the deployment of the all-IP UMTS architecture based on IPv6. Operators would thus wait until there is a clear demand and business model for implementing IPv6.

The migration towards IPv6 could take a long time and it has been predicted that IPv4 networks will not have completely disappeared before 2030-2040. Hence the need for coexistence mechanisms in the mean time.

Three mechanisms can be used during the IPv4 / IPv6 coexistence period which are not exclusive and can be used together.

- The dual stack mechanism is based on the principle that a piece of equipment is enabled with both the IPv4 and the IPv6 protocols. It can therefore use the most suitable protocol depending on the capabilities of the correspondent device and of the transporting network. As an example, UMTS terminals will be dual stack.
- The tunnelling mechanism enables the hosts of clouds running a specific version of the IP protocol to communicate with each other through a network running the other version by the means of a tunnel. With this approach, IPv6 clouds could communicate through an IPv4 network (as it will most likely happen in the first stage of IPv6 deployment) or IPv4 clouds could communicate through an IPv6 network (later stage of IPv6 deployment). Tunnelling can occur between hosts, or between routers, or between host and router. For instance, in the case of tunnelling between two routers, the packets are encapsulated by an edge router at the edge of a network cloud and then de-capsulated by the border router of the receiving cloud.
- The translation mechanism enables the communication between IPv6 and IPv4 hosts when the dual stack mechanism has not been implemented. The translation between IPv4 and IPv6 is performed by a dual-stack border gateway.

3.1.4. Mobile networks migration

3.1.4.1. Current situation

Since the beginning of 2000, European countries have allocated 3G licenses to operators for the use of radio spectrum usage, either based on auctions, “beauty contests” or both. During this allocation the expectations for 3G were very high. Operators were backed by the financial market, and the prices of some licenses have thus reached very high amounts. This currently contributes to the financial problems of many operators. A number of large operators even wanted to secure a pan-European presence and have thus bought licenses in several countries. Mobile operators are now facing the challenges of recovering their investment costs but also to fulfil the commitments made in the bids for licenses.

One of the big challenges for 3G as opposed to the previous and very successful 2G standard GSM lies in the great complexity of the new environment. A number of other players are involved in the services provided by the mobile operator: content providers, service providers, roaming operators, billing aggregators, etc. In terms of standardisation, standardisation of 3G involves a greater number of standardisation bodies and fora such as 3GPP, 3GPP2, ITU, JAIN, GSM association, IPv6 forum, etc. In comparison the GSM standard was fully specified by ETSI, with further guidelines given by the GSM association.

In Europe, the 3G standardisation work has mainly been carried out by 3GPP. The resulting standard, UMTS, has been divided into 3 phases, corresponding to 3 stages in the roll-out of the mobile network: UMTS release 99, UMTS release 4, UMTS release 5 and UMTS release 6. The breakdown into such a number of stages shows that the evolution from 2G to 3G will be highly gradual, with overlaps between 2G and 3G technologies requiring multi-mode phones. The first release, UMTS release 99, strongly relies on the features introduced by the intermediary mobile generation (2.5G): GPRS. In Japan, NTT DoCoMo is already beginning the construction of an experimental 4G mobile network, but no standardisation has been made yet and no frequencies have been allocated for this further step. The milestones from 2G GSM to 3G UMTS release 6 are described below.

3.1.4.2. Migration milestones

> From 2G GSM to 2.5G GPRS

The GSM standard has been built with the same focus as the PSTN network: it is a circuit-switched network, primarily aimed at providing telephony services, with the added functionality of the radio access and of mobility. As in the case of PSTN networks, data sessions have been made possible, but with the constraints of a circuit-switched network, e.g. poor utilisation of the bandwidth utilisation from the operator's perspective and low data rates.

GPRS services have been launched from mid-2001 and provide a more efficient way of handling data traffic. As illustrated in the figure below, the innovation of GPRS is to introduce a packet-switched core network in the mobile operator's network, enabling the activation of packet-based "always on" data sessions.

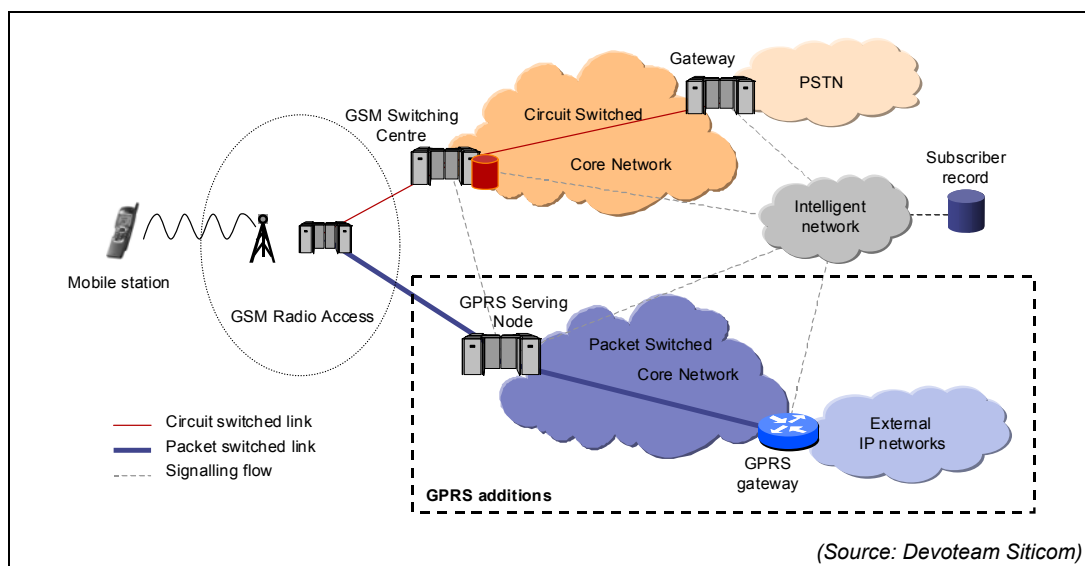


Figure 12 Migration from GSM to GPRS

> From 2.5G GPRS to UMTS release 99 and UMTS release 4

In the first release of UMTS, UMTS release 99, the core network principles remain the same: it uses a circuit-switched core network for voice service and a packet-switched core network for data services. The main change lies in the radio access part, called UMTS Terrestrial Radio Access Network (UTRAN), enabling higher data rates, as shown in the figure below. Release 4 offers further services and features such as Virtual Home

Environment (VHE) and Open Service Architecture (OSA). Release 4 also fully supports Location Services.

The first commercial offers based on UMTS release 99 will be launched in Europe starting from the end of 2002 and in the course of 2003, in selected areas only. In order to minimise investment, operators are actually limiting the rollout of UMTS and are focusing on dense urban areas only in a first stage, which requires the availability of "dual-mode" handsets in order to provide GSM/GPRS access where UMTS is not available.

Some mobile operators are also thinking of, or already launching, WLAN services as an additional service, which could be seen in the longer term as a complementary technology to UMTS when multi-mode handsets are available. There is however currently no standardisation in this area, although operators, the GSM association and 3GPP are currently working on this (see release 6 below).

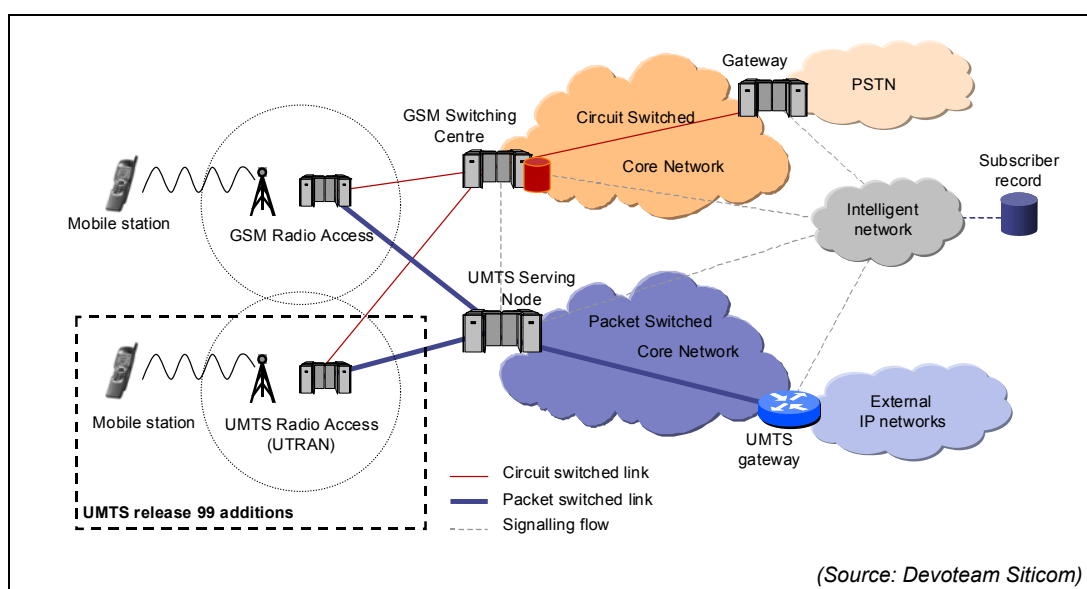


Figure 6 Migration to UMTS release 99

> From UMTS release 4 to UMTS release 5 and release 6.

UMTS release 5 introduces a major change in the core network: the circuit-switched core network disappears and the architecture becomes all-IP. Voice calls can be carried on the converged IP network, using SIP as the signalling protocol for call set-up, as has been specified by 3GPP, and IPv6 as the version of the IP protocol.

The access network and the packet-switched network remain the same, but UMTS release 5 introduces the concept of IP Multimedia Subsystem (IMS). The IMS includes all new control and signalling functions for the management of multimedia sessions, including voice calls and data sessions. It is functionally separated from the packet-switched core network, which provides the transport functionality.

The general principles of the IMS are:

- It performs control and signalling functions, capabilities and QoS negotiations, and media translation.

- It must support legacy services based on Intelligent Network, new services based on IP, SIP, JAIN, XML, etc, and third party services based on the Open Service Architecture (OSA) (see section 3.4 on Service provisioning).
- It is independent on the network access, which means that it can interwork with UTRAN, GSM, WLAN or PSTN access.

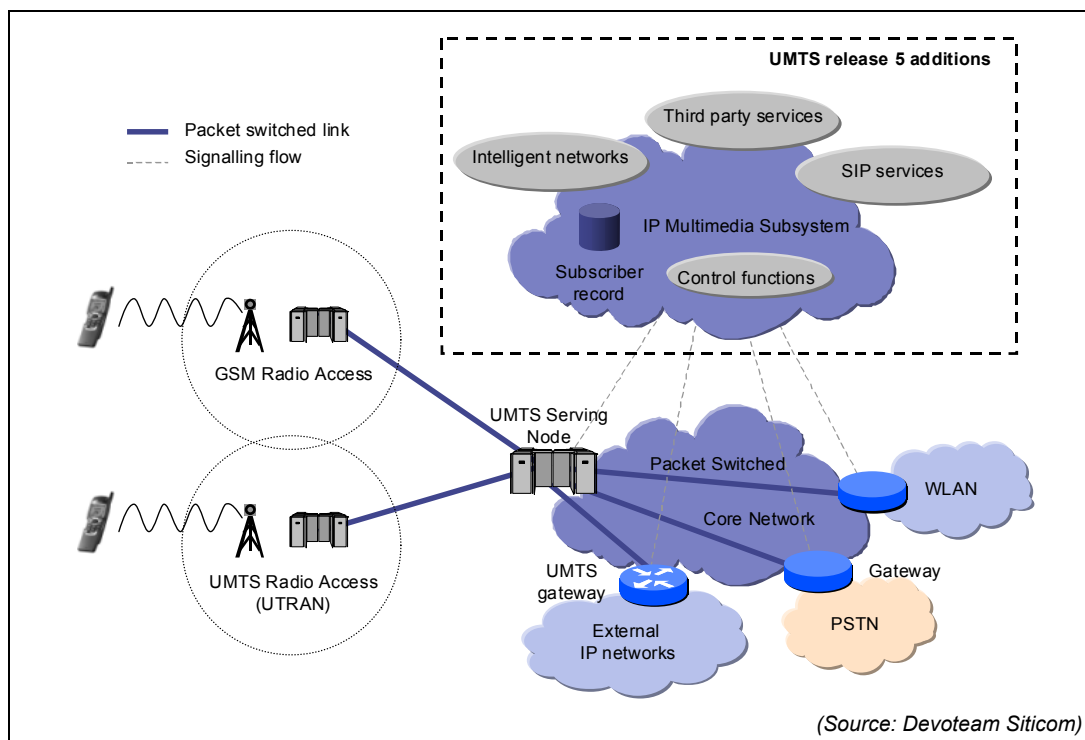


Figure 7 Migration to UMTS release 5

A release 6 is currently under development which should include further enhancements in the radio network, in multimedia messaging services, in security and consider further issues such as multimedia broadcast and multicast services, interworking between Wireless LAN and UMTS, etc.

3.1.5. Broadcast networks migration

Market players' views on the convergence of broadcasting and telecom and data vary and convergence is understood in several different ways:

- As the use of legacy broadcasting networks to provide any-to-any unicast services or datacast services on top of traditional broadcasting services;
- As the use of data networks such as the Internet to provide multicast services (that is one-to-many) on top of traditional unicast services.
- As a combination of broadcasting and cellular technologies to deliver services and content to mobile users.

This section focuses on the evolution of broadcasting networks to provide telecom and data services and on the possible combination of broadcasting and cellular technologies to deliver services to mobile users. Section 3.2 provides further information about cable TV, DTTV and satellite technologies in access networks.

3.1.5.1. Evolution of legacy broadcasting networks toward data and telecoms

Legacy broadcasting networks include terrestrial broadcast networks, cable TV networks and broadcasting satellite networks. These networks have been upgraded in a first stage to provide digital rather than analogue services in order to increase the quality and the number of channels provided for a given bandwidth. Digital broadcasting has been based on two main standards: DAB (Digital Audio Broadcasting) and DVB (Digital Video Broadcasting), standardised by ETSI.

Following further enhancements based on the technology HFC (Hybrid Fibre Coaxial, see section 3.2.5.2), cable TV networks have been able to carry interactive TV, two-way Internet and telephony services.

Internet access offers based on broadcasting satellites can support unicast transmission to the customer but require an alternative Internet access such as dial-up for the upstream data flows. The recent DVB-RCS (Digital Video Broadcast – Return Channel for Satellite) standard endorsed by ETSI can however enable two-way asymmetric communications and remove the need for another access mechanism for the upstream flows.

Digital terrestrial television (DTTV) networks could also potentially provide one-way unicast services and would therefore also require the use of another Internet access mechanism.

The recent DVB standard MHP (Multimedia Home Platform) is promoted by ETSI to enable the provision of digital television services and Internet services on broadcasting networks, including cable, fibre optic, satellite and DTTV networks. This requires the distribution of MHP-enabled set-top boxes that can be used together with either a television receiver or a personal computer. The penetration of MHP is however still limited as it requires broadcasting operators to replace all set-up boxes already supplied, which represent a very high investment.

3.1.5.2. Combination of technologies

Although satellite and DTTV generally only provide one-way communications, some argue that it could provide an alternative solution for providing high-bandwidth downstream transmissions to mobile users whereas usual mobile technology (GPRS or UMTS) could be used for upstream transmissions. In the short term, there are however some technical issues that prevent the implementation of such a solution:

- DTTV and mobile cellular technologies are very different in terms of cell coverage, transmission and terminal equipment and network planning. Therefore a mobile network operator could not take advantage of their actual mobile network to provide broadcasting services using DTTB technology (and vice versa).
- Various national power transmission restrictions on DTTV infrastructure are in place and the infrastructure is still used to provide both analogue and digital transmission. This implies that networks will be unable to transmit to smaller mobile receivers. If these limitations were overcome, handsets would need to be multimode to be able to receive communications using several technologies.
- The use of broadcast infrastructure to provide unicast transmission means that only lower data rate could be achieved (as opposed to broadcast transmissions) and

therefore the use of the broadcast infrastructure would not present an obvious advantage as opposed to the mobile network infrastructure.

- Lack of frequency harmonisation among broadcasting networks would require further complexity in the receiver to enable roaming across Europe. Moreover, in some countries, the frequency bands available can not provide sufficient bandwidth to deliver true broadband services.

During the interviews conducted as part of this study broadcast operators claimed that they have no plans for fully migrating broadcast services to IP and that they will continue to use broadcast specific non-IP networks.

3.2. Developments in access technologies

3.2.1. Introduction

The development of access technologies is an important driver for the development of NGN and will impact the development of broadband services and content.

This section provides an insight into the technologies under development today or technologies that are representative of market trends. The section does not describe directly comparable technologies in terms of functionality (in the sense of OSI layer functions) and it remains clear that new standards and technologies will emerge in the future. As such the section does not provide an exhaustive list of all access technologies.

Access network technologies usually refer to networking technologies providing connectivity between the end-user and the transport plane, what is usually called “the last mile” or the “local loop”. In this section, this definition is extended to include further technologies providing shorter- or longer-range connectivity:

- Personal area access technologies - providing only short-range connections;
- Local area access technologies - providing wider but still local access;
- Last mile or local loop technologies – providing full connectivity between the end-user and the transport network;
- Global area access technologies – providing connectivity up to the customer premises but potentially covering much wider areas than usually covered by “last mile” access networks and potentially also providing transport functionality.

3.2.2. Trends in access technologies

Figure 8 shows the positioning of some access technologies according to the categories given above. Some of these technologies are described in the following sections.

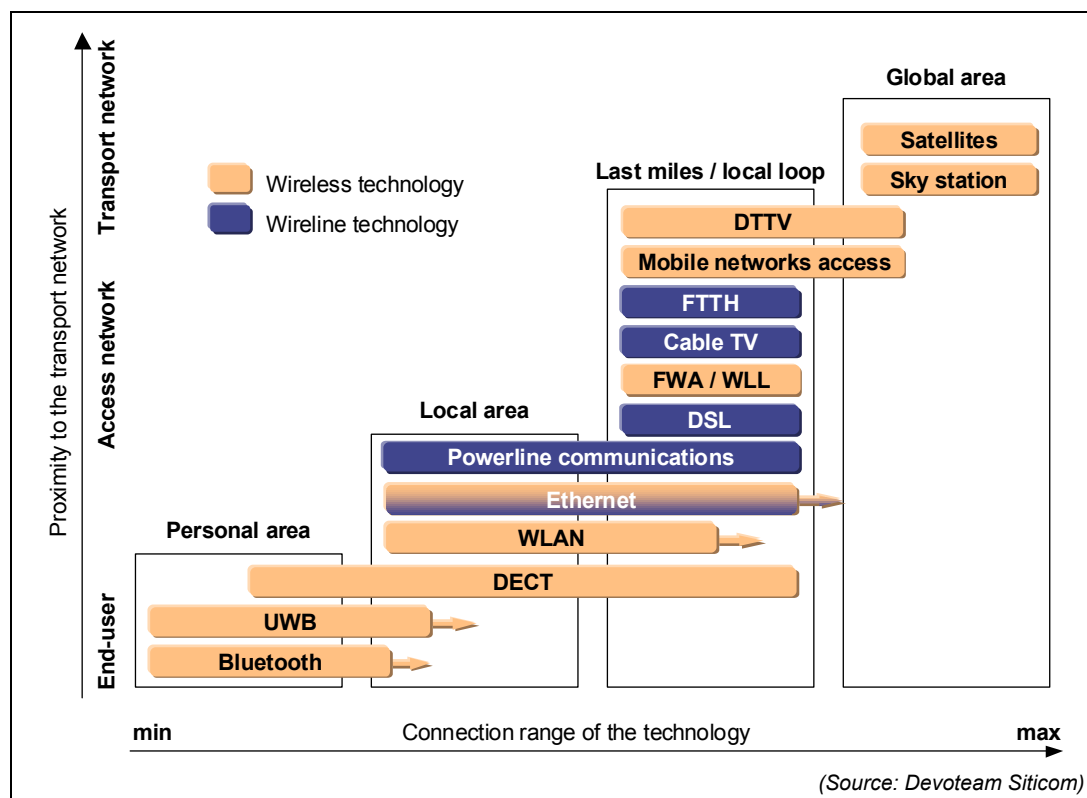


Figure 8 Positioning of access technologies

Trends in access technologies show that:

- Access networks will increasingly be based on broadband, packet-based and “always-on” technologies.
- Mobility and wireless connectivity is becoming increasingly important and is supported by Wireless Ethernet (WLAN), 3G networks, Bluetooth, etc.
- Technologies are developed to enhance the capabilities of existing wireline infrastructures and optimise the usage of existing assets providing wide or even ubiquitous coverage – e.g. DSL technologies, Cable TV technologies, etc.
- Previously dedicated networks such as electricity networks or broadcasting networks are perceived as potentially suitable means for providing data and/or voice services.

The massive development of access technologies reflects the strategic importance of the access market to keep or gain control over the customer.

3.2.3. Personal area access technologies

A number of technologies are providing short-range connection between personal devices forming a “Personal Area Network” (PAN). Two technologies are especially significant in this category: Bluetooth and Ultra Wide Band (UWB) technologies.

Bluetooth has been developed recently in 1998 by the Bluetooth SIG industry group and has already been widely implemented in the latest computers, mobile phones, Personal Digital Assistants, etc. It is a wireless protocol operating in the Ultra High Frequency spectrum (2.4GHz band, like the IEEE WLAN 802.11b standard, see below). Up to 8 devices enabled with Bluetooth transceivers can exchange voice and data traffic with each other at data rates up to 1Mbit/s both ways within 10 meters. With increased

transmission power, Bluetooth could evolve to provide higher-range wireless connections of the type of Wireless LAN. In such case, a Bluetooth-enabled device would serve as an access point to the wired network.

The Ultra Wide Band technology, or digital pulse wireless, has been developed much earlier, in the 60s, for military purposes in the US. It is a short-range connection technology operating in the frequency bands from 3 to 10GHz and in the band 26 / 28 GHz and in the band 43 GHz of the radio-frequency spectrum and which can carry broadband traffic with a very low transmission power. Such a low transmission power guarantees that no interference is possible as UWB signals are lower than usual background noise. The technology could provide data rates up to 10Mbit/s up to 50m and potentially 100Mbit/s for very short distances.

Although the UWB technology is older than Bluetooth, the FCC has only approved the use of UWB in wireless products for consumer and business usage in February 2002. Companies including JVC, Sonic, Panasonic, Intel, Motorola and Sharp are currently working on defining an industry standard based on UWB. The first products could be launched in 2003 for the linking of personal devices but also for instance for the supply of TV signals to other TV sets or computers within a house or building. Other applications of the technology include radar-type or tracking applications.

3.2.4. Local area access technologies

Local area access technologies provide longer-range connections but are still limited to a Local Area Network² (LAN). The connection of a LAN to a transport network generally requires an additional access link. In addition to the technologies explained below other technologies such as DECT might continue to play a role in local area access.

3.2.4.1. Ethernet

Ethernet protocols have been standardised by IEEE in the 802.3 series. Ethernet was designed as a wired LAN technology, connecting computers within a building or an office. Several protocols have been successively standardised since the 70s to provide greater functionality. The bandwidth provided has increased from 10Mbit/s initially, to 100Mbit/s with Fast Ethernet (1995), 1Gbit/s with Gigabit Ethernet (1998) and potentially up to 10Gbit/s.

Therefore, although Ethernet protocols have traditionally been used in wireline LAN (on copper wires or copper coaxial cables) they can now also be implemented in wireless LAN and in MAN³ or WAN⁴ on fibre optic cables for the fastest versions. This nominates Ethernet not only as an access technology but also as a transport technology.

Ethernet has been designed to carry IP traffic and can carry different types of services including voice, data and video.

² A LAN is a group of computers connected together within a building, an office or a home.

³ A MAN (Metropolitan Area Network) interconnects computers in a geographic area larger than a LAN, but still smaller than the area covered by a WAN (see below). Typical examples of MANs are the interconnection of city networks into a single network for efficient connection to a WAN or the interconnection of several LANs.

⁴ A WAN (Wide Area Network) interconnects geographically dispersed computers and networks.

3.2.4.2. Wireless LAN (WLAN)

WLAN standards have mainly been developed by two standardisation bodies, IEEE and ETSI. WLAN technologies enable the connection of devices to an access point, which is then connected to the Internet and/or Intranet.

The most popular of the WLAN standards is currently 802.11b as specified by IEEE in 1999. It enables data rates up to 11Mbit/s (about 6Mbit/s in usable bandwidth) in the frequency band of 2.4GHz (the same as Bluetooth, microwave ovens, etc). This standard has been widely supported by the Wireless Ethernet Compatibility Alliance (WECA) which includes major industry players (Apple, Microsoft, Cisco, Ericsson, Epson, HP, IBM, Infonet, Sony, etc). WECA has defined the Wi-Fi certification, which guarantees interoperability between 802.11b-enabled devices.

The IEEE 802.11a standard has been released after 802.11b and provides higher data rates up to 54Mbit/s in theory (about 30Mbit/s in usable bandwidth) at 5GHz. IEEE is currently working on other standards for WLAN, among them 802.11g that should provide the same data rates as 802.11a but in the 2.4GHz band and at a cheaper cost.

802.11a has however met compatibility issues in Europe due to the standard developed by ETSI, Hiperlan 2, which is using the same frequency band of 5GHz. ETSI had previously defined other WLAN standards such as RLAN (Radio LAN) and Hiperlan 1, also in the 2.4GHz and the 5GHz bands.

Both standardisation bodies are currently working on providing compatibility between their standards.

WLAN technologies can be used in both private and public environments. In the first case, a private entity (hotel, company, etc) provide the WLAN access to e.g. its members, customers, employees and requires the services of an access operator for the connection to Internet services or more generally transport services. In the second case, an operator set-up WLAN "hot spots" (access points) in public areas (coffee shops, airports, etc) and provide Internet access through these access points.

3.2.5. "Last mile" or "Local loop" technologies

Last mile or local loop technologies are used between the transport network and equipment at the end-user premises. This can be end-user devices (traditional telephone) or gateways such as modems and routers.

3.2.5.1. Digital Subscriber Line (DSL) technologies

Digital Subscriber Line (DSL) technologies are a family of modulation techniques enabling a wider use of the frequency spectrum available on copper wires.

These wireline technologies are particularly relevant to provide packet-based broadband services on the copper local loop of the public telephony networks and represent the incumbent's first choice to provide end-user with broadband services at a minimum cost.

The Local Loop Unbundling (LLU) process has aimed at making the local loop available to the incumbent's competitors.

Within the LLU process, third parties can have full or shared access to the local loop. Within the full access scheme, the copper wires arriving at the incumbent's Main Distribution Frame (MDF) are connected to the third-party's equipment and network.

Shared access to the local loop applies when the third-party does not want to provide the public telephony service. In this case, a splitter situated after the MDF splits the frequency spectrum of the copper wires into low and high frequencies, the low frequencies are then transmitted to the incumbent's network while the high frequencies are transmitted to the third-party's equipment. How interconnections are made in practical terms depend on the location of the third-party equipment (co-location or distant location).

The major limitation of DSL technologies is that the bandwidth provided decreases when the distance between the end-user and the operator's equipment increases. Depending on the topology of the incumbent's network, a certain (typically large) percentage of the population can thus not benefit from this solution.

DSL technologies can carry any type of higher protocols, such as ATM and IP.

The most well known DSL technology is Asymmetric DSL (ADSL) which is well-suited for asymmetric applications such as browsing. Typical data rates would be between 500 and 1.5Mbit/s downstream and 100Kbit/s upstream at 3km. This supports always-on data connection and can also support several voice channels. HDSL (High-bit-rate DSL) and SDSL (Symmetric DSL) are symmetric flavours of the technology, providing data rates above 2Mbit/s both ways. These technologies thus represent possible alternatives to leased line services. VDSL (Very high rate DSL) provides even higher data rates on smaller distances (51Mbit/s over 300m).

A new ITU-T standard, G.SHDSL (Symmetric High bit rate Digital Subscriber Line), is now being marketed. This supports symmetrical bandwidths up to 2 Mbit/s at a distance of 3 km, scaling down to 192 Kbit/s at 6 km.

3.2.5.2. Cable TV

Cable TV networks are wireline broadcasting networks used to broadcast television services. They are traditionally based on copper coaxial cable. In many areas, these networks have recently been upgraded to provide, digital television services, two-way broadband Internet access and telephony services.

Upgraded – and sometimes new – networks are based on Hybrid Fibre Coaxial (HFC) network technologies. HFC networks use a combination of fibre and cable technologies on different parts of the network to optimise performance with limited investment cost, as the fibre network is not deployed up to the customer premises. The network upgrades are however expensive as they still require operators to dig up the roads. Such networks can carry a wide range of technologies such as SDH, Ethernet, Frame Relay, ATM, and IP.

The standard MHP (Multimedia Home Platform) has also been released recently aimed at broadcasting networks – this does not only include cable TV networks, but also fibre optic, satellite and DTTV network (see below) – and enables the provision of digital television and Internet services.

Internet over cable TV usually provides around 1.5Mbit/s downstream, but neighbours in an area compete for bandwidth, which decreases significantly the data rates down to typical values of 500kbit/s downstream. Services offered can either be symmetrical or asymmetrical.

Access through cable TV benefits from the wide coverage of such networks across countries, although this coverage is heterogeneous depending on countries. For this reason, it has been seen as the main competitive technology to DSL technologies.

3.2.5.3. Power Line Communications

Power Line Communication (PLC) technologies are also referred to as Power Line Transmission or Digital Power Line technologies. They rely on the addition of equipment to the electricity network to provide two-way transmission services (Internet, voice and data service) at speeds of 1Mbit/s and beyond. Depending on the specific PLC technology, “taps” installed in the electricity network will interface with communication equipment and inject and extract communication signals into and from the electricity wires. Broadband communications can therefore be provided without new wiring by plugging a PLC modem into a normal power supply plug.

PLC technologies take advantage of the coverage and ubiquity of electricity networks to provide access in areas not covered by other technologies. They can therefore represent a broadband alternative to DSL over the copper local loop or to broadband over cable networks and increase competition in broadband infrastructures. PLC technologies can also be used for connecting buildings or for LAN or home networking purposes within buildings.

The PLC forum⁵ is an active forum for the exchange of PLC-related information between members and for various other purposes such as lobbying on regulatory issues, interoperability and standardisation issues, marketing and business case development. Members include utility companies (Electricité de France, Tokyo Electric Power Company, Scottish and Southern Energy, etc.), universities (Ecole Polytechnique Fédérale de Lausanne, University of Dortmund, etc.), manufacturers (Mitsubishi Electric Corporation, Schaffner, etc.) and others.

In terms of development, there is an increased level of activity in PLC trials in Europe (e.g. in Sweden, the UK, the Netherlands, Belgium, France, Italy, Portugal, Austria and Switzerland) whereas a number of utilities have already launched services for instance EnBW in Germany and Endesa in Spain.

3.2.5.4. Fixed Wireless Access (FWA) technologies

Fixed Wireless Access (FWA) or Wireless Local Loop (WLL) can be provided based on different technologies, one of them being Local Multipoint Distribution Services (LMDS) technology, which supports ATM and IP.

Depending on the band of the frequency spectrum on which it is implemented (e.g. 3.5GHz, 26GHz, 28GHz, 32GHz), FWA technologies can provide coverage from 3 to 10km. FWA carries different types of traffic, including voice, data, video and streaming traffic. Typical data rates offered for Internet access are similar to the ones provided by ADSL, but the technology supports symmetrical services and much higher bandwidth.

FWA technologies provide the opportunity for alternative operators not to rely on the incumbent for the provisioning of the local loop. However, the roll-out of expensive base-station and terminal equipment can still represent a major inhibitor.

⁵ www.plcforum.org

3.2.5.5. Fibre to the Home (FTTH)

“Fibre-to-the-home” (FTTH) is the term used to refer to an access network based on fibre optic cables. Fibre optic is a more recent generation of transmission media providing very high capacity, in the order of several hundreds of Gbit/s on a monomode fibre. Fibre optic cable can carry any type of upper layer technology.

