2

Evolution from GSM to UMTS

Evolution is one of the most common terms used in the context of UMTS. Generally it is understood to mean the technical evolution, i.e. how and what kind of equipment and in which order they are brought to the existing network if any. This is partly true but in order to understand the impact of the evolution, a broader context needs to be examined. Evolution as a high-level context covers not only the technical evolution of network elements but also expansions to network architecture and services. When these three evolution types are going hand in hand the smooth migration from 2G to 3G will be successful and generate revenue (Figure 2.1).

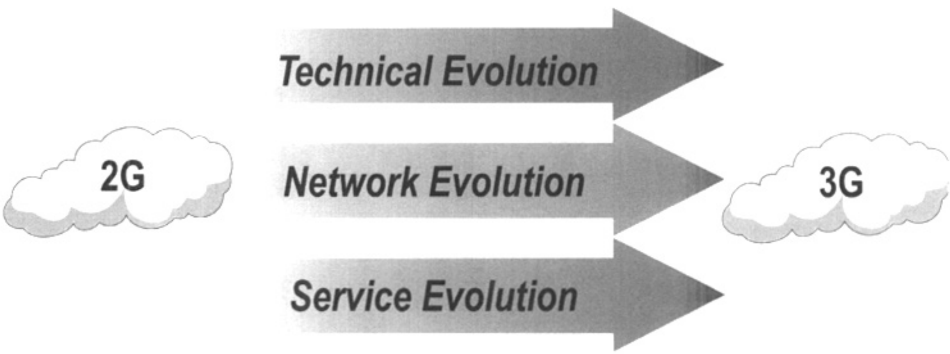

Figure 2.1. Evolution types

Technical evolution means the development path of how network elements will be implemented and with which technology. This is a very straightforward development and follows strictly the general, common technology development trends. Because network elements together form a network, the network will evolve accordingly, in theory. In this phase one should bear in mind that a network is as strong as its weakest element and due to the open interfaces defined in the specifications many networks are combinations having equipment provided by many vendors. Technical evolution may proceed however with different paces in association with different vendor's equipment and when adapting evolution-type changes between several vendors' equipment the result may not be as good as expected.

Service evolution is not such a straightforward issue. It is based on demands generated by the end-users and these demands could be real or imagined; sometimes network operators and equipment manufacturers offer services way beyond subscriber expectations. If the end-users' needs and operators' service palette do not match each other, difficulties with cellular business can be expected.

The main idea behind the GSM specifications was to define several open interfaces, which determine the standardised components of the GSM system. Because of this interface openness, the operator maintaining the network may obtain different components of the network from different GSM network suppliers. Also, when an interface is open it defines strictly how system functions are proceeding at the interface and this in turn determines which functions are left to be implemented internally by the network elements on both sides of the interface.

As was experienced when operating analogue mobile networks, the centralised intelligence generated a lot of load in the system, thus decreasing the overall system performance. That is why the GSM specification in principle provided the means to distribute intelligence throughout the network. The above-mentioned interfaces are defined into places where their implementation is both natural and technically reasonable.

From the GSM network point of view, this decentralised intelligence is implemented by dividing the whole network into four separate subsystems being Network Subsystem (NSS), Base Station Subsystem (BSS), Network Management Subsystem (NMS) and Mobile Station (MS) (Figure 2.2).