Fibre technologies deployed to the customer premises provide high capacity on the last mile(s) (implementations exist at 10Mbit/s both ways). This supports high-bandwidth real-time multimedia services such as Video-on-Demand, Internet, Telephony, Video-telephony, and any-to-any video conferencing.

The roll-out of such an infrastructure is however very time consuming and expensive.

3.2.5.6. Mobile networks access

The GSM technology for mobile communications has met a considerable success in Europe. GSM networks ("2G" networks) are cellular digital narrowband networks relying on a radio connection between the end-users' terminal and the base stations of the mobile networks, connected to the mobile operator's transport network. With the specification of the 2.5G and 3G mobile networks, this radio access is being upgraded to provide further functionality. All these upgrades have been specified by the 3GPP and the 3GPP2 global partnership projects.

As GSM networks were principally designed for telephony services, GPRS access networks provide a first enhancement to GSM access networks by enabling "always-on" packet-based connections for data services. The usage of different coding schemes and multiple timeslots for data transmission enables the data rate to increase from 9.6kbit/s with GSM to about 50kbit/s with GPRS.

The UMTS access network (called UTRAN) relies on a different multiple access technology than GSM and GPRS, called CDMA (Code Division Multiple Access). The effective bandwidth experienced will however remain limited to between 100 and 300 Kbit/s.

The expected limited coverage of UMTS in a first stage and the fact that GSM/GPRS and UMTS operate in different frequency bands (900MHz and 1900 - 2100MHz respectively) will require multimode handsets to allow interworking and roaming between GSM/GPRS and UMTS access networks.

Further evolutions are expected, with NTT DoCoMo already beginning the construction of an experimental 4G mobile network enabling speeds of 100Mbit/s downstream and 20Mbit/s upstream⁶. In the UK, the Mobile Virtual Centre of Excellence defines long-term research programmes “beyond 3G” involving industry players (including a growing number of broadcasting and media companies) and universities in the UK.

3.2.5.7. Digital Terrestrial Television (DTTV)

Digital Terrestrial Television is a wireless broadcasting technology based on a terrestrial radio network. It can be provided on an upgraded version of the legacy radio network used to broadcast analogue television, or on a new network built from scratch.

⁶“NTT DoCoMo starts 4G experiment”, vnunet, 20.03.2002

Such networks could be used to provide datacast services on the downstream path only but require the use of an alternative access technology for the upstream path.

Section 3.1.5.2 discusses in more details the possible combination of broadcast and mobile technologies to provide services to end-users.

3.2.6. Global area access technologies

Global area access technologies provides connectivity up to the customer premises but can also cover much wider areas than usually covered by “last mile” access networks, potentially also providing transport functionality.

3.2.6.1. Geostationary satellites

Geostationary satellites have traditionally been used to provide broadcasting services, including radio and TV broadcasting services using the DAB and DVB standards. Satellites are however increasingly involved in the provisioning of communication services.

Broadcast satellites operate at microwave frequency bands and require copper coaxial cable between the dish used to receive the signal and the end-user equipment. When used for communication purposes, these satellites can usually provide one-way communication at data rates between 200 and 800kbit/s downstream and require an alternative access technology for upstream traffic (also referred to as the “return channel”).

The recent DVB-RCS (Digital Video Broadcast-Return Channel for Satellite) standard endorsed by ETSI allows two-way asymmetric communications at data rates up to 8Mbit/s downstream and up to 2Mbit/s upstream.

3.2.6.2. Mobile services satellites

Inmarsat, which is owned by a coalition of telcos and which already runs a number of geostationary satellites for communication services, has announced plans for the launch of further new geostationary satellites as part of a project of “Broadband Global Area Network” (B-GAN). The service supported by such satellites should include Internet access, Video on Demand, video-conferencing and telephony, at speeds up to about 400Kbit/s. These B-GAN services are expected to be compatible with 3G networks, when using compatible equipment and thanks to the use of open telecom standards⁷.

⁷ “Inmarsat awards Ericsson \$55 million contract for core network infrastructure”, Inmarsat Press Releases, 05.12.2001

3.3. IP network interconnect

NGN will rely on a number of heterogeneous networks of different size and different coverage linking together servers, workstations and terminal devices, as the Internet does today. In the long term, IP will be the common interworking technology for Next Generation Networks.

If NGNs are to provide convergent services and to become multi-service networks, a number of interworking issues between networks has to be overcome. Different sections of this tutorial are specifically dealing with some of these issues (interworking of addressing systems, interworking of signalling systems and roaming).

This section mainly describes the exchange of routing information between networks and QoS issues across networks.

3.3.1. Routing between networks

In IP-based networks, interconnection between networks means that routing information needs to be passed between networks. This routing information enables routers to forward a packet to the next network point (router or host).

The information is maintained in routing tables held by routers, and which contains details of available routes and further associated details.

How this information is maintained and the mechanisms used to update it depend on the routing policy established within each administrative domain or Autonomous System (AS), that is within one or several networks under the administrative control of a single organisation.

This section describes the basic features of a routing table, the different types of routing and the differences between intra-domain and inter-domain routing.

3.3.1.1. Features of routing tables

The content of a routing table varies depending on the routing model and routing protocols used (see below).

The entries of the routing tables are address prefixes corresponding to aggregated routes. Each prefix or route entry can represent one or several networks, with several hosts attached to each of these networks. The shorter the prefix is, the higher the level of route aggregation is. On the opposite, the longer the prefix is, the more information is provided about the network address. In the case where several prefixes match the recipient's address on a packet, a router will need to forward the packet on the aggregated route corresponding to the longest – therefore more detailed – prefix.

Depending on the routing protocol used, further information associated with each route is stored, such as number of hops required to deliver the packet, number of autonomous systems crossed, bandwidth available, delay, etc.

For the efficiency and performance of the Internet, it is important that routing tables do not grow faster than the switching capability of the routers: too many entries in routing tables compared to the processing power of the router could have the effect of delaying the forwarding of packets. It is therefore important to control the growth of the routing tables.

In the early 90s, the inefficient structure of IPv4 addresses into classes had indirectly led to an exponential growth of the routing tables. The main mechanism implemented to slow down this exponential growth is CIDR, "Classless Inter-Domain Routing", which provides a better solution to aggregate hosts and sub-networks addresses. CIDR discards the original class structure of IPv4 addresses and enables a more accurate allocation of addresses to organisations, depending on the actual needs of host addresses. It is currently used for address allocation in IPv4.

Multi-homing, the principle of having several interconnections with ISPs, is currently seen as a strong factor increasing the growth of routing tables. Multi-homed ISPs have blocks of addresses reachable through different routes; this prevents the use of address aggregation above a certain level and requires that longer prefixes are used as entries of routing tables, therefore contributing significantly to their growth.

3.3.1.2. Intra-domain versus inter-domain routing

> Differences between intra-domain and inter-domain routing

Two types of routing protocols can be used, either for intra-domain routing or inter-domain routing.

- Intra-domain routing is based on "Interior Gateway Routing Protocols" (IGP) - these protocols enable the exchange of routing information within a single administrative domain or AS.
- Inter-domain routing is based on "Exterior Gateway Routing Protocols" (EGP) – these protocols enable the exchange of routing information between different administrative domains or AS. They enable the exchange of traffic between interconnected ISPs.

The figure below illustrates that different interconnected AS can have different intra-domain routing protocols implemented within their respective administration domain. They need however to implement a common inter-domain protocol to exchange relevant information for interconnect purposes.

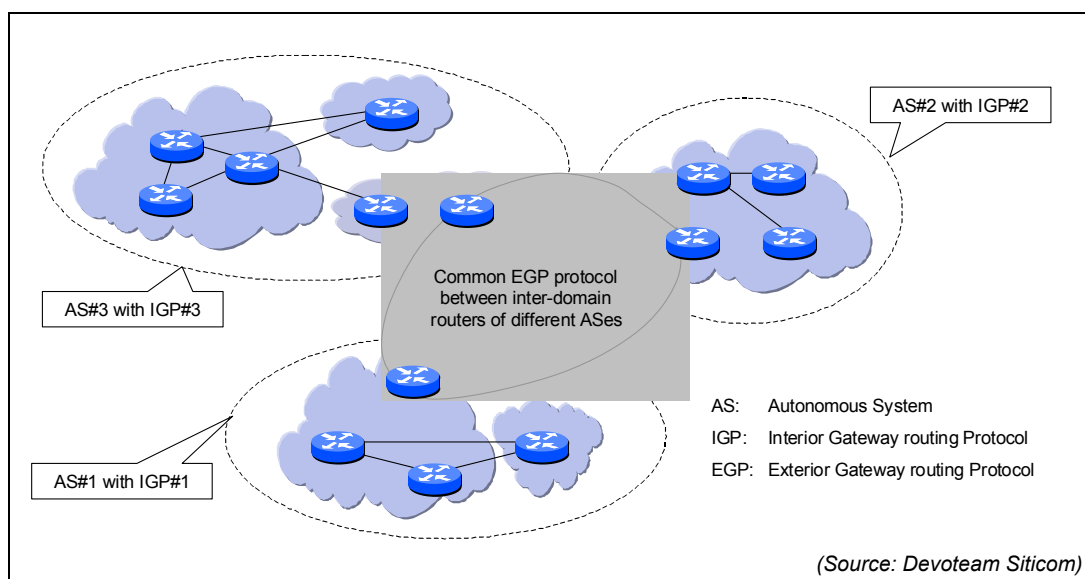


Figure 9 Implementation of Interior and Exterior Gateway routing Protocols

Both EGP and IGP have the purpose of making the routing of packets as efficient as possible. IGP however also have to take into account more strategic conditions related to

interconnect agreements. Such conditions would have to be implemented manually on the routers as a separate process. Conditions could for instance require that traffic never crosses a domain with a specified AS number, or never leaves the country if it is aimed at terminating in the same country as the one it originates in, etc.

3.3.2. Quality of Service across networks

3.3.2.1. Importance of QoS in NGNs

By default, IP networks provide "best-effort" quality of service, which has allowed the Internet to remain as simple as possible while the intelligence and complexity was relying on the end-devices. The challenge of provisioning QoS in IP-based networks is to add sufficient intelligence in the network while keeping the simplicity and robustness of the Internet.

Moreover, the ability to provide QoS does not involve only a single network but all the networks involved in the provisioning of a specific service. End-to-end QoS therefore requires that the different QoS protocols implemented across networks can be mapped to each other.

The following paragraphs introduce some of the QoS protocols and mechanisms that can be implemented in IP networks and briefly discusses the interoperability challenge between these different protocols. It also describes other mechanisms that can have an impact on the perceived QoS.

3.3.2.2. Possible QoS mechanisms in IP networks

Different mechanisms can be used to provide QoS in IP networks. Increasing the bandwidth available in a network or increasing the processing power of hardware components constitute simple mechanisms that do not require the implementation of further protocols in the network. Within a single administrative domain, the efficiency of routing protocols and network congestion policies also have a significant impact on the experienced QoS. This section only describes some protocols and mechanisms which have been defined to directly improve QoS in IP networks, that is by guaranteeing or optimising the value of QoS parameters for certain types of traffic. QoS policies have to be implemented along with QoS protocols in order to define which flow are entitled to which type of service enabled by the QoS protocol.

> Resource Reservation Protocol

The Resource Reservation Protocol (RSVP) defined by the IETF is inspired by telecoms and ATM networks in which a signalling phase takes place to reserve the necessary resources along the path from the sender to the receiver(s) before the communication is set-up. RSVP is a signalling protocol only and does not carry information packets – the transport of information is undertaken by the Real-time Transport Protocol (RTP). The reservation message contains information about QoS parameters and the description of the packets involved in the communication. The routers along the path of the communication are responsible for allocating the necessary bandwidth to the traffic flow and need to remember the QoS parameters of the traffic flow as long as the communication lasts.

RSVP enables Integrated Services (IntServ), that is the provisioning of two types of CoS: the Guaranteed Service CoS, which provides strict guarantees on parameters such as

bandwidth or delay, and the Controlled Load CoS, which provides a CoS similar to a best effort service under unloaded conditions whatever the condition of the network.

This protocol provides high QoS guarantees but also increases significantly the complexity of routers.

> Differentiated Service

Differentiated Service (DiffServ) is a simpler and more scalable protocol also defined by the IETF. Instead of reserving resources, DiffServ consists in classifying packets depending on their priority in the network and mark them with an index defining its Class of Service. The marking (and unmarking) of packets occurs at edge ingress points (and edge egress points) while core routers process the packets according to their priority.

The priority can be given in the ToS (Type of Service) field of the IPv4 header or in the Traffic class field of the IPv6 header depending on the IP protocol implemented. The two main CoS used in DiffServ are Expedited Forwarding (EF), which minimises delay and jitter, and Assured Forwarding (AF), which minimises packet loss.

As opposed to the RSVP protocol, no signalling is required and the routers involved in a communication using DiffServ are stateless while the intelligence remains at the edge of the network. DiffServ however only provides statistical guarantees.

> Label Switching

In label switching protocols, packets are processed depending on how they have been labelled by the ingress edge device. MPLS (Multi-Protocol Label Switching) is the label switching protocol defined by the IETF following the definition of other label switching protocols by Cisco and other manufacturers. It is similar to DiffServ in the sense that packets are processed depending on how they have been marked at the ingress point.

MPLS is primarily a traffic engineering protocol used to define fixed bandwidth pipes within IP networks or other types of networks such as ATM. This traffic engineering mechanism results in increased QoS performance.

The label is not used to establish the priority of packets as it is the case in DiffServ but to determine the next router hop. The router therefore forwards the packet to a known destination depending on the label (switching), instead of having to compute a route depending on a destination address (routing). Routing is only done once at the ingress point where the label is allocated.

A further protocol has to be implemented between routers in order to enable the exchange of label information between routers.

3.3.2.3. End-to-end QoS

End-to-end QoS is only as good as the weakest link – therefore each link needs to be able to provide the required QoS. Problems usually do not occur within networks, but rather at interconnection points. They can relate to equipment interoperability issues but also to lack of SLAs or lack of coordination of QoS policies.

This section shows how the mapping of Classes of Services (CoS) between interconnected operators can be achieved on a technical level. Later in the report it is further explored how the business models and the characteristics of the market players may also affect the provisioning of end-to-end QoS. This includes the example of an

operator with large coverage who could be reluctant to provide another operator with QoS facilities across a network.

> Mapping between Classes of Service

Mapping needs to be done between the different Classes of Service (CoS) defined within the different QoS protocols in order to enable interoperability between e.g. DiffServ-enabled or RSVP-enabled IP networks, ATM networks and MPLS-enabled networks. Such mapping requires interconnected networks to agree on a way of exchanging and presenting information at interconnect points. Each network would therefore be required to shape its traffic as agreed at the egress point before handling it over for transit or delivery.

Different fora and standardisation bodies are working on defining a mapping between the CoS of different types of networks. The setting-up of SLAs between networks is also a commercial issue that will need to be solved between interconnected networks. A few examples of how mapping is solved between certain types of networks are given below.

- Interoperability between DiffServ-enabled networks has been taken into account in the definition of the DiffServ architecture. DiffServ assumes the existence of SLAs between interconnected networks which details the traffic profile and policy criteria used.
- In UMTS networks, 3GPP has defined 4 Classes of Service that will be applied depending on applications used by end-user and which mainly describe how delay sensitive the traffic is and represent a trade-off between delay requirements and error rate requirements. These classes are conversational class, streaming class, interactive class and background class. UMTS operators are responsible for determining the appropriate mapping between the UMTS CoS defined above and the DiffServ CoS used in their networks.
- The IETF is working on interoperability between RSVP and DiffServ networks – where RSVP is mainly implemented in the edge networks and DiffServ in the core network.

> QoS management and negotiation

While network operators hold the main capabilities in providing quality of service guarantees, a number of additional players have the capabilities to indirectly affect the actual quality of service provided.

In NGN, application Service Providers will increasingly be involved in QoS issues, as they represent the main interface with the end-user for the services provided. The QoS required for a given type of service provided would therefore first be negotiated between the end-user and the ASP depending on a number of criteria such as the type of service required, the type of terminal used, the type of access line used, the credit available, etc. The ASP could thereafter be in charge of negotiating the required QoS with the various network operators involved in the transport of the data streams.

The negotiation of QoS could be done directly by the ASP either with all the network operators involved in the transport of traffic, or just a few of them provided that they have established SLAs with the remaining operators.

This could also be done via a third-party acting as a QoS broker between the ASP and the network operators.

A bandwidth broker could also be associated with each network domain and could interact with other bandwidth broker to communicate service agreements (SLAs) to enable end-to-end QoS.

3.4. Service provisioning

3.4.1. Migration to the NGN services environment

Traditionally, content providers and network operators are different commercial entities. This is due to history; the core business of telecommunication providers has always been communications services. However, in the eighties, the concept of Intelligent Network (IN) was invented in the US. From then on, telecommunications providers saw a business opportunity in offering their infrastructure to content providers. This could generate more traffic in their networks, and more turnover, as the billing of the content was also done by the telecommunications provider.

IN standards were developed to provide standardised and modular ways of creating services, which was independent of the underlying network. A basic idea was to open up service provision to 3rd parties, and to strengthen the market for content services and other services. But so far IN capabilities are mostly used internally by network operators to provide advanced telecommunications services.

In the service convergence environment of NGNs, one could also use the same types of tools and the same technologies to implement any type of services, e.g. telephony services or Web Services. A number of standards have already proven to be successful in this area and will play an important part in the NGN migration: Java, standardised APIs such as JAIN, Parlay and OSA, XML, etc.

A direct consequence of the emergence of such standards and technologies is that an NGN environment offers opportunities for third party service providers. In traditional telco networks such as PSTN and GSM networks, the services accessible by the end-users are limited by what has been implemented by the operators on their networks, most often by using proprietary development tools. User can “control” these services (like for instance by enabling or disabling call barring) but only within the scope of what has actually been allowed by the operator. In an NGN environment based on IP and the tools and technologies described above, third parties can have a greater importance and the end-users will also have more choice with regards to services and environment personalisation – provided that there are not exclusionary practices and that NGN are open to the third-party service providers. Examples of developments include:

- XML and CORBA based APIs for instance can move service development and deployment away from network operators, and into open software development communities such as Java and .NET communities.
- More and more of the intelligence and services can be moved from the current dedicated networks into third party's servers connected to the Internet. Users can therefore potentially access any service provided by an application server on the Internet. Peer-to-peer voice services using SIP for call set-up could for instance be provided by service providers without the knowledge of the carrying operators, provided that the QoS supplied is sufficient.

- Languages such as XML-based Call Processing Language (CPL) provide an easy interface for end-users enabling them to create personalised environment from their terminal, such as call redirection.

3.4.2. Perspectives on APIs and Web Services

XML based APIs, known as Web Services, can change the way Internet is being used. Software applications can be published on the net, to be identified and called from other software applications. Using this technology, basic capabilities of circuit switched and packet switched networks may also be “published” or be made available on the Internet as Web Services. This would be capabilities such as call control, messaging, mobility, billing and security. Software applications on servers connected to the Internet may access such Web Services, to create new higher level services. There are several interesting perspectives to this:

- Customised communications and content services can be offered seamless across network borders and across national borders.
- Basic network capabilities (such as those mentioned above) are made available to service providers across the Internet in a standard fashion, using web technology. This makes the community of service developers much bigger than today.
- New services may be composed by combining existing services.
- Services may be offered across access devices of different types, and adapted to each device as necessary.
- It will be much easier to customise services for individual customers
- Capabilities offered by public networks may be integrated in private applications, at a much lower price than today.

Such developments could develop into a scenario where lower level network capabilities are published as Web Services, while higher level services are developed and deployed in software applications that use these Web Services. Also, there could be changes in market roles, e.g. software vendors could become network service providers, e.g. if they start publishing telecom related Web Services.

However, there are some issues that need to be resolved for the above perspectives to be realistic:

- The developments may interfere with current telecom regulation, e.g. where regulation requires different behaviour of a specific service in different countries.
- The telecom industry must have an incentive to accept and deploy the emerging Parlay and JAIN APIs, as well as using them to publish Web Services
- A credible framework for managing access, use and billing of Web Services must be established
- Most network operators are focussed on, if not restricted to, providing either fixed, mobile or Internet services. Multi network service integration may be a challenge, because few single players have expertise or capabilities in all areas.

3.4.3. Scope of services provision

As the telecommunications networks are developing quickly towards broadband capacity to more and more end users, content provision has gained a whole new perspective, which is delivery of multimedia services through the World Wide Web. The potential for content, that end users are willing to pay for, is much greater than ever before. This also means that a growing number of different kinds of content providers could benefit from delivering paid content through the web.

The notion of telecommunications services becomes more and more blurred, as end user devices, network technologies and software applications continue to supplement each other in still new combinations. Telecommunications services are not just PSTN, ISDN, GSM, Frame Relay etc. Content, messaging and e-commerce continue to increase, and the involved technology spans fixed and mobile networks and access devices, Internet, PBX's, PCs, client/server architectures, directory services, Authentication, Authorisation and Accounting servers, firewalls, VPN technology, and more.

In this section, 'telecommunications services' includes the following generic elements:

- Communications (person to person)
- Contents (content delivery system to access device)
- Remote system integration (system to system)

These service classes can be subdivided into lots of sub classes. The term 'content delivery system' refers to anything that can generate contents in the form of voice, data or video. 'Access device' can be anything that has a man-machine interface. The term 'system' can be anything that runs software applications.

Within telecom terminology, the term 'application' is often confused with 'service'. In this section, the term 'application' strictly refers to a software application running on a server. This software application may provide a 'service' to the end user.

The two most promising technologies that are being applied to enhance service provision are:

- Application Programming Interfaces (APIs)
- Web Services.

These are closely related, and are further described in the following sections.

3.4.4. APIs for telecommunications services

APIs are predefined interfaces between software programs, which enable 3rd party programmers to develop applications on top of existing software platforms. As an example, among other system integration tools, Microsoft provides APIs for the Windows platform, thus enabling 3rd party programmers to develop applications that run under Windows. The programmer only need to include the API definitions in his source code, and then use procedure calls to access the Windows application.

APIs in telecommunications networks are historically related to the concept of Intelligent Networks (IN). The 'intelligence' in telecommunications networks can be defined as the ability to establish specific services between endpoints in a network. As an example, the ability to set-up a voice call resides in each telephone switch, which then provides the

'intelligence'. However, with the introduction of Intelligent Networks (IN) in the eighties and early nineties, it became possible to implement more advanced services in a central IN-node. This has required that a set of standards are developed for the signalling between the IN-nodes and the telephone switches (the IN terminology for these are TCAP, INAP, etc.). Part of the 'intelligence' was moved out from the telephone switches to a central node, which could run service scripts made by the network operator.

Carrier grade IN platforms have remained expensive and time consuming to handle for the creation of new services. On the other hand, network operators need to reduce costs of service development, speed up deployment of new services, and eventually to provide more content by opening up their networks towards 3rd party service and content providers. At the same time, the demand for Internet based services has made it relevant to integrate services across circuit switched networks and IP networks.

The purpose of telecommunications APIs is to enable 3rd party programmers to develop new services on top of existing networks. Standard APIs make it possible to create services that utilise all the capabilities of different networks, as opposed to development within the proprietary service creation environments associated with IN platforms. A key benefit is that services can be made across different networks using the same standard APIs. An important feature is that such APIs can pass on network functionality to 3rd parties in a safe and controlled manner, while still hiding much of the network complexity from service developers. APIs are therefore a key element for providing access to the network infrastructure capabilities and could therefore represent a key control point for network operators and/or service providers.

The most significant standard APIs for next generation services creation on telecommunications networks are those made by the Parlay Group, often referred to as the Parlay APIs, and those made by Sun Microsystems, called Java Advanced Intelligent Networks (JAIN). Both are further described below.

3.4.4.1. The Parlay APIs

Some of the main telecom equipment vendors initiated the Parlay Group in 1998 for the development of APIs for easier development of services on telecom networks. More members have joined the group, which now includes all major network equipment vendors and several major operators as well. The development efforts have been continued. The first products offering Parlay APIs have been released recently.

Several standardisation bodies have adopted the Parlay specifications, and co-operation efforts are being done. Thus parts of the Parlay APIs have been adopted by 3GPP and ETSI, which means that the 3G concept of Open Services Access (OSA) is identical to parts of Parlay. However, the Parlay APIs are aimed at all kinds of networks, while 3GPP and ETSI only focuses on 3G mobile networks and fixed NGN architectures, respectively.

Parlay API specifications are also being aligned with the JAIN specifications (see below). These developments make the APIs realistic to work with for a large number of developers within both telecom centric service providers, and web/Internet oriented service providers.

The Parlay APIs and their environment are illustrated in the figure below:

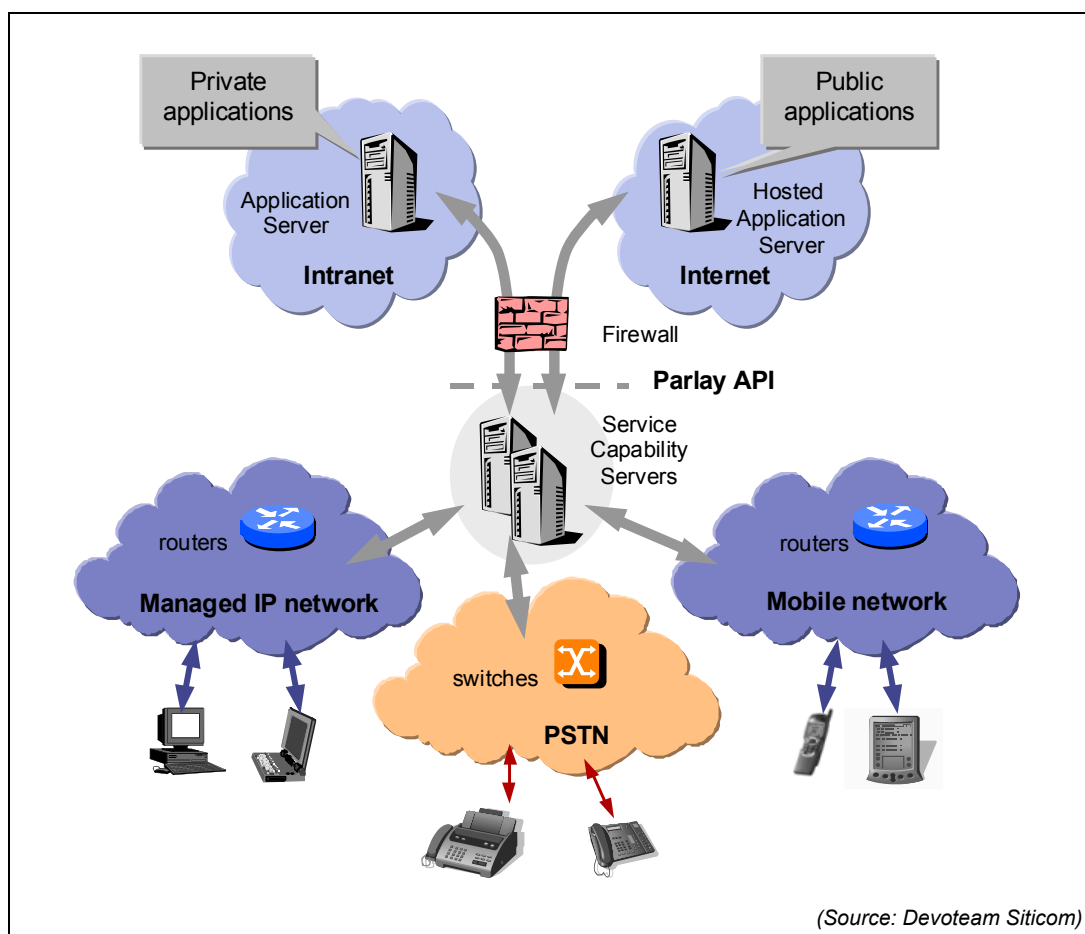


Figure 10 Environment of Parlay APIs

The Service Capability Servers (may also be referred to as Parlay gateways) provide sets of 'Service Capability Features'. The Service Capability Features makes features such as call control, messaging etc. of the network elements available for Parlay applications running on external servers, protected by a firewall. The client software may utilise Service Capability Features to implement all kinds of services across different underlying networks, without having to interface directly to each network. The Service Capability Features are provided through a framework that ensures the integrity of the underlying network. The framework provides security measures, which is a fundamental prerequisite for the success of Parlay.

The Parlay APIs are based on middleware (see below) and provide a unified view of the underlying networks towards the programmers, i.e. it saves them from worrying about specific details about call set-up, user location, messaging, charging, etc. in either fixed, mobile or IP networks. The Parlay APIs are being developed both as CORBA middleware and as Web Services (see below). The CORBA based APIs are usually implemented on UNIX based systems, while Web Services works across all software platforms and networks.

A stripped-down version called Parlay X is being made, which is aimed at web programmers not used to working with telecommunications services. It provides a simplified interface that still covers a substantial part of the functionality. Parlay X is expected to be finished Q4 2002. Despite of this, Parlay gateway vendors are already implementing it, and adjusting the interface according to standard adjustments.

An example of a service using Parlay could be a conference call service, where the customer can set-up and control the conference call by clicking hyperlinks on a PC connected to the Internet. The functionality of Parlay based services will be restricted only by the offered network facilities, while the form and design of user interfaces is entirely up to the service provider.

Some prerequisites for the success of Parlay APIs are:

- Vendors of Parlay gateways should support a wide range of network equipment such as voice switches, SMS-centres, multimedia messaging centres, 2G, 2½G, 3G network elements, VoIP network elements, softswitches, media gateways, core IP switches, billing systems, operations & support systems etc.
- Network operators should use Parlay to open up their networks towards 3rd party service providers and content providers.

The current status is that Parlay gateways are being marketed, and the first implementations have been announced. Service providers are keen to exploit this new technology; they see it as a new revenue source, and not as a replacement for existing services. An example could be to develop integrated services across web based services and Java enabled handsets. Experiments and trials are taking place this year, and commercial implementations in larger scale can be expected in 2003.

3.4.4.2. JAIN

The purpose of the JAIN APIs is identical to the purpose of the Parlay APIs: to open up telecommunications networks towards 3rd party service providers and Independent Software Vendors (ISVs). The JAIN APIs are being developed under the Java community led by Sun Microsystems. They are only applicable from Java based software platforms, which means that JAIN is basically a supplement to the Parlay APIs, rather than a competitor. As mentioned above, they are being aligned with Parlay, allowing programmers to work from either software platforms without being concerned about functional differences.

The JAIN architecture includes:

- A software component library (procedure calls that accesses the underlying networks)
- A set of development tools
- A service creation environment
- A service logic execution environment (for executing the service applications)

Down towards the network elements, the JAIN APIs includes support for the following signalling protocols:

- TCAP and INAP for IN Capability Set 2 (CS2) services
- ISUP for circuit switched call control (ISDN and PSTN)
- MAP for GSM and 3G MAP for UMTS
- MGCP for controlling calls across PSTN and IP networks
- SIP for Voice over IP gateways
- MEGACO for H.323 call control of Voice over IP calls

The JAIN APIs also support the ENUM protocol for mapping of addresses.

The service provider can make use of the following functions provided by JAIN:

- Trust and security management
- Integrity management and event notification
- Call Control
- User location & Status
- Generic user interaction
- Presence availability Management
- Payment facilities

Parlay gateways can be implemented today using JAIN, since some of the signalling protocols mentioned above are already available as software packages with JAIN APIs.

3.4.4.3. Implementation of Parlay and JAIN using middleware

In reality, the Parlay and JAIN APIs described above are implemented using middleware, which is a term for software tools that allow applications to run on multiple hardware or software platforms. Middleware handles the transfer of input/output data between the two platforms. All management of code execution and variable transfer across different platforms is transparent to the programmer, thus allowing him to focus on the core logic in the applications.

The figure below illustrates the location of middleware within a standard business application architecture (please note that in some cases middleware solutions include the business logic):

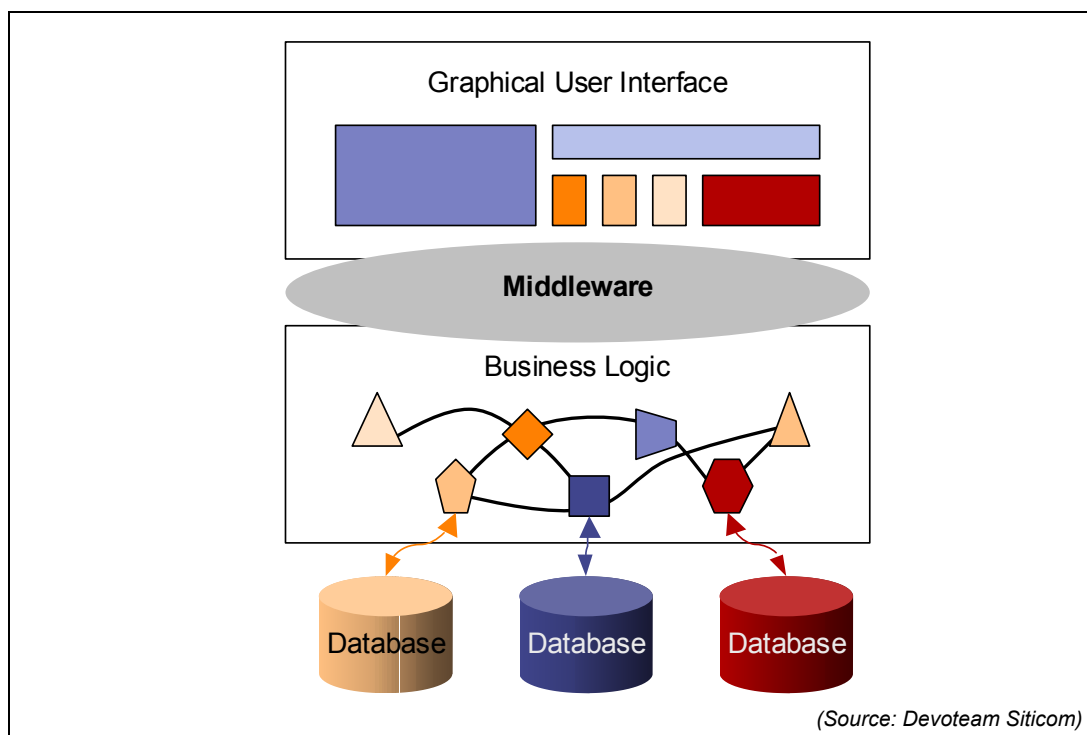


Figure 11 The location of middleware within a standard business application architecture

In the Parlay and JAIN environment, the middleware positions are placed as indicated in the drawing below:

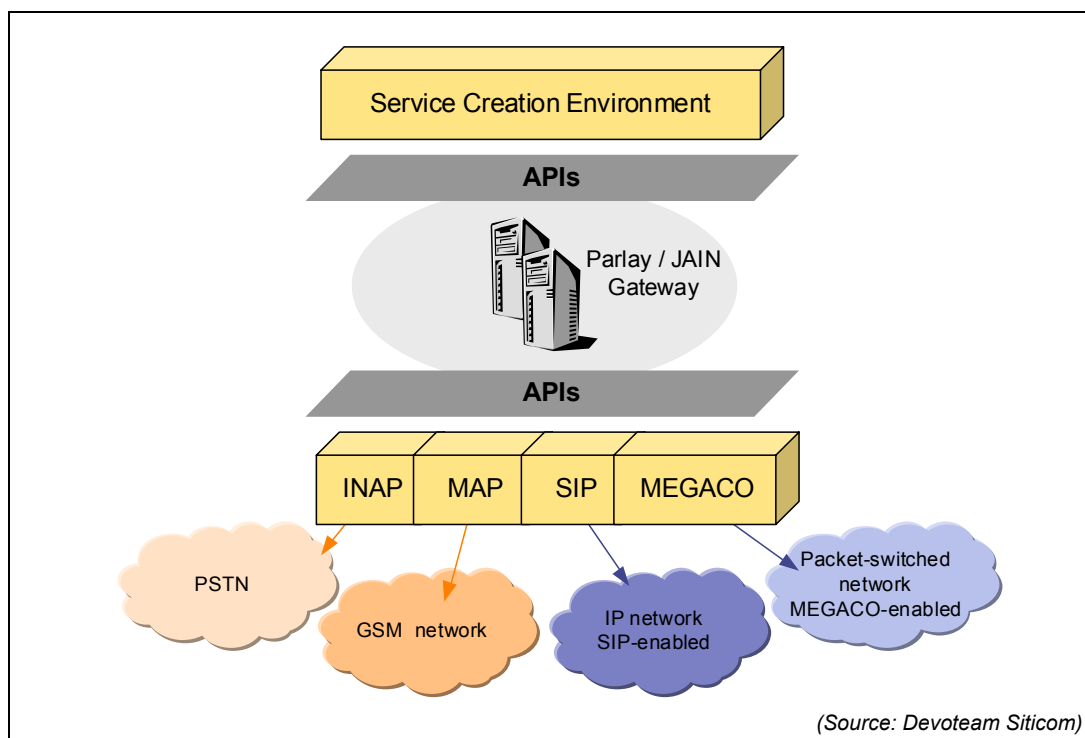


Figure 12 The middleware in a Parlay / JAIN environment

3.4.4.4. Evolution of Middleware towards Web Services

The use of Middleware has increased heavily through the last 10-15 years, mainly as part of client/server applications. Middleware has originally been defined by proprietary standards, which were not compatible. Therefore industry standards for middleware emerged. The most successful ones were the Common Object Request Broker Architecture (CORBA), defined by the Object Management Group (OMG) in the early 1990s, and Microsoft's DCOM (Distributed Component Object Model). Within the Java environment, the Remote Method Invocation (RMI) is available. Using these standards, it is possible to implement client/server systems based on software components from different vendors. The standard software interfaces make it feasible to distribute system development over multiple vendors. These standards are still commonly used. But software applications have to be written specifically to each middleware type, and it is not straightforward to convert software applications from one type of middleware to another.

As web technology came around in the 1990s, new ways of interconnecting computers appeared. The HyperText Markup Language (HTML) and HyperText Transfer Protocol (HTTP) emerged as tools that facilitated the publication of websites. HTML defines the contents of the website using commands written in characters, not binary codes, and HTTP is a simple protocol used to transfer the HTML pages from server to client.

It became clear that the principles used in HTML would be very well suited for the exchange of documents and information in general. The *eXtensible Markup Language* (XML) was defined using the same principles as HTML. Important properties of XML include that it is self-descriptive, and that it separates data from the presentation of data. These properties make XML a powerful tool to exchange information between software applications. And thus XML has become the basis for a new standard for middleware

often referred to as *Web Services*. The term 'Web Services' indicates that the XML-based middleware is capable of using standard Internet protocols such as HTTP, and thus can be used seamlessly across servers and clients on the Internet.

Parlay supports both CORBA and Web Services, and this choice can be made in each project. Operators and service providers will tend to use CORBA towards service gateways, while enterprises will use the web version. The web version runs on HTTP which is very common everywhere. Big enterprises may still use CORBA, though.

Thus we see an evolution from different, incompatible middleware solutions, to one XML based standard middleware solution, generally referred to as Web Services. All major players in the software industry accept this technology as a good solution. The main perspective is that applications can be created within different programming languages, different development environments, different operating systems etc. and still be parts of integrated systems. This is recognised across the industry as one of the biggest achievements within software architecture and open systems in the past few years.

Market forces have apparently made a steady progress towards open interface standards in the computer hardware industry as well as the software industry.

3.4.5. Web Services

Web services is a term for XML based middleware. Web Services are composed of the following building blocks:

- **Simple Object Access Protocol (SOAP)**. This is the protocol used to connect applications, e.g. convey procedure calls from an application on one server to an application on another server.
- **Web Services Description Language (WSDL)**. This is the standard for describing how to access an application (i.e. input and output parameters, functionality etc.)
- **Universal Description, Discovery and Integration (UDDI)**. This is a standard for building up databases of WSDL descriptions, making it possible to identify and access registered Web Services across the Internet.

There are at least two domains in which these standards apply:

- **Enterprise applications**

SOAP is already the preferred middleware standard for integration of software applications, e.g. from business logic to Graphical User Interface (GUI). SOAP may run over HTTP or other protocols within Local Area Networks.

- **Internet domain**

These standards may represent a whole new way of using the Internet: Dynamic integration of applications located across the Internet in a standard fashion. These standards can make it much easier than today for businesses to integrate systems as needed. Services for end users can be made more flexible and dynamic than today. The following section will focus on Web Services as Internet technology.

3.4.5.1. Applications of Web Services

With SOAP being used on HTTP, applications can be located anywhere on the net and still be able to access each other through existing firewalls etc. UDDI makes it possible to locate other Web Services that fulfil the needs in a specific situation. UDDI databases can be used at development time and in run time:

- In *development time*, developers may use UDDI to identify existing Web Services that can solve a specific task. If they find one, it will save them the time necessary to program the facility himself.
- In *run time*, an application may automatically connect to other applications using UDDI and SOAP. This run time query may be more or less intelligent, requiring more or less parameters to be specified on beforehand.

Example: A mobile application provides a map of a person's location within a city. A service provider offers local maps on the mobile device wherever you are. This is based on an application on the mobile device that is able to look for Web Services that can provide maps based on the current geographic location of the device. When in Paris, the mobile application looks for a Web Service via UDDI that fulfils the required characteristics, e.g. area to be covered, resolution of the map, colour or b/w, size of display etc. It may find a suitable map from a Web Service on a French server, since a French provider will be most likely to have updated maps. When in New York, the application may find a suitable Web Service on a US server, for the same reasons. The necessary specifications for the UDDI registration of the map service are to be agreed in relevant industry fora, e.g. OASIS (Organization for the Advancement of Structured Information Standards, an organisation that works with e-business and XML standards), OMG (Object Management Group), or other industry fora.

At a more abstract level, Web Services could be described as an evolution from web browsing by humans to web browsing by software applications. Web Services thus has the potential to turn the Internet into one big software/content infrastructure.

A long-term perspective of Web Services across Internet and telecom networks could be that the current distinction between network services, software applications, and Internet applications become more and more blurred, and eventually disappear. Everything becomes 'services' in different areas. Each 'service' is internally composed of Web Services published by network operators and other providers of Web Services. Users may not necessarily notice a difference, but development and deployment conditions have changed completely, so that new services are created and deployed in a fraction of the time necessary today.

3.4.5.2. Performance issues of Web Services

Web Services will imply generic requirements in terms of bandwidth, latency and coverage for each service. Applications must be able to adapt to the offered network service, e.g. by negotiating SLA levels. For example, if the map service from the example above is called in a place where there is little bandwidth available, the map may be downloaded in B/W instead of colour, to save time. Some of these parameters may also be resolved at protocol level. Regarding billing, the Parlay interface may provide billing data for the network operator.

There is currently a general discussion on whether performance requirements imply that the ASCII based XML protocols must be changed to binary XML. There is no consensus

yet, and one standpoint is that it could be dangerous to introduce a new XML format, since the ubiquity of XML could disappear with several incompatible versions around. With EDI, every business pair had to agree on which specifications to use. The value of ASCII XML is that there is no need to negotiate. There is a trade off between CORBA and XML: CORBA offers more control facilities of the network to the programmer, but uses also more interactions, while XML offers less control, but also fewer interactions.

3.4.5.3. Development tools for Web Services

> Two competing platforms

The two best known products for development of Web Services are Microsoft.NET and Java 2 Enterprise Edition (J2EE), which is a standard developed in an industry forum led by Sun Microsystems.

Both platforms include tools designed to allow non-Web Service specialists to build Web Services at a high abstraction level, thus hiding the more basic infrastructure from the programmer.

Both platforms use a compilation principle that was introduced by SUN Microsystems in the mid-nineties, currently well known under the name of Java. The principle is described below.

There is a market 'battle' going on between Sun and Microsoft, which has recently resulted in a Microsoft announcement that Microsoft will withdraw support for Java in their solutions. However, this does not affect the ability of the Java solution to create Web Services that can interact with Web Services created on a Microsoft platform. The two platforms can currently be perceived as equally strong in the market place. Each has its own strengths and weaknesses within different areas, but there is no clear 'winner'.

However, an important aspect is, that Web Services generated from those two platforms, or other platforms will be compatible due to the ubiquity of the SOAP, WSDL and UDDI standards.

> Use of intermediate program code

The Web Service is written in a programming language supported by the development platform, e.g. Java or C#, the new programming language used in the Microsoft.NET platform. The application is then compiled from source code to an intermediate format, which is in between source code and executable machine code. This compilation may actually be postponed until the Web Service is invoked the first time. This intermediate format is called *bytecode* in the Java environment, while it is called *Microsoft Intermediate Language* in the .NET environment.

When the application has been compiled to the intermediate format, it can be executed using an interpreter that translates it into executable machine code. This interpretation is done in a so-called runtime environment, which may be located on the local system, or on a remote system. If it is a remote system, the intermediate code is transferred via the network to this system before execution. The runtime environment is called *Java Runtime Environment (JRE)* and *Common Language Runtime (CLR)* in either environment. A basic advantage is that by sending program instructions instead of screen graphics, the bandwidth requirements can be reduced dramatically. Only the initial transfer time is depending on the network capacity, not the application itself. The Java implementation is

widely used on the World Wide Web, e.g. for animations, games and other small applications available from web pages.

3.4.5.4. Deployment of Web Services

Currently, only very few Web Services have been deployed. Microsoft's Passport.NET is the most well known example of a live Web Service. Apart from that, other implementations are more or less at an experimental stage. There are more live implementations of SOAP and WSDL in enterprise software applications, where this technology has gained full acceptance in the market.

The example of the Passport.NET Web Service is a repository hosted by Microsoft that contains user identity information. The purpose of Passport.NET is to make it possible for users to store login credentials, credit card numbers, bank account numbers and other personal details in one place, thus saving the user from typing in details again and again on different web sites. This is called *Shared Context*. However, shared context can be implemented in other ways than the Microsoft approach, such as using several independent shared context services, that each store different kinds of information for the user. One such alternative is the specifications produced by "The Liberty Alliance" who are aiming to establish the necessary means of authenticating and identifying a user in relation to e-commerce.

3.5. Interworking of addressing systems

The convergence towards Next Generation Networks requires that customers of different market players, using different network technologies, can communicate with each other and access resources on another market player's network. This requires the interworking of naming, addressing and numbering systems.

Two main standardised solutions are being proposed, lead by the IETF (Internet Engineering Task Force) and TIPHON (Telecommunications and Internet Protocol Harmonisation Over Networks) respectively. These solutions are described below, following a brief description of different naming and addressing schemes.

3.5.1. Terminology

It is important to differentiate the two concepts of "name" and "address".

- An address identifies the specific termination points of a connection and is used for routing purposes. It carries information on the network topology in order to identify the location of the resource within the network.
- A name is a combination of characters that is only used to identify end-users and does not carry any network-related information. It thus needs to be translated into an address for routing purposes. A name can be ported between service providers whereas the address associated with the name will change to reflect the change towards a different location within the network. Examples of names are Internet names such as peter.smith@ecme.com, public telephony numbers or instant messaging identities.

3.5.2. Addressing and naming schemes

3.5.2.1. IP addresses

> Description of IP addresses

Although IP addresses are not commonly used by users to reach other users or resources connected to the Internet, they represent the fundamental addressing scheme for all applications running on the Internet. The sender and recipient addresses are included into two specific fields in the IP header. Their format can differ depending on the IP protocol implemented: IPv4 or IPv6.

- IPv4 addresses are coded over 32 bits. This could provide theoretically more than 4 billion addresses, but structure constraints reduce this limit to less than 1 billion addresses. The demand for Internet connectivity has however been much higher than expected and is still expected to grow (due to emergence of 3G, home networking or the possible connection of any piece of equipment to the Internet in the future). This has led to a shortage of IPv4 addresses. Different mechanisms have been used to get round the problem, such as dynamic addressing⁸ or private addressing using NAT (Network Address Translation) or CIDR (Classless Inter Domain Routing)⁹.
- One of the enhancements provided by the IPv6 protocol lies in the huge number of possible addresses, as they are coded over 128 bits. The IPv6 address is divided into several hierarchical levels that reflect the proximity of one network to the Internet backbone.

> Management of IP addresses

The Internet Registry system is responsible for the management of IP addresses. Its role is to ensure that the address space is managed in a globally fair and consistent way in order to minimise wastage and prevent an uncontrolled growth of the Internet routing tables.

At the top of this system, the ICANN (Internet Corporation for Assigned Names and Numbers) is responsible for co-ordinating the allocation of IP addresses on an international level. It has delegated the management of IP addresses to three regional non-profit organisations: the Regional Internet Registries (RIRs).

- RIPE NCC (Réseaux IP Européens, Network Coordination Centre) responsible for Europe, Middle East, Africa and parts of Asia;
- ARIN (American Registry for Internet Numbers) responsible for America;
- APNIC (Asia Pacific Network Information Centre) responsible for Asia.

⁸Dynamic addressing consists in allocating a temporary address from a pool of available addresses to a user for as long as the Internet connection lasts. If the user closes the connection and reconnects later on, the address allocated would have changed. This mechanism enables an ISP to own less addresses than actual subscribers, considering that not all subscribers are connected at the same time.

⁹CIDR mechanism discards the original class structure of IPv4 addresses and enables a more accurate allocation of addresses to organisations, depending on the actual needs of host addresses. CIDR is the mechanism currently used for address allocation in IPv4. On top of providing further addresses, it has enabled a decrease in the growth of the Internet routing tables. This is described in IETF RFC 1519.

Within their respective zones, the RIRs are responsible for allocating IP addresses to their members (mainly network operators and ISPs). Network operators and ISPs will then allocate IP addresses to their clients.

3.5.2.2. The public telephony numbering scheme

Public telephony numbers follow the ITU-T Recommendation E.164. This recommendation specifies the international public telecommunication numbering plan and the structure of public telephony numbers, referred to as "E.164 numbers".

This numbering plan was initially an addressing scheme but has become a naming scheme due to services such as number portability and non-geographic numbers. In the PSTN environment, the routing is thus done by associating a routing number, containing the required network information, to E.164 "names".

The ITU is responsible for the co-ordination of the international numbering plan and allocates the country codes to each Member State. Each Member State, in practice the national regulatory authority, is responsible on a national level for the allocation and the usage of numbering resource. They thus guarantee the integrity and the consistency of the national numbering plan. Furthermore, coordination of naming and numbering also occurs at the regional level.

3.5.2.3. Internet names

Apart from E.164 numbers, Internet names represent a very commonly used naming scheme in an IP environment. They are based on the concept of domain name (see section 3.5.3.1, "The Domain Name System (DNS)") and are usually of the form [user@domain](#) where domain represents the user's home network or host.

The same Internet name (e.g. [peter.smith@ecme.com](#)) may be used for different services, e.g. email and SIP voice call.

3.5.2.4. Instant messaging identifiers

End-users can also be given further identities depending on the applications they use. For instance, Instant Messaging providers can allocate other IDs to users.

Instant Messaging (IM) is a popular software communication tool that enables users to communicate in real-time. Initially, it consisted only of chatting services on PCs, but now includes further services such as telephony and gaming services and can be supported by other devices such as mobile phones. An attractive feature of IM is to offer the possibility for users to create a "list of buddies". When a user runs the client application of the software on a computer, he/she can then be identified by other users, with additional information, such as his/her willingness to communicate, possible communications services available (email, phone calls, etc) and further personal indications (photo, age, etc).

The identities provided to IM subscribers can either be strings of characters or numbers and are generally given on a first-come, first-served basis. Examples of Instant Messaging services are msn Instant Messenger, which provides an ID based on an email address or AOL's ICQ (for "I seek you"), based on numbers.

> Mobile identifiers

Although mobile phone numbers are part of the E.164 numbering plan, mobile phone users can also be identified through the mean of their International Mobile Subscriber Identity (IMSI). The IMSI number is coded in the SIM (Subscriber Identity Module) card. It is unique for each specific subscriber and does not vary as long as the subscriber keeps the same SIM card.

The IMSI number has been specified in the ITU E.212 recommendation and is divided into three parts:

- The Mobile Country Code (MCC) which identifies the country of the mobile operator;
- The Mobile Network Code (MNC) which identifies the network of the mobile operator within the country specified by the MCC;
- The Mobile Subscriber Identification Number (MSIN) which identifies the subscriber within the mobile operator's network.

The number provides all the information necessary to find a subscriber's record within a specific mobile network.

3.5.3. Mapping methods

3.5.3.1. The Domain Name System (DNS)

> Purpose of the DNS

As Internet users more often use Internet names rather than IP addresses to find resources (either people or documents) on the Internet, a mapping system had to be created to map these names into addresses and thus obtain information on the user's or resource's location within the Internet. The Domain Name System (DNS) provides this function and translates domain names into IP addresses. It was first described in 1983 in IETF RFCs 1034 and 1035 and has been enhanced and modified in several RFCs since then.

> Structure of the DNS

* The inverted tree structure

The DNS can be seen as a database whose entries are domain names (such as siticom.com or cullen-international.com) associated with a number of resource records, e.g. IP addresses. Each of these records can be retrieved individually through a query to the DNS.

The database is structured into an inverted tree called "name space". The tree has a unique root node, and is then divided into top-level nodes, second-level nodes, third-level nodes, etc., each node having a label. Considering this structure, a domain name is a sequence of labels, from a node to the root, e.g. siticom.com as shown in Figure 13. The owner of a domain name can create subdomains beneath its own domain and possibly delegate the administration of a subdomain to someone else. This process leads to the creation of administration zones within the DNS system.

* The name servers

Technically, name servers are used to store the database information relating to one or more administration zones. Name servers are said to be “authoritative” for a specific zone when they effectively maintain the data associated to this zone, as opposed to “caching” when they only store data obtained from authoritative servers. For security reasons, there are usually several authoritative servers for one specific zone: one is the master and can be dynamically updated; other servers are slaves containing the replicated data from the master server. There are 13 authoritative servers at the root level spread across the globe.

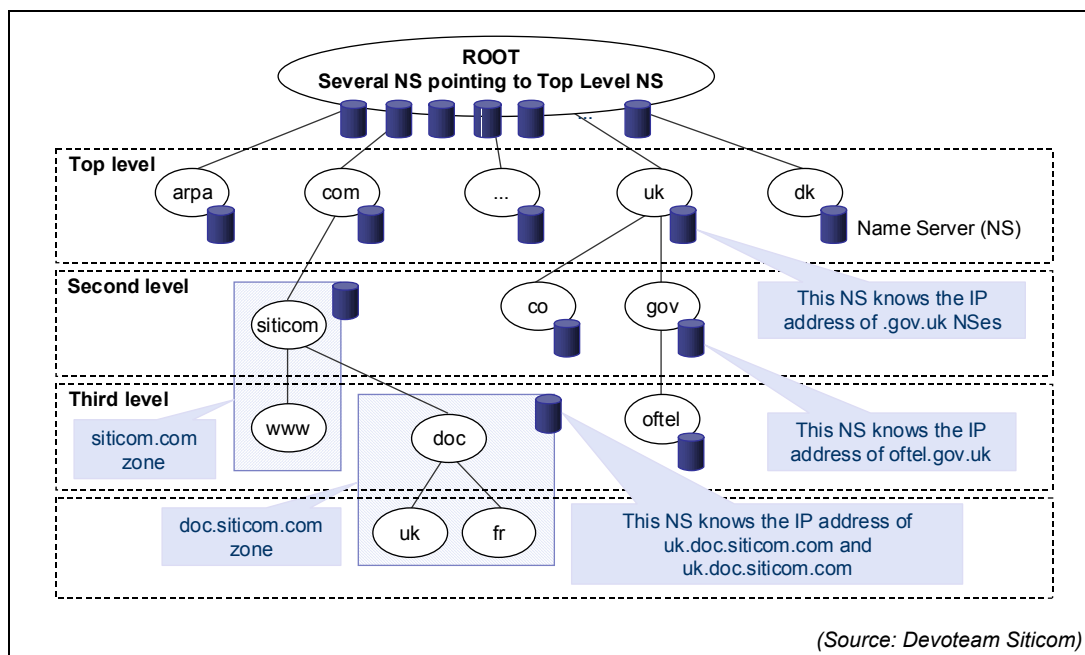


Figure 13 The DNS tree structure

> Organisation & management of the DNS

Within each administration zone of the DNS tree, three types of functions are performed by different organisations:

- The administrative management of the zone is performed by a manager or administrator. The administrator is in charge of the delegation of the sub-domain to other entities, the supervision of the name servers associated with the zone, and other administrative tasks. It is responsible for the consistency of the zone it administrates.
- The technical operation of an authoritative name server is undertaken by an organisation referred to as the “registry” for the associated administration zone. Typically, the registry will manage the resource records associated with its domain name and hold pointers to its delegated administration zones.
- The registration of domain name users into the DNS database is performed by the registrar. Further tasks undertaken by the registrar include validity/availability check, billing, etc.

The ICANN (Internet Corporation for Assigned Names and Numbers) is the administrator of the root level of the DNS. The technical administration of the root name servers is voluntarily ensured by a number of organisations such as ICANN, RIPE-NCC, U.S. Dept.

of Defense, Universities, etc. Each of these name servers points to the name servers of all the top-level name servers.

There is no registrar at the root level.

ICANN has created a number of top level domains of different types (generic, country code and others) whose management has been given to specific organisations.

Below the top-level domains, the administrator of each zone can delegate the responsibility for managing a subdomain to anyone else, in which case, the registry retains the links to the delegated subdomain.

> The name resolution process

Within the DNS context, the name resolution process is the process by which a domain name is mapped into a usable resource record, e.g. an IP address. The resolution of a name located anywhere in the DNS database can be achieved on the basis of two principles:

- all local name servers know the names and IP addresses of the root name servers;
- all name servers know the names and IP addresses of the name servers of their delegated zones.

In practical terms the name resolution is done through a number of queries to relevant server names.

> The registration process

The registration process involves the allocation of a domain name but also of an IP address and possibly other information.

- The subscription to a domain name is made through a registrar, which checks the availability of the domain name with the relevant registry and confirms the allocation of the domain name to the user.
- The allocation of the IP address is made by an ISP (possibly also being the registrar) which has been allocated a range of IP addresses by its Regional Internet Registry. The ISP then provides the Registry with all information required for populating the name server.

> Strengths & weaknesses of the DNS

The DNS is a hugely scalable system, offering the possibility to register an unlimited number of records into the tree-structured database. It can also handle high number of queries per second. Experiences from DDoS¹⁰ (Distributed Denial of Service) attacks have actually shown that root name servers could handle 50,000 queries per second whereas they usually only get 3,000 queries per second¹¹. A recent DDoS attack (22 October 2002) on the root servers has crippled nine out of the thirteen root servers for one hour but has shown that most Internet users did not notice anything and that the registry have been able to respond to the attack quickly.

¹⁰ DDoS attacks are attacks in which a number of compromised computers would attack a website, a company's e-mail server or a root server in order to force it to shut down and therefore deny access and service to authorised users.

¹¹ Source: Nominum

The selection of the registries is obviously very important for the functioning and reliability of the system. The most important registries (root zone and top level domains) must provide sufficient guarantees in terms of quality, efficiency, reliability and accessibility. For instance, they must provide guarantees in terms of processing power, throughput, interconnection capabilities, resiliency, etc. They must also commit on the reliability and accuracy of the information provided in their database.

At the top levels, several name servers maintain replicated information of the same administration zone: data is automatically replicated from the master server to slave servers. However, only the master server is dynamically updated, and this can obviously create a single point of failure. On the lower levels of the hierarchy, the usage of master and slave servers is strongly encouraged, but the reliability of these name servers is obviously not guaranteed.

There have been security concerns relating to the first specifications of the DNS protocol ("spoof" attacks were possible). Enhancements have been brought to the protocols, such as DNSSEC (DNS Security Protocol from the IETF) to remedy these gaps, but other concerns remain, like DoS¹² (Denial of Service) attacks or other operational considerations.

> Current status of the DNS

The DNS is a key element of the Internet and it will also have a key role to play in the migration from IPv4 to IPv6 as the name servers need to be upgraded to be able to handle both IPv4 and IPv6 addresses.

3.5.3.2. ENUM (tElephone NUmber Mapping) protocol

> Description of ENUM

The ENUM protocol has been described by the IETF ENUM Working group in 2000 in RFC 2916 called "E.164 number and DNS". The purpose of ENUM is to use the already existing DNS and expand its functionality to provide a mapping between E.164 numbers and a number of other addresses/names such as email address, fax number, mobile number and website address. These addresses can then be used to contact a resource associated with that E.164 number.

As part of the DNS, the ENUM system follows the same principles as DNS in terms of organisation and processes.

The ENUM protocol describes:

- how E.164 numbers can be mapped into domain names and inserted into the DNS;
- how E.164 numbers can then be mapped into "preferred" identifiers (email addresses, voice over IP, SIP addresses, voice mail servers, fax machines, etc).

The IETF also makes recommendations relating to security issues and the organisation of the root ENUM zone within the DNS.

¹²DoS attacks are attacks in which an authorised user is denied access to resources he should have access to.

> Structure of the ENUM zone within the DNS

The structure of the ENUM zone within the DNS reverse tree is derived from the mapping of E.164 numbers into domain names, as described by the IETF recommendation. This is done as explained below.

- The starting point is the E.164 number written in its full form, including the country code, e.g. +44 20 7960 4040.
- All characters except digits are removed, and dots are inserted between digits, e.g. 4.4.2.0.7.9.6.0.4.0.4.0.
- In order to respect the hierarchy levels of the DNS, the domain name is derived from the reverse order of the above string, completed with the ENUM root domain name, e.g. 0.4.0.4.0.6.9.7.0.2.4.4.[root].

In terms of administration zones, in order to protect the integrity of E.164 numbers, the IETF and the ITU have recommended a tiered architecture for the ENUM zone, divided into 3 tiers.

- Tier 0 corresponds to the ENUM root zone within the DNS. The tier 0 entity maintains the name servers for the ENUM root zone. It contains the authoritative name server records for the domain names corresponding to the country codes or portions of it of the E.164 numbering scheme.
- Tier 1 corresponds to the country code zone or portions of the country code zone. It can thus include several digits of the domain name, e.g. .4.4 for the UK, .3.3 for France or .3.3.1, .3.3.2, etc. for France if it is based on portions of the country code. It contains the authoritative name server records for the domain names corresponding to national zones.
- Tier 2 corresponds to the national zone and includes the remaining digits of the domain name. It contains the resource records associated with E.164 numbers.

> Implementation of the ENUM protocol

There are some divergent market views about the implementation of the ENUM protocol and the involvement of the ITU. Some argue that the integrity of the E.164 numbering scheme is at stake in this protocol, as poor management of the ENUM zone in the DNS system could lead to the emergence of inconsistencies within the E.164 numbering plan (e.g. emergence of new country codes). Others consider that, although ENUM involve E.164 numbers, it is in practice a separate mapping system providing a means of pointing at E.164 numbers and that therefore the underlying E.164 numbers and hence their integrity are not affected by this additional layer of mapping.

Practically however, cooperation between the IETF and the ITU on ENUM-related issues is taking place and the ITU has been involved in the implementation process.

* Tier 0: ENUM root level

The administration of the root zone is made on a global level. The IETF and IAB (Internet Architecture Board) have agreed that the ENUM root domain would be .e164.arpa under the .arpa domain (arpa stands for Address and Routing Parameter Area). At the time of writing, this decision has not been confirmed and agreed and is still subject to discussion with the ITU.

In order to enable the implementation of ENUM trials in different countries, an agreement has been reached on the fact that the registry on the root level would be operated by RIPE NCC registrar and that the ITU would be the registrar at the root level. ITU will thus receive requests from Member States for the administration of tier 1 zone(s). The ITU will for instance ensure that the Member States are allocated the proper country codes. This situation is however still not definitive.

* Tier 1: country code zones

The administration of tier 1 is left to the ITU Member States. The Member States will thus be responsible for the selection (or the operation) of a reliable registry. In line with the protection of the E.164 numbering plan integrity, it is likely that the registry will be either operated or under the control of the National Regulatory Authority

* Tier 2: national zones

As for tier 1, the administration of tier 2 is left to the Member States. As the resource records will be stored at this tier, it will involve the interaction of a greater number of players: not only the administrator, the registry and the registrar but also the Application Service Providers (ASPs) providing the services associated with the resource records, the Telephone Service Provider, etc.