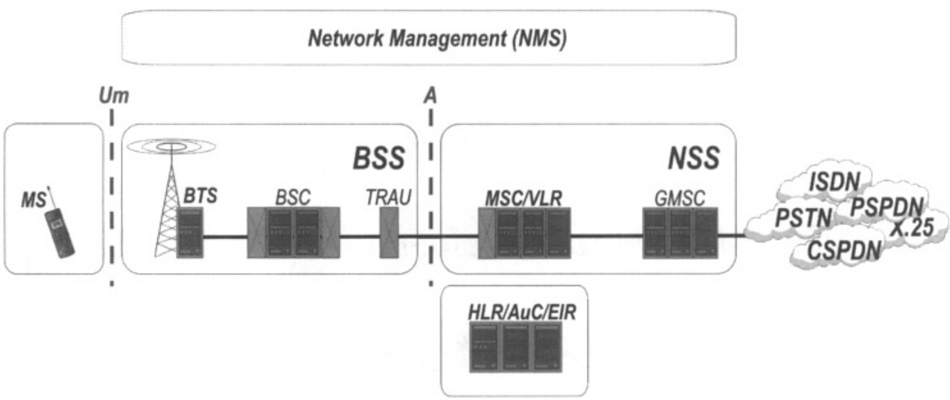

Figure 2.2. Basic GSM network and its subsystems

The actual network needed for call establishing is composed of the NSS, the BSS and the MS. The BSS is a network part responsible for radio path control. Every call is connected through the BSS. The NSS is a network part taking care of call control functions. Every call is always connected by and through the NSS. The NMS is the operation and maintenance related part of the network. It is also needed for the whole network control. The network operator observes and maintains the quality and services of the network through the NMS. The open interfaces in this concept are located between the MS and BSS (Um interface) and between the BSS and NSS (A interface). The interface between the NMS and the NSS/BSS was expected to be open, but its specifications were not ready in time and this is why every manufacturer implements NMS interfaces with their own proprietary methods.

The MS is a combination of terminal equipment and a subscriber's service identity module. The terminal equipment as such is called Mobile Equipment (ME) and the subscriber's data is stored in a separate module called the Service Identity Module (SIM). Hence, ME + SIM = MS.

The Base Station Controller (BSC) is the central network element of the BSS and it controls the radio network. This means that the following functions are the BSC's main responsibility areas: maintaining radio connections towards the MS and terrestrial connections towards the NSS. The Base Transceiver Station (BTS) is a network element maintaining the air interface (Um interface). It takes care of air interface signalling, ciphering and speech processing. In this context, speech processing means all the methods BTS performs in order to guarantee an error-free connection between the MS and the BTS. The Transcoding and Rate Adaptation Unit (TRAU) is a BSS element taking care of speech transcoding, i.e. it is capable of converting speech from one digital coding format to another and vice versa.

The Mobile Services Switching Centre (MSC) is the main element of the NSS from the call control point of view. MSC is responsible for call control, BSS control functions, interworking functions, charging, statistics and interface signalling towards BSS and interfacing with the external networks (PSTN/ISDN/packet data networks). Functionally the MSC is split into two parts, though these parts could be in the same hardware. The serving MSC/VLR is the element maintaining the BSS connections, mobility management and interworking. The Gateway MSC (GMSC) is the element participating in mobility management, communication management and connections to the other networks. The Home Location Register (HLR) is the place where all the subscriber information is stored permanently. The HLR also provides a known, fixed location for the subscriber-specific routing information. The main functions of the HLR are subscriber data and service handling, statistics and mobility management. The Visitor Location Register (VLR) provides a *local* store for all the variables and functions needed to handle calls to and from mobile subscribers in the area related to the VLR. Subscriber related information remains in the VLR as long as the mobile subscriber visits the area. The main functions of the VLR are subscriber data and service handling and mobility management. The Authentication Centre (AuC) and Equipment Identity Register (EIR) are NSS network elements taking care of security-related issues. The AuC maintains subscriber identity-related security information together with the VLR. The EIR maintains mobile equipment identity (hardware) related security information together with the VLR.

When thinking of the services, the most remarkable difference between 1G and 2G is the presence of the data transfer possibility; basic GSM offers 9.6 kb/s symmetric data connection between the network and the terminal. The service palette of the basic GSM is directly adopted from Narrowband ISDN (N-ISDN) and then modified to be suitable for mobile network purposes. This idea is visible throughout the GSM implementation; for example, many message flows and interface handling are adapted copies of corresponding N-ISDN procedures.

The very natural step to develop the basic GSM was to add service nodes and service centres on top of the existing network infrastructure. The GSM specifications define some interfaces for this purpose, but the internal implementation of the service centres and nodes are not the subject of those specifications. The common name for these service centres and nodes is Value Added Service (VAS) Platforms and this term describes quite well the main point of adding these equipment to the network (Figure 2.3).

The minimum VAS platform contains typically two pieces of equipment; Short Message Service Centre (SMSC) and Voice Mail System (VMS). Technically speaking the VAS platform equipment is relatively simple and meant to provide a certain type of service. They use standard interfaces towards the GSM network and may or may not have external interfaces towards other network(s).

From the service evolution point of view, VAS is the very first step in generating revenue

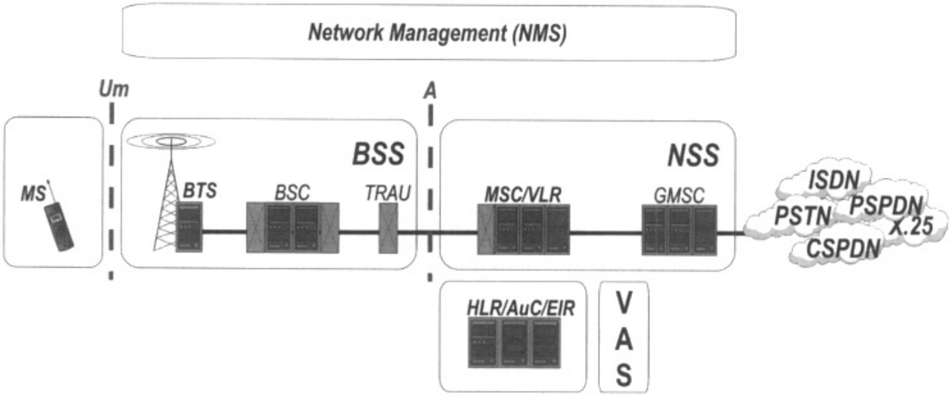

Figure 2.3. Value added service platform

with services and partially tailoring them. The great successor in this sense has been the SMS, which was originally planned to be a small add-in in the GSM system. Nowadays it has become extremely popular among GSM subscribers.

Basic GSM and VAS are basically intended to produce "mass services for mass people" but due to requirements raised from end-users, a more individual type of services is required. To make this possible, the Intelligent Network (IN) concept (Figure 2.4) was integrated together with the GSM network. Technically this means major changes in switching network elements in order to add the IN functionality and, in addition, the IN platform itself is a relatively complex entity. IN enables service evolution to take big steps towards individuality and also with IN the operator is able to perform more secure business, for example, pre-paid subscriptions are mostly implemented with IN technology.

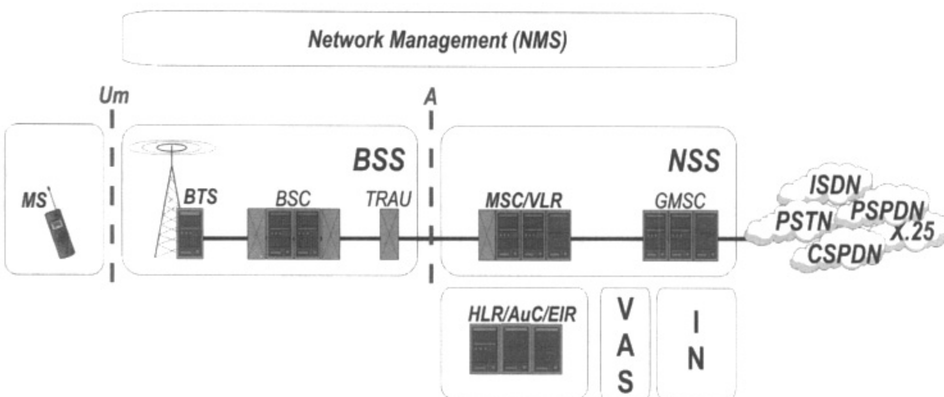

Figure 2.4. Intelligent network

In the beginning, GSM subscribers have used the 9.6 kb/s circuit switched symmetric "pipe" for data transfer. Due to the Internet and electronic messaging the pressures for mobile data transfer have increased a lot and this development was maybe underestimated at the time when the GSM system was specified. To ease this situation, a couple of enhancements have been introduced. Firstly, the channel coding is optimised. By doing this the effective bit rate has increased from 9.6 kb/s up to ≈14 kb/s. Secondly, to put more data through the air interface, several traffic channels can be used instead of one. This arrangement is called High Speed Circuit Switched Data (HSCSD) (Figure 2.5). In an optimal environment a HSCSD user may reach data transfer with 40–50 kb/s data rates. Technically this solution is quite straightforward but unfortunately it wastes resources and some end-users may not be happy with the pricing policy of this facility; the use of HSCSD very much depends on the price the operators set for its use. Another issue is the fact that most of the data traffic is asymmetric in nature, i.e. typically a very low data rate is used from the terminal to network direction (uplink) and higher data rates are used in the opposite direction (downlink).