> Implementation issues and alternatives

A number of options are left open to the choice of the Member States for the implementation of ENUM within their country. The drivers for the choice between options mainly include the protection of customer data, the promotion of competition (that includes the prevention of control points harmful to the competition) and the protection of the E.164 numbering scheme.

The selection of the various entities within the ENUM architecture must ensure that the system will be scalable, robust and secure at all the hierarchical levels. As the ENUM system inherits the limitations of the DNS, and thus also security issues, the IETF recommends the use of DNSSEC within each administration zone. The possible options left to the Member States' decision include:

- Decision on the number of tier 1 registries (single or multiple) and their identification.
- Delegation of names from the tier 1 registry(ies): individual numbers or blocks of numbers.
- Decision on the number of tier 2 registries and registrars and their identification: public telecommunications operator, ENUM service provider, etc.
- Validation and authentication of ENUM requests and records.
- Telephone Service Provider records, regarding how service providers will include specific information on e.g. routing for their exclusive use (and not for end-users' usage) in the ENUM records.
- Relationships between different entities: user, registrar, registry, telephone service provider, number portability administrator, etc.

A number of Member States have already or are in the process of launching workshops and trials to assess their choice in ENUM implementation and identify further issues.

Beyond the decisions relating to the organisation of the ENUM system and interactions of the different entities, there are a number of further issues that need to be addressed, such as:

- Abuse of data stored for spamming¹³ and slamming¹⁴ purposes and other privacy issues;
- Handling of shared E.164 numbers, that is numbers being shared by a community, e.g. a household;
- Handling of numbers which are not registered with the ENUM system;
- Initiation criteria of an ENUM query;
- Incentive for users to update their records.

Moreover, the implementation of ENUM should guarantee the continuity of existing services such as number portability, Carrier Pre-Selection (CPS), Calling Line Identification (CLI), etc.

> The name resolution process within ENUM

The ENUM protocol offers the possibility to select the preferred method of communication depending on the capabilities and characteristics of the calling party such as traditional PSTN voice call, VoIP, email, etc. Each record is associated with order and preference values, which allow flexible communication strategies.¹⁵

3.5.3.3. TIPHON

> Scope of TIPHON

In their third release, TIPHON provides a procedure for “determining IP addresses for routing packets on interconnected IP networks that support public telephony”.

Their approach is different from the IETF ENUM approach in that:

- With ENUM, the IETF provides an interworking solution between the PSTN and the Internet by associating E.164 numbers with Internet names that can be used for telephony over the Internet.
- TIPHON aims at supporting the public telephony service on IP technology. As a result, interworking between the PSTN and the Internet is possible only if the Internet telephony user has also been allocated an E.164 number. In this situation, the Internet name is not sufficient to provide public telephony services, although it can still be used on its own for VoIP services. At a certain stage though, the E.164 number has to be mapped into an Internet name that will be resolved by the Internet Telephony provider.

¹³ Spamming is the principle of using users' records (email address, postal address, phone number) to send them unsolicited emails, mails or make unsolicited calls.

¹⁴ Slamming is the principle of switching users to another service providers without their knowledge or authorisation.

¹⁵ The regulatory implications of convergence of numbering, naming and addressing are the subject of a separate study by Political Intelligence under the same title.

Within this scope, TIPHON produces standards in order to enable public telephony and other advanced services, such as multimedia services, over IP technology and thus enable telecom operators to migrate their circuit-switched network into an IP-based packet-switched network. The control of these services still remains within the network operator, as opposed to the ENUM model where the service is selected by the users using application protocols. The TIPHON approach implies that public telephony over IP is provided by a specific service provider to its customers and that communication with the customers of another service provider requires interconnection between the service providers.

As an example, the case of an address resolution between two SIP telephony users is described below:

In the ENUM approach:

- The called user does not need to have an E.164 number associated with his IP phone, although he could.
- The calling user dials an E.164 number, which is not necessarily the number of the IP phone but which is associated with a number of identifiers in the ENUM system, including the IP phone identifier.
- The calling party initiates an ENUM query, as it would do with a usual DNS look-up.
- Based on the preferences entered into the ENUM by the called user and the capabilities of the calling party, the SIP address is selected as the best option for the communication set-up. It could also have been a mobile phone number, a fixed legacy phone number, etc.

In this approach, the address resolution is performed by the calling party, without control from the transporting network.

In the TIPHON approach:

- The called user needs to have an E.164 number associated with his IP phone. This number has been provided by a public telephony service provider.
- The calling party dials the E.164 number associated with the SIP phone.
- The address resolution is performed by the operators of transporting network(s), with a step-by-step approach (e.g. network A will first resolve that the call needs to transit via network B, then network B will resolve that the call needs to transit via network C, etc). For the call to be set-up, there needs to be service level interconnection between the operators of the different transporting networks.

In this approach, the address resolution is performed by, and remains within the control of the network operators. In the US for instance, Telcordia performs the address resolution on behalf of the Bell Operating Companies.

> Resolution types

In a TIPHON compliant system, three types of resolution may be used in a public telephony service using E.164 numbers as identifiers over SCN or IP technology.

- A search resolution.

The purpose of the search resolution is to obtain a usable E.164 number from what has effectively been entered by the user (a local number without country code or

national code, a voice signal, etc.). The search resolution is performed only once by the calling party.

- A service resolution.

The service resolution is aimed at retrieving information on services associated with the E.164 number, e.g. number portability, personal numbering, non-geographic services, etc. It resolves the E.164 number into a destination network name and thus provides information about how the call should be completed.

When needed, this resolution is performed by a control element within the traversed network(s).

For instance, in the case of an E.164 number associated with a SIP Internet telephony service, the service resolution will consist in the provision of a URI, which can then be used with the E.164 number to produce a SIP address <E.164_number>@domain.

- A routing resolution.

The routing resolution is aimed at obtaining routing information towards the next hop and is performed by call control elements with the transportation network(s).

3.5.3.4. Other methods

The implementation of a standardised ENUM architecture within the DNS will of course not prevent the development of alternative ENUM-like structure operated by individual companies, which would thus appear under other root domain names (e.g. e164.com, enum.org). Some players favour these competitive developments whereas others think it may endanger the consistency and integrity of the numbering system.

Some service providers have already developed proprietary mapping methods enabling them to provide ENUM-like services to their customers.

> Example: Microsoft msn Messenger

Microsoft offers an Instant Messaging online service, msn Messenger, supported by PC and mobile devices with Internet access. Depending on the information provided by their “buddies” to the msn Messenger server, users can chat, make voice calls, video conferences, send SMS, emails, files, etc. Addressing and naming information are stored on a Microsoft remote server and depending on the service required by the user (SMS, voice call, etc), the required mapping is completed.

Privacy is guaranteed through a number of mechanisms:

- Personal information such as phone numbers is only made available to the user’s buddies.
- Users can see who added them to their buddies’ list.
- Users can create a “block list” of the users whom they don’t want to receive messages from or provide status information to.

> Interworking between Instant Messaging services

Interworking between Instant Messaging services would require that an IM service provider has access to its competitor server to retrieve real-time information on the status of their customers and map their customers' IM IDs with other naming and addressing information.

Interworking is not currently being provided by the players (although some companies have sought regulatory action to ensure that competing IM services can interconnect).

The IETF is also working on standards to enable communications between different IM systems. The IMPP (Instant Messaging and Presence Protocol) working group is working on common service definitions for a minimum set of messaging and presence functionality with security functions appropriate for end-to-end security. The resulting specification is called CPIM for Common Presence and Instance Messaging. The SIMPLE (SIP for Instant Messaging and Presence Leveraging Extensions) working group is working on the application of SIP (Session Initiation Protocol) to IM services, following the CPIM specifications. The PRIM (PResence and Instant Messaging) working group is building protocols for presence and instant messaging compliant with CPIM specifications. The APEX (APplication EXchange) working group is also involved in standardisation for presence and Instant Messaging.

3.6. Interworking of signalling systems

3.6.1. Introduction and perspectives

One of the big challenges of convergence is to enable circuit-switched services, mainly telephony, to be migrated onto packet-switched technology. Signalling (in particular Signalling System 7) plays a major role in PSTN voice networks, as it not only ensures the initiation and termination of voice communications, but also the provision of further advanced services in the case of Intelligent Networks.

A number of signalling protocols has been developed by different standardisation bodies and consortia in order to enable voice and multimedia services on packet-switched networks. Although they are regularly being upgraded in order to overcome the limitations of their previous releases, each of them has been developed within a specific context (telecoms, Internet, etc) and for this reason does not fit all technologies or strategies equally well.

The BICC (Bearer Independent Call Control) protocol has been created by the ITU and is biased towards ATM backbones. This protocol suits for instance operators who want to carry telephony services originated and terminated on the PSTN on a packet-based network ("trunking") but still provide PSTN-like quality of service.

For many reasons – some of which are given in the table below – the IETF SIP protocol will most likely be the most successful protocol for providing voice and multimedia services over IP. SIP has been chosen by the Third Generation Partnership Project (3GPP) as the call control protocol in 3G networks. However, the alternative protocol from the ITU, H.323, was released before SIP and has already been widely implemented and deployed, thus slowing down the deployment of SIP.

	SIP	H.323
Standardisation body	IETF, Internet background	ITU, telecom background (derived from ISDN protocols)
Architecture	Modular: does only signalling and integrates easily with other Internet protocols fulfilling other functions such as capability negotiation, registration	Monolithic: defines all aspects relating to call set-up, including capability negotiation, registration, etc.
Call control devices	Proxy / redirect server	Gatekeeper
Encoding	Text	Binary

Both SIP and H.323 protocols have a distributed architecture, which means that both protocols rely on the intelligence being distributed between the end-points and the call-control devices. For various reasons detailed later in this section, some network operators could prefer a centralised approach where relatively dumb end-points are controlled by intelligent devices in the network. Centralised architectures are usually associated with the MGCP, MEGACO or H.248 protocols for media gateway control in which centralised gateway controllers (also called call agents or softswitches) take charge of call control, using one of the signalling protocols described above. This approach is favoured by telecom operators, as it is very close to the centralised approach they are used to in circuit-switched networks.

Based on these considerations, networks will remain highly heterogeneous in the foreseeable future, using different signalling protocols and a different approach in terms of architecture (distributed or centralised). This requires the use of interworking mechanisms between the different signalling protocols involved, including the PSTN signalling system, SS7.

This section describes the different signalling protocols that have been developed over different types of technologies and the steps towards the interworking of these protocols.

3.6.2. Signalling protocols for call set-up

3.6.2.1. SS7, the signalling system in PSTN

The Intelligent Network (IN) architecture implemented in the PSTN has enabled the introduction of computing into telephony networks, and therefore the introduction of advanced services.

In this architecture, the signalling is transmitted “out-of-band” as opposed to “in-band”, which means that signalling and voice are transmitted over two physically separated channels. The transport network only carries voice traffic whereas the overlay control network carries the signalling traffic. This enables simpler and more efficient network management and the provisioning of further advanced services, such as interacting with databases for non-geographic numbers and billing.

SS7 (Signalling System 7), defined by the ITU, is the signalling system used in the IN architecture and has been

The SS7 network is made of packet switches, called Signal Transfer Points (STP), that route the SS7 messages over the network, and of databases, called Service Control Points (SCP), containing information used in services such as non-geographic numbers

and number portability. The SS7 network is linked to the SCN through Service Switching Points (SSP).

The SS7 network can be divided into several layers, with a number of protocols performing the functions of the application-oriented layers of the OSI model (higher layers). Some of these protocols are described below.

- ISUP (ISDN User Part) protocol is used for all types of calls and performs the functions related to the management of circuits, that is mainly call set-up, call management and call release.
- TCAP (Transaction Capabilities Application Part) protocol is used for exchanging other information than the one related to circuits between applications. It is used for instance for retrieving the routing number associated with a non-geographic number and authenticating a PIN (Personal Identification Number) of a mobile subscriber. In the case of a mobile network, TCAP messages carry the MAP messages (Mobile Application Part), used for instance in roaming.

3.6.2.2. Bearer Independent Call Control protocol (BICC)

In 1999, the ITU started working on the Bearer Independent Call Control (BICC) protocol. The purpose of this work was to define a protocol for the communication between call servers, independently of any specific bearer and thus to enable operators to migrate telephony services from SCN to packet-based networks.

In order to enable a smooth migration between the networks in terms of services, delay and cost, BICC has been inspired from the signalling protocol used in SCN: ISUP.

The first version of BICC (Capability Set 1), issued in 2000, was principally covering the control of calls transiting on ATM networks. The second Capability Set (CS2) was completed in July 2001 and supports call control on IP networks and interworking with a number of other signalling protocols (ISUP, INAP (Intelligent Network Application Protocol), H.323). Interworking with SIP has been postponed to the following version of BICC.

In practical terms, however, BICC is biased toward ATM networks, and it is most likely that BICC will be limited to implementation on ATM networks whereas other protocols more focused on IP will be implemented on IP networks.

3.6.2.3. H.323 standard

> Background

In 1996, several software providers and manufacturers were already offering Voice over IP products (VoIP) and Netscape took an early lead in market acceptance of its standard. Its competitors, Microsoft and Intel therefore took part in various standardisation working groups and influenced the work of the ITU and the IETF. Following this, in 1996 the ITU proposed the recommendation H.323 for "Visual telephone systems and equipment for local area networks which provide a non-guaranteed quality of service", whose signalling concepts are based on ISDN access signalling and therefore on circuit-switched signalling.

H.323 standard describes the handling of multimedia sessions involving telephony in the case of peer-to-peer connections between intelligent end-points (distributed architecture),

and includes mechanisms for call routing, call signalling, capabilities negotiation, media control, etc.

It serves as an “umbrella protocol” and involves several other protocols relating to different aspects of telephony and multimedia applications, as for instance:

- H.225 is for call signalling and media stream packetisation for packet-based multimedia communications systems
- H.245 is for the control of communications between visual telephone systems and terminal equipment (media channel set-up, capabilities negotiation, flow control, etc)
- A number of standards on audio codecs in the G-series, including for example G.711
- A number of standards on video codecs in the H-series
- Other optional standards (e.g. for confidentiality and authentication)

The Registration Administration Status (RAS) protocol is optional. It is used for the registration of a terminal with gatekeepers.

> Network elements

The network elements involved in H.323 signalling include:

- Endpoints – Endpoints can either be terminals or gateways interconnecting PSTN and IP networks and performing protocol and media conversion.
- Gatekeepers – Gatekeepers most often manage the communications of H.323 terminals or gateways within an administration zone: they enable address translation, terminal registration, admission control, etc.
- Multipoint Control Units (MCUs) – MCUs are used as central servers in the case of multipoint conferences.

> Call set-up in an H.323 network

The communication in an H.323 network can be broken down into the following stages:

1. Signalling between the calling endpoint and the gatekeeper to register and obtain admission to the network.

The call starts either from an H.323 terminal. The terminal contacts the gatekeeper using the RAS protocol (registration and admission protocol), requesting permission to establish a link with the called party.

If permission is granted, the gatekeeper sends the call signalling channel address of the called party to the calling party.

If the call starts from a PSTN phone, the call goes through a gateway that authenticates the user. The gateway then contacts the gatekeeper. If the called party is a PSTN phone, the gatekeeper sends the address of the relevant gateway to the calling party (gateway or H.323 phone).

2. Signalling between the calling and called endpoints to establish the call. This signalling may go either direct or via a gateway (gateways) if relevant.

The calling and called parties use H.225 protocol during this phase.

3. Establishment of the media control channel using the in-band H.245 protocol. This signalling may go either direct or via a gateway (gateways) if relevant.

During this phase, the features are negotiated (such as codecs, etc).

4. The media communication using the same transport addresses as the media control channel.

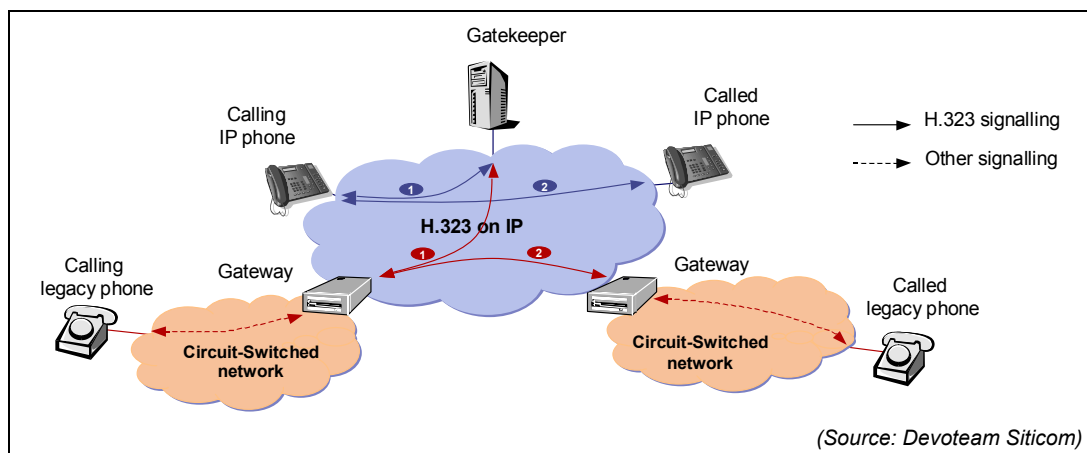


Figure 14 Elements of an H.323 network

> Adoption of the H.323 standard

H.323 standard was welcomed in 1996 by industry, which started developing and implementing H.323 solutions, as H.323 was seen as the best signalling standard for developing telephony applications over IP networks.

However, the standard also presents several drawbacks, in particular the fact that it was derived from ISDN signalling. It is not especially suited to IP networks and it involves a great number of protocols which makes its implementation heavy and expensive.

3.6.2.4. SIP, the Session Initiation Protocol

> Background

In March 1999, the IETF proposed their own signalling standard (in RFC 2543), called Session Initiation Protocol (SIP). SIP is an application-layer signalling protocol for initiating, modifying and terminating media sessions, including Internet telephone calls and multimedia conferences. As H.323, it is based on a distributed architecture.

However, as opposed to H.323, SIP is largely inspired by the Internet in order to facilitate both the integration with other protocols such as HTTP, SAP (Session Announcement Protocol), SDP (Session Description Protocol), RTSP (Real Time Streaming Protocol), but also the development of further services, which could use SIP. As examples of its Internet background, SIP is a text-based HTML-style signalling protocol, whereas H.323 is based on binary coding. SIP has also incorporated HTTP protocol and the DNS system and uses an email style addressing. Its implementation on an IP network is thus much simpler and lighter than in the case of the competing H.323 standard.

SIP supports redirection services. This enables personal mobility, that is the ability for a user to use different terminals on different locations to access subscribed services.

> Network elements

SIP is based on a client / server architecture, whose main components are described below.

- The SIP User Agent, which includes the User Agent Client (UAC) that sends SIP messages and the User Agent Server that receive SIP messages;
- Proxy servers that pass requests and responses to and from other servers;
- Redirect servers that accept a SIP request, map the address into zero or more new addresses and return these addresses to the client or proxy;
- Location servers that are used by proxy and redirect servers to obtain information on alternative addresses to reach the called party.
- Registrar servers that accept registration requests.

> Functions performed by SIP

SIP performs a number of tasks during a session between two parties:

- User location: determination of the end system to be used for communication
- User capabilities: determination of the media and media parameters to be used. This enables the calling party to reach terminals with different capabilities: SIP phone, mobile phone, etc.
- User availability: determination of the willingness of the called party to engage in communications
- Call set-up: "ringing", establishment of call parameters at both called and calling party
- Call handling: including transfer and termination of calls.

Other aspects of the session are handled by other IETF protocols with which SIP integrates very easily, as for instance SDP for the description of capabilities, DNS for service location and URLs for addressing.

> Examples of SIP operations

A User Agent Client can let a SIP server (proxy or redirect server) or a registrar server know at which addresses it can be reached by sending a registration message. If a user is logged on at several locations, then several different addresses will be registered for the same SIP address.

A calling party that wants to make a call will send a request for a called party to join a call. In most cases, this request is sent to the locally configured proxy server (step 1 in the figure below).

This request can then be redirected by a number of redirect servers or sent through a series of proxy servers. The figure below shows a simplified case where the proxy server contacts the location server to obtain a more precise location of the called party (step 2), and then forwards the request to the address(es) returned by the location service (step 3). If the called party is available, its user agent will alert it and return a success indication to the proxy server (step 4), which will be passed back to the calling party (step 5). At the receipt of the indication, the calling party will send an acknowledgement to the called party, not necessarily through the proxy (step 6).

The request includes a session description in SDP format, that is information on the media to be sent. If the called party accepts the call, it will also include a session description in its response, depending on what has been required by the calling party and the called party's capabilities.

The media session can then be established directly between the two parties using another route than the SIP packets (step 7).

The proxy servers that are traversed by the signalling messages must add their identity in a record-route field contained in messages so that all signalling messages can take the same route.

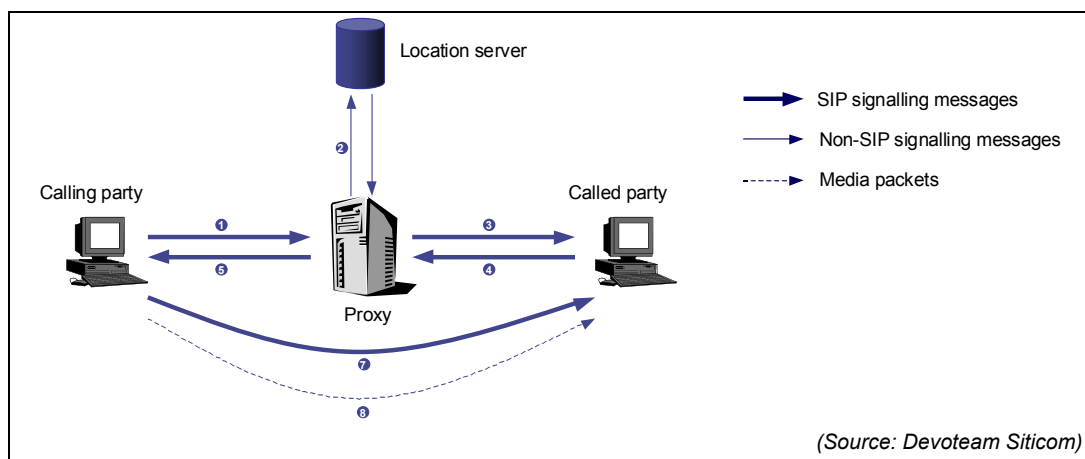


Figure 15 Example of SIP scenario

> Adoption of SIP standard

Despite the advantages offered by SIP in terms of its simplicity and easy integration with other Internet protocols, the standard has been slowed by the fact that the industry had already largely invested in the competing H.323 standard.

The advantages of SIP have nevertheless started to be taken seriously into consideration from 1999. Its simplicity suggests that it is the best alternative in a number of situations as for instance:

- Signalling protocol between terminals, used in conjunction with another simple protocol for the centralised control of gateways: MGCP / MEGACO / H.248 (see below).
- Implementation of a "light" signalling protocol in mobile handsets for next generation mobile networks based on IP.

SIP products and services are thus being developed and implemented in a number of contexts:

- 3GPP and 3GPP2 have selected SIP for multimedia call control in 3G networks;
- Windows XP, Microsoft's latest operating system, incorporates SIP in its Windows Messenger capabilities;
- A number of operators are either deploying, carrying out trials or implementing equipment compatible with SIP.

3.6.3. Signalling for gateway control

H.323 and SIP described above are based on a distributed architecture, meaning that both protocols rely on the intelligence (service provisioning, billing, call handling, etc) being distributed between the end-points and the call-control devices (phones, gateways, media servers, etc). This distributed implementation offers a flexible architecture, where for instance voice and multimedia applications can be treated like any other IP application. The service can be set-up between the end-user and the service provider without the knowledge of the transporting operator. The major drawback of this model is the cost of the intelligent handsets.

A centralised architecture relies on an intelligent network where call and multimedia sessions are managed centrally, as it has been the case in legacy PSTN network. The protocols introduced in this section enable the introduction of a centralised management in a SIP or H.323 network. This architecture is most likely to be implemented by incumbents and competitive voice carriers, which are more used to the traditional centralised PSTN approach than to the distributed Internet approach.

MGCP (Media Gateway Control Protocol), MEGACO (Media Gateway Control) and H.248 recommendation are three protocols for the control of media gateways. MGCP was first developed by a private consortium to be completed by the IETF MEGACO working group into RFC 2885 on Media Gateway Control. The IETF has since worked together with the ITU to publish the H.248 recommendation.

In the centralised architecture defined in these protocols, the role of endpoints and gateways is limited to basic functions, such as coding and media translation, whereas the intelligence is centralised in pieces of equipment called call agents, media gateway controllers or softswitches. MGCP, MEGACO and H.2448 protocols only describe the signalling protocol between the gateways and the call agent, and other signalling protocols such as SIP and H.323 still need to be used for call and multimedia session signalling.

The figure below shows the signalling protocols that can be implemented in an MGCP / MEGACO / H.248 architecture and provides an idea of the functioning of this architecture in different types of scenarios:

- If a call remains with the packet-based network in the case of an IP-phone to IP-phone call (within the dark blue cloud in the figure), the signalling will use usual SIP or H.323 standards.
- In other cases (call involving SCN terminal and IP-phone or call between two SCN terminals transiting through a packet-based network), then call agents and gateway control protocols will be involved.

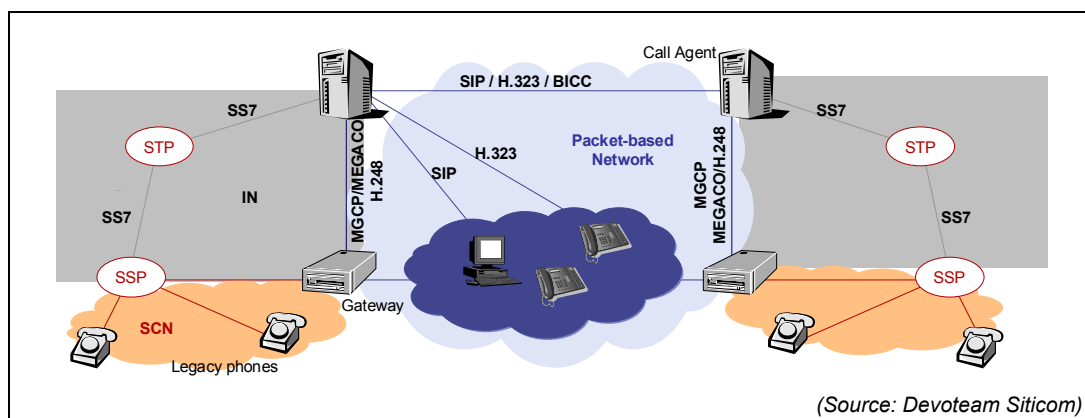


Figure 16 MGCP / MEGACO / H.248 architecture

3.6.4. Interworking between signalling protocols

3.6.4.1. Meta protocols: the TIPHON project

There is a strong need for interworking between the different signalling standards because different standards have already been implemented in different networks and this will remain the case in the foreseeable future. Work is currently being done in order to allow interworking. Translation mechanisms between these protocols are however hard and complex to define considering the differences between the protocols (due to their background) and the level of details of some of them (H.323 is specified on more than 1,000 pages). In the short term however, translation can be made between packet-switched signalling protocols using two successive translation mechanisms between packet and circuit switched signalling protocols. For instance, for translation between the packet-switched protocol PS1 and the packet switched protocol PS2, operators can use a first translation mechanism to translate PS1 into a circuit-switched signalling and then another mechanism between the circuit-switched mechanism and PS2.

TIPHON is an example of work being done to enable interworking between protocols. TIPHON develops “generalised communications protocols” (that is “meta-protocols”) to support voice over IP services, based on public telephony in a first stage (that is telephony using E.164 numbers, see section 3.5.2.2, “The public telephony numbering scheme”). The meta-protocols are then mapped into actual protocols such as SIP, H.323 and H.248.

The current release of TIPHON, release 3, addresses for instance SIP and H.323 networks interworking. The following releases will address further issues such as VoIP support through NAT and firewalls.

3.6.4.2. Sigtran, protocols for the transport of signalling

Sigtran (Signalling Transport) is a working group of the IETF currently working on the issues relating to the transport of packet-based PSTN signalling over IP networks.

The work being done by Sigtran aims at providing interworking of PSTN and IP networks by enabling the transport of PSTN signalling in IP networks, typically in the case of voice over IP involving both IP and PSTN networks.

The work of the group is limited to the transport of signalling between gateways (e.g. between the signalling gateway and the media gateway controller), so that the media

gateway controller can then allocate network resources based on the requirements and on the network's local policy. It therefore needs to take into account functional and performance requirements of PSTN signalling.

The goals of Sigtran are to define the performance requirements necessary to support PSTN signalling over IP and to identify a method of encapsulation of different signalling protocols. This method needs to take into account the fact that different protocols will be carried, and that some components of these protocols may have been translated or terminated at the signalling gateway. Sigtran does not aim at defining translation mechanisms, neither further QoS nor security protocols.

3.7. Roaming and mobility

3.7.1. Introduction

Roaming is a facility that enables a subscriber to use the infrastructure of another operator when he/she is located beyond the coverage of his/her home network. Roaming requires the setting up of commercial agreements between the involved operators.

International roaming has been supported in mobile networks standard and the GSM association has published guidelines and reference documents on technical and business aspects of roaming between mobile operators. With the advent of 2.5G and 3G mobile networks, the GSM association has widened the scope of its responsibility and produced further documents for roaming in GPRS networks.

As opposed to mobile networks, roaming within other types of networks such as ISP's networks or WLAN has not been the topic of specific standards and a number of proprietary solutions have emerged, offered by what could be called roaming operators or roaming brokers. These solutions are generally based on the same principles although they may differ in the details of the implementation.

The convergence towards NGN and the provisioning of associated services (such as seamless service provision) introduces new challenges: the ability of roaming between networks based on different technologies (inter-technology roaming) and the ability of maintaining a packet session while roaming between two networks (mobility support). Mobility support solutions could actually also solve inter-technology roaming issues depending on how they are specified. As an example, the IETF has described a standard for IP mobility support (referred to as "Mobile IP") that could solve both issues, but which has not been widely deployed so far. This could be explained by the fact that the implementation of such a standard in an IPv4 environment introduces a number of constraints, as for instance the requirement for tunnelling through the home network. These limitations can however be overcome in an IPv6 environment (see section 3.1.3.2 on IPv6). The market is showing a growing interest in this standard, with for instance 3GPP considering the implementation of the standard in the next stages of the deployment of next generation networks.

The major issues in roaming and mobility lie in the definition of a commercial model, the setting-up of service level agreements and billing processes. As these issues are most often not tackled by the standardisation work, they could represent a critical bottleneck and significantly delay the commercial launch of roaming solutions.

This section describes the solutions offered for roaming in different types of networks and the solutions offered for IP mobility support. In particular, the role of the home operators in roaming and mobility scenarios is highlighted and the extent to which they are retaining control over communications.

3.7.2. Roaming in mobile networks

3.7.2.1. Introduction

Although basic international roaming support was already part of the GSM standard, the GSM association has played a further role in the specification of roaming between mobile networks and in the publishing of guidelines and reference documents on technical and business issues related to roaming.

Roaming in GSM networks is based on bilateral agreements set up between operators of different countries. A GSM operator therefore needs to set-up hundreds of bilateral agreements to provide roaming services to travelling subscribers.

GPRS networks, as well as UMTS networks in their first release, are and will still be based on a GSM-based circuit-switched core-network for voice services. It is planned that circuit-switched and packet-switched links will coexist in mobile UMTS networks for a long period of time before a move to an all-IP UMTS version. In terms of roaming, this means that, for voice services, GPRS and UMTS networks can still rely on GSM roaming agreements, which involve the Home Location Register (HLR) and SS7 based signalling for location tracking, authentication and authorisation purposes.

GPRS technology introduces a new challenge relating to packet-oriented sessions. Mobile users can not only initiate voice calls from abroad but also packet sessions, which requires new GPRS-specific roaming functions in addition to those already provided by GSM. In April 2000 the GSM Association therefore published the GPRS roaming guidelines, highlighting the principles involved in roaming for packet sessions. As far as UMTS networks are concerned, the first release will be built on an enhanced version of GPRS network architecture, as mentioned in 3.1.4.2. The roaming principles defined for GPRS networks will thus still be relevant for UMTS networks.

This part provides an overview of GPRS roaming as introduced by the GSM association. The GPRS architecture is shown in the figure below.

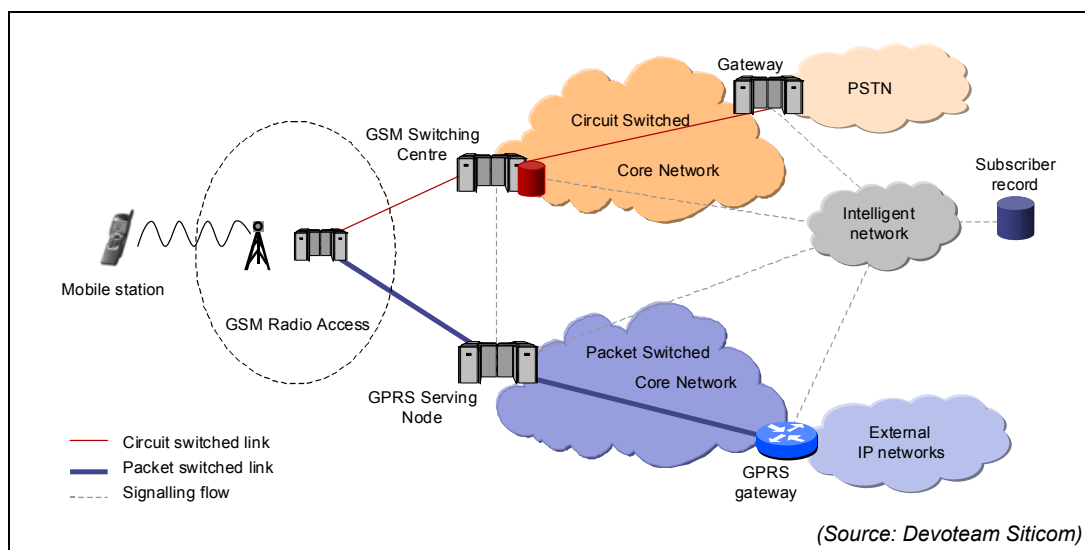


Figure 17 GPRS architecture

3.7.2.2. Registration and attach procedure

> Registration procedure

Before being able to make any communication, the Mobile Station (MS) needs to perform a “GPRS attach” procedure in order to register with the GPRS serving node. In a roaming situation, the MS performs the attach procedure with the visited serving node. During this procedure, the visited serving node exchanges information about the mobile station with the HLR of the home mobile network containing the subscriber’s record. At the end of the procedure the visited mobile network knows whether the MS is allowed to roam within its coverage and passes this information to the MS.

> Attach procedure

If the Mobile Station is granted permission to roam, it then needs to perform another procedure, called “PDP (Packet Data Protocol) context activation” procedure before being able to send any data. The information gathered by the visited GPRS serving node during this procedure will determine which roaming scenario will be applied during the packet session. The information gathered can come from three types of sources:

- the information sent by the user during the PDP context activation;
- the information from the subscriber’s subscription record, gathered during the attach procedure;
- the default data from the visited GPRS serving node.

3.7.2.3. Possible roaming scenarios

Two basic roaming scenarios can be supported in a GPRS network, based on a connection through either the visited GPRS gateway or the home GPRS gateway as shown in Figure 18 and Figure 19.

As explained above, the visited GPRS serving node determines the roaming scenario from the information gathered from different sources:

- the information sent by the user during the PDP context activation, and whether or not he/she has explicitly specified that the data session should use the home GPRS gateway (by providing the home mobile network identifier).
- the information from the subscriber's subscription record, and whether or not the home mobile operator has explicitly specified in this record that the data session should use the home GPRS gateway.
- the default data from the visited GPRS serving node, and whether or not the data service requested by the mobile user is available from the visited mobile network.

An operator may want to provide roaming services via the home GPRS gateway, and therefore provide the corresponding notice in the subscriber's record, so as to provide "home" services, including support of specific services, filtering, secure access to corporate networks, etc. On the other hand, for other specific services, such as location based services, access through the visited GPRS gateway would be preferred.

The "GPRS attach" procedure is the same in both scenarios. The following paragraphs describe the steps for both scenarios starting from the "PDP context activation".

> Roaming scenario using the visited GPRS gateway

1. The MS initiates the scenario by performing a "PDP Context Activation"
2. If possible, the visited GPRS serving node queries its local internal DNS server to resolve the data service requested by the user into the IP address of a GPRS gateway in the visited mobile network. This is possible because:
 - the user has not explicitly specified that he/she wanted to use the home mobile network, by providing the home mobile identifier;
 - AND the user subscription record specified by the home operator allows the usage of the visited mobile network for the requested data service.
3. The local internal DNS server returns the IP address of a GPRS gateway in the visited mobile network to the visited GPRS serving node. This is possible because the service requested by the user is available from the visited mobile network (e.g. the service requested is available through the public Internet or is a location-based service).
4. The visited GPRS serving node uses this address to create a tunnel and connect the user to the visited GPRS gateway.

In this roaming scenario, there is therefore no connectivity between the two mobile networks.

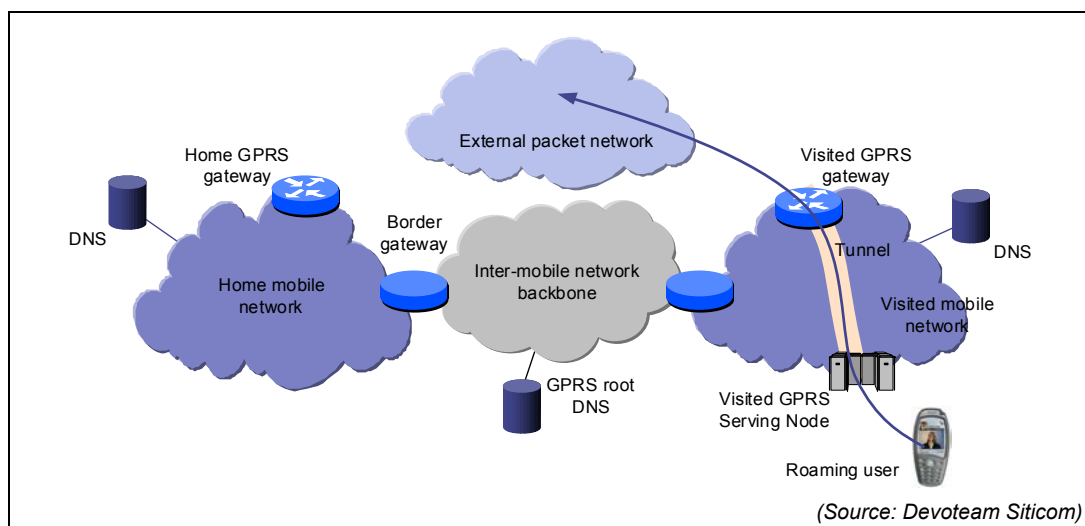


Figure 18 Roaming scenario involving the Visited GPRS Gateway

> Roaming scenario using the home GPRS gateway

1. The MS initiates the scenario by performing a “PDP Context Activation”
2. If the subscription record does not allow the usage of a visited GPRS gateway or if the user has explicitly specified that he/she wanted to use the home GPRS gateway, the visited GPRS serving node directly queries the DNS in the home mobile network via the root DNS server in order to be provided with the IP address of the GPRS gateway in the home mobile network.

Otherwise, the visited GPRS serving node first queries its local DNS server as in the previous scenario. If the local DNS can not find a GPRS gateway in the visited network to provide the requested service, then it will query the DNS in the home mobile network via the root DNS server.

3. The root server returns the IP address of a GPRS gateway in the home mobile network to the visited GPRS serving node.
4. The visited GPRS serving node uses this address to create a tunnel and connect the user to the visited GPRS gateway.

In this roaming scenario, connectivity between the two mobile networks is required in order to transmit data packets. The home operator retains control on the services accessed by the end-users.

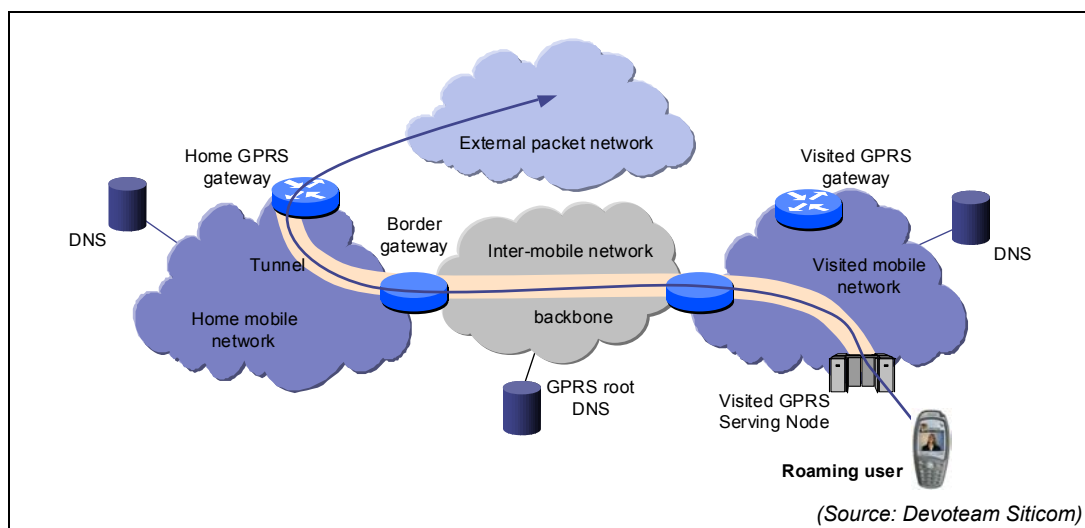


Figure 19 Roaming scenario involving the Home GPRS Gateway

3.7.2.4. Inter-mobile networks Backbone: the GRX approach

The roaming scenario based on the home GPRS gateway involves the connectivity of the visited and home mobile networks. The inter-mobile network backbone is defined as the network carrying the traffic between the border gateways of these two networks.

There are two alternatives for connecting the mobile networks to each other:

- by setting up direct connections between the operator's backbones – either for instance over dedicated lines or over the public Internet by using encryption mechanisms;
- by setting up one or several inter-operator GPRS roaming network(s) managed by GPRS roaming exchange operators.

In the reference document IR.33, the GSM Association has specified a centralised IP routing network for interconnecting GPRS networks as shown in the figure below. This document describes the GPRS roaming network made of GPRS Roaming eXchanges (GRXs).

The GPRS roaming network is a virtual network implemented on global backbones: it is logically fully separated from the Internet backbone but can physically use the same trunks and routers.

The GRXs are routers that can be connected to the Border Gateways of the mobile networks or to other GRXs. The GRXs also provide DNS services and maintain a root DNS name server.

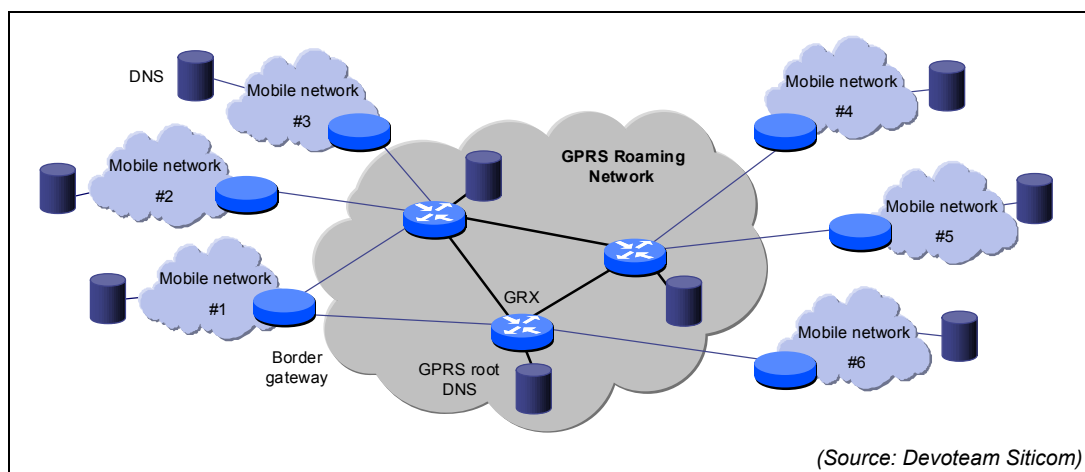


Figure 20 Inter-PLMN backbone implemented as a GPRS Roaming network

This implementation of inter-mobile network backbones involves a new category of players, the GRX service providers, which could for instance be a mobile operator or a data carrier. A number of players (several dozens) provides GRX services, such as Aicent, Sonera, Cable & Wireless, Telenor, Belgacom and France Telecom. Some of these have already set-up peering agreements with each other, e.g. C&W with Sonera and Aicent with Sonera. On the top of these GRX peering agreements, GPRS roaming agreement between mobile operators need to be set-up. However, the main challenge facing GPRS roaming today is the billing part, which is the reason why operators have only provided roaming services to their customers within the coverage of their own group (e.g. Telia). One of the first examples of GPRS roaming based on GPRS roaming agreements and GRX peering agreements is the MMS (Multimedia Message Service) roaming service between Sonera in Finland and CSL in Hong Kong launched in June 2002 and which is based on GRX services provided by Sonera and Aicent.

A major advantage of the GPRS Roaming network and GRX approach is to avoid the requirement for dedicated connections between each roaming partner: each operator only requires a small number of connections to the relevant GRX(s). Further advantages include scalability, simpler DNS configurations, faster implementation of new roaming agreements and, of course, cost efficiency.

3.7.3. Roaming in fixed networks

As opposed to mobile networks, no specifications or guidelines have been agreed on or implemented for roaming between Internet Service Providers (ISPs). Roaming is however an important service for such players as none of them can offer sufficient coverage to provide local access anywhere to their travelling business users. This is particularly true in the case of Wireless ISPs (WISPs) which only provide broadband access in a number of hot spots such as hotels, airports, etc. Access is allowed free of charge or can be paid locally. The main advantages provided by ISP roaming is the ability of accessing the Internet from anywhere and still be charged by the user's own ISP.

3.7.3.1. General roaming principles in ISP's networks

Considering the large number of ISPs, the option of setting up agreements with each individual ISP is not viable. A number of companies, including GRIC communications and i-Pass, have taken advantage of this situation and launched roaming offerings for ISPs.

The roaming solutions offered vary in terms of practical implementation and processes but are generally based on the same principles as shown in the figure below.

They involve three types of players: the home operator/ISP, the foreign operator/ISP and a roaming operator. The service providers compliant with a specific roaming offering only need to have relationships with the roaming operator and avoid the burden of having to set-up roaming agreements with each other service provider.

The roaming user registers with the visited ISP, for instance by dialling the local number of the local ISP in the case a dial-up access or automatically in the case of e.g. DSL or WLAN. He/she provides the visited ISP with his/her Network Access Identifier (NAI) of the form [user@domain](#), where the domain name identifies the home ISP (a). Very often, the domain name is not provided directly by the user, but added to the user name by the client roaming software used for the connection. From the NAI provided, the foreign ISP can detect whether the user requesting access is a local user or a roaming user and in the latter case forwards the registration request to the roaming operator (b). The roaming operator can then use the domain name part of the NAI to forward the request to the home ISP (c). The home ISP can then eventually authenticate the roaming user and authorise roaming (d).

After this registration procedure, the roaming user can access the Internet using the visited ISP connection.

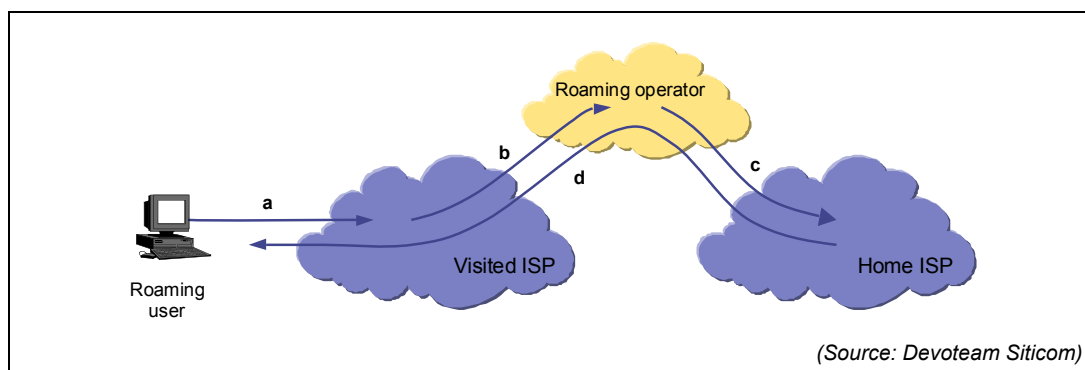


Figure 21 Roaming principles between ISPs

3.7.3.2. Roaming between Wireless ISPs (WISPs)

As for roaming between ISPs, roaming between Wireless ISPs has not been standardised in WLAN specifications but this will most likely be based on the same principles. WISPs may however want to provide their roaming subscribers with better service levels than general ISPs and be more cautious on the roaming service provided by roaming operators. Several trends in the market indicate that roaming will be of essential importance in WLAN.

- Established dial-up roaming operators such as GRIC and i-Pass have upgraded or are upgrading their offer to provide authentication mechanisms for WLAN access and roaming between WISPs.
- A new WISP association called Pass-One was created in June 2002, with the goal of specifying both the technical details associated with roaming between WISPs but also industry agreements on service levels in order to provide the most seamless experience between different service providers.

3.7.3.3. Inter-technology roaming

A further step in roaming services is to offer inter-technology roaming services, that is roaming services between operators of networks based on different technologies. An inter-technology roaming user would thus be offered the possibility of entering into a single relationship for a general service to be provided (e.g. Internet or intranet access) and use several types of access infrastructure (such as WLAN, GPRS, UMTS), possibly belonging to different service providers, to access this service.

Proprietary solutions combine the features of the different technologies. A solution will for instance be offered at the end of 2002 by Performance technologies, Transat and t-net. This solution is initially based on WLAN-enabled devices enhanced with specific software and a standard SIM card. Roaming is then provided between WLANs and mobile networks using specific gateways.

A critical aspect of inter-technology roaming lies in the setting up of commercial agreements, including billing and payment issues.

As previously mentioned in section 3.1.4.2, in the medium term, mobile operators are also considering inter-technology roaming between UMTS and WLAN as some of them have already launched WLAN services as an additional service. There is however still no standardisation in this area.

3.7.4. The IP mobility challenge

The roaming solutions described above assume that the roaming user starts a new data session each time he/she roams to a new network. A user may however want to start a data session like a file transfer at the office, switch automatically to the UMTS or GPRS network while on his way back home, and finish the file transfer at home with a DSL or WLAN connection. Such service would require the implementation of macro-mobility protocols, that is mobility between different IP-subnets.

None of the access network standards for 3G access networks, WLAN or DSL access is inherently supporting inter-system mobility.

Several mobility solutions, either standardised or proprietary, are being worked on, which could solve at the same time intra-technology and inter-technology roaming issues. They are based on the principle that a mobile user who wants to maintain a connection needs to keep the same IP address as higher-layer protocols such as TCP use the IP address to identify the end user.

3.7.4.1. IETF IP Mobility standard

> Mobile IP principles

The IETF has been working on standards for IP mobility support since 1996. They have defined mobility support for both IPv4 and IPv6. The standards are designed to allow mobility between IP subnets, possibly based on different technologies (e.g. WLAN, 3G) and without interruption of the session.

Mobile IP is based on the principle that a mobile device needs to be allocated two addresses: a home address, which is allocated by the home network and which remains unchanged whatever the location of the mobile device, and a care-of address, reflecting the actual location of the mobile device.

In an IPv4 environment, IP mobility will be enabled by the implementation of two mobility agents, which are in fact routers in practical terms.

- A home agent situated in the home network that maintains an association (“binding”) between the home and care-of addresses of a mobile node.
- A foreign agent situated in the network visited by the mobile device. It generally provides registration services to the mobile device as well as routing services and de-tunnels the packets sent by the home agent.

The implementation of mobility in IPv6 networks does not require foreign agents, as will be described later on.

> Acquisition of a care-of address and registration with the home network

When away from home, the mobile device needs to be associated with a care-of address. In an IPv4 environment, this address is most often the same as the foreign agent address, and is “discovered” by the mobile device using a mechanism based on router discovery. In such cases, the foreign agent would thus receive all packets delivered to the mobile device and pass them on to it. However, the care-of address can also be acquired through other means than the foreign agent, for instance in the case where the mobile device owns another address to be used when visiting a specific foreign network. The later case presents the advantage that the mobile device can function without the help of any foreign agent.

In an IPv6 environment, this process is simplified as a mobile node can automatically create its own care-of address by combining the visited network’s identifier with the device identifier, e.g. its MAC address. This guarantees that a care-of address is always available and removes the need for a foreign agent within the visited network.

The mobile device regularly sends updates to the home agent to keep it informed of its actual care-of address so that the home agent can maintain an updated binding association. The update messages can be sent via the foreign agent depending on how the care-of address has been acquired by the mobile device (see above). The home agent authenticates the mobile device on mechanisms based on the home IP address or digital certificates.

> Routing packets to and from the mobile device

The routing of packets between the different entities (mobile device, home agent, correspondent and possibly the foreign agent in an IPv4 environment) has evolved between the different versions of the standard and is different depending on whether IPv4 or IPv6 is used, as described below and on Figure 22.

All the packets to and from the mobile device need to contain the home address of the mobile device in the relevant address field of the IP header. This means that the home agent needs to intercept the packets sent to the home address of the mobile device (a) and tunnel them to the care-of address (that is either to the foreign agent or directly to the mobile device) (b). A tunnel is required in order to hide the mobile device’s home address from the routers handling the packets, as they want to be routed to the care-of address. In the same line, the packets are sent by the mobile device using a “reverse tunnelling” mechanism: they are also relayed by the home agent, a tunnel being used in order to hide the sender’s IP address contained in the IP header (1c). The major drawbacks of this solution are the additional bandwidth usage and latency issues due to the increased number of hops.

The IPv6 environment presents a number of advantages over IPv4 for the implementation of IP mobility. They are based on the fact that IPv6 packets can use extensible headers in which they can provide both the home address and the care-of address. This removes the requirement for reverse tunnelling (2c) and also provides the contacted device with binding information. Therefore, tunnelling will only be used in a limited number of cases, such as for the first packets sent to a mobile device. The devices will then be able to communicate directly with each other without having the packets transiting through the home network (2d).

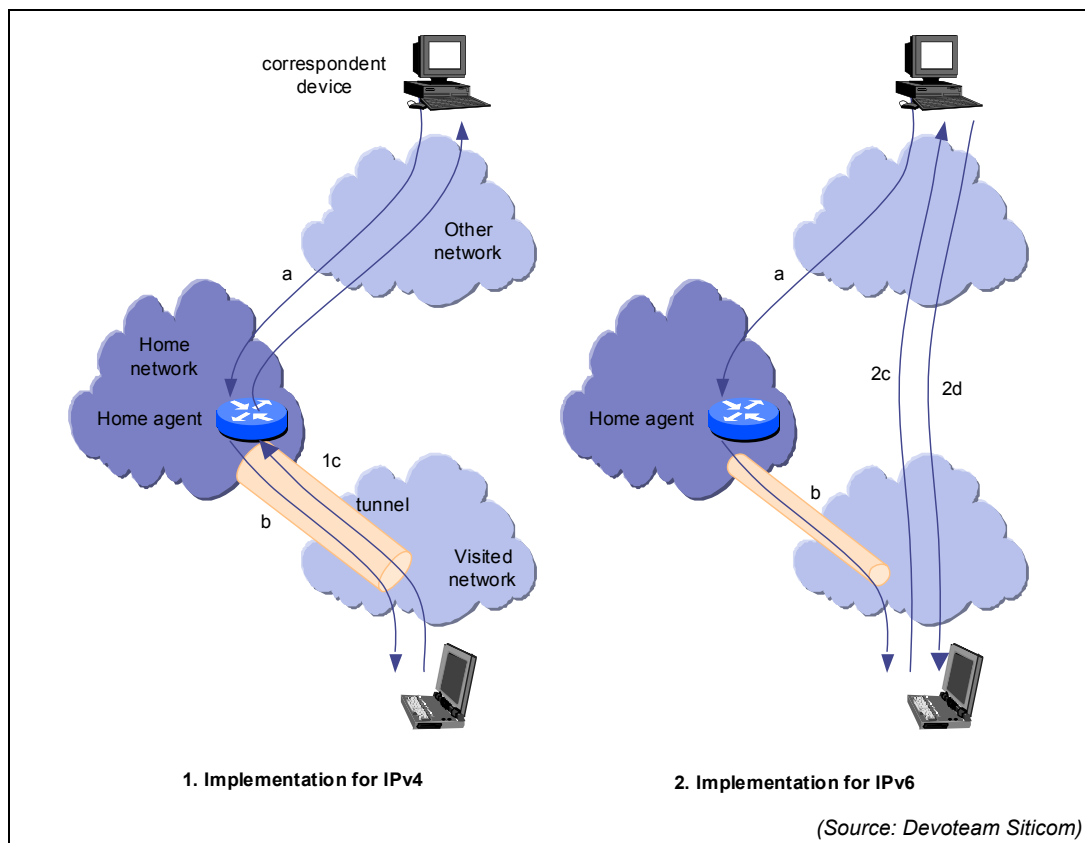


Figure 22 Implementation of Mobile IP

IP mobility implemented in an IPv4 environment also inherits the drawbacks of IPv4 and the issue of IPv4 addresses scarcity. NAT mechanisms, for instance, represents a major constraint in peer to peer applications, e.g. Voice over IP: devices that have only obtained a local address can not be reached by external applications if the device has not established a session first, as the external application does not know the actual address of the device. This constraint, as well as the further advantages offered by an IPv6 environment and described above, implies that mobile IP will only show its worth when IPv6 is further deployed. However, the implementation of mobile IP over IPv4 can provide home network operators with more control over the services accessed by end-user, as the data streams need to go through the home network.

An additional major challenge for the commercial exploitation of IP mobility technology is the billing part, which has not yet been explored.

3.7.4.2. Proprietary solutions

Proprietary solutions for providing mobility between networks are also being worked on, based on different mechanisms.

Some mobility management systems have for instance been based on a client / server application which requires data packets to be routed to a server. PacketAir has launched a solution, which is based on mobility routers and does not require any modification of or software installation in the mobile client. They claim that this solution provides the best alternative to mobility issues as it removes the need for a client in the mobile device to run specific mobility management software, as it is the case in client / server application and in the IETF standard.

4. Market developments

This chapter looks at some of the key market developments identified in this study. Interviews were conducted with representatives of 30 companies from diverse segments of the industry, including incumbent fixed operators, wireless operators, new entrants, Internet service providers, manufacturers, software companies and content providers.

Company representatives were asked for their views on a number of issues, including regulation in an NGN marketplace, and 'control points' that may, or may not, warrant the attention of regulators. Views and opinions obtained from the interviews are summarised below.

4.1. Phasing out circuit switched equipment

The overall technology trends predicted by the industry are the following:

- Circuit switched equipment will gradually be phased out and replaced by packet-based equipment based on IP.
- Access networks will increasingly become packet-based, both through deployment of 3G networks but also through developments in the fixed local loop.

Many players have a vision, which, in principle, corresponds well with these trends and the outline architecture. This includes manufacturers who develop hybrid products that support a migration strategy towards IP while still utilising existing circuit switched infrastructure.

4.1.1. Speed of transition

The speed with which the transition will happen is not clear, but many believe that while most of the technologies associated with NGN are available today, the full vision will take more than 5 years to be realised.

Several operators have stated that within the next two years they no longer expect to be investing in traditional circuit switched equipment. The scope of such decisions is mainly the operator's backbone, however these could also push developments in the local loop. Circuit switched equipment will be replaced by packet-based equipment based on IP – including potentially IPv6 for both mobile and fixed communications.

A key step in the evolution of the NGN architecture is that access networks will ultimately become fully packet-based. This would mean that all services, including voice over the local loop, are being delivered end-to-end based on packet-based protocols.

There is a general perception that for mobile access, this evolution is likely to happen faster than for the fixed local loop, where the speed of development is less certain. Alternatives will be offered by cable companies and providers of digital television, but for the foreseeable future the local loop is likely to constitute a technology mix of packet-based and circuit switched solutions offered by both incumbents and alternative providers.

The speed of transition will be affected by the operator's financial situation and the requirement to leverage past investments in circuit switched technologies. For incumbent operators with vast infrastructures the transition could thus take a considerable time, and it is likely that newer entrants on the market will complete the transition before the incumbents.

Experience, however, also suggests that once a transition has started, a conversion could happen faster than anticipated and faster than planned for. One reason for this could be the change in skills requirements where technical people both within the manufacturers and the operator are re-educated to the new technologies. It becomes difficult to find people who still master old technologies as young people will not educate themselves in technologies that are being phased out. In this context manufacturers however point to the fact that often existing support and maintenance contracts could run for up to 20 years and that they will obviously need to honour such commitments.

In any case, a transition will be highly gradual and, for the next 3-5 years, network backbones will constitute a mix of circuit switched and packet-based technologies. Some have even suggested a timeframe of up to 10 years.

The "transition" discussions and views outlined above are not shared by green-field players establishing new infrastructures. Several of these organisations believe that the NGN architecture as described in this study is too complex and that a much simpler architecture could meet the requirements of a future NGN market. There have even been claims that the investment needed to upgrade a legacy (cable) network to a newer (cable) network is as expensive as building a new from scratch. For green-field players, it only makes sense to deploy IP based multi-service fibre optic networks utilising simple protocols such as Ethernet and IP.

4.1.2. Drivers for transition

Industry views suggest that the burst of the ".com" bubble and the impact on the financial situation of key players in the IT and Telecoms industry has led to a stronger focus on ensuring that technology developments are driven by *cost savings* and *customer demand* rather than by technology itself. These two drivers will thus affect the speed of the transition towards NGN architectures.

A majority of the interviewed industry representatives claimed that while "new revenue channels" are an obvious driver, the services that can generate additional revenue have not been clearly identified. Some even claim that the reason for manufacturers pursuing and promoting an "open services environment" is the lack of clear killer applications.

Broadband rollout will play a key role in driving NGN developments and while the demand for "broadband to the home" is generally expected to grow, a number of applications have so far failed to attract the demand initially expected.

One example includes Voice over IP which has so far not succeeded as main driver for technology change. Reasons for not succeeding yet include lack of general interest due to the fact that no new functionality is being offered and that voice over IP is not seen as a credible substitute to the proven PSTN voice services (lack of QoS, unclear call set-up mechanisms etc.). Furthermore the use of Network Address Translators (NATs) and firewalls in the current Internet environment prevents seamless implementations of any-to-any VoIP. Finally, circuit switched voice is seen by some as a lucrative business meaning that the desire for operators to offer VoIP could be limited.

There are claims that currently no services or applications have been proven as main drivers for NGN developments, but a large number of *potential drivers* have been identified including multimedia messaging, location based services, mobility and entertainment.

4.2. Standardisation and interoperability

A key element in driving the deployment of existing network infrastructures (telecoms, data and broadcast) has clearly been the fact that certain standards have become globally adopted (as for example current PSTN standards, IP, MPEG2 and DVB).

Many of these standards will continue to support the basic infrastructure of NGN but network convergence and the new service environment mean that new standardisation and interoperability requirements emerge – mainly in relation to interfaces between networks and services.

The study has shown that there is some concern that proprietary standards, succeeding as *de facto* standards, could become potential control points requiring regulatory attention.

4.2.1. The multi standards environment

A move away from a monolithic standards structure has been taking place for some time, and is expected to continue. NGNs will be multi-standard environments with several standards solving the same issues.