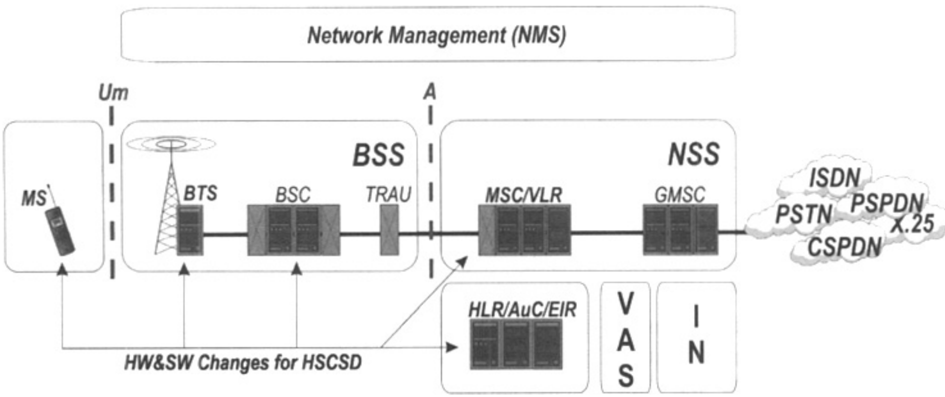

Figure 2.5. Channel coding and HSCSD effects

The circuit switched symmetric Um interface is not the best possible access media for data connections. When also taking into account that the great majority of the data traffic is packet switched in nature, something more had to be done in order to "upgrade" the GSM network to make it more suitable for more effective data transfer. The way to do this is General Packet Radio Service (GPRS) (Figure 2.6). GPRS requires two additional mobile network specific service nodes: Serving GPRS Support Node (SGSN) and Gateway GPRS Support Node (GGSN). By using these nodes the MS is able to form a packet switched connection through the GSM network to an external packet data network (the Internet).

GPRS has the possibility to use asymmetric connections when required and thus the network resources are utilised better. GPRS is a step bringing IP mobility and the Internet *closer* to the cellular subscriber but it is not a complete IP mobility solution. From the service point of view, GPRS starts a development path where more and more traditional circuit switched services are converted to be used over GPRS because those services were originally more suitable for packet switched connections. One example of this is Wireless Application Protocol (WAP), the potential of which is to be discovered when using GPRS.

When packet switched connections are used, the Quality of Service (QoS) is a very essential issue. In principle the GPRS supports the QoS concept but in practise it does not. The reason here is that GPRS traffic is always second priority traffic in the GSM network: it

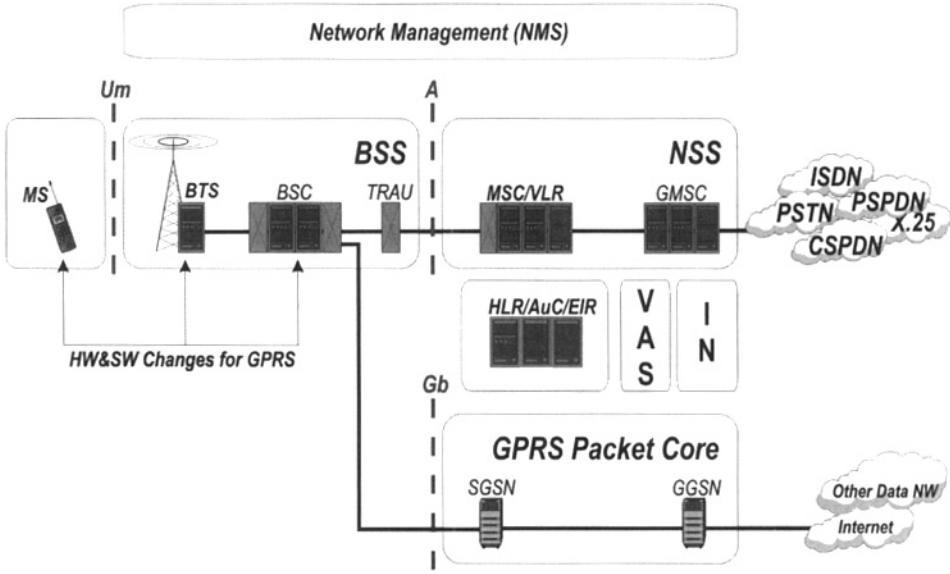

Figure 2.6. General packet radio service (GPRS)

uses otherwise unused resources in the Um interface. Because the amount of unused resources is not exactly known in advance, no one can guarantee a certain bandwidth for the GPRS continuously and thus the QoS cannot be guaranteed either.