Standards will be a mix between:

- Government enforced or “*de jure*” standards, where a standard is made compulsory by public bodies. For the telecoms sector such initiatives have typically emerged from the ITU and the national regulators. Development of such standards might not have been initiated by public bodies but are adopted at a later stage, adjusted and then made compulsory by government bodies. These standards are sometimes referred to as “Government driven open standards”.
- Standards emerging from industry fora, where commercial interests drive industry players into collaboration. This category could be referred to as “industry driven open standards”.
- “*De facto*” standards, where a standard emerges through market adoption and critical mass. For new products and services, interoperability and compatibility with “*de facto*” standards is unavoidable. One example of such a standard is Microsoft Windows. This type of standards could be referred to as “Market driven standards”.

A trend identified during interviews is that the important standards in NGN are the protocol standards and the Application Programming Interfaces (APIs) in the service creation environment.

4.2.2. Industry driven standardisation

A key standardisation trend raised by many industry players is that the successful NGN standards are likely to be largely driven by the industry and the market. Government driven standards are expected to play a lesser role and the significance of “the old

standardisation industry” is diminishing. The high number of current standardisation fora and the current cost of standardisation could be pointing to a need for modernising the way standards are developed.

4.2.2.1. Rationalisation of standardisation fora

The emergence of a multi standards environment (as described above) increases demand for interoperability. In addition, many of the agreed standards are not perceived as specific enough to guarantee interoperability and this has led to a large number of industry fora. Examples of such fora and alliances are:

- the Open Mobile Alliance (OMA) working towards interoperability (for instance between mobile devices) and open standards. The OMA includes Nokia, IBM, and HP.
- the Wireless Ethernet Compatibility Alliance (WECA), including organisations such as Apple, Microsoft, Cisco, Ericsson, Epson, HP, IBM, Infonet and Sony. WECA has defined the Wi-Fi certification, which guarantees interoperability between 802.11b-enabled devices.
- the World Wide Web Consortium (W3C), including about 450 member organisations, and working towards a technical evolution of the web with interoperable solutions. W3C is pointed out by the industry as likely to have a crucial role in future services development.

These “non-official” bodies may agree on standards specifications and, in some cases, promote them as approved standards by presenting their work to official standardisation bodies.

The move from voice to data as well as the convergence of fixed, mobile and data applications means that the current standardisation landscape looks overwhelming (as seen in Figure 23).

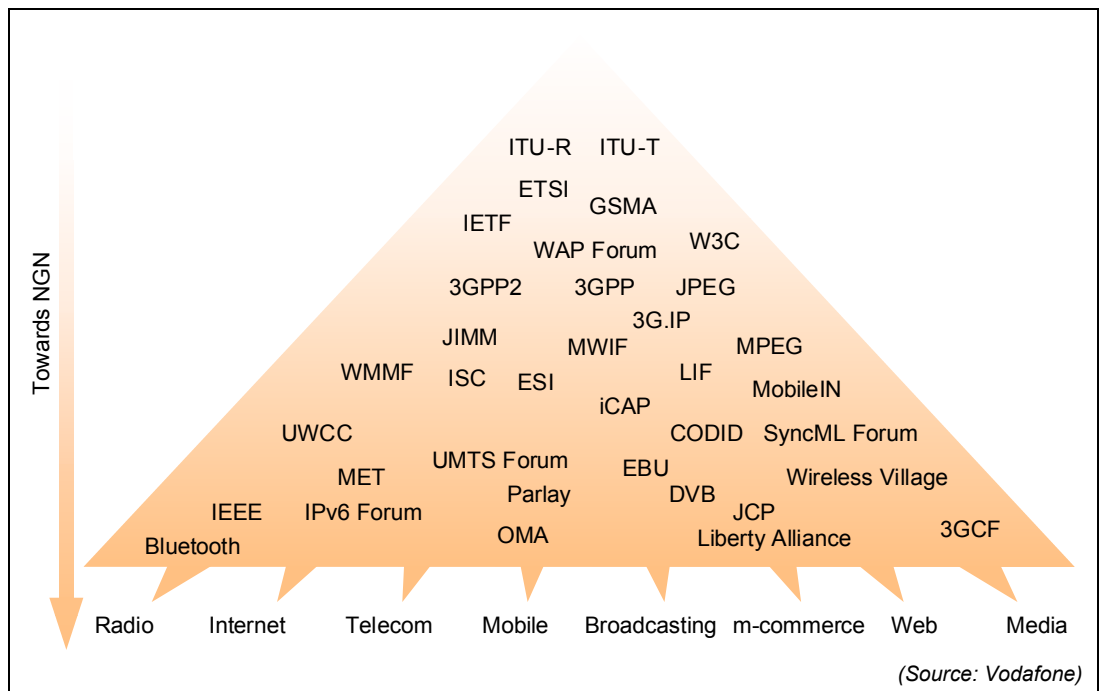


Figure 23 Overview of industry fora

As shown, the number of NGN related standardisation bodies has increased significantly over the last years. This is seen as a natural consequence of the increasing complexity of technology and convergence.

Many industry players however now see a need to modernise the standardisation process - initially by rationalising the number of individual and independent fora. A rationalisation will help industries in prioritising their standardisation efforts and spending.

Rationalisation is seen mainly as an action for the industry itself rather than for government.

One example of such a rationalisation is the recent announcement of the Open Mobile Alliance, which will incorporate the WAP Forum and the Open Mobile Architecture Initiative.

4.2.2.2. Drivers for standardisation

Industry players mention that top-down regulation, as seen in the past, takes too long and a key driver for pushing industry-driven standardisation is thus to cut-down the time required to define an acceptable level of standardisation.

Speeding up the standardisation process, however, requires a certain level of mutual trust between industry players and there is a growing tendency that standardisation issues are handled by a sub-set of significant industry players. Generally the industry do not accept postulates that certain standards are enforced by single players, however both vertical and horizontal "circles of interest" are becoming increasingly important in terms of adopting standards.

A specific example mentioned is that network operators increasingly tend to leave standardisation work to groups of hardware manufacturers.

While it is generally accepted that industry driven standards are necessary, the trends outlined above also raises a few concerns mentioned by several players:

- There is a risk that some industry players might be excluded from the standardisation process and strong industry players might use standards fora as a vehicle for pushing proprietary standards into becoming de-facto standards.
- While competing standards might support innovation, the fact that we have multiple solutions for the same problem might prove a barrier to end-users. Even when interoperability is provided, a multi standards environment still locks customers into certain technologies. I.e. if a large corporate organisation bases their VoIP infrastructure on H.323 it might be difficult to change to a provider using the SIP standard.
- The cost of participating in standards development can be significant and with multiple potential standards to support, it is not clear where industry players should invest their time and money.
- Industry driven standardisation, combined with the significant overlap of standardisation bodies, can lead to "forum shopping" where industry players are trying to push the same standard through several channels.

4.2.3. Interoperability and proprietary standards

Traditionally, standardisation in the telecommunications area has been concerned with the components along the transmission chain. In NGN, standards in the higher layers of the architecture become increasingly important. This includes the software layers that enable communications and services, such as operating systems, middleware, communications applications (including browsers) etc. Many industry players claim that development and acceptance of standards seems to work better on a hardware level and on the "lower layers" of networks. This implies that network specific standards have been more successful in gaining industry-wide acceptance than "user related" standards for software and services.

One reason for this could be the fact that software components are more easily replaced than hardware components and that "standards" on the software level thus tend to be much more dynamic.

In any case, development of open standards and interfaces does take time, and the successful deployment of NGN requires interoperability in a number of areas including Quality of Service, content provisioning, authentication, accounting and billing. Based on this, there is a growing recognition among the industry that *interoperability* is the goal rather than standardisation – and software based components such as proxy servers and gateways offer more flexibility in achieving such interoperability. Interoperability can thus be achieved at a much lower cost than that of developing open standards.

Some players do however maintain that while interoperability is the goal, open and unlicensed standards will always remain an important tool.

4.2.3.1. Quality of Service

Interoperability in Quality of Service is mentioned by many as a challenge, which is less of a technical issue than an issue of reaching relatively simple agreements between industry players.

The close relationship between NGN and the concept of "mobility" could pose a dilemma when trying to attract users. Many of the challenges identified in this study will imply that the NGN environment could initially be highly heterogeneous and that early instances of NGN could be "service and technology islands". To offer users true "mobility", seamless quality of service across multiple operators becomes a requirement and an environment with multiple conversions and gateways could prove a serious obstacle.

As already described in section 3.3.2 different QoS protocols support the delivery of multi-media content across a network, and the key challenge is to agree how such traffic is handed over between network operators without changing the traffic parameters (delay, packet-loss etc).

While no standard has yet been universally accepted for QoS exchange, it is generally accepted that interoperability is achievable on a case-by-case basis between two specific operators.

The issue is being addressed by several initiatives (GMPLS Forum, "the LION project", as well as IETF and ETSI activities) but no standards have emerged yet.

In relation to QoS, more far-reaching standardisation is being debated among operators. This relates to allowing 3rd parties control of network performance on another operators network. Some operators are not in favour of such developments as, at this point in time,

controlling network performance and maintaining a high quality backbone is seen as one of the differentiating propositions of operators.

4.2.3.2. Services and content provisioning

Having multiple standards in the services and software areas is seen by many as healthy as it increases competition. It is however also a general concern that 3rd party service development should be encouraged and supported.

A key element in ensuring interoperability is the Application Programming Interface (API) which is further explained in section 3.4.

Two main paradigms exist for supporting 3rd party service development:

- The “standard API” approach where standard APIs are being developed for certain functions in the network. Examples include Parlay and JAIN APIs described elsewhere in this report and the Multimedia Home Platform (MHP). This approach is criticised by some as being too expensive and too time consuming to develop.
- The “proprietary API” approach where APIs are published and made available to 3rd party services. In this scenario tools for “generic applications development” could allow applications to be ported to several proprietary platforms.

Different market and technical conditions may determine what approach to pursue.

As an example, industry players in the domain of “set-top boxes” claim that pursuing a standard API approach (as the Multimedia Home Platform) might be more expensive in terms of development cost than the cost of porting applications to platforms. This is partly due to the very broad scope of functionality which needs to be supported by the standard API across different types of set-top boxes. As a result, some of the major television content providers claim that the focus should not be on ensuring *portability* through standard platforms and APIs but rather on ensuring *interoperability*. Smaller players, like independent cable operators, maintain that open standards are crucial for success.

In other domains, such as telecommunications services, developments of standard APIs have already proven successful, as is the case with the Parlay and JAIN standards described elsewhere. Some state that one reason for these developments succeeding could be that relatively few proprietary APIs were developed before the standardised approach was initiated.

APIs in the telecommunications services domain are currently being implemented by equipment vendors, in response to requirements from operators and service providers, wanting to build open markets for service- and contents provision.

In a regulatory context, it has been claimed that, while certain elements of the service layer might be outside the scope of telecoms regulation, it must be ensured that no players are able to abuse a dominant position in these layers. This means that interoperability and open interfaces could become a requirement, with governments enforcing (or encouraging) 3rd party access through an open service creation interface. This should not prevent proprietary standards, but rather ensure that these are interoperable with other standards for those functions where interoperability is required.

4.2.3.3. Authentication

The open services environment created by NGN, and in particular the Web Services environment described elsewhere, requires interoperability for functions such as user authentication, single log-on and user profile management.

A number of somewhat overlapping initiatives are currently trying to address these issues.

As mentioned in section 3.4.5, "Microsoft Passport" offers a proprietary solution to user authentication, single log-on, authentication and user profile management. Many see MS Passport as "unavoidable" due to the widespread use of other Microsoft products.

The "Liberty Alliance" (with broad industry representation) addresses some of the same functional areas in an effort to establish the necessary means of authenticating and identifying a user in relation to e-commerce.

Both approaches thus support authentication of users across different services by passing on user credentials to other organisations (as agreed with the user in question). However, both initiatives also allow individual organisations (Microsoft or a Liberty Alliance member) to build significant and valuable repositories of user data that are *not* shared with other organisations even if the user demands so. While such closed repositories of user data makes good commercial sense in developing service innovation and competition it also serves as a powerful tool for controlling the market for certain services.

Whatever the approach, many see specific players such as portals, instant messaging providers and operators as having a good position in terms of providing authentication features. Also already existing mobile authentication schemes (in the SIM toolkit) and the likely extension of SMS into the fixed line environment means that mobile operators can play a significant role in authenticating users.

4.2.3.4. Addressing

Recently the ITU, the Internet architecture board (IAB) and the IETF announced their approval of the interim procedures for ENUM. ENUM, as described in section 3.5.3.2, is mainly focused on the mapping of user identities.

The role of ENUM is widely debated in the industry with some players seeing it as rather insignificant while others see it as a good framework in need for real commercial implementations.

Those who are more critical of ENUM claim the key challenges to be that:

- So far, no service or application has been identified which is hugely dependent on a standard like ENUM.
- ENUM principles work only under proprietary control where filtering and anti-spam mechanisms can be put in place.
- ENUM will never work in the public space due to privacy concerns and regulation. The only application for ENUM would be large private corporate intranets.

4.2.3.5. Billing and accounting

In addition to the authentication challenges described above, other areas where interoperability should be further encouraged include:

- User authorisation
- Accounting information (use of resources)
- Billing information

The successful exchange of customer billing information and information on a customer's use of network and services could be a requirement for the successful deployment of roaming functions and functions related to 3rd party service and content providers.

In this context the issue of market control through closed repositories of user data discussed in 4.2.3.3 also becomes relevant.

4.3. Powerful NGN roles

The study has shown that NGN could potentially lead to a consolidation of traditional roles into new partnerships which could become powerful entities in controlling both the NGN infrastructure and the services offered over NGN. As shown later in this section, the *exclusivity* of such partnerships will vary and will also be limited by current regulation on open access requirements and non-discriminatory pricing.

As highlighted in a recent report published in the UK¹⁶, there are a number of means by which control in one market (or set of markets) could be levered into other markets. Depending on the precise circumstances, these may either breach competition law or sector-specific regulations. The mechanisms include the following:

- Cross subsidies or price squeezing
- Differentiating product availability
- Differentiating level of service
- Differentiating availability of information
- Bundling different products
- Leverage of customer base¹⁷

This section describes potential developments in the consolidation and integration of the various roles of the value chain. This includes the relationships between content providers, service providers, operators and manufacturers.

4.3.1. Characteristics of main players

The NGN environment could lead to a battle between network operators, traditional ISPs, service providers, content providers and manufacturers. This battle is initiated by the fact

¹⁶ Electronic Networks, Challenges for the Next Decade published by The Cabinet Office, Strategy Unit Report, December 2002.

¹⁷ List extracted from the report mentioned above.

that each party, by default, will be in control of important parts of the value chain. This will however also work as a driver for establishing new strategic relationships.

Figure 24 below summarises some of the relationships likely to develop between various players in the value chain of NGN and next generation services.

Potential developments in these relationships are described further below.

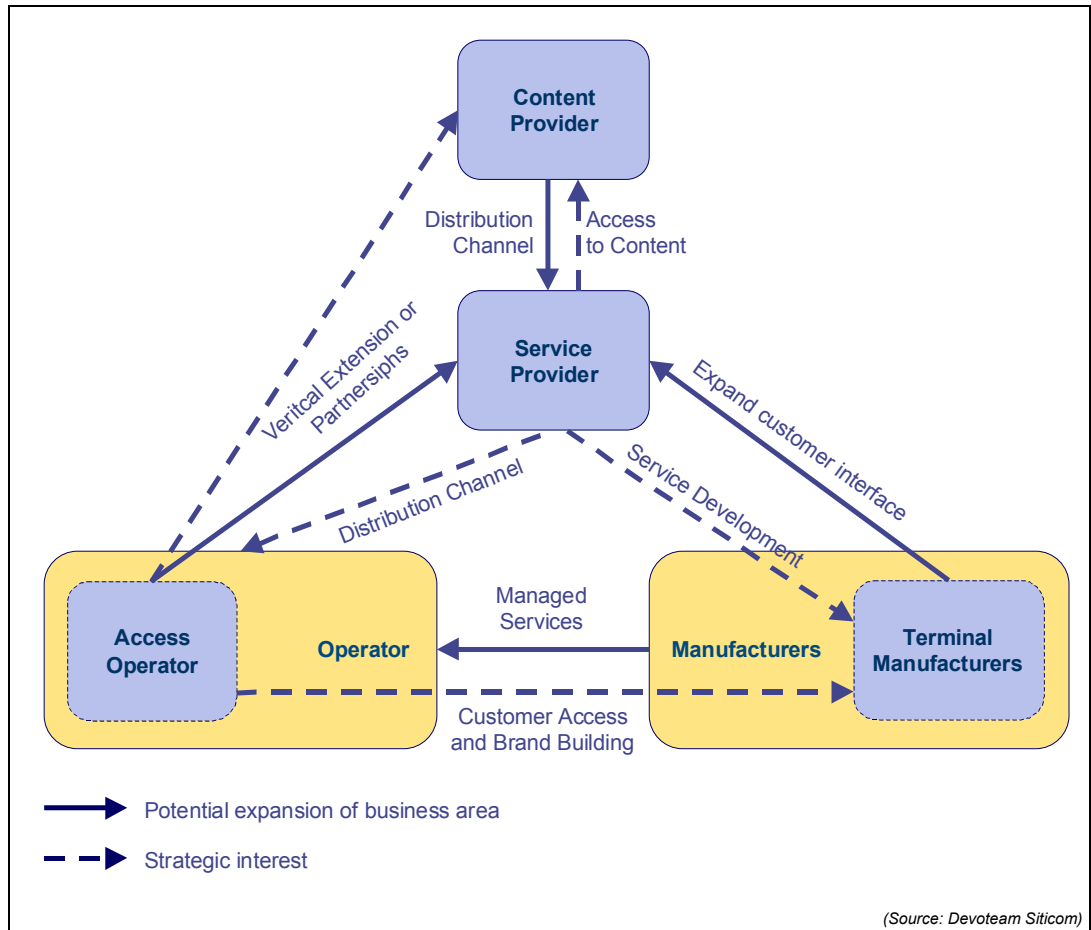


Figure 24 Examples of potential industry integration and bundling

4.3.1.1. Operators

Operators are currently facing a difficult financial situation, which is made even more difficult by expected investments in NGN technologies and the fact that it becomes increasingly difficult to make margins on backbone services.

For access operators the main business opportunities in NGN thus lie in the vertical extension of market power (based on existing direct customer relationships). For many incumbent operators this implies horizontal alignment of the organisation, introducing a clear split between “access”, “transport” and “services & content”.

Operators are seen by other players, such as service and content providers, as attractive partners. Main reasons for this are:

- Access operators could become the main distribution channel for many types of content including software and music. Seen from the content provider’s view a partnership with an operator or carrier is thus valuable as it provides better control with future distribution channels.

- Access operators control huge repositories of customer data and in most cases they own the customer relationship through the billing interface.
- For some types of content (real-time video and multimedia) the cost of delivering the content could be significantly higher than the perceived value. In this situation operators are in a better position than the content owners to optimise delivery and get the price/value balance right.
- Quality of Service on the network is critical for some applications (such as multimedia messaging) and this aspect is purely controlled by the network operators.

There are however claims from certain industry players that the limited success of local loop unbundling and competition could lead to a monopoly situation where content providers would only get very small revenues from access providers. In such a scenario the "attractiveness" of access providers will be limited and ultimately this could lead to less content being made available online.

Somewhat related to this discussion, other industry players suggest that mobile operators are more willing to open up towards content providers through content-related revenue share arrangements than incumbent fixed line operators. Some players even see a requirement for some sort of incentives to operators who support 3rd party content and service development.

These developments also point to the issue that interconnecting infrastructures in a technical sense might be far less of an issue than the interconnection of fundamentally different business models. Aligning revenue channels in new value chains could prove to be a substantial hurdle.

4.3.1.2. Hardware manufacturers

In the past, equipment manufacturers have invested significantly in operators, mainly in the mobile sector. Cases of funding up to 250% of the value of an order placed by the operator have been seen.

The current market situation could lead to some of these operators failing, effectively forcing the manufacturers to take possession of the operator's assets. While some of the assets will be sold off, this also allows hardware manufacturers to expand their service portfolio towards other clients. Hardware manufacturers are increasingly offering managed services bundled with equipment sales – effectively leading to a situation where manufacturers are closely involved in the day-to-day operation of an operator.

In seeking new business opportunities, hardware manufacturers are however expected to be very careful in not stepping into the domain of their customers. This means that manufacturers are not likely to become operators in their own right.

Hardware manufacturers (in particular handset and terminal manufacturers) are seen as attractive partners due to the fact that "Handset ownership" is a very important part of brand building and customer ownership. This is due to the perceived link between user values and the capabilities and brand of the handset. For most users the handset constitutes the only tangible part of a service and the handset is directly associated with the brand of the service provider and/or the manufacturer.

As an example confirming this, some mobile operators claim that "handset relationships" between current GSM operators and manufacturers have proven much more valuable to operators than content deals.

4.3.1.3. Content and service providers

In an NGN market, content and service providers will be faced with new opportunities for both distributing traditional content and for deploying new multi-media services. It seems likely that service and content providers will look into both the operator and the handset/terminal space, to gain more control of distribution channels.

The expectation that the NGN market will be driven by availability of content and new services means that content and service providers are seen as attractive partners. Access operators will look for ways of sharing the risk associated with new infrastructure investments and this is likely to happen through partnerships or joint ventures with content providers. This trend has been confirmed by both content providers and mobile operators.

Some portals and content providers, however, point to the fact that producing content (like movie productions) is becoming an increasingly risky business with huge budgets and limited guarantees on the return on investment. It has been mentioned that content is arguably a riskier business than that of operating and upgrading network infrastructure. This could affect content providers willingness to share risk through joint ventures with operators or manufacturers, meaning that traditional distribution channels (such as cinema, TV and publishing) will remain the preferred means of content delivery.

In relation to this, some access operators claim that by sticking to traditional channels and release cycles, content providers prevent content from becoming the expected key-driver to NGN and broadband deployment. This refers back to the issue of interconnecting business models mentioned in section 4.3.1.1.

4.3.2. Integration of services and networks

As shown above, there are a large number of drivers for creating relationships between various roles or even for expanding certain roles into new business areas.

Vertical integration of companies or vertical expansion of a business mainly becomes a regulatory concern if the integrated business entity is dominant in parts of the value chain.

The market position of the parties will affect the types of relationships that will be established. The exclusivity of the potential partnerships indicated below will vary and could also be limited by current regulation on open access requirements and non-discriminatory pricing. Examples of possible relationships or expansions are listed below.

- Collaboration between various parties could lead to “unnecessary” bundling of software, elementary service functions, control functions and network infrastructure. In such a relationship, service providers could mandate certain infrastructure elements under the excuse of security or quality – even though the service and the infrastructure are technically independent. One example of this could be a service provider mandating the use of a certain browser and a certain access operator.
- In NGN, ISPs will increasingly seek to control (and be able to restrict) certain network centric functions as for example QoS. This means that the role of being an "ISP" becomes of strategic importance and many players confirm that the ISP function will increasingly be tied to infrastructure ownership. Examples include cable Internet and mobile IP where the ISP functions are always owned by the access provider.
- Established access operators with large customer bases will not necessarily seek exclusive partnerships but rather pursue partnerships to gain better control of their

revenue potential through access to certain types of services or end-user devices. These arrangements would be guided by a desire to drive traffic and maintaining a strong brand.

- Newer access operators, who are seeking to gain a stronghold in the market and managing the risk of investments, could be more focused on entering exclusive partnerships with both hardware manufacturers and content providers. These arrangements would be guided by a desire to drive traffic, build a brand and attracting customers.
- Handset and terminal manufacturers could be moving into the service space and seek to establish their own direct relationship with customers. Communities initiated by manufacturers, such as “Club Nokia”, could serve as a delivery channel for services and content delivered independently of the operator. However, as already mentioned, manufacturers will be very careful in not stepping into the domain of their customers.

It should be mentioned that many of the relationships suggested above have not been proven as viable commercial models yet. As an example mentioned, the current “exclusive” deals between content providers and operators are seen merely as “trials” rather than proven business models.

Also, some players state that one might not place too much importance on the vertical integration trend and that this should be seen merely as a current fashion or “flavour of the day”. In future business cycles horizontal integration might gain momentum again. This statement is underpinned by some incumbent access providers who place more importance on opening up the infrastructure to many service providers than on moving towards vertical integration.

4.3.3. Further industry consolidation

While the industry has seen significant horizontal consolidation over the past one or two years, further consolidation is seen by many as both necessary and unavoidable.

Consolidation will take place at all levels and this could create a difficult situation for the competition authority.

Competition, and the fact that bandwidth prices are still coming down, means that it is becoming increasingly hard for players in the transport/transmission segment to make money. These players could either move up the value chain (vertical integration) or seek further consolidation in the transport segment, to gain access to more volume. Some claim that surviving in the transport segment requires you to be one of a very few players and that within a couple of years the number of international carriers in Europe will be reduced to be between three and five.

Others state that on new efficient green-field infrastructures (such as “pure” Ethernet and Fibre infrastructures) operators could gain sufficient revenues from selling bandwidth alone. Many see it as likely that there will still be a wholesale market due to the fact that ISPs and larger service providers are unlikely to want to become European operators. These providers, as well as mobile operators, will be buying bandwidth from backbone providers – more or less as a commodity.

The commoditisation of bandwidth also creates an opportunity for companies to take on the role as “bandwidth traders” or “bandwidth brokers”. Such companies would enter into

contracts with ISPs and content providers offering bandwidth “on demand” and management of service level agreements.

4.4. Interconnecting NGN infrastructures

This section summarises market views in the area of interconnecting NGN infrastructures. This includes views in relation to peering arrangements, termination rates, quality of service and service level agreements.

This study (and others) has found that when exchanging traffic, global and large backbone providers have an advantage when negotiating charges and quality for interconnect. Furthermore it has become clear that, if necessary, this issue must be dealt with at a trans-national level as it makes little sense to look at national markets in isolation.

4.4.1. Interconnection in NGN

Transit and peering agreements between service providers have traditionally been a key topic when discussing the competitive environment surrounding the Internet.

Open standards such as IPv4 and IPv6 allow providers to provide transparent packet transport service to other providers, however a number of potential bottlenecks have been identified in the past, which could prevent the free flow of traffic. These bottlenecks included discriminatory pricing and restrictive practices such as the rationing of IP addresses or the blocking of certain protocols.

Traditional interconnect agreements on the Internet were based around a hierarchical structure meaning that requirements on “non-discriminatory peering” and interconnect were raised frequently by players who felt discriminated by larger players.

Many parts of the industry now claim that commercial mechanisms have been successful in resolving such IP interconnection issues and that the Internet has become much less hierarchical over the last 5-6 years - effectively meaning that the market power of large backbone operators has been reduced¹⁸.

Several operators claim that these developments mean that regulatory measures in the area of interconnection are becoming less necessary.

In NGN, issues of market power and dominant position are likely to become more related to higher-level specific services than to arrangements for simple traffic exchange.

4.4.2. Peering agreements

Peering arrangements are generally accepted as the most cost efficient way of exchanging traffic between “similar” parties. Since the model is based on a “sender keeps all” principle the requirements for exchanging billing information is minimal (see figure below). Increased focus on cost savings could thus drive market players towards peering agreements rather than towards complex interconnect agreements.

¹⁸ This trend is confirmed in WIK-Consult, “The Economics of IP Networks – Market, Technical and Public Policy issues relating to Internet Traffic Exchange”, May 2002.

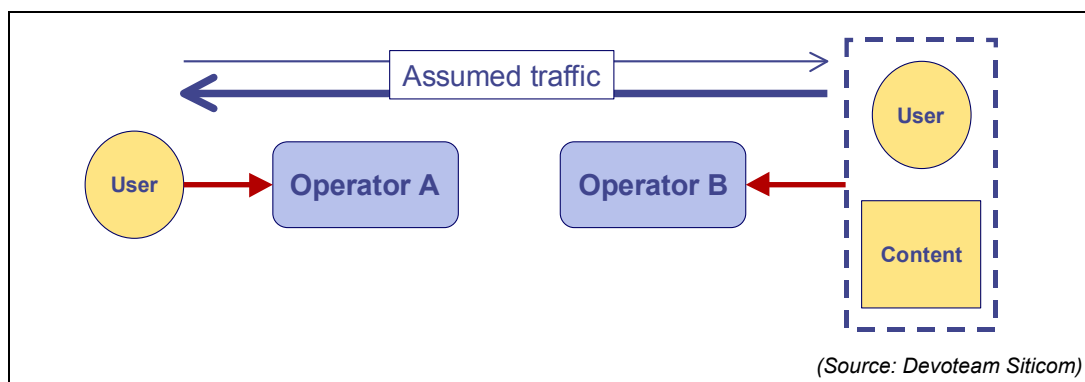


Figure 25 Revenue flow (red arrows) in simple peering

However, peering agreements only work well between parties with a similar cost base or between parties with complementary market positions. Peering agreements are typically based on valuation metrics such as network size, content volume hosted and relative flow of traffic. Most operators have published terms for interconnection (and peering) however the identification of a “true peer” will be different in each case.

Such peering is only one of several models for exchanging traffic and other settlement-based models are seen by the market as making more economical sense if the two parties are very different.

Ultimately, both fixed and mobile users and service providers will rely on the global Internet backbone providers to provide global connectivity. Selling transit to smaller players is one of the core businesses of these backbone operators and they do thus not see it economically feasible to offer free peering arrangements. As already mentioned, many see the transport/transmission segment becoming increasingly “commoditised” and in such an environment backbone providers need to move toward selling such a commodity rather than giving it away for free.

Critical voices suggest, however, that not only global backbone providers but also large national incumbent operators (often in control of broadband access) will continue to have the power to control prices and quality for interconnect arrangement and that this is an area which require some monitoring and potential regulatory action. This has been confirmed by the recent WIK study¹⁹, which concludes that it should be closely monitored whether any backbone or access providers have the capability to discriminate between on-net and off-net traffic as this could lead to an unstable situation where the largest network would dominate.

4.4.3. Transit and "Paid Peering"

Peering is seen as a barter relationship and “declining to peer” is not the same as blocking an ISP’s ability to reach customers or sites as other arrangements will be available. As an example smaller players have the option of joining up with other players to qualify for peering with a global backbone provider.

Another alternative will be to establish an agreement of transit or an agreement of “paid peering” between two parties.

¹⁹ WIK-Consult, “The Economics of IP Networks – Market, Technical and Public Policy issues relating to Internet Traffic Exchange”, May 2002.

In transit agreements, settlements between the parties are typically based on actual traffic measures and throughput.

Paid peering refers to the situation where a flat fee, based on access capacity or other metrics, is paid by one of the parties (see the figure below). Findings suggest that paid peering arrangements are relatively rare and only deployed in the few situations where transit arrangements are not practical²⁰.

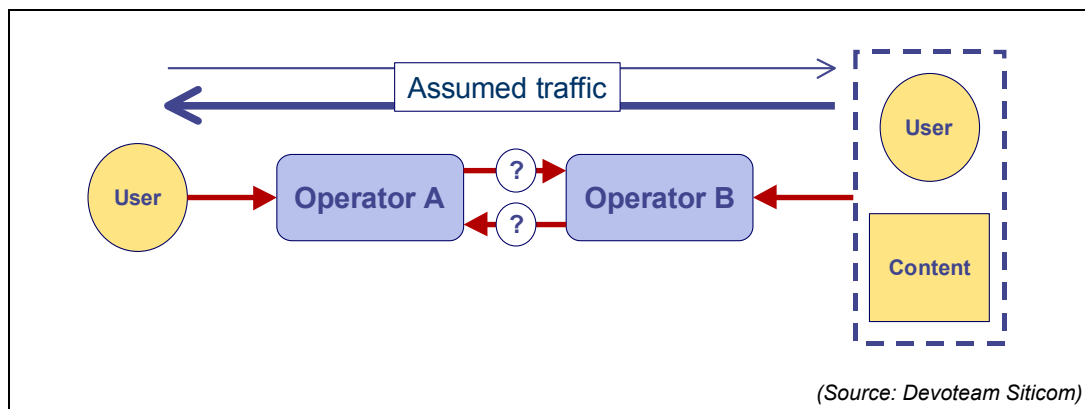


Figure 26 Transit or "Paid Peering" revenue flow (red arrows)

4.4.4. Termination rates in NGN

The traditional notion of "termination" might be very difficult to apply in an IP environment where it is not possible to easily identify whether traffic is "push" or "pull" and, as consequence, who should be billed for the traffic.