By applying a completely new air interface modulation technique, Octagonal Phase Shift Keying (8-PSK), where one air interface symbol carries a combination of three information bits, the bit rate in the air interface can be remarkably increased. When this is combined together with very sophisticated channel coding technique(s), one is able to achieve a data rate of 48 kb/s compared to conventional GSM which can carry 9.6 kb/s per channel and one information bit is one symbol in the air interface. These technical enhancements are called Enhanced Data Rates for Global/GSM Evolution (EDGE) (Figure 2.7).

Development of the EDGE concept is divided into two phases, EDGE Phase1 and EDGE Phase2. EDGE Phase1 is also known as E-GPRS (Enhanced GPRS). Also the BSS is renamed as E-RAN (EDGE Radio Access Network). EDGE Phase1 defines channel coding and modulation methods that enable up to 384 kb/s data rates for packet switched traffic under certain conditions. The assumption here is that one GPRS terminal gets eight air interface time slots for one connection, thus 8×48 kb/s $= 384$ kb/s. In addition, the EDGE terminal must be close to the BTS in order to use a higher channel coding rate. EDGE Phase2 contains guidelines on how this same speed is achieved for circuit switched services. EDGE Phase2 is also commercially known as E-HSCSD.

From a network evolution point of view, EDGE in general has its pros and cons. A good point is the data rate(s) achieved; these are almost equal to UMTS urban coverage requirements. The disadvantage with EDGE is that the data rates offered are not necessarily available throughout the cell. If EDGE is to be offered with complete coverage, the amount of cells will remarkably increase. In other words, EDGE may be expensive solution in some cases. The future of EDGE is still to be seen since it has to compete with the true 3G solutions.

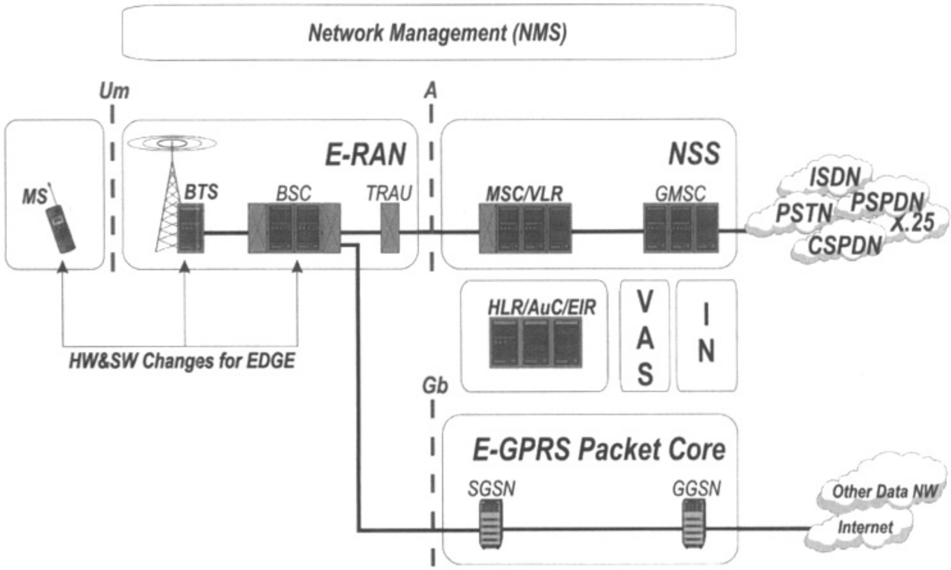

Figure 2.7. Enhanced data rates for global evolution

Figure 2.8 represents a scenario how a 3G network is implemented according to 3GPP R99 specifications.

3G introduces the new radio access method, WCDMA. WCDMA and its variants are global, hence all 3G networks should be able to accept access by any 3G network subscriber. In addition to globality, WCDMA has been thoroughly studied in laboratory premises and it has been realised that it has better spectral efficiency than TDMA (in certain conditions) and it

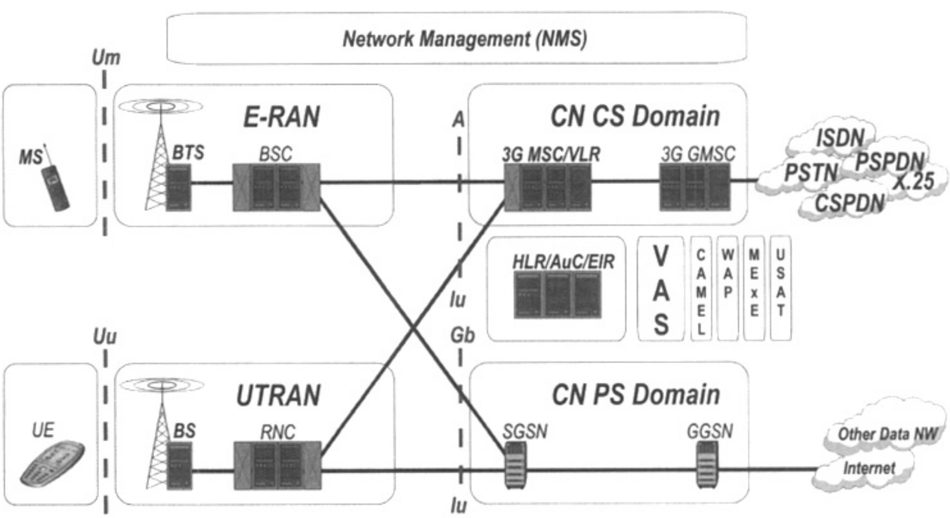

Figure 2.8. 3G network scenario (3GPP R99)

is more suitable for packet transfer than TDMA based radio access. WCDMA and radio access equipment as such are not compatible with GSM equipment, and this is why adding the WCDMA to the network one must add new elements: Radio Network Controller (RNC) and Base Station (BS).

On the other hand, one of the key requirements for UMTS is GSM/UMTS interoperability. One example of interoperability is inter-system handover, where the radio access changes from GSM to WCDMA and vice versa during the transaction. This interoperability is taken care of by two arrangements. First, the GSM air interface is modified so that it is able to broadcast system information about the WCDMA radio network in the downlink direction. Naturally the WCDMA radio access network is able to broadcast system information about surrounding the GSM network in the downlink direction, too. Second, to minimise the implementation costs, the 3GPP specifications introduce possibilities to arrange inter working functionality with which the evolved 2G MSC/VLR becomes able to handle the wideband radio access, UTRAN.

So far, the abilities provided by the IN platform have been enough from the service point of view. The concept of IN is directly adopted from the PSTN/ISDN networks and thus it has some deficiencies as far as the mobile use is concerned. The major problem with standard IN is that the IN as such is not able to transfer service information between networks. In other words, if a subscriber uses IN based services they work well but only within his/her home network. This situation can be handled by using ''evolved IN'' called Customised Applications for Mobile network Enhanced Logic (CAMEL). CAMEL is able to transfer service information between networks. Later on, the role of CAMEL will increase a lot in 3G implementation; actually almost every transaction performed through the 3G network will experience CAMEL involvement at least to some extent.

Transmission connections within the WCDMA radio access network are implemented by using ATM on top of a physical transmission media (3GPP R99 implementation). A pre-standardisation project FRAMES (1996–1998) discussed a lot whether to use ATM in the network or not. The final conclusion was to use ATM because of two reasons:

- ATM cell size and its payload are relatively small. The advantage here is that the need of information buffering decreases. If buffering a lot, expected delays will easily increase and also the static load in the buffering equipment will increase. One should bear in mind that buffering and thus generated delays have a negative impact on the QoS requirements of real-time traffic.
- The other alternative, IP, and its version IPv4 was also considered but IPv4 has some serious drawbacks, being limited in its addressing space and missing QoS. On the other hand, ATM and its bit rate classes match very well with QoS requirements. This leads to the conclusion that where ATM and IP are combined (for packet traffic), IP is used on top of ATM. This solution combines the good points of both protocols: IP qualifies the connections with the other networks and ATM takes care of the connection quality and also routing. Due to IPv4 drawbacks a compromise has been made. Certain elements of the network use fixed IPv4 types of addresses but the real end-user traffic uses dynamically allocated IPv6 addresses, which are valid within the 3G network. To adapt the 3G network to the other networks in this case, the 3G IP backbone network must contain an IPv4 ↔ IPv6 address conversion facility, because the external networks may not necessarily support IPv6.