NGN will offer a large variety of business models, which are not based on termination rates. These could include:

- Capacity based billing
- Service based charges
- Revenue sharing
- Risk sharing (Partnerships or ventures with specific agreements on risk sharing)

Some operators maintain that termination rates could continue to play a role in the settlement between operators. For traditional voice traffic, termination revenues have played a vital role in the business model for access providers, and also allowed the operators a level of control over the access link. This legacy could mean that many access operators will seek to implement traditional termination charges for IP traffic. IP termination charges would thus fit well with established business models and also allow the access provider to maintain a level of control similar to that for voice traffic.

At the moment no one can predict how the markets will look, but it has been suggested that, if necessary, termination rates could be differentiated by Quality of Service classes rather than by specific applications such as "voice termination".

²⁰ This trend is confirmed in WIK-Consult, "The Economics of IP Networks – Market, Technical and Public Policy issues relating to Internet Traffic Exchange", May 2002.

4.4.5. Monitoring Quality of Service

All players agree that Quality of Service interoperability between networks is a very important issue in NGN. Currently, operators define different QoS parameters and classes within their own networks and, with no interoperability in place, there is a danger that the QoS perceived by end-users across networks will deteriorate.

Some players even suggest that this is an area where regulatory requirements could be necessary and that it might be necessary for regulators to implement some form of Quality of Service surveillance. Such measurements could also be used in a broader sense to ensure that no parties abuse their market position by not meeting or neglecting interconnect commitments.

The practicality of such “interception or monitoring points” is however questionable, and others have highlighted that in trying to monitor interconnect points regulators might create artificial boundaries that might actually strengthen the position of operators with large networks.

Other players claim that no regulatory intervention will be required as such issues are purely contractual matters and that disputes can be solved legally.

4.5. The new services environment

The NGN architecture directly supports 3rd parties creating and offering services to end-users (as detailed in section 2.1).

The vision that mobile and fixed terminals should be able to share the same services over a common IP backbone network requires interoperability on many levels.

The study has shown that in reality the separation between service provisioning and network provisioning will not be as complete as suggested by the NGN architecture. However, it is also suggested that the main interoperability requirements will be linked to the services environment where many of the potential control points in NGN could reside.

4.5.1. Next Generation “Killer Applications”

Many claim that NGN developments are held back by the lack of next generation “killer applications”. These sources also claim that as of yet, no services have been associated exclusively with NGN and no services have been proven as main drivers.

This view is countered by others claiming that the applications that will drive NGN developments are here already. NGN will mainly serve as a vehicle for expanding the scope of existing successful applications. Examples include in-car data services, Internet access through hot-spot technologies for wireless LANs, enhanced location based service, enhanced messaging and services that are currently successful in the youth segments (SMS, chat and sharing of multi-media content).

In this context it is worth noting that current delivery platforms, such as SMS, allows operators significant control over the content delivery channel both in terms of cost, quality and content. While such controls are inherent in current architectures, the NGN architecture holds the potential to open up competition in these delivery channels providing more options and flexibility to content providers.

The development of Internet, SMS and WAP based services has shown that very traditional types of content can be successfully enhanced for new channels. One example is weather information, which was traditionally “packaged” for broadcast media.

This last view implies that existing (successful) services and content are enhanced for mobility. In this context it has been mentioned by some industry players that established content providers, such as the film industry and television producers, seem to be less convinced of the “multi channel opportunity” offered by NGN. This should be seen in relation to claims from content providers (see 4.3.1.3) that producing content is an increasingly competitive business and that utilising new distribution channels means introducing new risks in relation to revenue, finance, piracy, IPR and copyrights.

4.5.2. Service bundles

Service bundles offer service providers a degree of control by allowing them to cross-subsidise between services, ensuring that new features can be offered at competitive rates.

From a consumer perspective, service bundles are also seen as making good sense as they allow users to benefit from the new world of converged services without having to make too many choices.

There are claims that in a competitive and thriving market, service bundles will eventually become comparable. This builds on the assumption that customer demand is uniform across a given market segment and that bundles aimed for specific segments will contain the same services. Service providers will seek to differentiate their service bundles through special services and capabilities, but eventually a given service will become available in competing bundles.

While operators and service providers claim that market demand will ultimately create adequate transparency for users to make an informed choice, others have concerns that service bundles and the vertical integration of operators and service providers makes it increasingly difficult for consumers to compare service packages.

4.5.3. Semi-walled gardens

“Walled garden” is a term applied to a service package that provides customers sole, or preferred, access to certain pre-determined functions and content. It exemplifies business strategies designed to attract and retain customers.

A potential control point in the services space could be found where some industry players are leveraging their access capabilities with a “walled garden” services strategy. Whether or not this constitutes a *problematic* control point is subject to debate.

Interview observations suggests that at least a “semi-walled garden” (portals where the public Internet is available, but only through a number of menu choices) is seen as a viable business strategy, although service providers should be monitored closely. As an example, certain service providers with a de facto monopoly of domestic customers in a given area, can attempt to impose a ‘walled garden’ on those customers, restricting their access to services from other service providers.

It is also argued that walled gardens or semi-walled gardens could provide a good starting point for many users when first entering a new service space. This is particularly the case for services offered over mobile devices with limited bandwidth and smaller

screens. In this type of environment it is valuable for users to have quick access to a portal of selected services.

Examples have already emerged, where Cable TV and Digital TV operators are seeking to create a walled garden for a selection of services suitable for the TV interface (buying flowers, simple home banking etc.).

However, walled gardens could lead to loss of transparency in the market, as it is not clear to the user, which parts of “the Internet” will be available as part of the service.

In addition, customer demand of full open access and the concept of “personalisation” do not fit well with closed environments. These trends could eventually lead walled gardens and some service bundles to fail.

An alternative to semi-walled gardens controlled by service providers and operators could be user-defined portals (web pages) where easy non-restricted access is provided to any set of services or content defined by the user. Early examples of such user-defined portals include KPN's I-mode (for mobile phones) and the fully customisable Avantgo service for Personal Digital Assistants (PDAs).

In terms of regulation, many interviewees believe that walled gardens are not necessarily detrimental control points and that regulators should not regulate *against* walled gardens as the concept works very well for many users. Transparency does, however, become a requirement as users need to clearly understand what a service bundle or walled garden allows them to do.

4.5.4. Control through elementary service functions

The NGN marketplace will rely on a variety of development and delivery platforms located at different network layers to create a broad range of information society services. It is clear that the new services environment will rely on a number of “elementary services” being available in the network(s). When building and offering services in a Web Services environment (as described in section 3.4.5), access to certain core functions becomes critical.

While in principle the NGN architecture suggests that these functions are separate from the network infrastructure, it is generally suggested that a large number of service functions will be closely tied to the bearer network.

This section briefly describes those functions identified by the industry as important in relation to the further development of NGN. Many of these functions could become potential control points and it is foreseen that network operators and players in the service layers (e.g. ASPs) will compete for such control.

Some findings from the study suggest that regulatory attention should be directed towards such functions to ensure that open platforms and interfaces are provided to service and content providers. In this context it is also highlighted that the layered model makes it easier to pinpoint bottlenecks and put in place regulatory remedies such as unbundling.

Other findings indicate that it is still rather early for regulators to precisely identify what the candidate interfaces for regulation might be.

4.5.4.1. Basic call functionality

It is generally accepted that voice communications will remain a very important basic service in the new services environment.

In NGN, however, the control of voice and multimedia communications – as for other services – could move from the control of network operators to the control of pure ASPs offering voice services purely based on software running on a shared IP infrastructure.

For voice services on IP networks, as opposed to ATM and SS7, all signalling information will be embedded in IP packets (“in-band signalling”). This allows 3rd parties to provide such VoIP services over traditional IP infrastructures.

Network operators remain involved in the provisioning of the QoS necessary for the provisioning of the services.

Such developments could lead to a situation where large enterprises would eventually choose to build their own call servers and only use network operators for the transport of data.

Functions offered by the operator through standard APIs (Based on PARLAY or JAIN as described in section 3.4.4.3) would however be required when building services relying on caller line identification, initiation of calls, billing information, security etc.

4.5.4.2. Resolution of names and numbers

Voice services and multimedia-messaging services across networks rely on the ability of a service provider to resolve name and address of a called party. As described in section 3.5.3, access to Domain Name Servers and other name and number servers (based on ENUM, TIPHON or proprietary standards) is a prerequisite for such services.

The resolution of names and numbers has involved a number of players from different backgrounds so far:

- Public telephony operators, including incumbent and mobile operators, who own databases for routing numbers; public and non-profit organisations have also played a role in the administration of databases of routing numbers (e.g. to solve issues such as number portability between operators);
- Public and non-profit organisations, but also ISPs, end-users and various organisations such as universities for the administration of DNS servers.

The implementation of ENUM as part of the DNS would involve public and non-profit organisations as well and the cooperation between the different providers of addresses (access operators, ISPs) and identities (ASPs). Questions arise however on how to implement ENUM in practical terms. It remains unclear who should administer ENUM servers and whether ENUM-like domains are required to promote competition, protect the integrity of the existing schemes, and prevent fraud, misuse and abuse of power.

4.5.4.3. Quality of Service Tunnels

As explained in section 3.3.2.3, NGN must offer application service providers an interface to "Quality of Service tunnels" guaranteeing a certain end-to-end performance across the infrastructure. This allows service providers to allocate bandwidth and performance as required by individual applications and services.

4.5.4.4. Single log-on and customer identity management

In a Web Services environment single sign on and user authentication becomes crucial as several applications (belonging to different providers) might be involved in delivering a single service to a user.

Organisations providing single sign-on functionality today are ASPs and software developers such as Microsoft providing Passport.Net and the Liberty Alliance project led by Sun. In the future, however, ISPs or mobile operators could be in a good position to serve as authenticator on behalf of the user against ASPs, e-commerce organisations or content providers. Moreover, the principles for roaming authentication might prove useful when building more generic authentication features such as Single Sign On.

4.5.4.5. Functions for determining location

Location based services are still seen as being an important part of the future services environment and for 3rd party service providers the development of such services depends on open access to location information.

Operators claim that from a consumer-privacy perspective, it might not be easy for operators to share this kind of information. However, subject to these legal issues being solved, operators seem to be willing to make the information available on commercial terms.

Other service providers fear that this could prove to be a competitive barrier in deploying new services. This issue is highly related to the discussion in 4.2.3.3 on market control as location data might be kept in closed data repositories that are not shared with other service providers even if the user demands such sharing.

Mechanisms allowing the user to sanction that location information is shared with selected 3rd parties is thus seen by many as an urgent requirement. The feasibility of such mechanisms will however depend fully on a clarification as to whether such data is owned by the operator/service provider or by the end user.

4.5.4.6. Filtering mechanisms

While filtering mechanisms have traditionally been a small utility used for Internet browsing, many have highlighted that filtering becomes increasingly important in the new services environment.

Filtering functionality includes mapping domain names requested against a "black list" of forbidden domain names or scanning the content of accessed resources (looking for e.g. key words or virus files).

Both application service providers and network operators will increasingly get involved in offering such services as customers will expect services such as "parental controls" to be in place by default. Filtering becomes increasingly important as Internet content is being made available on mobile devices which could be "out of sight" from parents and teachers.

4.5.4.7. Digital rights management

The issue of digital rights management remains largely unresolved, and many claim that this should not be considered an "elementary service" as it operates on content on a

much higher level than the basic services. This also implies that network operators should not be held responsible for content transmitted over networks.

Having said that, it is widely accepted that functions for managing and protecting digital rights are crucial for the successful distribution of content. It has been suggested that as long as no standard approach for digital rights management has been identified, “walled gardens” (as described above) might provide the best channel for content delivery and payment.

Digital rights management could be seen in the context of more elementary functions restricting access to certain types of content and services through filtering mechanisms (described above). Eventually such mechanisms could work on the basis of a defined Digital Rights Management policy.

4.5.5. Development of software and services

Software applications available in an NGN marketplace may reside at different levels of the network and in terminal devices. This raises the issue of whether software platforms on which applications operate should be viewed as potential control points by regulators.

The study findings suggest that content providers, access operators and manufacturers will be focusing on service and software development. This also means that all three parties will have their own programmes for developing software components to end-user devices (Java based applications or applications for specific operating systems).

These developer's programmes will have different focus:

- Access operators will focus on developing services and software that utilises features that are specific to the network.
- Manufacturers will be developing services and software that utilises features that are specific to the handsets and devices.
- Other content providers will be developing components and services that are more generic between networks and platforms.

In such an environment access to systems and network through proprietary or standard Application Programming Interfaces (APIs), could become an important aspect (APIs are discussed further in section 3.4.4).

The study also suggests that software developers may claim intellectual property rights over functions, claiming that they are not only proprietary, but also protected in a way that could make it illegal for competitors to interpret them. Browsers were mentioned as a specific example where software could become a control point because dominance in this area gives the opportunity to control data exchange formats and thus what software can be used effectively on the Internet.

4.6. Customer control through billing

Managing the customer relationship is an essential part of any business. In the NGN marketplace, the billing relationship and access to customer information is viewed by some as a control point which could potentially require regulatory attention. Access operators control huge repositories of customer data, they are closest to the subscriber *and* they are more experienced than most service providers in managing the customer

relationship. In addition to these points, the access operator is granted additional control through the ability to ultimately disconnect a subscriber.

4.6.1. Third party billing and payment services

Even though the NGN architecture (as interpreted for this study) does not directly address billing and payment, the ongoing debate on this subject will remain relevant in an NGN context.

The importance of customer ownership and control through billing remains clear, and a great interest in entering the billing space as a “gatekeeper” for payment systems has been predicted.

Although billing could be provided by each single player providing a service, some players or even third-parties could provide billing services on behalf of other players. The drivers for such an approach include:

- The ability to provision a single, simpler and more consistent bill to the end-user;
- The experience of specific players in billing and customer relationship management for similar types of services (e.g. network operator or ISP);
- The experience of an organisation in financial-related matters and the trustworthy brand image it provides (e.g. financial institution).

Developments in billing and mediation platforms create new opportunities as these platforms become capable of handling content as well as traffic. Concepts like “micro billing” allow new players to enter the market and effectively introduce a split between the billing of transport and the billing of content. In the Internet Web environment micro billing has already been implemented in the form of pre- or post-paid accounts, which are debited when users click a button on a content provider's web site.

In this context, it is however important to note that the ability to collect information on a user's use of resources is provided within or close to the devices providing the resource to be measured. Such functions are for instance carried out in authentication servers (e.g. RADIUS servers) but also in application servers. Interoperability, open interfaces or “clearinghouses” and “roaming brokers” providing access to usage information is thus a key requirement for providing 3rd party billing services.

So far, only few proprietary solutions have managed to establish themselves as providers of 3rd party payment services.

Financial institutions are expected to be seeking ways to leverage their position through billing and payment services. So far, however, little development has been seen.

One reason for the cautious approach is that the case for payment in the open wireless world remains unproven:

- There are claims from established credit card companies that it is very difficult to build a business case for providing low-value transaction payment services (below 5 Euro). An average transaction value of 20-30 Euros has been mentioned as a requirement for being able to make revenue from such services. This view could be seen as a sign that the existing payment transaction infrastructure for these companies is not suitable for micro payment transactions.
- “Cash on your device” technologies (like eCash, CyberCash and DigiCash) has not proven successful yet (Double SIM-card solutions, “cash” on your PC etc.).

- The scope of “m-commerce services” is not clear, and some of the billing might be catered for through existing solutions such as premium numbers, premium charged SMSes etc.
- Some players claim that consumers are not yet interested in using their phones as a payment device in shops, super markets etc. People might be more willing to use the mobile phone to pay for content and services delivered to the device (like I-mode in Japan). Another problem highlighted is that of reimbursement in case a product does not comply with specifications.
- It has been remarked that the real challenge in 3rd party billing is collection rather than the actual billing itself. Incumbents are rather powerful because they can ultimately cut off a subscriber’s access.

In addition, developments in the handset market will provide better and more powerful handsets. Once the Internet browsing capabilities of a handset become similar to those of a PC, existing (Internet based) payment mechanisms might meet most of the requirements of (mobile) payment services.

Some industry players also highlight the fact that telecom operators are very good at billing smaller amounts which will allow them to play a significant role in micro billing (Vodafone and Omnitel have been mentioned as examples of companies offering billing and payment services to other parties)

4.6.2. NGN charging models

There is considerable uncertainty about future billing models. The speed at which the transition away from per minute or per-package charges will happen is not clear. As an example it could be that billing for voice will continue to be based on per-minute charges for a long time.

In addition there are trends for operators and service providers to differentiate between business pricing and consumer pricing despite the fact that the underlying technology is the same. Such trends could reflect the fact that consumers prefer “predictable” flat-fee pricing whereas business users are willing to pay for volume-based transactions.

This section summarises market views on some of the potential billing models.

4.6.2.1. Volume based charges

Operators need to recover costs and make profits and the current Internet billing model does not generate sufficient revenues.

It remains clear that billing models will not be driven by technology but by competitive pressures and by the value perceived by end users.

Some operators are currently planning to introduce or keep a volume component in their billing model for the foreseeable future. For mobile telephony, one reason for this is that frequency is a scarce resource.

For business users the usage of a connection/subscription might vary considerably over time, and such users might be willing to face charges for high volumes when they occur rather than an overall flat-rate. Congestion charging and bandwidth auctions might be viable billing models towards business users.

One reason for operators *not* to adopt a volume based billing model is the fact that volume based charging is less cost-effective in terms of the billing process. Operators might see a cut in the operational overhead by moving away from per-minute or byte volume charges.

4.6.2.2. Flat-rate access

There are claims that flat-rate access arrangements are essential for the further development of Internet and Internet services. Volume or time based mechanisms could constitute a psychological barrier to usage, even if the actual cost are reasonable. Volume based charging (for data in particular) suffers from that fact that the model does not reflect the actual value delivered to the customer.

Charges are thus likely to be service based in combination with a charge based on access bandwidth. Initially, this model will apply for data but increasingly also for voice services.

Content will be charged separately and I-Mode has been mentioned as a good billing model, where operators charge a content commission of around 10-15%. Such revenue sharing models are seen as a result of a horizontal alignment of the services market. There are however also claims that, at least in the past, the I-Mode model was difficult to adopt in Europe mainly due to many operators wanting a disproportionate share of the revenue.

Finally it has been claimed by some that a market based on flat-rate billing for access and transmission leads to better price transparency. This statement does however not necessarily imply that flat-rate billing will be the most prevalent or successful billing model.

4.6.2.3. Advertising

Today, television and broadcasting are largely funded by advertising. With the increased focus on content and services, advertising is expected to become a significant part of the revenue stream for operators and service providers. For the I-mode service in Japan, advertising revenue constitutes 20-25% of total revenues.

4.7. Market views on regulatory approach

The rationale for regulatory intervention during the current transition from monopoly to competition is ensuring a level playing field for all market participants by regulating the main control points passed forward from the monopoly era – residual market power and network bottlenecks. Learning from these experiences, regulators may view their job in the NGN marketplace as something similar: identifying control points and applying the appropriate regulation. Regulators will have to distinguish between potential control points that promote normal competitive activity, and those that may harm competitive activity.

In conclusion, the interviews carried out as part of this study suggests that regulators must be cautious when putting in place ex-ante measures, and that ex post tools are seen by a large part of the industry as sufficient in a mature market.

While the following chapter will reflect further on the views gathered during interviews, this section briefly summarises the main observations.

4.7.1. General approach

Most respondents had an opinion about the general approach to regulation appropriate for NGN. Several expressed the need for a go-slow approach. There was little enthusiasm in general for a heavy hand, although some of the advice offered to regulators about how to proceed would entail considerable regulatory involvement.

Many claim that NGN developments and an increasingly mature and competitive marketplace call for some regulatory objectives to be redefined. New objectives should recognise the fading power of previous monopolies, a need to encourage further investments, and the expectation that services will converge into completely new service propositions.

Many respondents agree that local and national government have an interest in seeing an efficient broadband infrastructure deployed to help attract new industry, and generally to make life better for citizens.

4.7.2. Ex Ante versus Ex Post

A number of those interviewed commented on the 'ex ante versus ex post' trade-off for regulatory action.

Ex ante tools include regulations put in place in anticipation of a problem, such as those to ensure cost-based, non-discriminatory access to the incumbent's bottleneck facilities by competitors, and the blocking of alliances seen as anti-competitive. A few respondents have indicated that ex ante tools would be preferred in an NGN environment as competition law has proven to be too slow. Cases taking up to three years to reach a conclusion were mentioned, by which time the outcome might be irrelevant.

Ex post tools are mainly antitrust remedies under competition law applied after an abuse of market power is alleged or demonstrated. Many respondents suggested that regulators should be careful of not regulating too early and that the "long-term risk and reward aspect" must be taken into account. Players should be allowed a certain level of freedom when building services and gaining market. In mature market, competition law (ex post) is seen by many as sufficient although it is recognised that operators need to be careful in such an environment. Ex post could represent a financial risk to players in a fast moving market like NGN where a significant part of an operator's revenue could be put at risk if a dominant or abusive position is alleged or proved.

5. Regulatory implications

While previous chapters have elaborated on the technical implications of implementing the NGN architecture as well as potential market developments, this chapter will look at the main regulatory implications identified by the study.

5.1. Introduction

NGN will be subject to the regulatory framework for electronic communications networks and services that has recently been established in the EU. This regulatory framework is technology neutral and seeks to ensure that all communications technologies will be able to compete under the same conditions.

NGN will present a range of challenges for regulators. The first and obvious challenge arises from new technology. The 2003 package is intended to be technology neutral, but current regulatory practices are strongly related to the CS environment. Some of the traditional regulatory requirements may become less relevant with NGN, other will remain but change character. Fresh analysis and assessments will be required to apply the regulatory principles of the 2003 package to NGN. Some of the regulatory implications that follow from technology differences are discussed in section 5.4 below.

A second set of challenges comes from understanding that NGN is more than new technology applied to an existing paradigm. It represents a paradigm shift where the electronic communications market becomes heavily integrated with information society services. As a consequence, regulatory concerns will shift upwards to the higher layers of the network hierarchy. This is further discussed in section 5.5 below.

A third set of challenges is associated with trying to understand sources of market power and how that power can be abused in the NGN environment. So far, market power has been associated with a certain market share related to networks and transmission services. With NGN, market dominance may also be derived from controlling more limited sets of functions and capabilities that are necessary for the provision of services to end users. One example is the system of domain name servers which perform translations from domain names to IP addresses. The ability to use domain names is fully dependent upon this function and in this sense, the number translation function constitutes a *control point*. In a hypothetical situation where the domain name translation would be controlled by a commercial profit seeking organisation, there would be obvious dangers of abuse of dominant position. With the new regulatory framework, the solution would be to recognise that number translation services constitute a relevant market and to designate the dominant operator as having significant market power. Following this procedure, ex ante conditions could be defined to ensure non-discrimination and cost-orientation.

NB. The administration of the root level of the DNS is managed by ICANN. The example is only intended to illustrate the notion of a control point, not to cast any doubt on ICANN as a fair and neutral administrator.

The notion of control points and potential sources of market power in the NGN environment is further discussed in section 5.2 below.

A fourth main challenge is to understand the implications of regulatory action (or inaction) in cases where some form of dominance over control points occurs. In the NGN

marketplace, regulatory action may have unintended consequences that will retard, rather than promote investment and competitive activity.. This important challenge is discussed further in section 5.3 below.

Finally, regulators will face new challenges with regard to interconnection. The vision of NGN reflects a unified backbone network which can encompass both PSTN, Mobile, Internet and Cable TV type of traffic. In a long transition period legacy networks will continue to exist, but over time they are expected to be transformed and upgraded to enable additional types of services. When considering what type of regulatory principles to apply to a given interconnection task, it seems therefore more logical to focus on the category of service to be interconnected, than on the origin of the network. This is further discussed in section 5.6 below

This study has a clear focus on the main competitive aspects of NGN. There are many other regulatory concerns that also could be discussed as NGN will raise significant issues, in particular in the areas of data protection, privacy, universal service and consumer protection. There are also broader questions about the relationship between the new regulatory framework for electronic communications networks and services on the one hand and the regulatory framework for information society services and electronic commerce on the other hand. These provide substantial scope for further study.

5.2. Control points

5.2.1. Introduction

The NGN environment includes a wide range of special facilities or server functions²¹ that in theory could have open interfaces and/or could be provided competitively. Each of these could in theory be seen as candidates for being considered as relevant markets, but most would normally be subject to competitive pressures without a need for regulatory intervention.

Only where there are significant bottlenecks or control points that threaten to block the development of the market can ex ante regulation be justified as a means to achieve fair competition. The real challenge posed by NGN is to understand where these can appear in an environment that will be quite different from the PSTN and mobile environments, which are familiar to regulators. They could well be related not to the transmission layer but to some other aspects of service creation. The potential control point need not be owned by the operator of an electronic communications network or service. It could equally well be a critical software platform controlled by a software vendor.

There is no consensus in the market today that such control points will emerge with NGN nor a common understanding of what the potentially harmful control points would be. On the other hand, all players in the market try legitimately to achieve competitive advantage that can build or sustain market power and provide some degree of customer control. Given the complexity of the environment in which technology can provide a seemingly infinite range of possibilities, it is possible to construct a number of theoretical scenarios under which important NGN function(s) can come under control by a single commercial organisation.

²¹ See Figure 2

It would be important for regulators to understand where to look for those control points that can become sustainable and/or irreversible sources of dominance and which would lead to market failure unless addressed by ex ante regulations. Such control points would probably involve ownership of elements that would be necessary in order to provide certain services and that could not easily be replicated. The local loop is a classic example of such an element. In an IP setting, there could be other elements that could be equally critical for the provision of a certain set of services. They would probably be strongly related to individual customers, such as customer identity information or information on customer preferences.

Control points could also take the form of market power that would enable an operator to impose bundling and/or interoperability limitations which reduce customer choice and/or competitive alternatives.

Below is a number of examples of functions that are expected to play important parts of the future NGN environment and which therefore could provide basis for control points.

5.2.2. Control points relating to network capabilities

"What can be done with the infrastructure and how can dominant operators limit infrastructure capabilities for competition?"

- **Network Address Translators and firewalls**
As long as IP V4 remains the main protocol for IP communications, address translation mechanisms between networks constitute points where control on communication capabilities, traffic and quality can be exercised (see sections 3.1.3.2, 4.1.2 and 4.4.5).
- **Routing tables**
Operators that control routing tables are in a position to control the flow of data and to which network data is forwarded (see section 3.3.1).
- **Quality of Service capabilities and interconnect**
If large backbone operators refuse to offer Quality of Service tunnels or Quality of Service termination to other operators, these capabilities become important control points which could effectively force users to use specific network providers for services requiring certain QoS parameters (like voice and multi-media services). Currently large backbone providers have an advantage when negotiating charges and quality for interconnect (see sections 2.2, 3.3.2.3, 4.2.3.1 and 4.4).
- **Network coverage**
Currently large backbone providers have an advantage when negotiating charges and quality for interconnect. If such providers refuse to deal with other operators on equal terms without commercial justification, this could become an important control point. Furthermore it has become clear that if necessary, this issue must be dealt with at a trans-national level as it makes little sense to look at national markets in isolation (see section 4.4).
- **Termination capabilities**
Access operators will naturally have a certain level of control over the access link relating to both capabilities and termination rates. In a scenario where it is possible to differentiate the traffic transported over the access link, this control could extend to the termination capabilities for certain types of traffic (as is currently seen for SMS).

In a scenario where services depend on a single access link, the access operator has an important control point (see sections 2.2, 3.2.2 and 4.4.4).

5.2.3. Elementary Services and NGN capabilities

"What services can be built and how can a dominant operator restrain competitive service alternatives?"

- **Call set-up capabilities**

In NGN, the control of voice and multimedia communications could stay with network operators or move to the control of pure ASPs offering voice services based solely on a shared IP infrastructure. In both cases controlling call set-up functions becomes an important control point. It is a question of market power in the control plane where a dominant operator may have control over the call set-up, bandwidth, quality of service, etc. (see sections 2.2, 3.1.2.1 and 3.6.3).

- **Proprietary standards**

If proprietary standards succeed in becoming "de facto standards" by large players (or "circles of interest" constituting a small group of powerful players) these standards may dictate what functions and services can be supported in an NGN environment (see section 4.2.3).

- **Non-proprietary standards**

Even non-proprietary standards may be biased in the sense that they influence the balance of power between network operators and service providers. For example, the TIPHON concept, under development by ETSI, has the implication that network operators can retain control over the service provided (see section 3.5.3.3).

- **Interoperability**

The degree of interoperability between the transport/control planes and the service planes will be controlled by network operators. They can therefore exert some control over the capabilities of service providers. APIs are therefore a key element for providing access to network capabilities and could be used to exercise control (see sections 3.4 and 4.2.3).

- **Application Programming Interfaces**

If a proprietary API gets significant market share, this would allow involved software or service providers to control what functions and services can be supported in an NGN environment. This will particularly be the case where software developers claim intellectual property rights over functions, claiming that they are not only proprietary, but also protected in a way that could make it illegal for competitors to interpret them (see sections 2.2, 3.4.1, 3.4.4 and 4.5.5).

5.2.4. Control points related to accessing services and content

"What services can a user access and how can operators and service providers limit his choice?"

- **Unnecessary software and service bundles**

If various powerful parties successfully market bundles of software, elementary service functions, control functions and network infrastructure, they may be able to mandate certain infrastructure elements under the guise of security or quality. Such

vertical integration may prevent open access on various levels of the NGN infrastructure (see section 4.5.2).

- **Walled Gardens**

If access providers choose to control the content and services that can be accessed by a customer, such a walled garden could become an important control point in differentiating the availability of information and services, unless rendered irrelevant by competitive alternatives (see section 4.5.3).

- **Tunnelling**

Mobile operator may keep control over their roaming subscribers and handle all types of services except outgoing voice calls by the home operator. In an IP environment, this will be done by tunnelling, whereby all traffic is sent back to the home operator before being serviced. Tunnelling is an important concept for retaining and controlling end-users (see sections 3.1.3.2, 3.7.1, 3.7.2.3 and 3.7.4.1).

- **Filtering Mechanisms and Digital Rights**

In a scenario where content providers and users need to ensure the legal status of content, filtering mechanisms and digital rights mechanisms in the network become important points for controlling what services and content can be accessed (see section 4.5.4.6).

- **End-user devices**

In a scenario where popular services and content are closely linked to the capabilities of specific handsets, manufacturers of such handsets will be in a strong position to control development and deployment of new services (see section 4.3.1.2).

- **Content**

Certain types of content would become an important control point in a scenario where content providers would favour certain delivery channels over others. As an example if providers of movie content stick to traditional channels and release cycles this could prevent new channels enabled by NGN to succeed (see section 4.3.1.3).

5.2.5. Control points relating to user information

"What can providers know about a user which facilitates service provision?"

- **Authentication, single log-on and user profile management**

If certain proprietary solutions for authentication, single log-on and user profile gain wide market share, software and service providers behind these solutions will gain significant control over how and what user data can be shared in an NGN environment (see section 4.5.4.4).

- **Customer billing information**

If major operators refuse other operators access to customer billing information (incl. use of network resources) on reasonable commercial terms this could constitute an important control point for the implementation of roaming functions and functions related to 3rd party service and content providers (see section 4.3.1.1 and 4.6).

- **Access to customer information systems**

In a scenario where content providers rely on access operators or other service providers to gain access to specific groups of customers (through technical or

commercial mechanisms), the access operator will be in a strong position to control how this customer base can be accessed by content providers (see section 3.4.5.4 and 4.2.3).

- **Resolution of names and numbers through Customer Identity systems**

In order for telephony traffic to work across IP and PSTN networks, it is necessary to perform number translation between E.164 numbers (traditional telephone numbers) and IP numbers. This raises several control point issues. For instance, the allocation of ENUM registries to operators could provide them with the ability to control the routing of the communication and thus decide which network(s) and operator(s) to use. In a scenario where traditional voice services as well as multi-media communications services rely on an IP infrastructure, the ability to resolve names and numbers may become an important control point (see sections 3.5.3.2, 3.5.3.3 and 4.5.4.2).

- **Functions for determining location**

Operators that collect user location data, such as the mobile operators, are in a strong position to control services that depend on this information. In a scenario where location information becomes an important revenue driver in the services market, the ability to provide such information becomes a potential control point (see section 4.5.4.5).

5.3. Regulatory intervention on control points

The section above identifies a number of potential control points that might enable operators to exercise significant market power in some NGN markets. It is therefore important for regulators to assess whether potential control points indeed create market power sufficient to warrant regulatory intervention.

If dominance over a given control point is achieved, a next step for the regulator is to consider whether some type of ex ante action is necessary. Under the new regulatory framework, such action can only be contemplated after having followed the procedure for definition of a new relevant market. The European Commission will play a strong role in this process through its recommendations on relevant markets as well as its power²² to veto markets defined at the national level.

Leaving the procedural issues aside, the assessment of whether to introduce an ex ante regulatory requirement can be extremely difficult because of the complexity of the NGN environment and because the consequences of regulatory action cannot be entirely foreseen. This is the basic and difficult challenge for regulators in the NGN environment where there are risks associated with any course of regulatory action or indeed inaction.

There are several good reasons why regulators should be reluctant to step in:

- search for success is a natural business objective that may lead to a strong but not necessarily dominant market position warranting regulatory action. Inappropriate intervention would alter the risk and reward calculations that drive investments and could create the impression that success will be punished by regulators. This could chill investment in the NGN sector and retard NGN development;

²² 2002/21/EC - Framework Directive – Art. 7.4

- the control point could be part of a new service which, in principle, should not be regulated²³;
- the business models of the service environment may not yet be stable and regulatory intervention could freeze commercial arrangements and market structures that are not efficient or viable in the longer term;
- inappropriate regulatory requirements would in affect mean that the regulator would pick winners and losers;
- the importance of the control point could fade away over time with new technology or other service alternatives. This means that the consequences of regulatory inaction over time could be less important than they appear at the outset;

There is thus significant risk that regulatory intervention could be counter-productive in the sense that the regulator would in effect be micro-managing the market instead of letting the market find its own solutions. Furthermore, the potential negative consequences of inaction could be remedied in time by other mitigating market developments. All of this suggest that regulators should avoid the temptation to intervene unless significant negative consequences of inaction are clearly foreseen.

On the other hand, it cannot be ruled out that dominance over certain control points could lead to serious barriers to market entry and thus justify ex ante regulation. In such extreme cases, lack of regulation could actually hamper development of the information society and the associated benefits. Ideally, there should be competition in as many domains as possible and market capture of a critical control point that can limit a range of competitive alternatives will require serious consideration.

The authors of this report, however, do not wish to identify any particular potential control point at this time as candidate for ex ante regulation, because this would be highly speculative in terms of how NGN will evolve.

5.4. Possible regulatory implications of the technical differences between CS and IP

It is widely expected that NGN will be based upon Internet Protocol (IP). This technology differs from Circuit Switching (CS) in a number of ways that have regulatory implications. Some of the major differences between CS and IP as communications networks are summarised below, along with the implications these differences may have for regulation.

5.4.1. SS7 signalling vs. IP addressing

The provision of advanced services in a CS network is based on an “Intelligent Network” framework, which is reached through the SS7 signalling network. Services in an IP network are based on servers which are accessible through normal IP addressing (IP addressing is described in section 3.5.2).

²³ 2002/21/EC - Framework Directive - Recital 27

Regulatory implications:

- There are more opportunities for open interfaces in IP networks. This provides a potential for wider participation and more competition in advanced communications services.
- Interoperability and open interfaces will become important issues as there will be many more forms of interconnection and access than in the CS environment.
- There is a potential for more geographic independence as a server can be located anywhere in the global IP network, whereas in the SS7 network only interconnected operators can have access.

5.4.2. Network centric service creation vs. network independent service creation

The creation of network services in a CS network is largely under the control of operators of the CS network. In an IP environment services can be created by integrating service components from various independent service providers distributed across networks and geography – independently of the underlying network infrastructure ("Distributed Service Creation") (described in section 3.4).

Regulatory implications:

- Physical presence can no longer be used to determine where a service is "located" in terms of networks and geography.
- It may be difficult to determine the country in which a service originates and thus to identify the applicable national law.
- Regulators will have to assess whether open interfaces to service components are necessary or if there will be sufficient interoperability being implemented through gateways that perform conversion between interfaces.

5.4.3. Switching vs. routing

Basic components of a CS network are the switches that perform many service- and management related functions in addition to establishing the transmission path from caller to receiver. The basic components in the IP environment are routers, which forward packets toward their destination, and servers which perform service and management functions. Servers can be addressable units in the network and communicate using the same protocols as user terminal equipment. There is therefore a potential for clearer separation between the transmission and control layers on the one hand and the service creation layer on the other, as functions in these layers are carried out by different types of equipment (described in sections 2.2 and 3.3.1).

Regulatory implications:

- In open IP networks (such as the Internet or global IP networks open to 3rd party service providers), there is the potential for more geographic independence between the user location and service creation. This could lead to more international competition for communications services, for example, through sophisticated forms for call-back services capable of dynamic optimisation based on real-time tariffs and currency rates.

- In many situations, it may be difficult to determine from which country a service is provided.

5.4.4. Clear vs. blurred boundaries

In CS networks there is a distinction between the network and the user equipment attached to it. Functions carried out in user equipment are normally not covered by telecommunications regulations beyond technical compliance requirements.

CS networks provide a fairly clear boundary between network and user terminal equipment, and intelligence is mainly centralised, meaning that calls are managed centrally by equipment belonging to the network. The boundary between network and terminal equipment is less clear in the IP environment, and intelligence is distributed between end-points and control devices within or at the edge of the network (described in section 3.6.3).

Regulatory implications:

- New criteria may be required for determining whether a given function belongs to the network and therefore may or may not be covered by its regulations.

5.4.5. Connection-oriented vs. connection-less

CS is connection oriented, meaning that a signalling phase takes place before the communication starts in order to trigger preparatory actions at the receiver end as well as along the transmission path. Signalling can also occur during the communication to adjust some of the parameters, and occurs at the end of the communication to release resources. In the lower layers, IP operates in a connection-less manner meaning that packets are sent out and received without any prior warning or preparations (routing and signalling in IP networks are described in sections 3.3.1 and 3.6 respectively).

Connection-less communications are better suited for network browsing and distributed service creation, which are part of a new service paradigm different from PSTN.

Regulatory implications:

- Connection-less communications do not have clearly defined start and end times. One consequence is that time based tariffs become unsuitable where cost orientation is a regulatory requirement.

5.4.6. Dedicated circuits vs. different routes

CS is based on the set-up of a dedicated circuit for the duration of a call, while IP will transmit packets using several different routes and infrastructure for a single communication (described in section 3.3.1).

Regulatory implications:

- In CS, time can be related directly to network usage while in IP networks this is not possible. Time based tariffs become unsuitable where cost orientation is a regulatory requirement.

5.4.7. Numbering vs. names and addresses

The CS numbering system is based on ITU Recommendation E.164 numbers. The numbering system is controlled jointly by ITU and national authorities.

The IP addressing system is based on IPv4 or IPv6 addresses in combination with names such as [name@domain](#). The IP addressing system is controlled by ICANN, which is a non-governmental international organisation (as described in section 3.5.2).

Regulatory implications:

- With the transition to IP, control over numbering resources is being transferred from governmental to non-governmental organisations.

5.5. Information society services

A comparison between CS and IP is not just a question of communications technologies. The corresponding networks can also be compared in terms of what types of communication they are intended for. NGN and IP based networks are production platforms and distribution mechanisms for information society services.

The development of communications networks can be described as going through three major phases:

1. CS systems were constructed years ago to set up a circuit between two parties to allow person-to-person voice communications.
2. IP in the Internet environment has enabled an effective environment for person-to-machine interaction. "Surfing" the net is a typical example of this capability.
3. The next level of communication is expected to include machine-to-machine communication across multiple technology platforms. One example may be intelligent agents searching for the best price for a defined product or service.

These developments have some obvious implications:

- It is significant for the overall growth potential of telecommunications. Person-to-person communication has limitations with regard to growth based on how long each person is prepared to spend talking to other people. Person-to-machine communication represents a significant additional growth potential, partly because each person may spend more time communicating and partly because a person can require much more bandwidth when communicating with a machine than with a person. But person-to-machine communications also has growth limitations represented by how much time each person is prepared to sit in front of a communication device. Machine-to-machine communication, however, represents a significant additional growth potential, which is not limited by human constraints.
- As communications migrate toward person-to-machine and machine-to-machine communications, regulators may have to turn their attention to higher protocol layers in order to ensure interoperability in the layers above the transmission layer.

The complexity of the interoperability requirements is further augmented by the fact that IP enables many alternative strategies and options for deployment of technologies and creation of services. The NGN concept, based on IP, provides tremendous flexibility in terms of options, where intelligence and functionality can be distributed in many different

ways among many different operators. Intelligence may be anywhere in the network as well as in the user equipment. For example, in IP a PBX function can be implemented in the traditional centralised piece of equipment on user premises, as a softswitch in a network server (a CENTREX type of solution), or as a distributed function in the terminals themselves. It is expected that NGN will create a communications market where different network implementation strategies will compete against each other to provide the same services.

These different approaches will be pursued by different types of actors. For example, a terminal based approach may be pursued by the terminal equipment manufacturers, while a softswitch option may be marketed by a network operator or a service provider (See also section 4.5.5 on Development of software and services).

The feasibility of potential implementation strategies will depend largely on the degree of openness with regard to necessary interfaces. Not surprisingly, there are conflicting views on which interfaces should be open and which may remain closed. In one absolutist view, interfaces between all functional layers and across all server functions including the control plane would be open to enable competition for a maximum number of separate service elements. This view would require a distinct separation between the transmission plane, the control plane and the service plane (see the NGN architectural concepts in section 2.2) and thus also between electronic communications networks and the services provided over them.

In a more traditional view, many interfaces will remain closed at least for some years. In the beginning this may be a necessity because only proprietary standards may be available and solutions for interoperability are yet to be implemented. Over time, when standards or solutions for interoperability have been developed, network operators, who own the control plane, may seek to carefully select which interfaces in addition to transmission they will choose to open and under which terms and conditions.

Within the NGN framework, it is foreseen that technical solutions, typically in the form of software driven proxies or gateways, can be found for all types of interoperability requirements for which there is market demand and commercial will on both sides of the point of interconnection.

5.6. Interconnection

5.6.1. Interconnection of networks

The Access Directive²⁴ defines interconnection as the “*logical and physical linking of public communications networks*”. It is therefore logical to consider what type of regulatory regime²⁵ to apply when NGN networks interconnect with each other and with other types of public communication networks. The interconnection “of public communications networks” seems to imply that all types of traffic are covered when two networks interconnect. Two different approaches can be considered:

1. Uniform regime for all kinds of traffic

²⁴ 2002/19/EC - Access Directive – Art. 2

²⁵ The term “interconnection regime” is not intended to imply a set of regulations. An “interconnection regime” may be unregulated or subject to a set of regulatory requirements.

If there were a set of regulatory requirements for an interconnection service such as call termination, and in particular if these included cost orientation requirements, one logical consequence could be to require that all services over an interconnected network to be terminated at the same cost because IP packets would cost the same regardless of their content.

If this approach were chosen there would be many interconnection issues. It would not only be necessary to interconnect CS and IP networks in a technical sense, but it would also require the interconnection of different business models. Essentially, the network for volume or time based PSTN services would have to interconnect with the network for flat rate or bandwidth related Internet business model. Taking this approach one step further, there could be significant transition problems where operators that benefit from voice termination revenues would object to a requirement to terminate IP voice packets under a general flat rate arrangement. On the other hand, a general volume based IP termination solution could jeopardise the Internet market.

2. Different regimes for different types of traffic

Another interpretation of a network related interconnection approach would be that an SMP designation for operators offering interconnection to a given network would cover all types of traffic, but that the regulators would have the option to differentiate interconnection regimes for different categories of services. This corresponds to current practice, where normal PSTN traffic is interconnected under a time based tariff regime, while interconnection to the Internet through PSTN in some countries is based of a flat monthly rate²⁶. Another example is mobile network interconnection where the main focus is on call termination of voice traffic, but where some regulators have used the designation of network interconnection to also regulate SMS interconnection charges.

This approach allows more flexibility for regulators compared with the uniform regime described above, but there is still the risk that new and innovative services would be subject to interconnection regulations when there is an SMP designation for operators offering network interconnection even if this were contrary to the intention of the Framework Directive²⁷.

5.6.2. Interconnection markets

With the new regulatory framework, the type of interconnection obligations that can be imposed will depend on what is established as relevant wholesale markets for interconnection.

In contrast to the more uniformly structured PSTN offerings, the variety of service offerings expected over NGN may complicate efforts to define and designate markets which may be subject to ex ante regulation. If the market definitions are too broad, many different types of services could be forced into the same interconnection regime. It will be necessary to characterise potential markets carefully (and sufficiently narrow) in terms of their service characteristics and to recognise that many of the services may represent new offerings with the presumption that ex ante conditions should only be imposed under exceptional circumstances.

²⁶ E.g. FRIACO (Flat Rate Internet Access Call Origination) in the UK.

²⁷ 2002/21/EC - Framework Directive - Recital 27

NGN will be a general-purpose network where a “telephony service” may be seen as one particular option along with Internet, broadcasting and other categories of service. Different network operators are likely to have different views on how such services should be made available and to what extent the services offered should be under central control. Some may choose a fairly open approach comparable to the Internet as it is currently known. They could offer a “converged access solution” with telephony, Internet and video services under the same subscription. Early implementations of such converged subscriptions are already available on the market. Others may offer a more limited service with superior quality of service and security for voice communication and a selection of Internet type services, but with constraints on outside Internet connectivity depending on the service package selected by the subscriber.

Within this wide range of possibilities, there will be many different service markets. For example, NGN service markets could include:

- the market for telephony services, characterised by the same quality and functionality as for PSTN;
- the market for Internet, characterised by best effort IP packet transmission without any quality of service guarantees;
- the market for video reception services, characterised by a quality of service and bandwidth enabling reception of high quality video and sound;
- the market for video conferencing, characterised by a quality of service and bandwidth enabling two-way simultaneous transmission of high quality video and sound.

Each of these broad service markets could be further divided into sub-markets for interconnection purposes and each could be regulated differently. It makes more sense to accept different types of network interconnection agreements for different market categories defined by their service characteristics rather than by network technology²⁸. Some, but not all of these agreements may be subject to significant market power conditions, depending on whether there is dominance in the corresponding market for interconnection.

5.6.2.1. Interconnection of telephony services

Network interconnection arrangements for telephony services in an NGN environment could if necessary retain many of the principles upon which the present reference interconnection offerings are based. There would, however, be certain additional considerations such as which party should perform the necessary protocol conversion between CS and IP in the transition stage and how to interconnect the signalling functions.

A new potential problem would be to maintain voice grade quality of service over the interconnecting IP networks. This can be achieved either through over-engineering the networks so that there is always sufficient capacity for the telephony traffic, or through some form of managed “quality of service” solution.

²⁸ This has already been recommended by the Commission in its Recommendation on Relevant Markets, where “Voice call termination on individual mobile networks” is one of the relevant wholesale markets.

In the latter case, it would mean that the interconnection nodes would have to distinguish between different categories of service and their corresponding quality of service requirements. It would also mean that the providers would have to map these requirements against their own quality of service implementations and traffic management systems. Bearing in mind that end-to-end quality of service is no better than that of the weakest link, this would obviously be an important problem to solve.

Although there are different approaches to achieving a required quality of service in a packet switched environment, there is no consensus yet for a given approach or standards for different quality of service categories. Quality of service in this environment is often about different probabilities of transmission delay and packet loss instead of the traditional noise and blocking problems associated with CS communications. The actual level of service achieved is therefore likely to differ between different networks. While there is some confidence that an acceptable quality of service can be achieved in a single network with a common architecture and single overall network management co-ordination, it is more difficult to achieve quality of service over interconnected networks through the exchange of quality of service parameters (see also section 4.4.5 on Quality of Service and Service level agreements).

From a regulatory point of view quality of service raises the following questions:

- Will interconnection agreements have to include service level agreements? If so, what are the required parameters?
- Is there a need to ensure that the quality of service provided by a dominant operator does not discriminate between operators?

Lack of a good solution for quality of service interoperability across networks would generally work in favour of the bigger networks because it would mean that good quality of service can be assured for on-net but not for off-net communications.

Another challenge for regulators will be to reconsider the cost-orientation requirements for call termination in view of the efficiency gains expected from transition to IP telephony. Furthermore, the organisation of an IP based network, regardless of whether the packets contain voice or not, is likely to be quite different from that of traditional telephony in a hierarchical network with local exchanges, transit exchanges and international exchanges. This may be another reason to revise the structure of cost oriented call termination fees.

5.6.2.2. Interconnection of Internet

“Wholesale broadband access” is recommended by the European Commission²⁹ as a relevant market for significant market power. This does not include the Internet market itself in terms of interconnection. So far these types of interconnection agreement have been concluded by commercial negotiation without the involvement of regulators.

This discussion of Internet interconnection is intended to indicate some of the difficulties or at least new challenges that may arise if steps are taken to apply the 2003 acquis to the Internet. It should not be understood as a recommendation for regulation of this market.

²⁹ Commission Recommendation of 11/02/2003 - C(2003)497.

The basic interconnection arrangement between Internet service providers today is “peering” with traffic exchanged between “peers” and terminated without financial settlements in “sender keeps all” arrangements. In addition, transit interconnection arrangements may be offered whereby backbone network operators carry traffic to be terminated in a third operator’s network. Not all operators are “peers” in the sense of being equal in terms of traffic flow, geographic coverage and other possible criteria. Basic peering is therefore supplemented by many other types of arrangements³⁰, which often combine elements of peering with payments from the smaller to the larger network operator.

The regulatory challenge would be to ensure that interconnection with a potentially dominant operator in a relevant Internet service market would satisfy requirements for non-discrimination. This could require the establishment of criteria for non-discriminatory arrangements, for example, based on peering.

Another consideration is that the Internet business model is probably not yet stable in the sense of providing sustainable and fair payments to all participants in the value chain, including content providers. Many business models are likely to be tried out in the market, with different dependency on termination charges. Dissuasive termination conditions may prevent certain business models from being tested. In this context, premature interconnection regulations could have the effect of regulating one part of the value chain and prevent the market from finding a workable solution (see also section 4.3.1.3 on Content and service providers).

5.6.2.3. Interconnection of other services

Telephony is today the main service category (in addition to transmission and access services) that is regulated with ex ante conditions for interconnection. The other main service category, Internet, is not regulated in this way.

Other types of services, such as SMS, MMS, WAP and mobile Internet access, are being offered by mobile operators. As NGN will be the production platform and distribution system for information society services, it is likely that also other types of services will appear, both in the fixed and the mobile environments. Many of these will combine elements of transmission and content. Such services may require new billing models with different formulas for sharing risks and rewards.

Billing models that are viable in the longer term are more likely to develop as a result of market forces than from regulatory requirements. Premature interconnection regulations could interfere in this process and prevent alternative billing models from being tested in the market.

This view is in agreement with the presumption in the new regulatory framework³¹ that new and emerging services should not be subject to inappropriate regulations. The conclusion is to avoid broad and sweeping categories for defining markets for the purpose of regulating interconnection and instead target narrowly only those specific service markets where it is evident that dominance could create permanent bottlenecks.

³⁰ See WIK-Consult study on The Economics of IP Networks – Market, Technical and Public Policy Issues Relating to Internet Traffic Exchange

³¹ 2002/21/EC - Framework Directive – Recital 27

6. Concluding recommendations

The concluding recommendations presented below highlight several of the points covered in the report. They do not represent an exhaustive list of implications drawn from the analysis, but are offered primarily to supplement the main purpose of the study, which is to raise the awareness of regulators to NGN developments and to stimulate debate and discussion on policy issues and regulatory options that may make sense.

6.1. Control Points

Essentially, NGN has the potential of providing a more open and competitive service environment. Paradoxically, this potential may also lead to additional sources of market power. Whereas traditionally, all elements of service creation were controlled by a single operator, the NGN environment enables many of these elements to be provided competitively. Where such elements have to be chained together in order to create an end-user service, it follows that control over any single element in the chain would provide control over the whole chain.

The potential control points that have been identified in this report indicate that the battle for market power will be fought on many different fronts. For regulators seeking to ensure a fair and well functioning market, it is of obvious interest to understand these mechanisms.

The regulatory framework for electronic communications and its (prospective) application will play a significant role in the development of the information society. "Playing a significant role" should not be interpreted however to mean that overt regulatory action will be required in all cases. Leaving the market alone is also a regulatory option, which in many of the cases that were identified would probably be appropriate.

To this end, the Framework Directive³², warns against applying ex ante regulation in newly emerging markets, such as new service markets within the NGN environment. Regulators must avoid creating the impression that success will be "punished" by regulation. At the same time, the new framework does not preclude the possibility of applying ex ante regulation to potential "control points" identified in NGN if they were to meet certain criteria. These criteria are set out in the Commission Recommendation on relevant product and service markets:

In identifying markets in accordance with competition law principles, recourse should be had to the following three criteria. The first criterion is the presence of high and non-transitory entry barriers whether of structural, legal or regulatory nature. However, given the dynamic character and functioning of electronic communications markets, possibilities to overcome barriers within a relevant time horizon have also to be taken into consideration when carrying out a prospective analysis to identify the relevant markets for possible ex ante regulation. Therefore the second criterion admits only those markets the structure of which does not tend towards effective competition within the relevant time horizon. The application of this criterion involves

³² 2002/21/EC - Framework Directive – Recital 27

*examining the state of competition behind the barriers of entry. The third criterion is that application of competition law alone would not adequately address the market failure(s) concerned.*³³

It is important for the European Commission as well as for national regulatory authorities to understand where new sources of market power or control points can be found. This will require an understanding of new technologies and how they interrelate. The technical sections in this report have been developed with this objective in mind. A long list of potential control points or sources of market power has been identified, but it is premature to assume that they will emerge with NGN in a sense that would require regulatory action.

With the shift of competition from the lower transport and network layers to the higher layers of services and applications, the European Commission should carefully consider the role played by customer information in NGN, including location data, and whether access to and use of this information should be controlled by the entity compiling it or the customer to whom it pertains. Further study on Customer Identity Systems (CIS) is recommended to better understand the extent to which the development of information society services and its market structure will depend on non-discriminatory access to customer information gathered by entities with significant market power and to what extent CIS could be used as a control point.

In addition, further study into the behavioural patterns in the marketplace is recommended with a view to creating awareness among authorities and to analyse the risks of leveraging a position of dominance into control over the delivery of future multimedia services.

The complexity of dealing with control points in an NGN environment from a regulatory perspective suggests that there is a need to develop a special type of expertise for analysis and risk assessment of regulatory action and inaction. Such expertise requires familiarity with NGN technologies, market structures and service opportunities, an understanding of investment decision making by commercial enterprises, economics and the business mind-set. It also requires knowledge of the application of competition law, the new framework for electronic communications and other relevant EU legislation. Since the European Commission has a primary role in the procedures leading to the definition of a control point as a relevant market for ex ante regulation, it must also ensure that the required expertise is available. This is likely to involve several EU authorities, including competition authorities and suggests a particular development programme, perhaps to pool expert resources and to build new circles of competence.

6.2. Interconnection

NGN will provide a converged approach where services traditionally carried over different networks and subject to different interconnection regimes (including unregulated regimes) will in the future share the same network. To define interconnection obligations for an NGN network would seem to mean that regulations could affect all services carried over it.

³³ COMMISSION RECOMMENDATION of 11/02/2003 On Relevant Product and Service Markets within the electronic communications sector susceptible to ex ante regulation in accordance with Directive 2002/21/EC of the European Parliament and of the Council on a common regulatory framework for electronic communication networks and services – Recital 9

Services that were previously carried over separate networks may represent different service options within NGN. Therefore, it would be more flexible to allow different network interconnection regimes to co-exist within the same network. For example, interconnection of the traditional voice circuit switched services as is currently provided in PSTN could, if necessary, continue to be regulated as before, while the interconnection of Internet services could continue to be unregulated even if the two service categories share the same network.

It is important to recognise that stable business models for information society services, where each participant in the value chain receives sustainable revenues, have not yet developed and that interconnection requirements, particularly if they involve price regulation, could therefore be premature because they could prevent certain business models from being tested in the market.

In practical terms, this recommendation affects the definition of relevant markets for the purpose of ex ante regulation. It suggests that as NGN develops, the wholesale markets for interconnection should be narrowly defined in terms of the services requiring interconnection, rather than broadly defined in terms of networks.

6.3. Other implications

As indicated above, this study has focused primarily on the regulatory implications with regard to competition. It is clear that there are important consequences also in other areas that deserve further attention.

In addition the study on Customer Identity systems suggested in section 6.1, the following are candidates for further study at the European level:

1. Data protection and privacy

Efforts to protect data and safeguard privacy will face even greater challenges in the NGN marketplace, where users will leave behind long, content-rich data trails with both communications operators and content providers.

These issues should be subject to further study with a view to better understanding the privacy exposures associated with NGN, and to investigate whether additional steps should be taken by policy makers or regulators.

2. Universal service

The ongoing development of NGN and information society services warrants periodic reviews of the scope of universal service, where any proposed expansions of services covered should be based on stringent criteria of common use and convincing justification based on social inclusion.

These reviews should, however, take a broad view of the barriers that may prevent information society services to be available to all citizens and to consider steps to remove them. The concept of affordability that includes low usage tariff options and mechanisms to control expenditure for telephony services may be possible to extend to a much broader range of services.

Appendix A. List of acronyms

3G	Third Generation
3GPP	Third Generation Partnership Project
AAA	Authentication, Authorisation and Accounting
ADSL	Asymmetric Digital Subscriber Line
AF	Assured Forwarding
APEX	APplication EXchange
API	Application Programming Interface
APNIC	Asia Pacific Network Information Centre
ARIN	American Registry for Internet Numbers
AS	Autonomous System
ASCII	American Standard Code for Information Interchange
ASP	Application Service Provider
ATM	Asynchronous Transfer Mode
B-GAN	Broadband Global Area Network
BICC	Bearer Independent Call Control
CDMA	Code Division Multiple Access 2000
CIDR	Classless Inter Domain Routing
CLI	Calling Line Identification
CLR	Common Language Runtime
CORBA	Common Object Request Broker Architecture
CoS	Class of Service
CPIM	Common Presence and Instance Messaging
CPL	Call Processing Language
CPS	Carrier Pre-Selection
CS1 / CS2	Capability Set 1 / Capability Set 2
CS	Circuit-Switched
DAB	Digital Audio Broadcasting
DCOM	Distributed Component Object Model
DECT	Digital Enhanced Cordless Telecommunications
DiffServ	Differentiated Services
DNS	Domain Name System
DSL	Digital Subscriber Line
DTTV	Digital Terrestrial TeleVision
DVB	Digital Video Broadcasting
DVB-RCS	Digital Video Broadcasting – Return Channel for Satellite
DWDM	Dense Wavelength Division Multiplexing
EDI	Electronic Data Interchange
EF	Expedited Forwarding
EGP	Exterior Gateway routing Protocol
ENUM	tElephone Number Mapping

ETSI	European Telecommunications Standards Institute
FRIACO	Flat Rate Internet Access Call Origination
FTTH	Fibre To The Home
FWA	Fixed Wireless Access
GMPLS	Generalised Multi-Protocol Label Switching
GPRS	General Packet Radio Service
GRX	GPRS Roaming eXchange
GSM	Global System for Mobile communications
GUI	Graphical User Interface
HDSL	High-bit-rate Digital Subscriber Line
HFC	Hybrid Fibre Coaxial
HLR	Home Location Register
HTML	Hyper Text Markup Language
HTTP	Hyper Text Transfer Protocol
IAB	Internet Architecture Board
IANA	Internet Assigned Numbers Authority
ICANN	Internet Corporation for Assigned Names and Numbers
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IGP	Interior Gateway routing Protocol
IM	Instant Messaging
IMPP	Instant Messaging and Presence Protocol
IMS	IP Multimedia Subsystem
IMSI	International Mobile Subscriber Identity
IN	Intelligent Network
INAP	Intelligent Network Application Protocol
IP	Internet Protocol
IPv4	IP version 4
IPv6	IP version 6
IPR	Intellectual Property Rights
ISDN	Integrated Service Digital Network
ISP	Internet Service Provider
ISUP	ISDN User Part
ISV	Independent Software Vendor
ITU	International Telecommunication Union
ITU-T	Telecommunication standardisation branch of the ITU
J2EE	Java 2 Enterprise Edition
JAIN	Java Advanced Intelligent Networks / Java APIs for Intelligent Networks
JRE	Java Runtime Environment
LAN	Local Area Network
LLU	Local Loop Unbundling
MAC	Medium Access Control
MAN	Metropolitan Area Network
MAP	Mobile Application Part

MCC	Mobile Country Code
MCU	Multipoint Control Unit
MEGACO	MEdia GAteway COntrol
MGCP	Media Gateway Control Protocol
MHP	Multimedia Home Platform
MMS	Multimedia Message Service
MNC	Mobile network Code
MPEG 2	Motion Picture Experts Group 2
MPLS	Multi-Protocol Label Switching
MS	Mobile Station
MSIN	Mobile Subscriber Identification Number
NAI	Network Access Identifier
NAT	Network Address Translation
NGN	Next Generation Network
NRA	National Regulatory Authority
NS	Name Server
OASIS	Organisation for the Advancement of Structured Information Standards
OMA	Open Mobile Alliance
OMG	Object Management Group
OSA	Open Service Architecture
OSI	Open System Interconnection
OSS	Operations Support System
PAN	Personal Area Network
PBX	Private Branch eXchange
PDP	Packet data Protocol
PIN	Personal Identification Number
PRIM	PResence and Instant Messaging
PS	Packet-Switched
PSTN	Public Switched Telecommunications Network
QoS	Quality of Service
RADIUS	Remote Authentication Dial In User Service
RAS	Registration, Admission and Status protocol
RFC	Request For Comment
RIPE NCC	Réseaux IP Européens, Network Coordination centre
RIR	Regional Internet Registries
RSVP	Resource Reservation Protocol
RTP	Real-time Transport Protocol
RTSP	Real Time Streaming Protocol
RTTE	Radio and Telecommunications Terminal Equipment
SAP	Session Announcement Protocol
SCN	Switched Circuit Network
SCP	Service Control Point
SDH	Synchronous Digital Hierarchy
SDP	Session Description Protocol

SDSL	Symmetric Digital Subscriber Line
SIG	Special Interest Group
SIM	Subscriber Identity Module
SIMPLE	SIP for Instant Messaging and Presence Leveraging Extensions
SIP	Session Initiation Protocol
SLA	Service Level Agreement
SMP	Significant Market Power
SMS	Short Message Service
SOAP	Simple Object Access Protocol
SS7	Signalling System #7
SSP	Service Switching Point
STP	Signal Transfer Point
TCAP	Transaction Capabilities Application Part
TCP	Transport Control Protocol
TIPHON	Telecommunications and Internet Protocol Harmonisation Over Networks
ToS	Type of Service
UA	User Agent
UAC	UA Client
UAS	UA Server
UDDI	Universal Description, Discovery and Integration
UMTS	Universal Mobile Telecommunications System
URI	Universal Resource Identifier
URL	Universal Resource Locator
UTRAN	UMTS Terrestrial Radio Access Network
UWB	Ultra Wide Band
VDSL	Very-high-rate Digital Subscriber Line
VHE	Virtual Home Environment
VoIP	Voice over IP
VPN	Virtual Private Network
WAN	Wide Area Network
WAP	Wireless Application Protocol
WECA	Wireless Ethernet Compatibility Alliance
Wi-Fi	Wireless Fidelity
WISP	Wireless ISP
WSDL	Web Service Description Language
WLAN	Wireless LAN
WLL	Wireless Local Loop
XML	eXtensible Markup Language