The core network nodes are evolved technically, too. The CS domain elements are able to handle both 2G and 3G subscribers. This requires changes in MSC/VLR and HLR/AC/EIR. For example, security mechanisms during the connection set-up are different in 2G and 3G and now these CS domain elements must be able to handle both of them. The PS domain is actually an evolved GPRS system. Though the names of the elements here are the same as that in 2G, their functionality is not. The most remarkable changes concern the SGSN, whose functionality is very different from that in 2G. In 2G, the SGSN is mainly responsible for Mobility Management activities for a packet connection. In 3G, the Mobility Management entity is divided between the RNC and SGSN. This means that every cell change the subscriber does in UTRAN is not necessary visible to the PS domain, but RNC handles these situations.

The 3G network implemented according to 3GPP R99 offers the same services as that of GSMPhase2+. This is, all the same supplementary services are available, teleservices and bearer services have different implementation but this is not visible to the subscriber; a speech call is still a speech call, no matter whether it is done through a traffic channel (GSM) or by using 3G bandwidth. In addition to GSM, the 3G network in this phase may offer some other services not available in GSM, for example, video call could be one of those. In this phase the majority of services are moved/transferred/converted to PS domain whenever reasonable and applicable. WAP is one of those candidates, because the nature of the information transferred in WAP is packet switched. The PS domain is taken into effective use and one service branch containing a variety of different services will be location based services utilising the subscriber location mechanisms built into the 3G network. The new positioning methods and related network elements are introduced in Chapter 7.

The development steps after 3GPP R99 are somewhat unclear in the level of details, but some major trends are visible. The main trends in the following development steps are separation of connection, its control and services and, at the same time, the conversion of the network to be completely IP based. From the service evolution point of view, these development steps also recognise that multimedia services should be provided by the 3G network itself. Multimedia means a service where at least two media components are combined, for example, voice and picture.

These trends are big issues as such and this is why they are implemented in phases, the first phase being 3GPP R4. The 3GPP R4 implementation introduces separation of connection, its control and services for the CN CS domain.

In the CN CS domain actual user data flow goes through Media Gateways (MGW), which are elements maintaining the connection and performing switching functions when required. The whole process is controlled by a separate element evolved from MSC/VLR called MSC server. One MSC server can handle numerous MGWs and thus the CN CS domain is freely scalable; when one wishes to add control capacity, a MSC server is added. When one desires to add switching capacity, MGWs are added.

When this kind of network has been set up, the pace of technology development and specifications set the next limit. The more momentum the IPv6 gains, the more of the 3G network connections that can be converted to IPv6, too. This decreases the need of IPv4 ↔ IPv6 conversions. In this phase the traffic relationship between circuit and packet switched will remarkably change. The majority of the traffic is packet switched and also some traditionally circuit switched services, like for example speech, will become at least partially packet switched (VoIP, Voice Over IP). For example, a conventional GSM call is actually changed to a VoIP call in the MGW where the BSS is connected. There are many ways to implement VoIP calls but the new CN subsystem called IMS (IP Multimedia Subsystem) is added here since it will offer uniform methods to perform VoIP calls. In addition to this the IMS is used for IP based multimedia services. Naturally the BSS part of the network could be implemented using IP but the time schedules of this change are unclear. The role of CAMEL will change, too. Because many of the services using CAMEL are converted from the circuit switched side to the packet switched side of the network, the CAMEL will now have connections also to the PS domain elements. In addition to this the CAMEL will be connecting elements between the service platforms and the network. This issue is explained in more detail in Chapter 7.

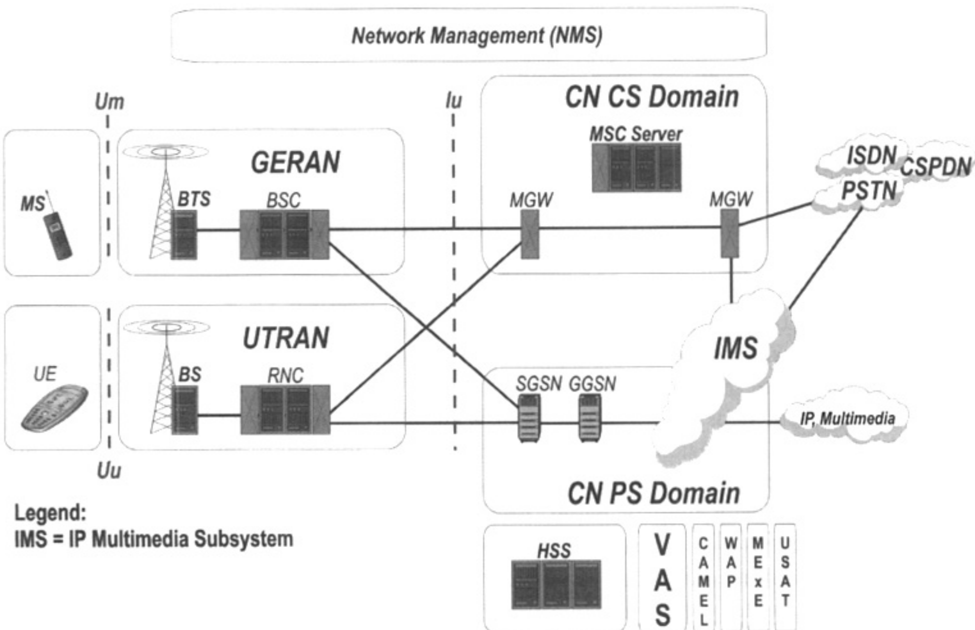

Figure 2.9. 3GPP R4 implementation scenario

In 3GPP R5 the evolution continues further and all traffic coming from UTRAN is supposed to be IP based. If we think of a voice call from UE to PSTN as an example, it is transported through UTRAN as packages and from the GGSN the VoIP call is routed to the PSTN via IMS, which provides required conversion functions.

From the UE point of view the network always ''looks'' the same in the development phases illustrated in Figures 2.8–2.10. Inside the network almost everything changes. The major change will happen in transport technology, which in 3GPP R99 implementation is ATM. The 3GPP R4 and R5 implementation scenarios aim to swap ATM for IP. Because the system must be backward compatible, the operator always has a choice whether to use ATM or IP as the transport technology or whether the solution contain both technologies. As explained earlier, the strength of ATM is its support for QoS at least in the beginning. As time goes by, the IP as a technology will contain QoS mechanisms implemented over various kinds of subnetworks, not only ATM.

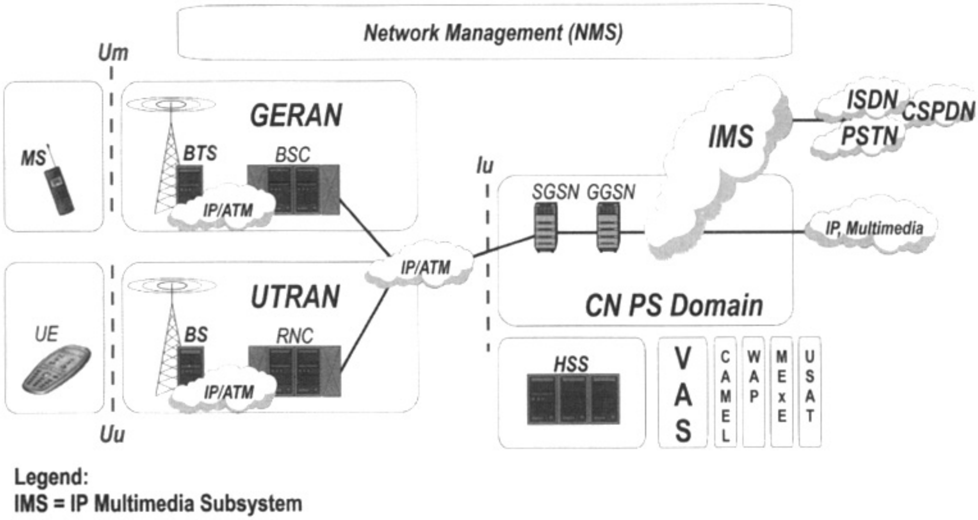

Figure 2.10. Vision of 3GPP R5 (All IP)

In this phase the services and use of the network are more important than technology itself and due to this the used radio access technology may become less important. The main selection criterion for the used radio access technology is to offer enough bandwidth for the service used. The future vision here is that the 3G core network has interfaces for several radio access technologies, for instance, GSM, EDGE, CDMA2000, WCDMA and Wireless Local Area Network (WLAN). Naturally this sets many requirements for terminal manufacturers and the terminals capable of handling different kinds of access technology combinations will be introduced according to the market needs. 3G is now a fixed part of life and the network offers services and connections which have traditionally used other media. The 3G terminal becomes a combination of personal belongings, like a phone, purse, ID card and passport. In other words, it is very difficult to state where the computer ends and phone starts.